ON THE PARAMETRIC AND NONPARAMETRIC PREDICTION METHODS FOR ELECTRICITY LOAD FORECASTING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ESRA ERİŞEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

SEPTEMBER 2013

Approval of the thesis:

ON THE PARAMETRIC AND NONPARAMETRIC PREDICTION METHODS FOR ELECTRICITY LOAD FORECASTING

Submitted by **ESRA ERİŞEN** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Murat Köksalan
Head of Department, **Indusrial Engineering** ________________

Assist. Prof. Dr. Cem İyigün
Supervisor, **Industrial Engineering Dept., METU** ________________

Examining Committee Members:

Prof. Dr. Gülser Köksal
Industrial Engineering Dept., METU ________________

Assist. Prof. Dr. Cem İyigün
Industrial Engineering Dept., METU ________________

Prof. Dr. İnci Batmaz
Statistics Dept., METU ________________

Assist. Prof. Dr. Fehmi Tanrısever
Industrial Engineering & Innovation Sciences Dept., Eindhoven University of Technology of the Netherlands ________________

Assist. Prof. Dr. Özgen Karaer
Industrial Engineering Dept., METU ________________

Date: ……….

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Surname: ESRA ERİŞEN
Signature:

ABSTRACT

ON THE PARAMETRIC AND NONPARAMETRIC PREDICTION METHODS FOR ELECTRICITY LOAD FORECASTING

Erişen, Esra

M.Sc., Department of Industrial Engineering
Supervisor: Assist. Prof. Dr. Cem İyigün

September 2013, 91 pages

Accurate electricity load forecasting is of great importance in deregulated electricity markets. Market participants can reap significant financial benefits by improving their electricity load forecasts. Electricity load exhibits a complex time series structure with nonlinear relationships between the variables. Hence, new models with higher capabilities to capture such nonlinear relationships need to be developed and tested. In this thesis, we present a parametric and a nonparametric method for short-term and long-term load forecasting, and we compare the performance of these models for different lead times ranging from one hour to one week. These methods include a modified version of Holt Winters Double Seasonal Exponential Smoothing (m-HWT) we present and a Nonlinear Autoregressive with Exogenous Inputs (NARX) neural network model. Using hourly load data from the Dutch electricity grid, an extensive empirical study is carried out for five different provinces. In the second part of the study, NARX is applied to long-term load forecasting in one Dutch province.

Our results indicate that NARX clearly outperforms m-HWT in one-hour ahead forecasting. Additionally, our modification to HWT leads to significant improvement in model accuracy especially on special days. Despite its simplicity, m-HWT outperformed NARX for 6 and 12-hours ahead forecasts in general. However NARX performs better in 24-hours, 48-hours and 1 week ahead forecasting. In addition, NARX performs superior to m-HWT in terms of maximum error and on short special days.

Computational results also indicate that with a well-trained closed loop NARX neural network model, electricity load can be forecasted successfully one and a half years ahead for hourly intervals. NARX can successfully capture nonlinear effects of special days and temperature. NARX has brought a performance improvement of 30% in terms of mean absolute percent error (MAPE) compared to the existing methodology; time shifting.

Keywords: electricity load forecasting, exponential smoothing, neural networks, NARX, HWT

ÖZ

ELEKTRİK YÜK TAHMİNLEMESİ ÜZERİNE PARAMETRİK VE PARAMETRİK OLMAYAN TAHMİN METOTLARI

Erişen, Esra

Yüksek Lisans, Endüstri Mühendisliği Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Cem İyigün

Eylül 2013, 91 sayfa

Hükümet denetimi kaldırılmış elektrik piyasalarında doğru yük tahminlemesi büyük önem taşır. Piyasa katılımcıları daha doğru elektrik yük tahminlemeleri sayesinde büyük finansal kazanımlar elde edebilir. Elektrik yük verisi, değişkenler arasında lineer olmayan ilişkiler ile birlikte kompleks zaman serileri yapısına sahiptir. Bu nedenle söz edilen lineer olmayan ilişkileri yansıtabilecek daha geniş kapasiteli yeni modellere ihtiyaç vardır. Bu tezde, kısa vadeli yük tahminlemesi için parametrik ve parametrik olmayan metotlar sunulmuştur ve model performansları 1 saat ile 1 hafta arasında değişen farklı tedarik süreleri için karşılaştırılmıştır. Karşılaştırılan metotlar, tarafımızdan modifiye edilmiş Holt Winters Üssel Yumuşatma (m-HWT) ve bir Dışsal Girdiler ile Lineer Olmayan Sinir Ağları (NARX) modelini kapsamaktadır. Hollandalı bir şebeke dağıtıcısının saatlik verisi kullanılarak, beş farklı bölgeyi içeren, kapsamlı bir deneysel çalışma yapılmıştır. Umut vadeden sonuçlar, bizi çalışmanın ikinci yarısında, NARX metodolojisini bir Hollanda bölgesi için uzun vadeli yük tahminleri türetmekte kullanmaya yöneltmiştir.

Sonuçlar göstermektedir ki, NARX 1 saat sonrası tahminlemede, m-HWT metoduna açık bir şekilde üstün olmuştur. Ek olarak, sunduğumuz HWT modifikasyonu, özellikle özel günlerdeki model doğruluğunda, kayda değer bir gelişme sağlamıştır. Sadeliğine rağmen, m-HWT modeli 6 ve 12 saat sonrası tahminlemesinde genelde NARX'e üstün gelmiştir. Ancak NARX performansı 24, 48 ve 1 hafta sonrası tahminlemede daha iyidir. Ek olarak, NARX maksimum hata ve kısa özel günlerde yük tahminleme açısından üstün bir performans göstermiştir.

Deneysel sonuçlar ayrıca göstermektedir ki, iyi eğitilmiş bir kapalı halka NARX sinir ağları modeli kullanılarak, elektrik yükü 1,5 yıl sonrası için saatlik olarak başarılı bir şekilde tahminlenebilir. NARX'in özel günlerin ve sıcaklığın yük üzerindeki lineer olmayan etkilerini başarılı bir şekilde yakaladığı görülmüştür. Mevcut model olan, zaman kaydırma modeline göre NARX %30 tahminleme doğruluğunu arttırmıştır.

Anahtar kelimeler: Elektrik yük tahminlemesi, üssel yumuşatma, yapay sinir ağları, NARX, HWT

To My Dear Family

ACKNOWLEDGEMENTS

I have received valuable support and motivation by many people throughout this masters thesis study.

First of all, I would like to express my sincere gratitude to my academic supervisor, Dr. Cem İyigün, Middle East Technical University, who deserves special recognition because of his wholehearted support, understanding, patience and help. I would not have considered a graduate study without his motivation and encouragement. He has truly made a difference in my life. He guided me not only in my graduate study but also in my whole life. With all his guidance and the great working environment he provided; he became more of a mentor, friend, than a supervisor. I would like to also thank my other supervisor, Dr. Fehmi Tanrısever, Eindhoven Technology University, for his interest in my work, valuable advices and support throughout my study. Without their guidance and persistent help this dissertation would not have been possible.

Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Gülser Köksal, Prof. İnci Batmaz and Dr. Özgen Karaer for their valuable comments and reviews.

I would also like to thank Eindhoven University of Technology for the financial support and Dr. Tarkan Tan for his valuable help in making me a part of the double degree program. I am also grateful to The Scientific and Technological Research Council of Turkey (TÜBİTAK) Scholarship for making this thesis possible.

I would like to express my deepest appreciation to Enexis B.V. for providing the opportunity of conducting my research project in the company. I really enjoyed being a part of Enexis, thanks to Reporting, Analysis and Quality Market Players Department team. A special gratitude I give to Guy van der Boom and Maikel Jonkers for being great mentors throughout the project. Their support and trust had been great motivation for me.

Furthermore, I would like to thank my dearest friends, Lina Baykan, Esen Kümbül, Sine Taymaz and Hasan Kartarı for their continuous support and motivation. They were always there, even at the hardest moments, easing the stress. I would also like to thank Gizem Sezen, Nicolas Lenoble, Simge Kaplan, Pablo Lezama Elguero, Efstathios Dimarelis and Nadine Gruchot. My experience in Eindhoven would not have been filled with such great memories without them.

Last but not least, I would like to thank my mother and father for their endless love and support during my whole life. Their trust in me that I can get through even the hardest things has been a great motivation for me. A special word of thanks also goes to my dear brother Eren Erişen, who inspired, encouraged and guided me in every trial that came my way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Energy security issues have been discussed all over the world during the last decade as the natural resources have been diminishing rapidly. In order to secure effective processing of energy markets, there have been efforts to convert energy markets into more free structures. The deregulation in Dutch electricity market is an example of conversions to free markets. Electricity load forecasting is very important in electricity utilities' processes because load forecasts are used in fundamental decisions in operations such as generation, infrastructure development, energy purchasing, maintenance, etc. Deregulation in electricity markets made load forecasting even more important due to the increased competition, generators, distributors, all means of suppliers need to provide higher level of service in order to endure. Unlike other commodities, electricity cannot be stored and has to be available on demand. There is no available stock, no customer inventory, no backordering, etc. to meet customer demand on time. Therefore, generators, grid operators and regulators need to project electricity demand with a very high accuracy and always keep the system in balance.

The prerequisite of efficient management of power systems is the generation of accurate load forecasts for both long and short-term planning [1, 2]. Electricity demand forecast horizon is divided mainly into three: **short-term**, **mid-term** and **long-term** forecasts. Short-term forecasts refer to forecasts between one minute to several weeks, where mid-term forecasts are one day to one year forecasts and long-term forecasts refer the forecasts for longer than one year lead times. **Short-term** forecasts are generally used in daily operations such as clearing electricity transactions, scheduling generation capacity, load flow analysis, etc. [4]. On the other hand, **long-term** forecasts are used in capital planning, in new generation and transmission capacity decisions [5]. In this study we mainly focus on short-term and long-term load forecasting.

Short-term demand forecasting helps the market participants to schedule transactions in real time [6]. After the deregulation in the market, the participants have developed the ability to adjust capacity and demand through short-term power transactions. In addition to the deregulation, technological improvements have also increased the ability to adjust capacity flexibly and meet the demand [7]. In this highly competitive setting, electricity load forecasting has become extremely important to provide uninterrupted, reliable, secure and economic supply of electricity [8]. For example from suppliers' point of view, they need to make sure having the necessary reserve to supply their consumers; in order to avoid facing very high costs. On the other hand, in the case of over prediction; i.e. keeping excess reserves, can end up with fines, or undervaluation of the available electricity. Hence both cases cost money to the market parties. Kyriakides & Polycarpou [4] summarizes functions of short-term load forecasting in:

(i) Actions such as negotiations of bilateral contracts between utilities and regional transmission operators
(ii) Studies such as economic dispatch, unit commitment, hydro-thermal coordination, load flow analysis and security studies
(iii) Operations such as scheduling of committing or decommiting generating units and increasing or decreasing the power generation.

On the other hand, long-term electricity load forecasting assists suppliers and grid operators in capital planning, in new generation and transmission capacity decisions [5]. Due to high investment costs, market parties try to avoid overestimation, which will lead to excess power facilities [9]. However underestimation means problems in supplying electricity; causes customer dissatisfaction and falling behind in the competition.

Efficient scheduling of electricity transactions is necessary to provide an economic and reliable supply of electricity, and is a key business competency of the firms in electricity markets [8]. Electricity load data includes both daily and weekly cycles. In addition, multiple factors, such as the **season** and **time of the day**, have complex and nonlinear effects on the electricity load [10]. These multiple seasonality and complex **nonlinear relationships** make it difficult to model electricity load with traditional regression models. In addition, as a confirmation of Kim's [11] claim on **idiosyncratic load patterns** on **special days**; during the data analysis, we observed that consumer behavior significantly diverge from the regular pattern on special days, namely national days, religious days, public holidays, school holidays, etc. He stated that on these days, because most of the businesses do not function, load is lower; however only effect is not only the decrease in the level, but also different pattern throughout the day.

In the literature, despite many studies are carried out on load forecasting there is no agreement on which method to apply for different lead times. In this study, a comparison of a parametric and a non-parametric model's performances are presented on electricity load data in short-term. We have derived forecasts for short-term electricity load using modification of Holt Winters Exponential Smoothing (HWT) and compared its performance to Nonlinear Autoregressive with eXogenous Inputs (NARX) Neural Networks. We utilized NARX also in long-term load forecasting. Long-term load forecasting results are compared to the performance of a regression model in a similar study [5] and to the existing model's performance, which is based on time shifting method. Considering effects of special days on electricity load, we relied NARX especially on capturing complex effects of special days. Kim [11] stated that in most of the studies special days are modeled by using dummy variables, which results as a failure in capturing intraday effects of the special days. He states that such models are unable to capture hourly effects, instead just reduces the level of consumption. By utilizing NARX's ability to capture complex effects of multiple variables on the output; we expect to obtain higher performance on special days' load forecasting.

In addition to the computational study, we believe another important contribution of the study is our modification to Holt Winters Exponential Smoothing (HWT) method. HWT is a commonly used methodology in short-term load forecasting, particularly due to its simplicity. We have modified a variation of HWT, enabling it to include effects of special days in electricity infeed forecasting. The **contribution** of this study is fourfold:

- **Modification of HWT that enables modeling special days.** We develop a modified version of Taylor's [6] exponential smoothing adaptation to consider for the impact of special days in short-term electricity load forecasting. The proposed model resulted in a dramatic improvement in forecasting accuracy especially on special days.
- **Proposition of NARX for short-term load forecasting.** Mostly conventional and classical approaches are adopted for short-term load forecasting, whereas computational intelligence based and fuzzy logic methods are mainly used for longer-term forecasts. There are only a few papers, which examine the effectiveness of NARX models for short-term load forecasting. However, none of these studies apply NARX to more than a few data sets and/or for few different lead times. In that respect, this study provides a very comprehensive analysis of NARX method for short-term electricity load forecasting applying to five different data sets and generating forecasts for six different lead times.
- **Proposition of NARX for long-term load forecasting.** There are only few studies applying NARX to electricity load data in the literature. Especially in long-term forecasting, there are no other papers investigating performance of NARX. Therefore for the first time – to our best knowledge- we applied NARX for hourly long-term electricity load forecasting.
- **A comparison of parametric and nonparametric methods.** A widespread empirical study is carried out in order to compare artificial methods to more conventional methods**.**

This thesis has mainly six sections. In Chapter 2, literature research on electricity load forecasting is presented. In Chapter 3, motivation behind this study and a case study are explained. Chapter 4 details proposed methods used in this study. Computational experiments are presented in Chapter 5 and Chapter 6 concludes the study with a conclusion section.

CHAPTER 2

LITERATURE SURVEY

In 80s electricity market deregulations were first started with privatization and some other market oriented reforms in Chile. During the 80s and 90s many other countries followed the reforms in order to build a competitive, free market that is expelled from any monopolistic power and external interventions. The aim of the deregulations was ensuring the highest customer benefit by increasing competition in the market. Former electricity market consisted of a few market players for every region, giving these players monopolistic power, as the customers had to receive electricity from market parties in that region. On the other hand, deregulated market structure allows customers to make deliberate decisions about their supplier and aims providing perfect market information to the customers to be used in their decisions. In 1998, Dutch electricity market deregulation activities were started. Turkish electricity market deregulation has taken place later in 2001. Since then, Turkish electricity market has been developing in the same direction as Dutch electricity market.

Efficient scheduling of electricity transactions is necessary to provide an economic and reliable supply of electricity, and is a key business competency of the firms in electricity markets. Efficient scheduling of electricity transactions can be achieved by accurate forecasting of electricity load. Due to the mentioned importance, which increased especially after the market deregulation in 1998, many studies have been carried out on electricity demand forecasting. In electricity market demand forecasting, two decisions should be taken very carefully, first, **the lead times of the forecast**, second, **time interval of the forecast**. These two decisions may affect both the methodology selection and the variables included in the model. Lead times of forecasts heavily depend on parties responsible from different operations on the electricity supply chain. For example, short-term models are mostly used in daily operations such as, decisions on electricity transactions, scheduling generator operation times, load flow analysis, etc. [4]. On the other hand, mid-term forecasts are mainly used to plan maintenance schedules, infrastructure adjustments, fuel purchases, etc. [12]. Lastly, long-term electricity demand forecasts are useful in capital planning, in new generation and transmission capacity decisions [5].

As mentioned, time interval of the forecasts is also very important. Unavoidably, there is a strong relationship between lead time and time interval of the forecast. In short-term forecasting, appropriate time intervals vary between quarter hour, half hour and one hour. On the contrary to considerably small intervals in short-term, in longer-term forecasts time intervals can vary from several days to one year. For example, among many studies on short-term forecasting in literature, Huang & Shih [13] focused on daily forecasting, where Taylor's [1] focus is half hourly forecasts and Soares & Medeiros [14] generated hourly forecasts. On the other hand Mirasgedis et al. [12] derived monthly predictions up to 12 months ahead and Von Hirschhausen & Andres [15] derived yearly forecasts for next decade in China.

Most of the papers in the literature are on short-term electricity load forecasting whereas number of papers on midterm and long-term forecasting also increases. Methods used in these papers vary from the simplest conventional methods to complicated and very recently introduced methods, such as fuzzy logic models, to a large extent. Hahn et al. [16] divide forecasting methods mainly into two categories: classical time series and regression methods and artificial intelligence and computational intelligence methods. However hybrid models can also be stated as a separate class of forecasting methods.

2.1. Classical time series and regression methods

Regression models are commonly used in electricity demand forecasting especially due to their ability to relate external variables to electricity load [16] There are many external variables that are believed to be effective on electricity load such as **calendar variables**, **meteorological variables** and **economic variables**. Relating them to load carries great importance particularly in midterm and long-term load forecasting. Although, linear regression models are dominant in the literature, effects of these variables are mostly nonlinear or complex to capture with linear models. However, easiness of application, understanding of relationship between input and output variables help these models keep their favorable position in the literature.

Al-Hamadi & Soliman [17] decomposed electricity load into many simple linear models and used these models in order to forecast weekly average load profiles for 24 hours of day up to several weeks to few years. They measured accuracy over weekly average daily load and obtained mean absolute percent error (MAPE) values lower than 3.8% throughout one year. Papalexopoulos & Hesterberg [18] presented a linear regression model building process which includes parameter estimation under heteroskedasticity by using weighted least squares and diminishing the effects of potential errors on forecasts by 'reverse errors-in-variables'. Different than many other studies, they studied not load shape forecasting, but peak load forecasting, by utilizing Pacific Gas and Electric Company. Ramanathan et al. [7] derived multiple linear regression models for short-term (up to 88 hours) hourly load forecasting. They derived one model for every hour of the day. In order to capture different patterns at the weekends, weekend load is modeled separately. By incorporating exponential smoothing of the forecast errors an adaptive version of the model is derived. The adjustment enabled them to compensate for systematic errors.

Bianco et al. [19] and Mohamed & Bodger [19] also used multiple linear regression modeling, however for long-term load forecasting; yearly forecasts for several years. They both utilized economic and demographic variables and derived countrywide forecasts for Italy and New Zealand respectively. Mirasgedis et al. [12] aimed capturing nonlinear effects of external variables by building logarithmic regression models with an autoregressive component to generate daily and monthly forecasts for 12 months. They applied the models to the historical data of Greek interconnected power system from 1993 to 2002.

Time series models can be mainly divided into univariate and multivariate models. Univariate models are mostly preferred for short-term load forecasting. A very simple class of time series prediction, autoregressive moving average (ARMA) models are used in Cancelo et al. [3] by dividing data into its components for generating daily and hourly forecasts several days ahead. The models were built for Spanish system operator using 2006 data. It is stated that daily forecasts are used for weekly network outage plan and lead times vary between 4 days to 10 days. The results showed that large errors are mostly sourced from weather prediction and special days' effects on load. In the meantime, hourly forecasts are used for determining the final dispatch schedule for the next day. Highest errors were observed at the weekends, however, their model performed better than two other hybrid benchmark models. Hagan & Behr [21] stated that a simple polynomial regression analysis combined with a Box and Jenkins transfer function model can result in more accurate forecasts. Additionally, they applied nonlinear transformation to the temperature variable. The models are built by using data of moderately sized southwestern utility with 450,000 customers. The results showed that nonlinear extension model has outperformed the other three methods in hourly forecasting. Goh & Choi [22] showed that introducing a stochastic component to time series models can be used in justifying variations that cannot be attributed to a time element. They derived day ahead forecasting for half hour intervals by using four-year electricity demand data. Taylor [1] modified Holt Winters exponential smoothing formulation to accommodate two seasonalities at the same time for up to day ahead forecasting with half hour intervals. He stated that results of the empirical study carried out on England and Wales load data outperformed not only traditional Holt Winters method, but also multiplicative double seasonal ARIMA model. Taylor favored time series methods in his many other works studies ([2, 24, 23]). In these studies different alterations of exponential smoothing methods are compared to generate short-term electricity load forecasts.

Soares & Medeiros [14] also included seasonality of electricity load series in their model through a two level seasonal autoregressive (AR) model for short-term load forecasting. They used eight-year data, which was obtained from a utility company from Rio de Janeiro, Brazil. There exists also some other modifications of time series models in the literature, such as Huang & Shih's [12] non Gaussian ARMA model and Huang et al.'s [25] particle swarm optimization to identify the autoregressive moving average with exogenous variable (ARMAX) model for short-term load forecasting.

2.2. Computational Intelligence Based Models

Recently, computational intelligence based models have been prevalent in the literature for electricity load forecasting. Al-Saba and El-Amin [26] developed an Artificial Neural Network (ANN) model for peak-load forecasting and compared the results to AR models. For the comparison, data of a Saudi Arabian utility, that provides services to large industrial, commercial and residential loads, was used. Results showed that ANNs provide accurate results in long-term electricity load forecasting. One of the first studies considering nonlinear autoregressive (NAR) neural network models for electricity load forecasting was introduced by Connor, et al. [27]. They compared a nonlinear autoregressive model to a recurrent nonlinear autoregressive moving average and a feed-forward nonlinear autoregressive models utilizing synthetic data on the Puget Power Electric Demand time series. They emphasize the importance of input configuration while presenting the superior performance of the recurrent networks. Efforts followed by investigation of NARX and nonlinear autoregressive moving average with exogenous variables (NARMAX) performances in short-term load forecasting [28]. The paper is mostly concerned with constructing scheme for moving average (MA) part. Espinoza, et al. [29] carried out a Kernel based NARX model identification study for lead times of one hour and 24 hours. Using electricity load data from student apartments Varghese & Ashok [30] compared performances of a feed-forward back propagation neural network, a NARX network and a radial basis function model. The only paper practicing NARX models in long-term forecasting – to our best knowledge, belongs to Awan, et al. [31]. They used thirty-year data of National Transmission and Dispatch Company of Pakistan. In their paper, NARX based feed-forward neural, support vector regression and neural network methods are compared.

Studies comparing the methods have been performed to illustrate whether it is worth dealing with more complex methods like Artificial Neural Networks (ANNs) or more conventional methods perform well enough. Abraham & Nath [32] is one of them. They compared three methods: Box Jenkins ARIMA model, a feed-forward ANN model and evolving fuzzy neural network (EFuNN) in terms of their performances in forecasting 2 days ahead half hourly forecasting. They carried out comparison on energy demand data for 10 months period in the State of Victoria. In order to make the results independent of data sample, they replicated training sample three times. Results indicate that EFuNN is superior to two other variations of ANNs and ARIMA model. Taylor et al. [33] also carried out a study on comparison of time series and computational intelligence based models. They studied 5 different models: ARIMA, exponential smoothing, ANN, components regression with PCA and seasonal version of random walk for up to one day ahead forecasting. The most unexpected result of the empirical studies, which was carried out on thirty-weeks load data, was the poor performance of ANNs. Tzafestas & Tzafestas [34] have presented a detailed study on computational intelligence based methods in their paper, which includes comparison of neural networks, fuzzy logic methods, genetic algorithms and chaos models. They also derived models that are combination of some of these methods for hourly load forecasting up to 7 days ahead.

2.3. Hybrid Models

Hybrid methods form a rising class of forecasting methods in load forecasting. Aim of hybrid models is modeling the base data with a comparably simple model. Then, assuming that the errors source form complex and nonlinear relationships existing in the data set, errors are modeled with a more complex method, such as computational intelligence models. González-Romera et al. [34] applied ANN and Fourier series methods for mid to long-term forecasting. For the Spanish monthly electric demand data they decreased forecast errors to values lower than 2%. Similarly, Amjady & Keynia [36] developed a hybrid model of neural networks and evolutionary algorithm for monthly load forecasting and tested their model on European Network on Intelligent Technologies (EUNITE) test data and Iran's load data. They derived forecasts for daily peak load for next month. Promising results put them on the track for developing an optimization method for adjustable parameters and preforecast structures. Desouky & Elkateb [37] developed another hybrid model that includes ANNs. They investigated possible ways of performance improvement for mid-term peak load forecasting. They had a huge data set consisting of seven years data for training and the model performances are compared using two years load data of Jeddah city. Azadeh et al. [38] applied another hybrid approach which includes preprocessing data with time series using moving average method and fed into a multilayer perceptron (MLP) network.

In this thesis NARX networks is adopted as a forecasting method, relying on neural networks ability of capturing nonlinear and complex relationships between variables and NARX networks ability of capturing time series structure in the data. Exogenous inputs part of NARX enables inclusion of external variables to the model. One of the most effective external variables on electricity load is special days. In Section 5.1.1.4 large effects of these days are explained in detail. In the literature there are not many studies focusing on effects of special days on electricity load, despite their clear effects. Kim [11] stated that in most of the studies special days are modeled by using dummy variables, which results as a failure in capturing intraday effects of the special days. He states that such models are unable to capture hourly effects, instead just reduces the level of consumption. He claims that intraday patterns of the special days tend to differ significantly from regular days. Therefore in [11] special days are represented with two variables and divided into two types not only in terms of special days but also in terms of hours of the special days. Another motivation in adopting this methodology is that despite promising results in the existing few literature studies, there does not exist deeper studies on effectiveness of NARX models in electricity load forecasting.

CHAPTER 3

MOTIVATION

In this chapter, the motivation behind this study, the project conducted is introduced. We have mentioned the gap in NARX modeling in electricity load forecasting in Section 2.2., which motivated us to study on this topic. Additionally, the company project has been another motivation in this study. The data is provided by a semipublic company in the Netherlands.

Rest of the chapter is organized as follows. Section 3.1 covers brief information about the company profile. Section 3.2. is for project background. In Section 3.3 a brief introduction about Dutch Electricity Market is presented in order to provide an understanding to the reader about how our proposed models can help in functioning of the market.

3.1. Company Profile

The company is one of the largest Distribution Network Operators (DNO), operating in six provinces, Noord-Brabant, Overijssel, Limburg, Groningen, Drenthe and Flevoland in the Netherlands. As a DNO, the company is responsible for construction, maintenance, management and development of the transportation and distribution network. The company is not only an electricity distributor but also a gas distributor, which operates with over 130.000 kilometers of electricity cables and 40.000 kilometers of gas pipes in order to distribute energy to 2.6 million customers.

The objective of the company is to facilitate the market with an affordable, reliable and sustainable energy network. It is a semipublic company that esteems customer benefit. The company is permitted to make profit but spends the profits in order to improve the service quality offered to the customers. As DNOs take a monopolistic role in the market, the tariffs are determined by market regulators considering that the company does not achieve higher returns than usual in economy. Profits are motivated to optimize network quality and earnings to run its operations. Despite no high profit targets, the company wants to become one of the most reliable players in the market. Therefore they value accurate forecasting of load in order to provide higher service levels.

3.2. Case Study Background

As mentioned before, load forecasting is very important for operations of electricity utilities, however load planning is not a simple task. There is a need for uninterrupted balance in the system at all times.

There exists the common misconception that distribution network operators are only responsible for transferring electricity from supplier to demand points, therefore system balance is not their concern. However, this does not reflect the reality. In addition to their responsibility for transferring electricity, DNOs are also responsible for connecting the demand points to the respective suppliers (as a well developed free market, Dutch consumers are free to choose whichever supplier they want), disconnecting the customers from the grid when they terminate their connection, detecting fraud in the grid, etc. Every failure that occurs in these operations returns to the DNO as loss in the system and forms one of the most important cost terms: distribution net loss. Distribution net loss consists of administrative losses and technical losses. Administrative losses are mostly fraud, measurement errors, faults in connecting to right suppliers, failures in disconnecting demand points in the case of termination of contract or in the cases of customers' failures to pay their bills, etc. Technical losses consist of network's own consumption during transferring the electricity, losses in the cables, losses occurring during conversion from high voltage to low voltage, etc.

Due to high uncertainty in terms of net loss, there is no statistically reinforced, accurate method for forecasting net losses. Therefore in the company, net loss is estimated as a percentage of total electricity infeed. Electricity infeed refers to the amount grid operator needs to distribute in order to operate the system appropriately. Accordingly, it is equal to the sum of the amount supplied from the national grid, the difference between the amount supplied from other grids and demanded by other grids and difference between the amount produced by generators of customers and consumed by them. Consequently, forecasting of electricity infeed carry great importance for companies generating, delivering and reselling electricity because even the slightest improvement in matching demand and supply can increase their profits significantly.

$$\begin{aligned}\begin{matrix} Electricity \\ Infeed \end{matrix} &= \begin{pmatrix} Supply\ from \\ national\ grid \end{pmatrix} \\ &+ \begin{pmatrix} Supply\ from \\ other\ grids \end{pmatrix} - \begin{pmatrix} Demand\ of \\ other\ grids \end{pmatrix} \\ &+ \begin{pmatrix} Generation\ of \\ customers \end{pmatrix} - \begin{pmatrix} Consumption\ of \\ customers \end{pmatrix}\end{aligned} \tag{3.1}$$

In order to balance the net loss, Customer Services, Reporting, Analysis and Quality Market Players Department provide net loss forecasts to the finance department and finance department makes the necessary transactions that are held in the electricity market. Our study consists of two phases; first is short-term forecasting and the second is long-term forecasting. During the preliminary meeting that were carried out with company representatives, they asked for a short-term load forecasting model in order to assist traders and finance department for their transactions in the spot market. Company is using a time shifting method based on similar years search and by using that they derive hourly forecasts for a full year in June of the previous year. Finance department, together with traders, make the transactions one and a half year ahead in the futures market, in order to minimize the cost. Throughout the year, deviations from the forecast are updated by utilizing spot market. Therefore in the first phase, short-term load forecasts are derived at different lead times in order to enhance short-term transactions. Two different methods are used in modeling of the first phase: a modification of exponential smoothing method, due to its simplicity and a time series neural network method - NARX, relying on its ability to capture complex and nonlinear relationships existing in the data set.

In the second phase of the study, due to the promising results of NARX in the first phase for longer lead times, we investigated if the initial forecasting performance of the company can be improved by using NARX in long-term forecasting.

3.3. Dutch Electricity Market

In 1998, Dutch electricity market was deregulated in order to build a competitive free market that is purged from any monopolistic power and external interventions. The aim of the deregulation was ensuring the highest customer benefit by enabling competition in the market. Former electricity market consisted of a few market players for every region, giving these players monopolistic power, as the customers had to receive electricity from market parties in that region. Deregulated market structure however, allows customers make deliberate decisions about their supplier and aims providing perfect market information to the customers to be used in their decisions.

Current Dutch electricity market consists of mainly five parties, suppliers, program responsible parties (PRPs, traders), national grid operator (NGO), regional distribution network operators (DNO), metering companies (MC). In Netherlands national grid operator is the semi-public company: TenneT that is the owner of high voltage grid network. Traders make forecasts for electricity load of next day for every 15 minutes and deliver it to TenneT. TenneT explores if there is an imbalance in the supply demand equilibrium and announces the price respectively. TenneT is not only responsible for administering transmission grid and maintaining the reliability and continuity of the electricity supply, but also for providing services and performing duties in order to improve functioning of electricity market. Suppliers' role is generating electricity on their own generators or making contracts with independent suppliers and providing electricity to the market. Fossil fuels, uranium, wind are some of the commonly used electricity generation resources in the Dutch market. Another responsibility of suppliers is reaching the consumers and making contracts with them. Moreover, they are in charge of reading meters of their customers and submitting these readings to the DNOs. DNOs' main role in the market is delivering electricity to the customers through their regional low voltage and mid voltage networks. Our company is, as the operator of 130.000 kilometers of electricity cables, one of the largest DNOs in Netherlands. In Figure 1, other DNOs and their dispersion in the country can be seen. In addition to distribution function, DNOs need to allocate the total electricity load to every single small customer considering the metering information of large customers provided by MCs and transmit this information to respective suppliers to be used in billing of customers and to PRPs. PRPs might be charged by TenneT if there are huge differences between their predictions and the actual load values.

Figure 1. Distribution of DNOs in Netherlands.

3.4. Applications of Load Forecasting in Turkish Electricity Market

In addition to Dutch market, this thesis is a valuable input for also Turkish electricity market. Turkish market is not as mature as Dutch electricity market yet, but following a similar path in terms of regulations and deregulations.

Similar to other electricity markets, load forecasting is divided into three different horizons in Turkish market: short, mid and long. Transmission parties use short-term forecasts in order to balance hourly supply and demand. Mid-term forecasts are used for transmission system maintenance operations and long-term forecasts are utilized in planning of transmission infrastructure. Electricity distributors employ short and mid-term forecasts in meeting large customers' hourly demand. Large customers are defined as the group of customers who buy electricity in large amounts with contracts, such as huge production facilities, plants, etc. Similar to the Dutch distributor, Turkish distributors also want the least dependency on the spot market in balancing activities, so they need to forecast the load accurately as early as they can. Distributors use long-term load forecasts while setting contract terms with large customers and planning the infrastructure. Traders utilize short, mid and long-term forecasts in making transactions for hourly electricity demands with the least costs. [39] provides detailed information on processing of Turkish electricity market.

On 4th of April 2006, in Turkish Electric Energy Demand Forecasts Regulation, it is mentioned that electricity load forecasting data may consist of economical, social, demographic, meteorological, environmental data and past load data in addition to the other required regional data. Another bullet states that the relationships between the variables and their effects on the load need to be supported by scientific and logical approaches. Considering our detailed data analyses on effects of different variables on load and comprehensive computational study on effectiveness of different methods on load forecasting, this study can be used as a guideline in load forecasting in the Turkish market.

In addition to the importance of load forecasting, with the liberalization and rising competition in electricity market, importance of price forecasting has also increased. One possible contribution of this study to electricity price forecasting might be using load forecasts as an element or a basis for price forecasting. Additionally, price and load patterns are very similar in electricity market; considering similarities in data sets, methods performing well in load forecasting can be expected to be effective in forecasting the price. Therefore we believe our study can be considered as a valuable outcome for also price forecasting in Turkish load forecasting.

As mentioned before, Turkish electricity market is an immature market, which is continuously developing with new regulations, legislations and laws. Most of the current models and applications being used are imported from the countries well developed in terms of electricity market. Even if the electricity load data structures of two countries are very similar, in order to use such applications there is a need for full understanding of the models. Exported models and applications are like black boxes, hence are had to interpret and difficult to intervene in when needed. We believe with studies similar to ours on the Turkish electricity load data, the dependency on external models can be reduced and better functioning of the market can be achieved. Additionally, such models costs a lot of money to the governments and these costs can be decreased tremendously with such studies.

CHAPTER 4

PROPOSED METHODS

In this section forecasting methods that are used in this study are presented. First method is a recently introduced modification of Holt Winters Exponential Smoothing method [2]. We have applied a modification to this method, in order to include effects of **special days**. Secondly, Nonlinear Autoregressive with eXogenous variables (NARX) neural networks is introduced.

4.1. Holt Winters Exponential Smoothing Method (HWT)

Exponential smoothing is a fairly simple forecasting method suitable for univariate time series data [40]. Despite its simplicity, it is known as one of the most successful methods in automatic forecasting. Its name origins from the fact that the method comprises weighted averages of all past observations, where the past observation's weight decreases exponentially as the observation gets older. It applies recursive updating schemes while smoothing and forecasting data [40]. The formulation of exponential smoothing for one-step ahead forecasting is provided in (4. 1).

$$\hat{y}_t(1) = \alpha y_t + (1-\alpha)\hat{y}_{t-1}(1) \tag{4. 1}$$

or equivalently in error correction form:

$$\hat{y}_t(1) = \hat{y}_{t-1}(1) + \alpha e_t \tag{4. 2}$$

$$e_t = y_t - \hat{y}_{t-1}(1) \tag{4. 3}$$

where is the observed time series, $\hat{y}_t(k)$ is k-step ahead forecast made at time t, α is the smoothing factor and e_t is one step ahead forecast error.

Holt Winters method is an extension of exponential smoothing, which is designed for series with trend and seasonality; therefore it is also referred as double exponential smoothing. Holt Winters method is a robust and easy way of forecasting that works especially well for short-term sales and demand time-series data, despite its simple structure [41]. Holt Winters method models the data by means of a local mean, a local trend and a local seasonal factor. There are two different formulations for multiplicative and additive seasonality. In this study we consider Holt Winters method with additive seasonality and without a trend term. Taylor [44] stated that including a trend term do not bring any improvements to forecast accuracy, as changes in demand level is not significant for short-term load forecasting. In addition to this, Taylor [1] developed an extension of regular Holt Winters method in order to accommodate for the presence of two seasonal cycles, which is typical in electricity load data. Taylor's extension is presented in equations (4. 4) - (4. 8).

$$\hat{y}_t(k) = l_t + d_{t-m_1+k_1} + w_{t-m_2+k_2} + \phi^k e_t \tag{4. 4}$$

$$e_t = y_t - \hat{y}_{t-k}(k) \tag{4. 5}$$

$$l_t = l_{t-1} + \alpha e_t \tag{4. 6}$$

$$d_t = d_{t-m_1} + \delta e_t \tag{4. 7}$$

$$w_t = w_{t-m_2} + \omega e_t \tag{4. 8}$$

where k refers how many steps ahead forecasts are being generated, m_1 and m_2 are the number of periods in the first and second seasonal cycles, which are in our case daily and weekly cycles. l_t is the smoothed level and d_t and w_t stand for seasonal indices for daily and weekly cycles, respectively. The smoothing parameters are denoted by α , ω and δ. By addition of autoregressive component, ϕ , Taylor [1] aimed correcting first order residual autocorrelation, which leaded significant improvement in forecast accuracies.

4.1.1. Modified Holt Winters Exponential Smoothing Method with Special Days

As it is presented above, Taylor's models do not account for special days. In Taylor's studies, periods that do not include any special days are used [1] or special days are smoothed out by averaging two adjacent weeks' load values for the corresponding period [33]. Similarly in many other studies, similar days approach is used such as treating special days as Sundays [42] , or switching with a similar day of the preceding week [43]. However, in real life, electricity load data consists of many special days such as celebrations, national and religious holidays, bridge days, etc. and these days create the biggest challenge for generating accurate forecasts. Therefore we believe special attention needs to be paid on such days in order to improve forecast accuracies. Hence, in this study we present a modified version of Taylor's Holt Winters extension, which considers special days. With our modification we allow the model to learn from its previous errors on special days, which brought significant improvement to the model performance. The updated model formulation is provided in (4. 9) - (4. 14).

$$\hat{y}_t(k) = l_t + d_{t-m_1+k_1} + w_{t-m_2+k_2} + \phi^k e_t + \left(\sum_{i \in S} s_{i,t+k} * \frac{\sum_{h=-L}^{L} s_{i,j} * \frac{e_j^T}{y_j}}{\sum_{h=-L}^{L} s_{i,j}} \right) * \hat{y}_t^T(k \tag{4. 9}$$

where $j = t + k - (365 + h) * 24$

$$e_t = y_t - \hat{y}_{t-k}(k) \tag{4. 10}$$

$$e_j^T = y_t - (\hat{y}_{t-k}^T(k)) \tag{4. 11}$$

$$l_t = l_{t-1} + \alpha e_t \tag{4. 12}$$

$$d_t = d_{t-m_1} + \delta e_t \tag{4. 13}$$

$$w_t = w_{t-m_2} + \omega e_t \tag{4. 14}$$

where $s_{i,t}$ is a binary variable that is equal to 1 if t is interval on a special day of type i, where i refers to the special day type time $t + k$ belongs to, S refers to the set of special days and e_j^T stands for the error when demand in time j is forecasted with Taylor's adaption of HWT [44]. In our modification, the model checks if time $t + k$ is a special day. If it is a special day, first the model goes to last year's data and checks a range of days $(\pm L)$ around $t + k$ to find a similar type of special day, and then updates the forecast by multiplying its forecast ($\hat{y}_t^T(k)$) by the percentage error of the last year's forecasting error on that special day (e_j^T / y_j). It should be noted that the modification is applied only to the special days. On the normal days, Taylor's proposed model outcome is used, in order to prevent peak values from rolling to the next days, which causes decrement in model performance on special days.

For one step ahead forecasting, i.e., for $k = 1$, formulations in (4. 9) - (4. 14) are used, but for multi-step ahead forecasts (4. 9) slightly differs. Taylor [44] formulized multi-step ahead forecasts for $1 < k \leq m_1$ as follows:

$$\hat{y}_t(k) = l_t + \frac{\alpha\phi(1-\phi^{k-1})}{(1-\phi)} e_t + d_{t-m_1+k_1} + w_{t-m_2+k_2} + \phi^k e_t \tag{4. 15}$$

where the first two terms sum up to the expected value of lagged level. We derived the formula for $k > m_1$, by computing the limit of the expected value of lagged value for very large values, as presented in (4.16):

$$\hat{y}_t(k) = l_t + \frac{\alpha}{(1-\phi)} e_t + d_{t-m_1+k_1} + w_{t-m_2+k_2} + \phi^k e_t \tag{4. 16}$$

4.2. Artificial Neural Networks

Artificial Neural Networks (ANNs) are highly interconnected simple processing units that are inspired by the biological neural nets transmitting signals via neurons and synapses. The method comprises capturing complex relationships between input and output information with the network structure. The biggest advantage of ANNs compared to other computational methods is their capability of providing information about nonlinear and hidden patterns in the data. Despite the network implementation is usually called a black box, ANNs' power simply sources from their execution; they implement linear discriminants, but in a space where inputs have been mapped nonlinearly. The key power of neural networks depends on implementation of fairly simple algorithms where nonlinearity can be trained from training data [45].

In order to comprehend processing of ANNs, fundamental components of ANNs, structure of nodes (a.k.a. neurons) is illustrated. As shown in Figure 2, a node receives inputs, multiplies each input by a weight value w, adds a bias value b_0 and uses a transformation function f to generate an output y.

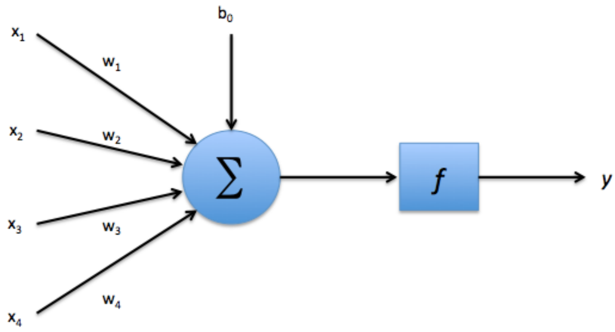


Figure 2. Structure of a single node.

Some of the commonly used transformation functions are log-sigmoid, hyperbolic tangent sigmoid, linear, etc. For example a sigmoid transfer function is as follows:

$$f(net) = \frac{1}{1 + e^{-net}} \tag{4. 17}$$

where *net* is the weighted input of the hidden layer and *f(net)* is the output of the hidden layer.

There are different types of neural networks present in literature and two major types with respect to connections between neurons and direction of data propagation are: feed-forward and recurrent networks [31]. In Figure 3 an example of three-layered feed-forward neural network is presented, the data is received through input layer and passed to hidden layer and transferred to output layer. The term **feed-forward** refers to the networks with interconnections that do not form any loops. Furthermore, **recurrent** or **non-feed-forward** networks in which there are one or more loops of interconnections also exist, so that input state is also combined with the previous state activation through an additional weight layer, an instance of recurrent networks is provided in Figure 4 [46, 47].

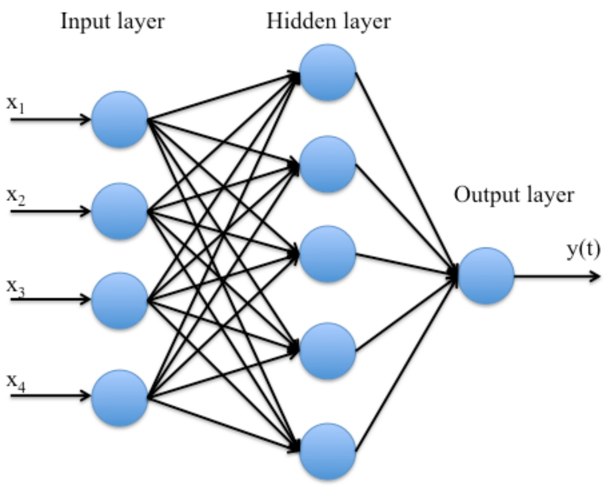


Figure 3. Example of three-layered feed-forward neural network structure.

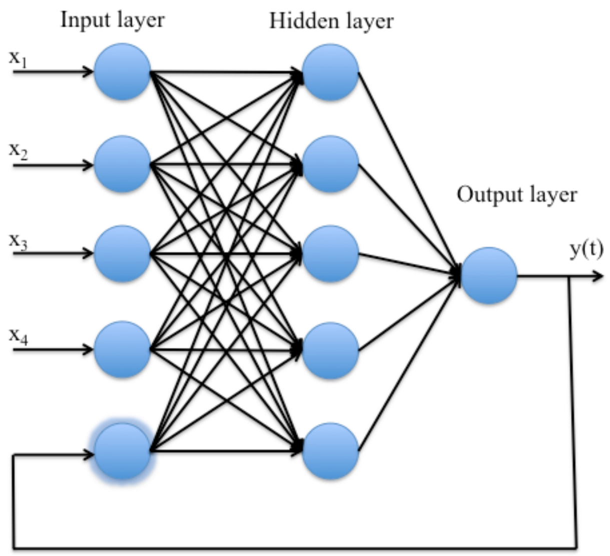


Figure 4. Example of three-layered recurrent neural network structure.

4.2.1. Nonlinear Autoregressive with eXogenous Inputs

Nonlinear Autoregressive with **eXogenous** Inputs (NARX) Networks is a type of recurrent dynamic neural networks with feedback connections between output and input layer. NARX networks are specifically used for time series forecasting. Another important property of NARX is represented under the name of 'eXogenous', meaning that it allows exogenous inputs to become network inputs. It is derived from Autoregressive exogenous (ARX) model. It can be mathematically written as:

$$\hat{y}_t = f(u_{t-D_u}, \dots, u_{t-1}, u_t \; ; \; y_{t-D_y}, \dots, y_{t-1}) \tag{4. 18}$$

where u_t and y_t are inputs and outputs of the model at time t and D_u and D_y are input and outputs delays. Delays are input and output memory orders respectively. f stands for the nonlinear transformation function [48]. In Figure 5 an example for NARX network with input delay of two and output delay of three is shown.

As mentioned NARX networks are all recursive, but there are two different types with respect to the information embeded into the feedback loop: **open-loop** and **closed-loop** networks.

The network in Figure 4 is an open-loop type, which means actual output values are fed back to the network. Another name for open-loop networks is series-parallel (SP) mode networks. In the other type of architecture, closed-loop, network's outputs, estimated values, are fed back to the network as inputs. These type of networks are also referred as parallel (P) mode. Therefore, the NARX model representation in (4.18) is for SP architecture, whereas a P architecture is mathematically as follows:

$$\hat{y}_t = f(u_{t-D_u}, \dots, u_{t-1}, u_t \; ; \; \hat{y}_{t-D_y}, \dots, \hat{y}_{t-1}) \tag{4. 19}$$

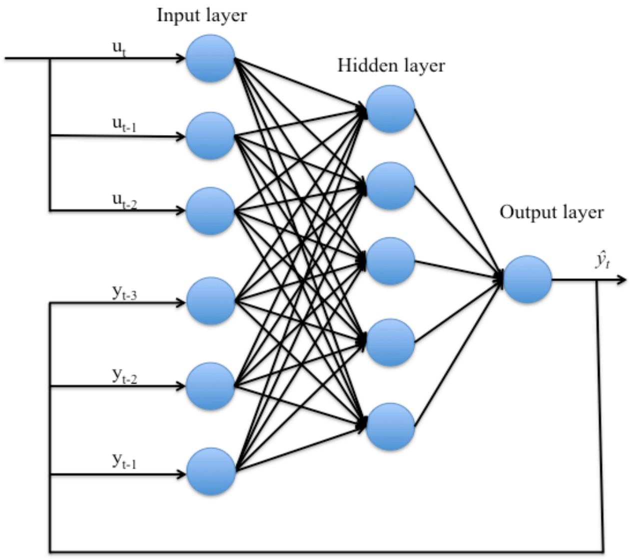


Figure 5. Example of a NARX neural network

4.2.2. Input Processing in Neural Networks

In a neural network, neurons in each layer are interconnected to each other by modifiable weights. In input layer no transformation takes place, it serves only for transferring the input. Transformation takes place only in hidden and output layers. For a network that consists of p inputs, each hidden node computes the weighted sum of its inputs, denoted by *net* in (4.20).

$$net_j = \sum_{i=1}^{p} x_i w_{ji} + w_{j0} \tag{4. 20}$$

where i stands for the nodes in input layer, j for the ones in hidden layer; x_i stands for any type of input (multivariate inputs or input from feedback loop), w_{ji} represents the input to hidden layer weights in hidden node j, and w_{j0} 's are the bias values.

After computing net value, a transfer function is used to generate the hidden node's output, *f(net),*

$$y_j = f(net_j) \tag{4. 21}$$

Behaviour of neural networks depends heavily on transfer functions specified for the units. Transfer functions are typically categorized into three:

- **Linear (or ramp)** functions generate an output that is proportional to the total weighted sum of inputs of that unit.
- **Threshold** functions examine level of total input value, if it is greater than a threshold value output is set equal to a number, if not, it is set equal to some other predetermined value.

- **Sigmoid** functions provide more similarity to real neurons than the other functions. In those units, the output varies continuously as the input changes.

Similarly for all hidden layers and output layer units, net values are calculated as weighted sum of the values received from the previous layer nodes plus the bias value and then transformation function of the node is applied. The general representation of the practice for a three-layered network with *k, j* and *m* neurons in the hidden layers is,

$$\hat{y}_t = f\left(\sum_{j=1}^{n_H} w_{kj}\, f\left(\sum_{i=1}^{p} w_{ji}x_i + w_{j0}\right) + w_{t0}\right) \tag{4. 22}$$

where $n_H$is number of hidden nodes in the output layer, w_{kj}'s are interconnection weights from hidden layer to output layer, p is number of nodes in input layer and w_{j0} and w_{t0}'s are bias values and each node applies a transformation function to the sum and transfers its product. It should be noted that for a single layer neural network with linear transfer functions in the output layer, the system can be interpreted as a linear regression model. Similarly, a network with logistic transfer functions is equivalent to logistic regression.

4.2.3. Training of Neural Networks

As a data driven model, training is fundamental in neural network modeling. During the training phase, network adjusts weight and bias values in order to produce the best predictive results. One of the most popular methods for network training is the back-propagation algorithm**,** which is a natural extension of the least mean squares (LMS) algorithm. Back-propagation algorithm is founded on gradient descent in error [45].

Gradient descent approach is based on finding a solution to the set of linear inequalities that defines a criteria function $J(\boldsymbol{a})$. Starting with an arbitrary $\boldsymbol{a}(1)$ weight vector, gradient vector $\nabla J(\boldsymbol{a}(1))$ is computed. Vector $\boldsymbol{a}$ is updated to $\boldsymbol{a}(2)$ in the direction of steepest descent of the gradient. General procedure is formulized in (4.23).

$$\boldsymbol{a}(k+1) = \boldsymbol{a}(k) - \mu(k)\, \nabla J(\boldsymbol{a}(k)) \tag{4. 23}$$

Here $\mu(k)$ is the learning rate. Learning rate sets the step size of the conversion to the solution minimizing $J(\boldsymbol{a})$.

The learning process can be summarized as, starting with an untrained network, a training data set is fed to the input layer, passes through the network and an output value is obtained. Then the obtained value is compared to the target value, which is the actual output in the data set. The difference corresponds to the error. In back-propagation criterion function is some scalar function of the network's weights. With respect to the learning rate, the weights are adjusted in order to minimize the error.

The training error is considered similar to LMS algorithm:

$$J(\boldsymbol{w}) = \frac{1}{2}\sum_{t=1}^{r}(y_t - \hat{y}_t)^2 \tag{4. 24}$$

where $\boldsymbol{w}$ stands for vector of all weights, y_t is the target value and $\hat{y}_t$ is the estimated value.

As back-propagation algorithm is based on gradient descent, the weights are updated in the direction of error reduction, starting from random values.

$$\Delta\boldsymbol{w} = -\mu\frac{\partial \boldsymbol{J}}{\partial \boldsymbol{w}} \tag{4. 25}$$

where μ is the learning rate, taking value between 0 and 1. Similar to gradient descent procedure, learning rate controls the amount of change in weights and bias values from one iteration to the other. Larger values can give a faster convergence to the minimum but also may produce oscillation around the minimum [46].

At each iteration the weight values are updated as follows:

$$\boldsymbol{w}(m+1) = \boldsymbol{w}(m) + \Delta\boldsymbol{w}(m) \tag{4. 26}$$

where m indexes iteration number referring to different networks with different weight and bias vectors. The weights are readjusted for every input-output pair of the training data set until an acceptable criterion for convergence is reached.

There are several criteria that may terminate training of the network. 'Maximum number of epochs' is one of them. While updating the weight and bias values, using all the observations in the training data set refers to one epoch. Maximum number of epochs limits number of processing of all training data set in the training phase. The simplest training stopping criteria is the 'performance goal', which terminates the training when the performance goal is reached. Another criterion is 'maximum number of validation failures'. Validation data set enables observing of error on a subset of data that is not part of the training data. The training process stops when the validation error increases for some number of iterations. 'Maximum performance increase' and 'minimum performance gradient' are two other training stopping criteria, where the first one stops the training if the performance has increased as much as maximum performance increase in an iteration. 'Minimum performance gradient' terminates training if the greatest rate of increase is smaller than the parameter's value. Lastly, 'maximum training time' is another criterion that keeps the training time to a limit.

NARX neural networks salient for being powerful, faster convergence and better generalization capability compared to other networks [49]. In this study, NARX neural networks with zero input delays and various outputs delays are considered. Outputs delays are selected with respect to autocorrelation values and seasonalities present in the data, whereas input delay is set equal to zero due to the fact that input variables include binary special day variables, so only effective on the forecasting period.

CHAPTER 5

COMPUTATIONAL STUDY

In this section, first analysis of the data set used in the computational study is explained in Section 5.1. Then in Section 5.2 the numerical results for short-term load forecasting are given. Using the proposed methods in Section 5.3 the computational results for long-term load forecasting.

5.1. Data Description

In empirical studies, we used the data set of hourly electricity load levels of five Dutch provinces for 256 weeks period from 1 January 2008 to 30 November 2012. For every province we had a data set of 43104 data points. Load data contains daily and weekly cycles, which can be clearly seen in Figure 6. It is seen that during weekdays electricity load patterns are very similar to each other, but at the weekends load decreases significantly.

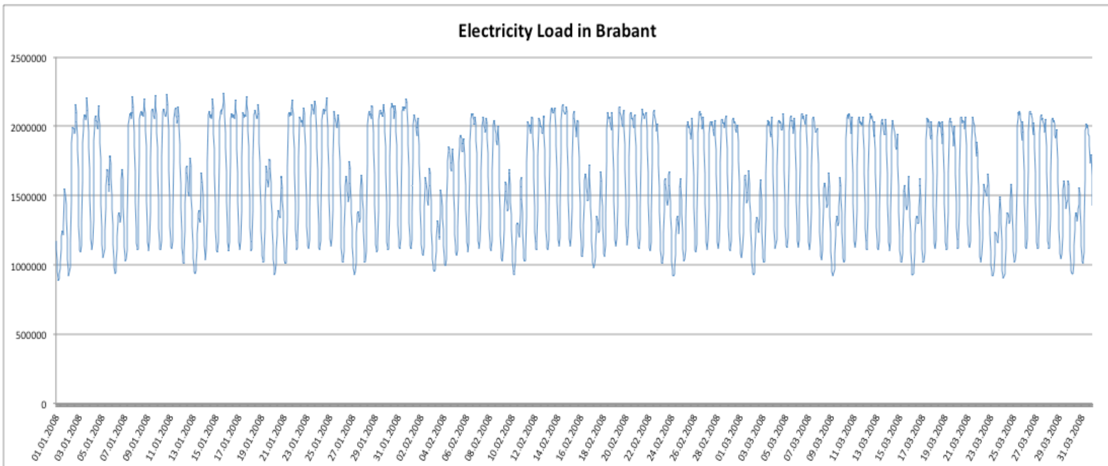


Figure 6. Hourly electricity load in Brabant between January 2008 and March 2008.

The computational study is carried out for five different data sets. However for easiness of reading the data analyses are carried out on Brabant data. Performances of the proposed methods are presented for five regions in Section 5.2.

5.1.1. General Data Analysis

In this section descriptive analysis of the data set will be explained, in order to give the reader a deeper insight about the difficulties and complexities involved in load forecasting.

General data structure throughout one year is presented in Figure 7. In this figure hourly intervals are aggregated into four-hour intervals for easiness of the analyses.

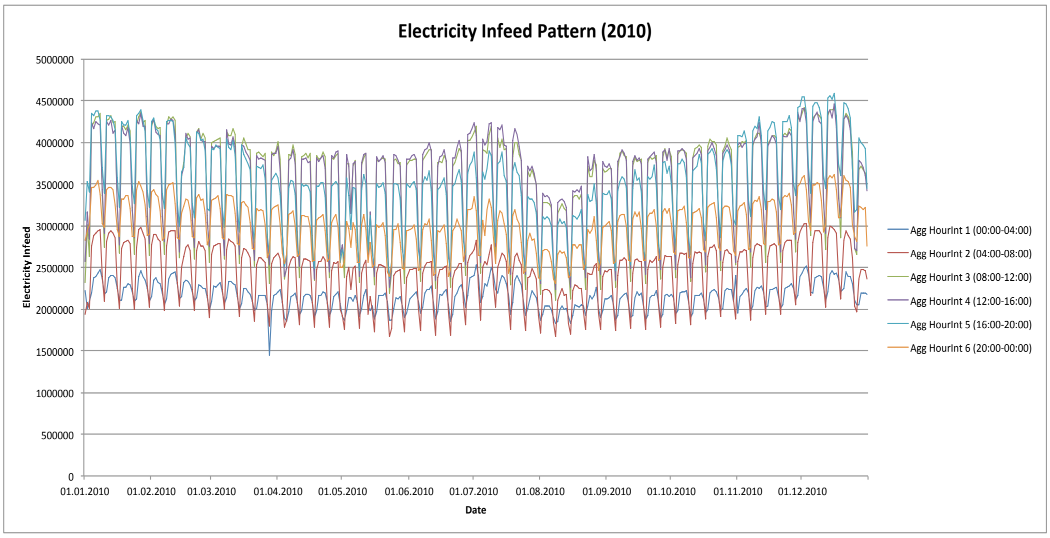


Figure 7. Aggregated hourly electricity load (GWh) in Brabant in 2010.

In Figure 7 it is seen that during late at night and early in the morning hours of the day, electricity load is at its lowest level. Morning and afternoon hours, accommodates the highest levels of electricity load and decreases gradually towards the evening hours.

If we take a closer look to one month, we can see the decrease in load levels at the weekends, see Figure 8. In this figure, the similarities of weekdays' load levels can be observed easily. Moreover, it is seen that despite during the week the interval between 00:00 and 04:00 has the lowest level, during the weekends, 04:00-08:00 interval is even lower and similarly load of the hour interval of 08:00-12:00 decreases significantly.

After aggregating load values four-hourly, we have plotted weekly aggregated load in Figure 9. We have observed every year load values decrease during summer weeks, resulting in a convex curve. In order to investigate if the fluctuations exist due to special days effect, in Figure 10 weekly aggregated values are plotted, but excluding the weeks that contain any kind of special days in year 2010.

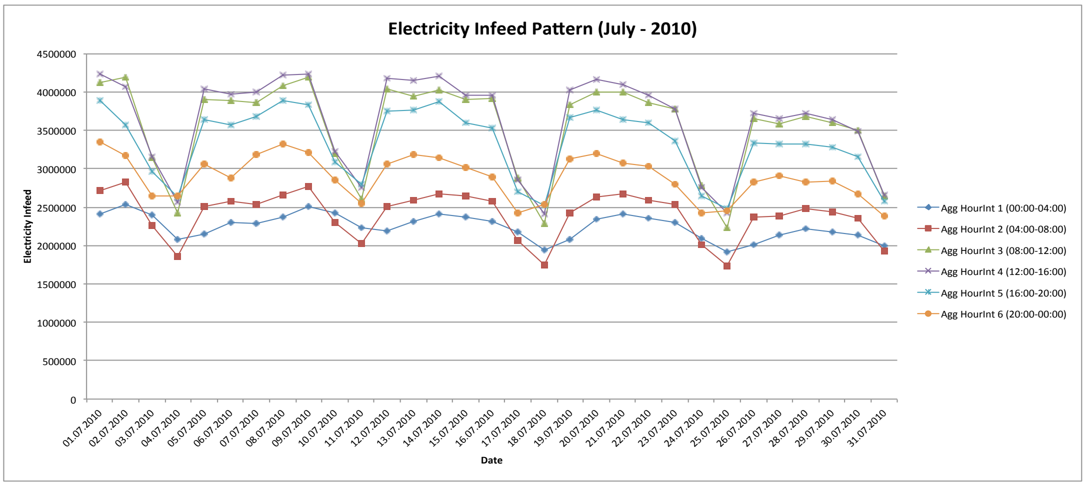


Figure 8. Aggregated hourly electricity load (GWh) in Brabant in July 2010.

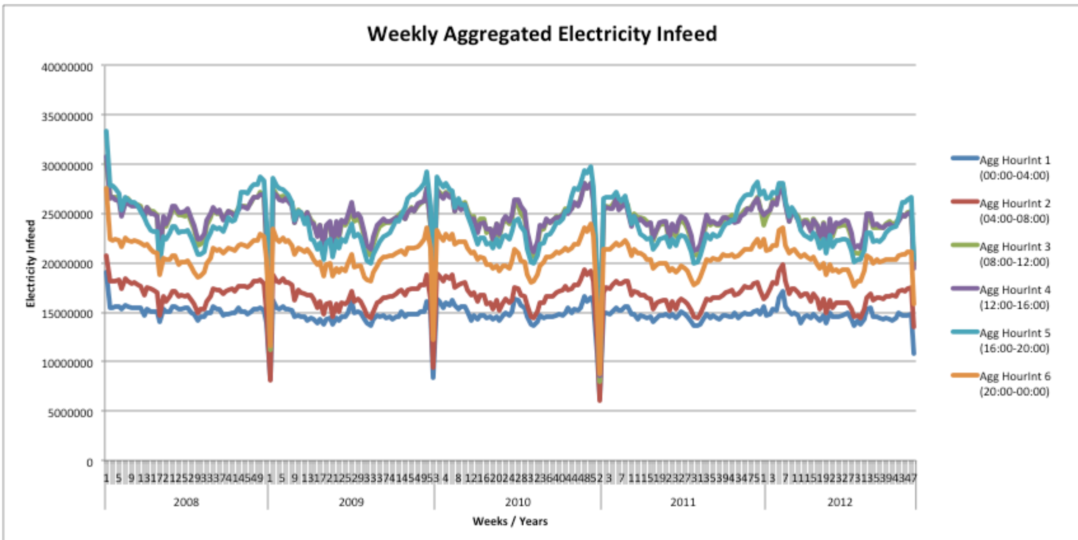


Figure 9. Weekly aggregated electricity load (GWh) in Brabant between 2008-2012.

In Figure 10, compared to Figure 9 smoother curves are obtained after exclusion of special days. The convex shape of load throughout the year can be seen more clearly. In addition to that for the hour interval 16:00 – 20:00 the effect of switching to summer-time is pointed out on this graph. During winter-time electricity load values are at the highest level, however, after switching to summer-time, electricity load values decrease. We have pointed out an unexpected increase in load values for a period during summer under the name of meteorological effect. This is due to the fact that in 2010 for that period of the year temperature values were higher than the average, so resulted as an increase in electricity load.

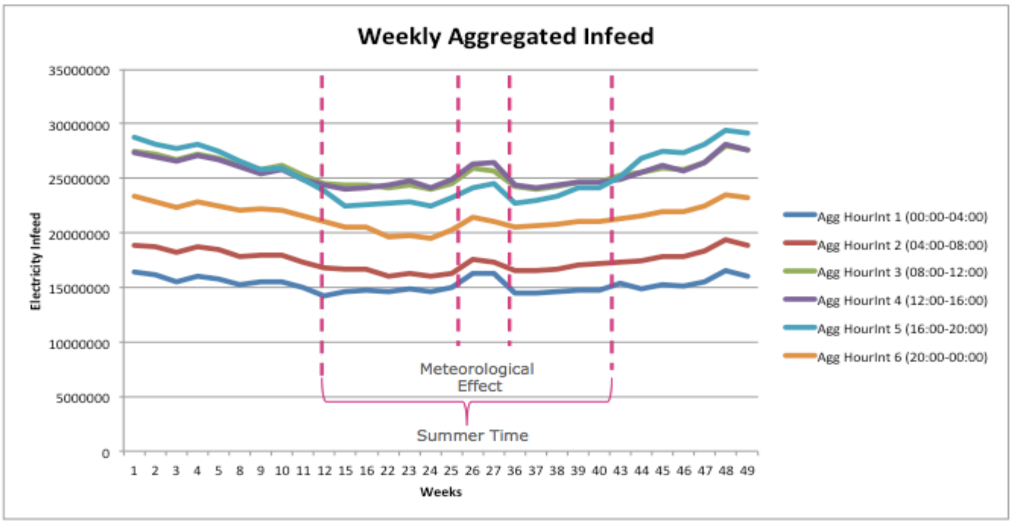


Figure 10. Weekly aggregated electricity load (GWh) in Brabant between in 2010; weeks including special days are excluded.

Additionally, average of hourly electricity load for seven days of week is presented in Figure 11. The values are obtained by averaging hourly load levels with respect to days through five years. Figure 11 points out the intraday cycle and the lower load level at the weekends. Lower load values during first hours of the Mondays also catch attention.

General distribution of hourly loads with respect to months is presented in Figure 12 with a box plot. Figure 12 shows that variation in hourly load values is higher in winter months compared to summer moths and the distributions are not skewed to either right or left.

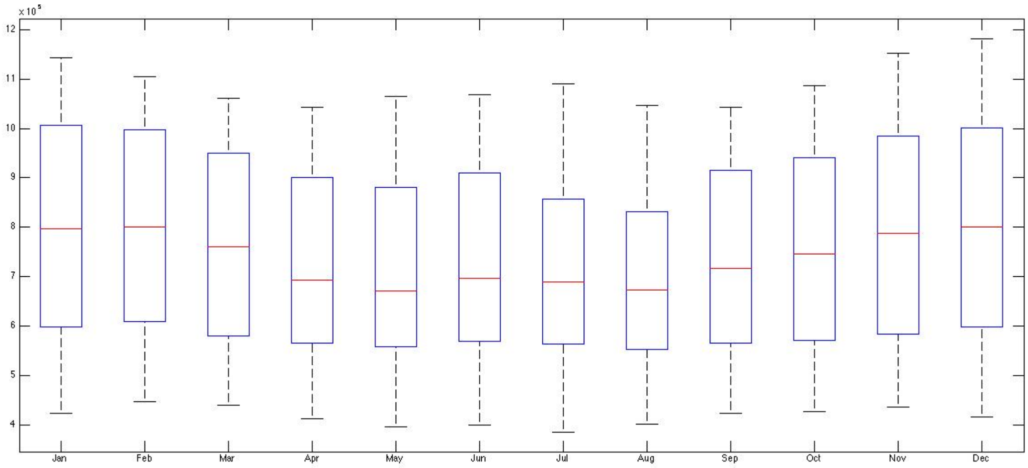


Figure 11. Average of hourly demand for seven days of the week.

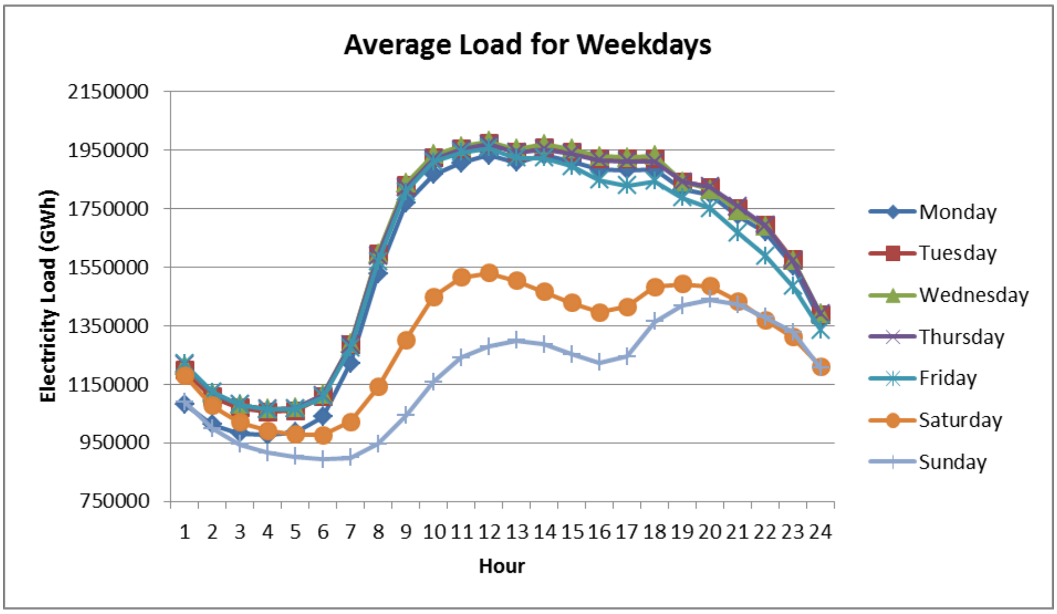


Figure 12. Box plot of hourly load values with respect to months in Brabant in 2010.

5.1.2. Analyses of Economical Variables

Possible effects of various variables are analyzed in order to investigate relationship between economical variables and total electricity load but no clear relationship is detected. Here only two of the variables are presented. Firstly in Figure 13, scatterplot of consumer prices, as an indicator of inflation, versus respective change of load compared to same month of last year is presented. No relationship is observed between the two.

Quarterly gross domestic product (GDP) values are extrapolated monthly and plotted versus percent change in electricity load, see in Figure 14. There seems to be a weak relationship, therefore a second graph is plotted, shown in Figure 15 in order to gain a better understanding. Figure 15 does not demonstrate any clear relationship between load and GDP. During the interviews with the analysts at the company and after comparisons with company analyses on the data, it is concluded that, GDP is effective on large customers' (plants, large facilities, etc.) load, however summed with small customers (households, public lighting, etc.) the effect disappears.

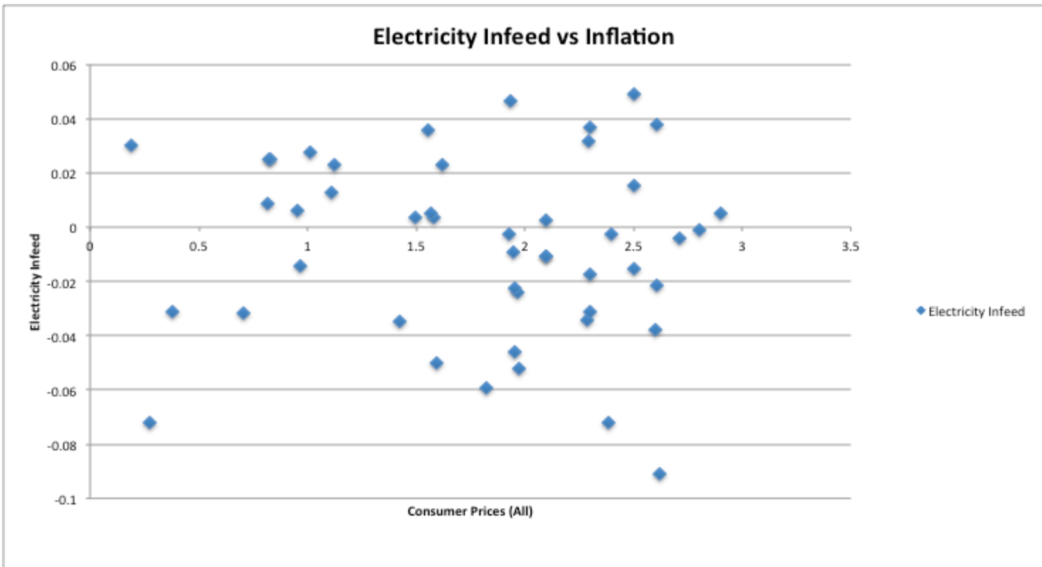


Figure 13. Scatterplot of change in total electricity load compared to one year ago vs. consumer prices index for all in Brabant.

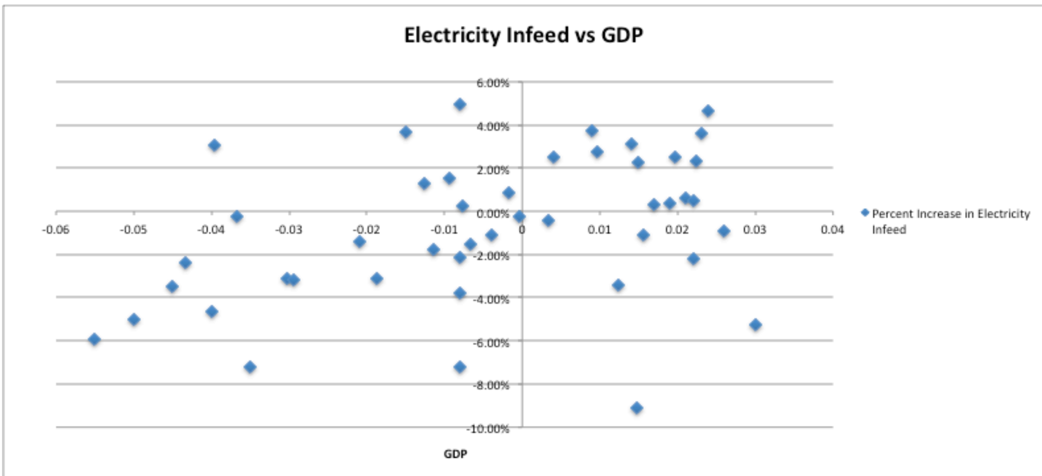


Figure 14. Scatterplot of monthly percent change in total electricity load vs. GDP in Brabant.

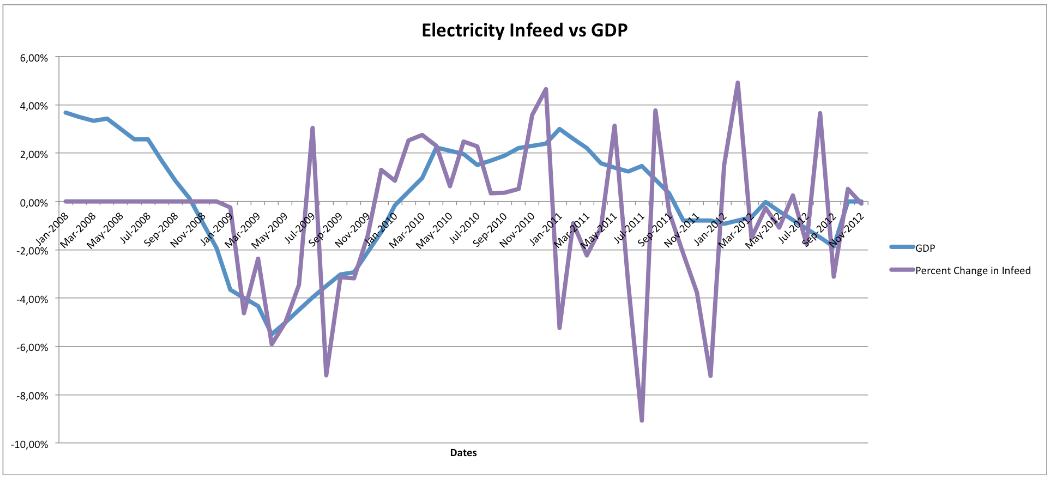


Figure 15. Graph of monthly percent change in total electricity load vs GDP in Brabant.

5.1.3. Analyses of Meteorological Variables

As expected, meteorological analysis starts with analysis on temperature's effect on load. In Figure 16, electricity load is plotted versus temperature and is grouped into two; weekdays and weekends. It is observed that as the temperature values get lower or higher than 17 ^{o}C, electricity load increases.

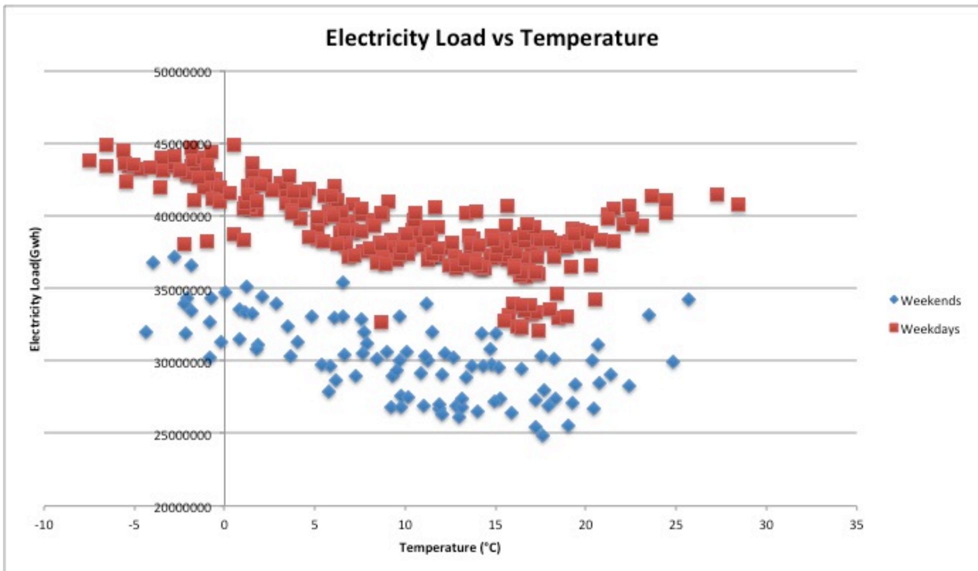


Figure 16. Scatterplot of electricity load vs temperature in Brabant.

Second meteorological variable that is discovered to be effective on electricity load is sunlight. We have represented sunlight in terms of percentage; for every hour of the day it is set equal to the percent of the hour that the sun is up. Figure 17 shows plot of a three-hours interval that is the period affected from the sunlight period the most. It is clearly seen that sunset leads people to more electricity consumption.

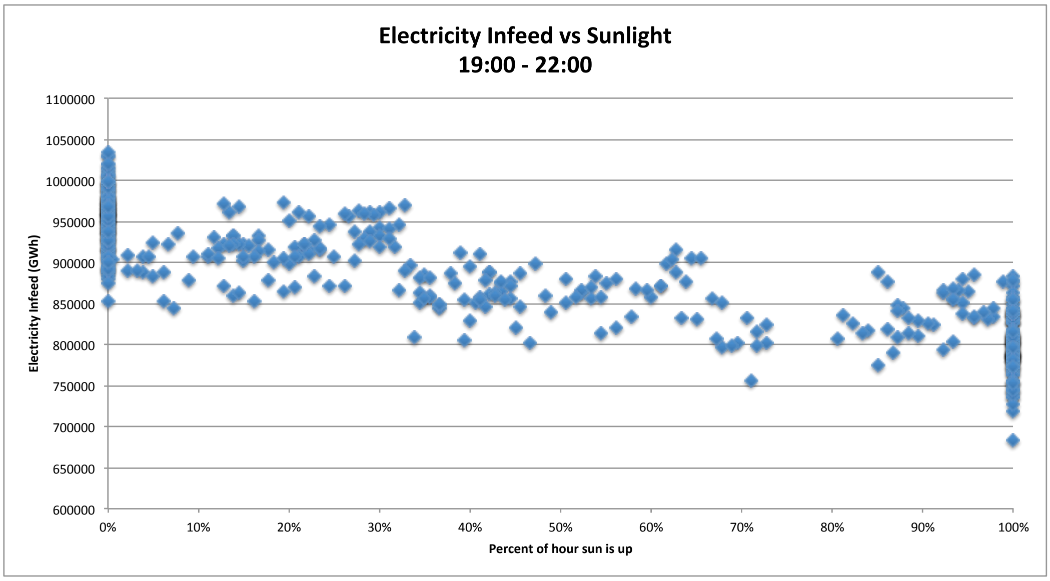


Figure 17. Scatterplot of electricity load vs sunlight in Brabant.

5.1.4. Analyses of Special Days Effect

Another important element of the data set is calendar variables, which covers special days, such as public holidays, national days, school holidays, etc. Before deciding on which variables to include, effects of all special days are analyzed. Below, you can find a list of special days:

- **School Holidays:** The period when the schools are closed. In Netherlands, schools are on vacation during public holidays as well as for Christmas, Spring break, May holiday, Summer break and Autumn break.
- **Bouwvak:** The period when companies in construction industry are not operating during the summer.
- **Liberalization day (Independence):** Public holiday, which is celebrated once in every four years.
- **Carnival:** Three days of celebrations taking place in southern part of the Netherlands. This variable is excluded from the data set of northern regions.
- **Christmas:** In the data set not only 25th of December but also one day before and after are also defined as separate variables; because deviations from regular pattern are observed on these days as well.
- **New Year's Eve:** 31st of December of every year.
- **New Year Holiday:** 1st of January of every year.
- **Queens Day:** A public holiday for celebration of Queen of Netherlands' birthday for one day.
- **Easter, Ascension Day, Whit Sunday and Monday, Good Friday:** Christian holidays in Netherlands.

Different special days have different effects on electricity load throughout the day, depending on the region, the day of week, etc. Therefore, we have carried out numerous analyses with respect to their effects on different hours of the day and with respect to the variation in their effects from year to year. We have calculated incremental effects of each special day using the following procedure:

Step 1 Expected values are calculated by averaging same day electricity infeed values of the weeks before and after for every hour of the day. **Step 2** Difference between actual infeed and expected value is calculated. **Step 3** Difference is divided by expected value calculated for every hour of a special day.

Figure 18. Algorithm for calculation of incremental effects of special days.

In Figure 19, yearly average incremental effects of the special days on different hours of the day are presented. It can be seen that some special days have very similar curve shapes and the values are very close to each other. Similarly, hourly effects are averaged and plotted with respect to years in Figure 20.

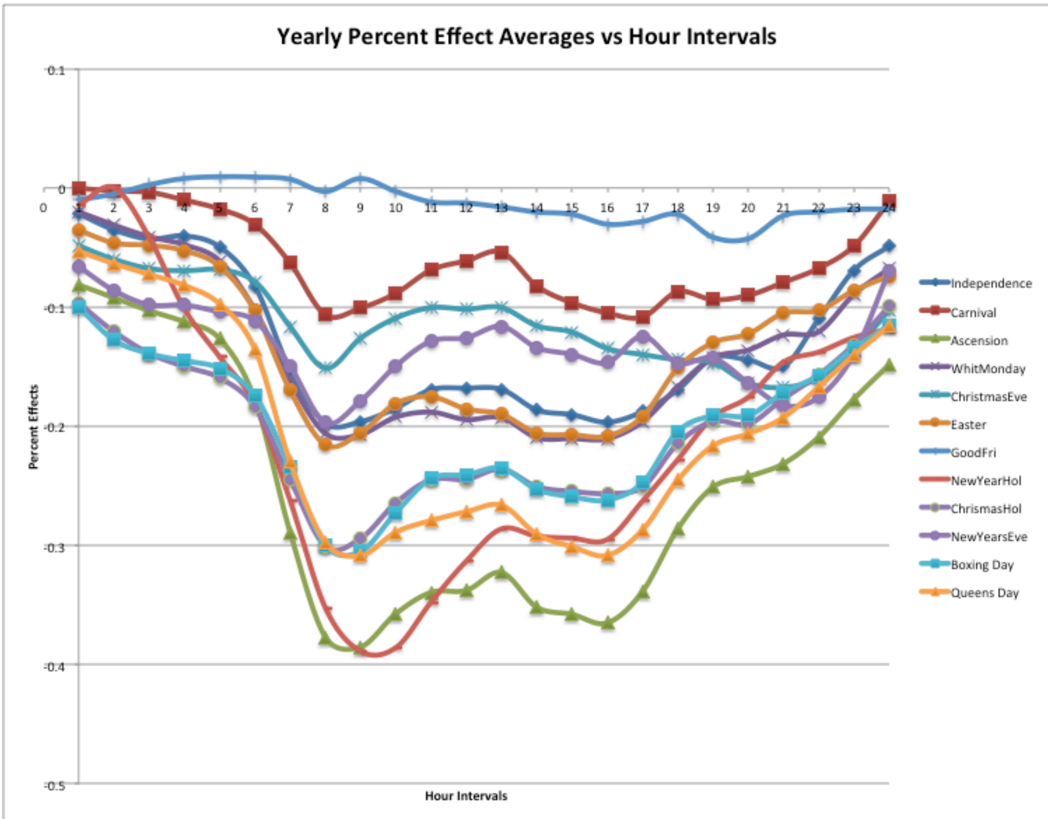


Figure 19. Effects of special days averaged for five years with respect to hours of day in Brabant.

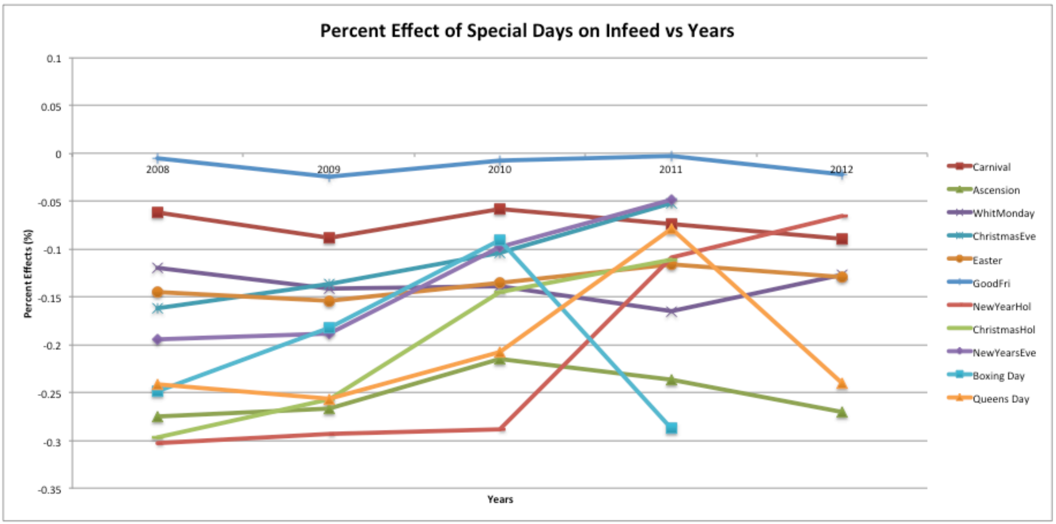


Figure 20. Average effects of special days in different years in Brabant.

Utilizing two graphs and least squared distance method between the curves; similar special dates are grouped. In Figure 21, how similarly Whit Sunday and Monday, Easter and Liberalization day affect electricity load is presented. Squared distance between these three curves are smaller than 0.01. Likewise, average effects in one day are plotted with respect to years in Figure 22. General shapes of curves are alike, however there is a significant deviation in 2011, for which higher temperature degrees during Easter and during the week before and after Whit Sunday & Monday in 2011 are accounted.

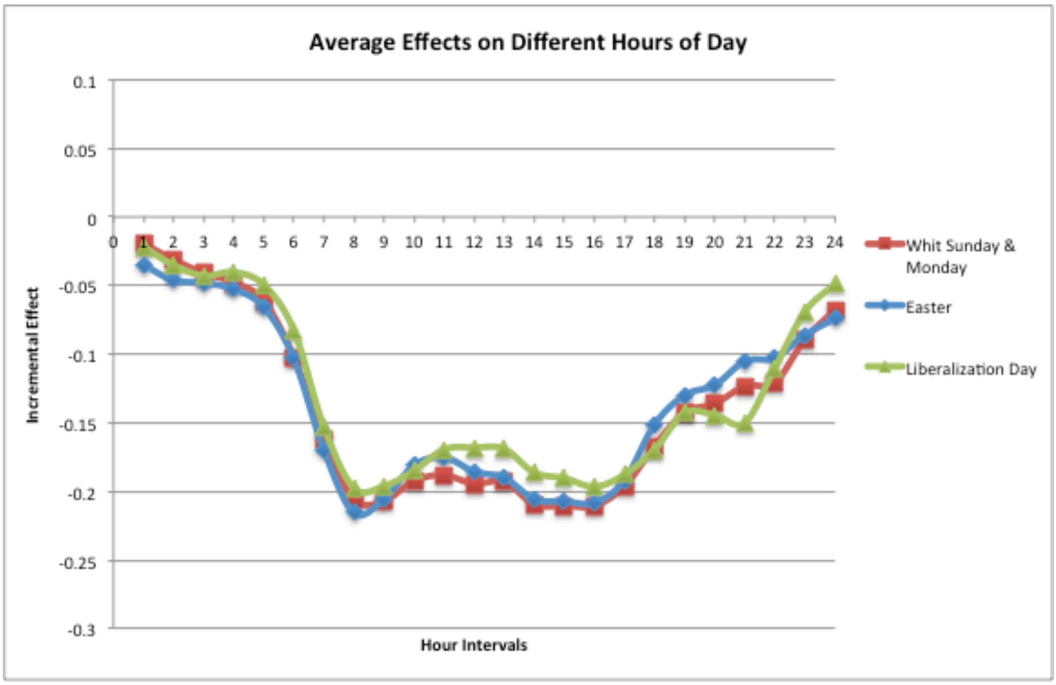


Figure 21. Effects of special days averaged for five years with respect to hours of day in Brabant (Whit Sunday & Monday, Easter and Liberalization Day).

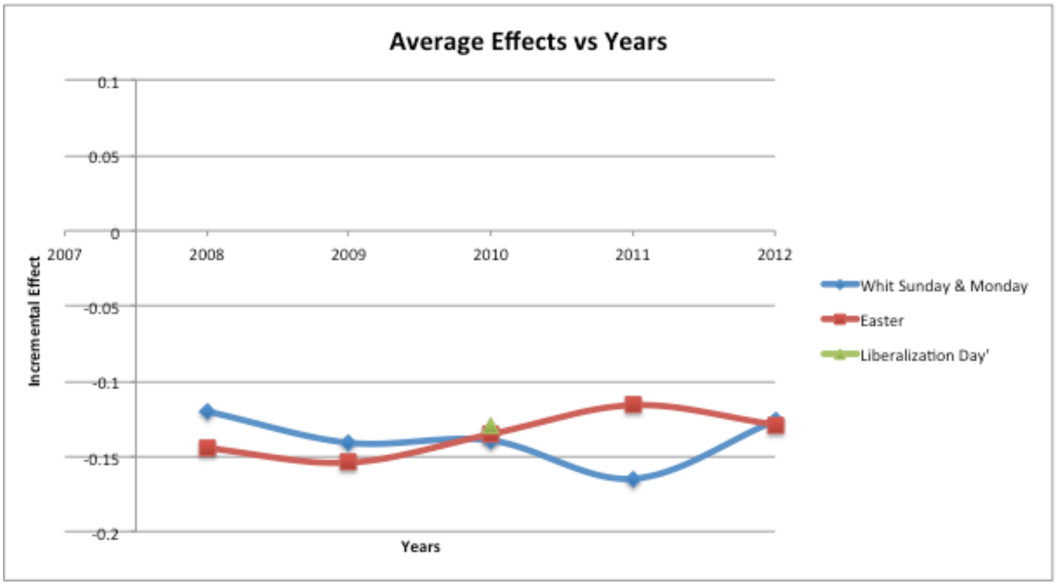


Figure 22. Average effects of special days in different years in Brabant (Whit Sunday & Monday, Easter and Liberalization Day).

Parallel to Whit Sunday and Monday, Liberalization Day and Easter, Boxing Day and Christmas Day curves also have very small squared distance that is lower than 0.01. In Figure 23, their effects with respect to hours of day and in Figure 24 their effects with respect to years are presented. Figure 24 presents also weekdays in order to emphasize special days' different effects on electricity load with respect to days. It is seen that Boxing Day and Christmas Day effect decreases at the weekends.

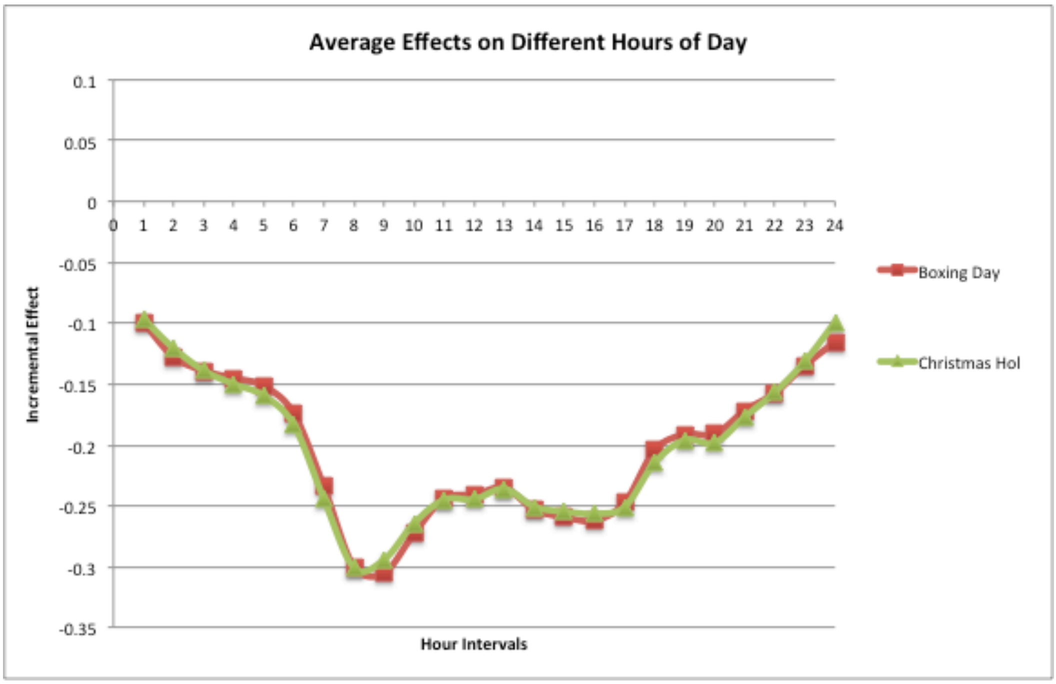


Figure 23. Effects of special days averaged for five years with respect to hours of day in Brabant (Boxing Day and Christmas Day).

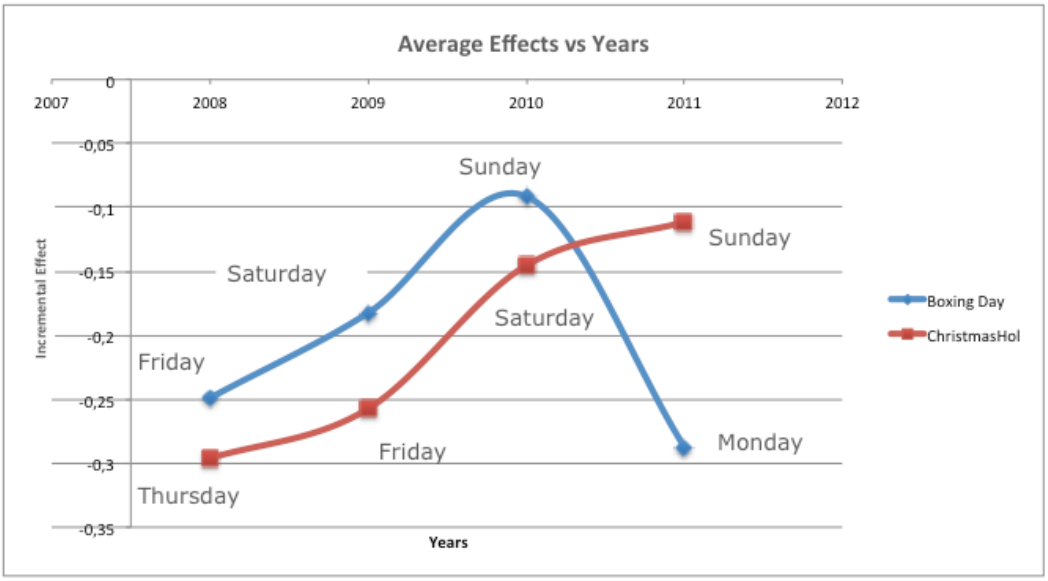


Figure 24. Average effects of special days in different years in Brabant (Boxing Day and Christmas Day).

As expected New Years Eve and Christmas Eve also have very similar effects on load, presented in Figure 24 and Figure 25. Similar to Boxing Day and Christmas Day, decrease in effect of New Years Eve and Christmas Eve clearly seen in Figure 25 at the weekends.

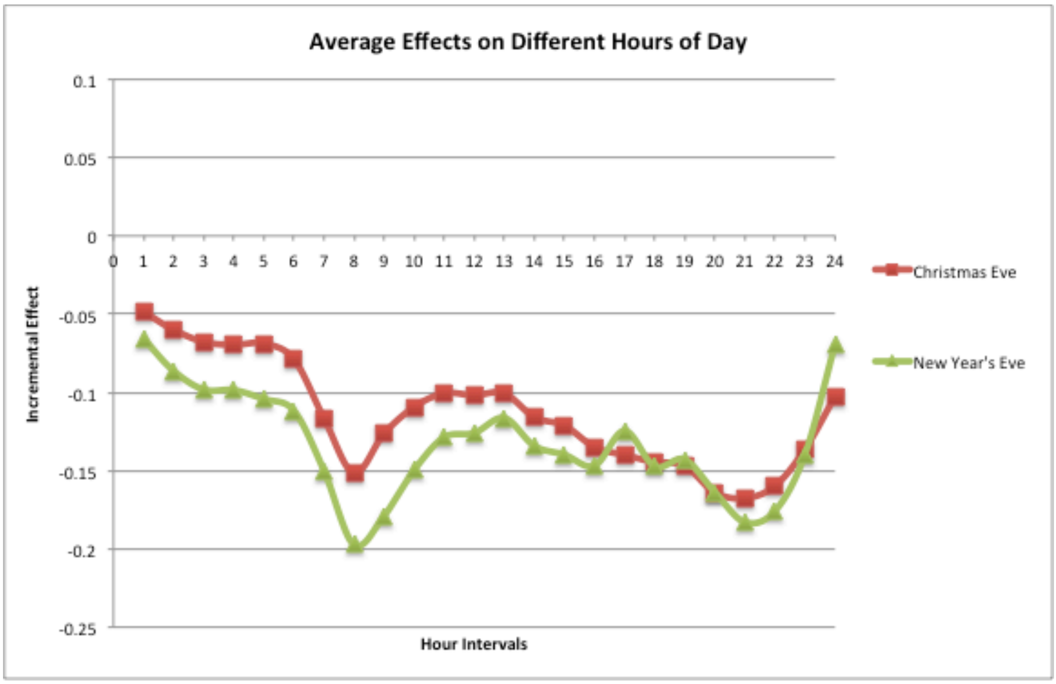


Figure 25. Effects of special days averaged for five years with respect to hours of day in Brabant (New Years Eve and Christmas Eve).

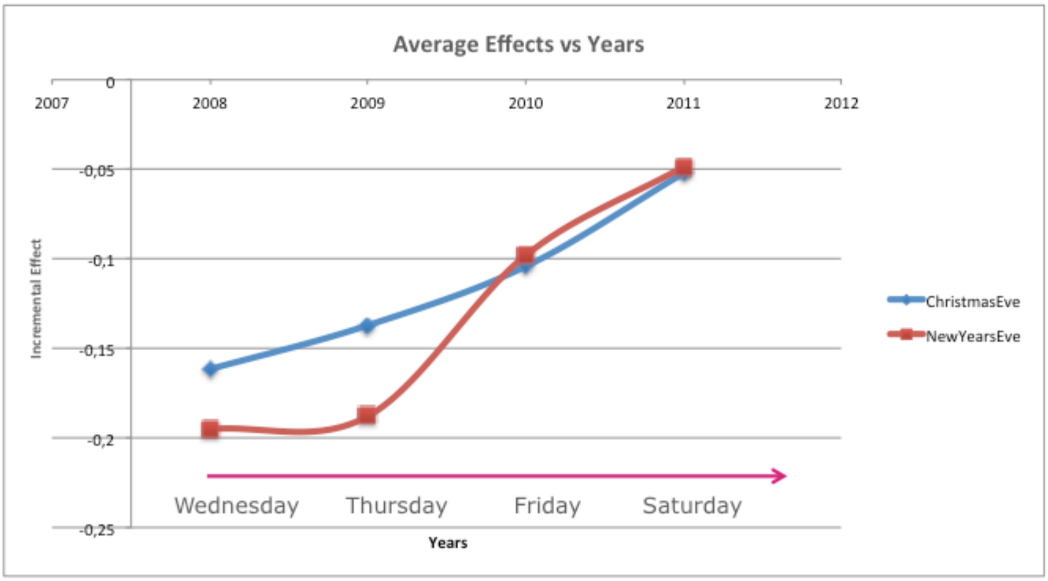


Figure 26. Average effects of special days in different years in Brabant (New Years Eve and Christmas Eve).

Among the special days mentioned above in the list, Good Friday is eliminated from the set of effective special days because of having very small effect on electricity load. Respective graphs showing the inadequate effect of Good Friday are presented in Figure 27 and Figure 28.

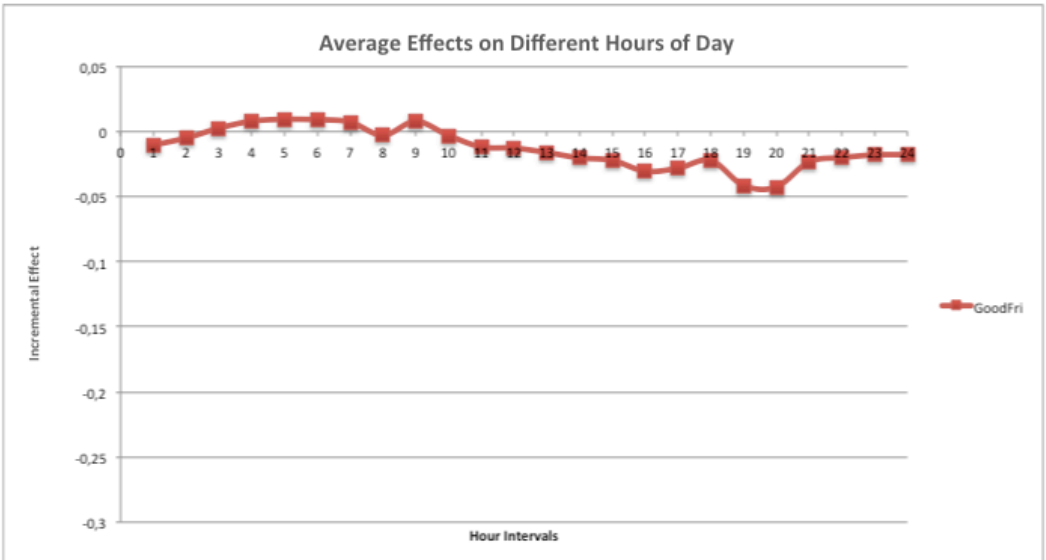


Figure 27. Effects of Good Friday averaged for five years with respect to hours of day in Brabant.

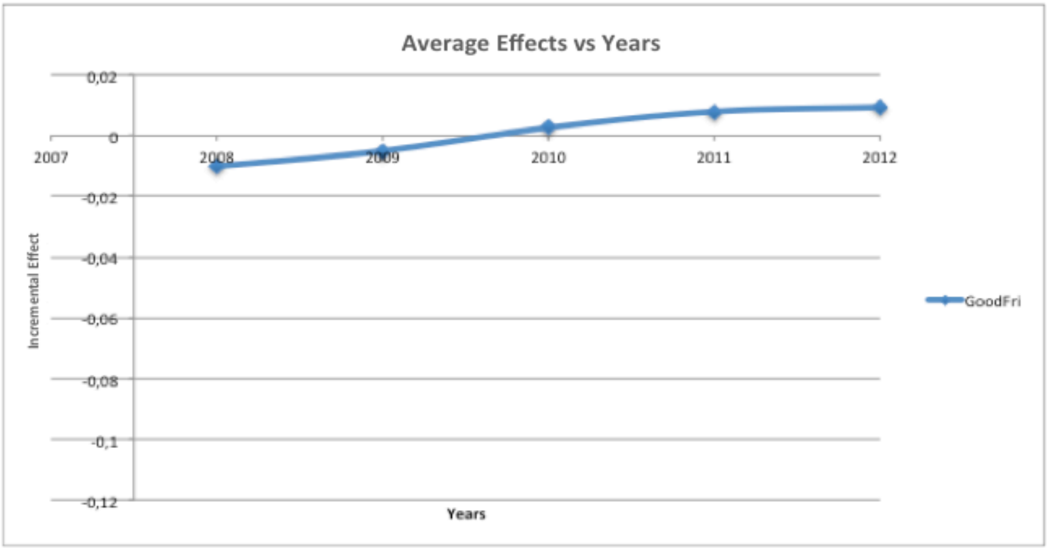


Figure 28. Average effects of Good Friday in different years in Brabant.

School holidays and Bouwvak holidays are not included in similar days analyses, because these holidays last for at least one week, so have different characteristics than the others. In Figure 29 their effects in Brabant data set in 2010 are presented with respect to days. Bouwvak's effect can be clearly observed in the figure, however school holiday seems to have only a slight effect on load. However, it is not always the case for school holidays, for example in Figure 30, school holidays' significance can be easily observed. Difference between two graphs is a good example to different effects of same special day in different periods of year. In Figure 30, school holiday is the 18th week; compared to summer a clearer decrease in load level is observed. Additionally, there is a significant decrease in level on Monday of 17th week, which is Easter Monday.

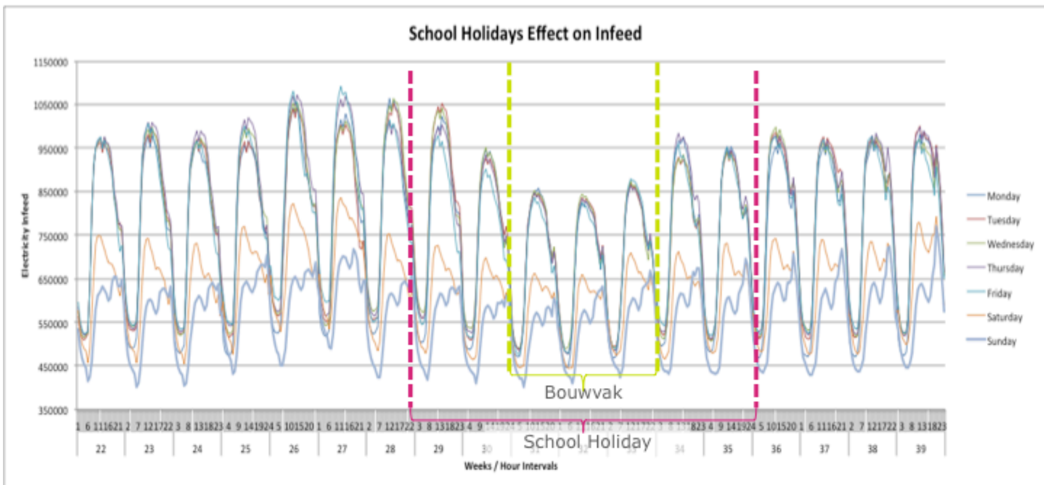


Figure 29. Electricity infeed (GWh) between weeks 22 and 29 in Brabant in 2010.

In addition to the exact dates, special days also affect the electricity load on the days close to them. Electricity load before and after holidays tends to decrease for most of the special days as shown in Figure 31. The figure shows load levels on Wednesdays of nine consequent weeks, where two Wednesdays, shown in dashed circles, are days before

Christmas and New Year's Eve respectively. Significant decrease in electricity load on these days can be easily seen. Similar tendency is observed before and after other special days as well. Therefore, variables named *'Day before holiday'* and *'Day after holiday'* are defined and added to the data set.

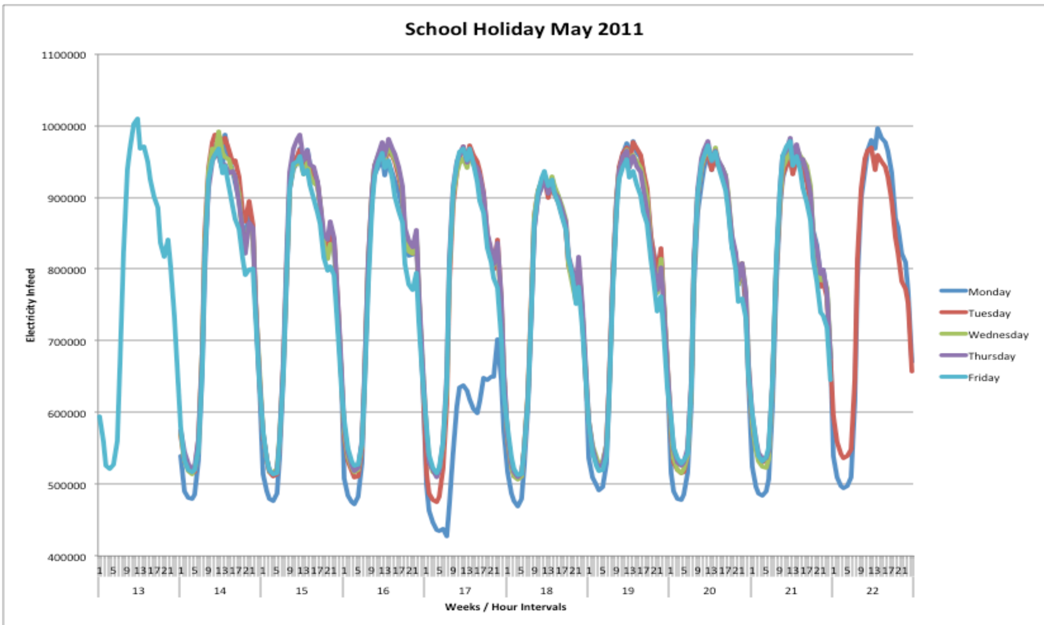


Figure 30. Electricity infeed (GWh) between weeks 13 and 22 in Brabant in 2011.

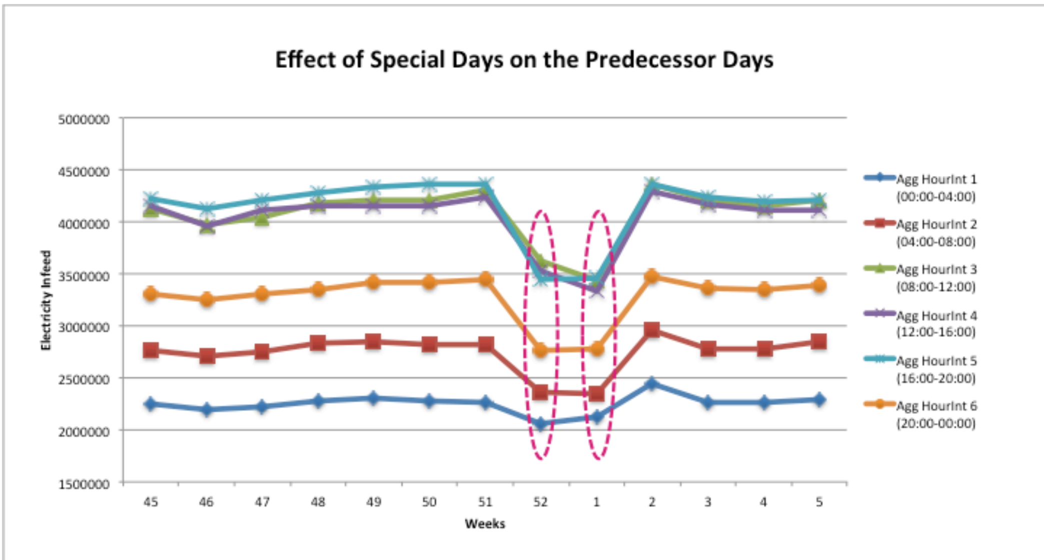


Figure 31. Electricity infeed (GWh) between weeks 45 in 2008 and 5 in 2009 on Wednesdays in Brabant.

More importantly, we observe that the effect is even more significant when the day before or after holiday falls on a Monday or Friday, due to the fact that on such days people are more willing to take one day off and have longer holidays. As an example, in Figure 32, Fridays are plotted for eight consequent weeks in order to point our decrease in electricity load day after Ascension Day. Hence, another variable named, *'Bridge day'* is added to the data set to denote these special days.

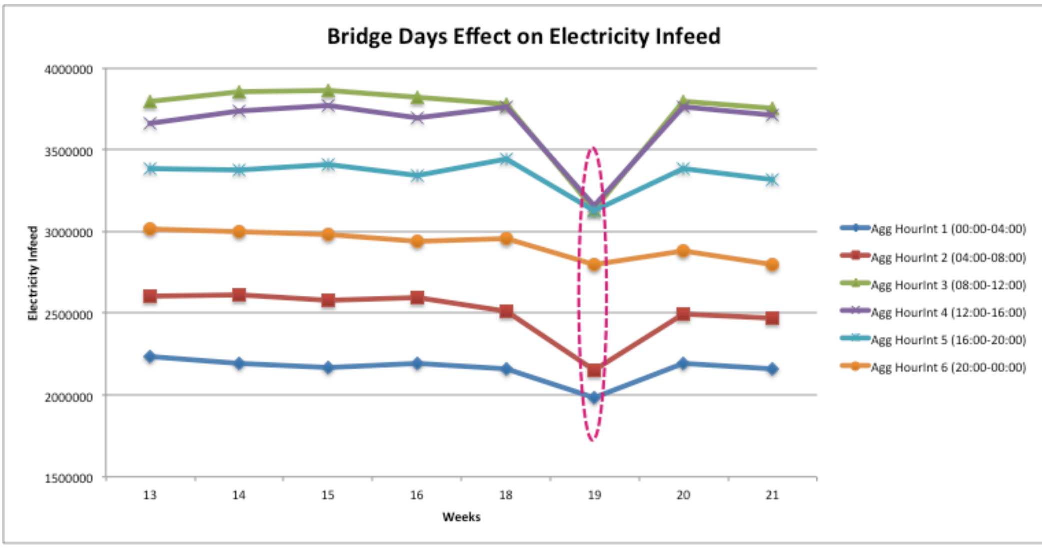


Figure 32. Electricity infeed (GWh) between weeks 13 and 21 on Fridays in Brabant in 2010.

In this section we have provided a summary of important special days analyses. Figures and graphs on the other special days are provided in Appendix A.

Analyses on special days effect on electricity load has shown us the concrete deviation from regular load pattern on these days. We have grouped the special days into two types: long holidays and short holidays. Long holidays that last at least one week, that are school holidays and Bouwvak. Due to taking longer time, these holidays have a leveling effect on the load. On the other hand, short holidays, which are the special days other than School Holidays and Bouwvak, have effects of abrupt drops in load. Short holidays result in changes in the general shape of load data. Due to such effects, special days, especially short holidays, are referred as most difficult times of the year to forecast electricity load. Therefore special attention needs to be paid on these days.

5.2. Evaluation Criteria for the Models

Prediction accuracies are computed by comparing the forecasts to actual values of the respective period. Our main performance indicators are mean average percent error (MAPE) and maximum average percent error (MaxAPE). Formulation of MAPE and MaxAPE are shown in (5.1) and (5.2) respectively.

$$MAPE = \left(\frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_{t-k}(k) - y_t}{y_t}\right|\right) * 100 \qquad (5.1)$$

$$MaxAPE = max_t\left(\left|\frac{\hat{y}_{t-k}(k) - y_t}{y_t}\right|\right) * 100 \qquad (5.2)$$

where $\hat{y}_{t-k}(k)$ stands for the forecasted load for time t, with a lead time k. The actual load value of time t is abbreviated by y_t. In MAPE, the proportion of error term to actual value is averaged, whereas in MaxAPE, maximum is searched.

5.3. Methods for Short-term Load Forecasting

In short-term load forecasting we present and compare a parametric and non-parametric method: modification of Holt Winters Exponential Smoothing and Nonlinear Autoregressive with eXogenous Inputs (NARX), respectively.

In this part of the study only historic data and calendar variables are used to forecast short-term electricity load. Taylor [23] also pointed out, meteorological variables have distinct effect on electricity demand in the long-term. However, for short-term forecasts, these variables can be excluded, as consumers' adaptation to changes in these variables takes some time; it is difficult to observe their effects and there are no immediately available forecasts for these variables.

5.3.1. Modified Holt Winters Exponential Smoothing

Considering the model specifications presented in (4. 9) - (4. 14), in order to model time series data with HWT method, initial values for level and seasonal components and smoothing parameters need to be estimated. For initialization of the state variables in the model, i.e., for I_t, d_t, w_t, we use a similar approach to the ones followed in [1] and [50]. We used two-week intervals that do not include any special days in order to prevent any unusual observations causing misleading fluctuations in initialization. The algorithm is presented in Figure 33.

In our model, lag value ($\pm L$), presented in (4.9), is set equal to 20, meaning that we searched for an interval of 40 days in the previous year, in order to find the effects of the respective special days. Lastly, model parameters, i.e. α, and ω, are derived by using a similar method in [44] [51] , as presented in Figure 34.

Different than Taylor's [44] approach, in our models, we derive best parameters for different forecasting horizons each time. In Table 1 model parameters for Brabant data set are presented, for the parameter values of the other regions, Appendix B should be referred.

Initialization:	Given two weeks electricity load data, $c_1 = 24$, $c_2 = 168$, referring to daily cycle and weekly cycle
Step 1	Initialize the level, , by averaging load data over the two week period.
Step 2	Initialize daily seasonal state variable, :
Step 2.1	Compute c_1 point centered moving averages
Step 2.2	Calculate differences between actual data and moving average values, that roughly gives seasonality effect
Step 2.3	Calculate average differences for c_1 hours of the day over two weeks
Step 3	Initialize of the weekly seasonal index, :
Step 3.1	Compute c_2 point centered moving averages
Step 3.2	Calculate differences between actual data and moving average values, that roughly gives seasonality effect
Step 3.3	Calculate average differences for c_2 hours of the week over two weeks

Figure 33. The algorithm for initialization of the state variables.

Iteration:	
Step 1	derive 10^5 vectors of four parameters which are uniformly distributed between 0 and 1
Step 2	for every vector, compute sum of squared errors (SSE) of the training data set
Step 3	define 10 vectors with the lowest SSEs as the set of possible model parameters
Step 4	generate all possible combinations of selected 10 vectors
Step 5	for every combination vector, compute sum of squared errors (SSE) of the training data set
Step 6	elements of the vector with lowest SSE are assigned as model parameters

Figure 34. The algorithm for derivation of the parameters.

Table 1 Parameters for models of province Brabant data.

Lead Times (hrs)	α_{best}	δ_{best}	ω_{best}	ϕ_{best}
1	0.5487	0.1832	0.2658	0.3399
6	0.0377	0.2354	0.1590	0.7750
12	0.0145	0.1820	0.1769	0.8330
24	0.0005	0.2649	0.1066	0.9034
48	0.0007	0.0715	0.0768	0.7265
168	0.0001	0.0225	0.1713	0.8379

HWT and m-HWT models of five regions for six different lead times are modeled on Microsoft Excel 2011. For the modification part, we have utilized Macro coding that searches for the respective special day in the lagged period of previous year and sets the correction factor.

5.3.2. Nonlinear Autoregressive with Exogenous Input Neural Networks

Computational study in NARX modeling consists of two stages: (1) training the model and (2) testing the model performance. Accordingly, each data set is divided into training (in-sample data), validation and test (out of sample data) data sets corresponding to 70%, 15% and 15% of the data, respectively. In addition to the historic electricity load data, special days are defined as binary variables and added to the data set.

Data analyses revealed that effects of some special days on load are very similar to each other. Therefore these special days are grouped together in order to decrease dimensionality of the problem and to enable the model to capture the difference in effects of special days on different weekdays. Consequently, calendar variables are grouped in seven categories as the following:

- Easter, Whit Monday and Liberalization day
- Carnival
- Christmas Eve and New Year's Eve
- Queens Day
- Boxing Day and Christmas Day
- First day of year (New Year Holiday)
- Ascension Day

As mentioned before, in addition to exact dates of special days, one-day prior and after these days and bridge days are also defined as variables and included in the model. Furthermore, *day of the week, hour of the day* and binary *summer time* variables are defined as inputs.

As mentioned in Section 4.2, a neural network model consists of three types of layers: input, hidden and output. Corresponding to the input variables, thirteen input nodes exist in our model's input layer. Additionally, input layer contains nodes of the feedback loop. Feedback delays are determined with respect to the autocorrelation values between different lags of infeed data. In this study autocorrelation values between 1 and 0,8 are identified as highly correlated. Hence, among these lag values, considering the seasonalities, the rational ones are selected as feedback delays, which are in our case: *1, 2, 23, 24, 25, 168 (1 week)* and *169 hours.* Therefore in total input layer consists of twenty input nodes. Obviously, output layer contains only one node which gives the electricity infeed forecast.

[52] states that one hidden layer architectures are sufficient for solving most of the forecasting problems but with a disadvantage of higher training times. Therefore, in the interest of keeping model architecture search to a reasonable limit, in this study single hidden layer recurrent networks are considered as candidate model architecture. For various lead times at every region, best performing architecture is searched. Due to the computational limitations, number of hidden nodes is kept between 5 and 85, tested with 5

node intervals, because it is believed that the networks with hidden nodes less than 5, would not be capable of modeling and learning the data and the networks with hidden nodes more than 80 are expected to fail at generalization and result in overfitting. Therefore, for every region, and every lead time 17 different architectures are run with 5 different initializations. The best architectures that are assigned to the models of every forecast horizon in Brabant are presented in Table 2. For each lead time best architectures are run 10 times to find the best weight and bias values and complete the network architecture. NARX neural network architectures for other regions can be found in Appendix C.

Table 2 Brabant model architectures for different forecasting horizons for Brabant data.

Lead Times (hrs)	Hidden Layer	Number of Hidden Nodes
1	1	30
6	1	35
12	1	75
24	1	35
48	1	30
168	1	35

Other components of neural networks, that are effective on network performance, are the training algorithm and transfer function. In this study, 'Levenberg-Marquardt' training algorithm, which is a modification of popular back-propagation algorithm, is used. Levenberg-Marquardt algorithm includes an approximation to Newton's method, which is considered to be more efficient up to a few hundred nodes [53]. In hidden layers, tangent sigmoid (tansig) function and in output layer linear transfer function (purelin) are used.

We have modeled the networks on MATLAB software, version 2012b. The training parameters are set to the default values, which are provided in Table 3.

Table 3 Training parameters of the NARX models used in this study.

Lead Times (hrs)	Hidden Layer
Maximum number of epochs to train	10
Performance goal	0
Learning rate	0.01
Maximum validation failures	5
Maximum performance increase	1.04
Minimum performance gradient	1.00E-10
Maximum time to train in seconds	inf

Considering all the effective components and different initializations of the network, a neural network is trained for various lead times for each region.

5.4. Results for Short-term Load Forecasting

In this section we present performances of all methods for short-term load forecasting. We first compared the performances of Taylor's HWT and our modified HWT, called m-HWT. Following the mentioned comparison, performance of NARX is compared to these two methods. Post-sample accuracies are measured in terms of mean absolute percentage error (MAPE) and maximum percent error (MaxAPE) for lead times up to one week ahead in five provinces. Post-sample data accounts for 15% of the total data set and corresponds to the period between 06.03.2012 and 30.11.2012. In Table 4 results are presented for Brabant data set. The performances are measured not only in terms of average forecasting performance, but also with respect to different day types.

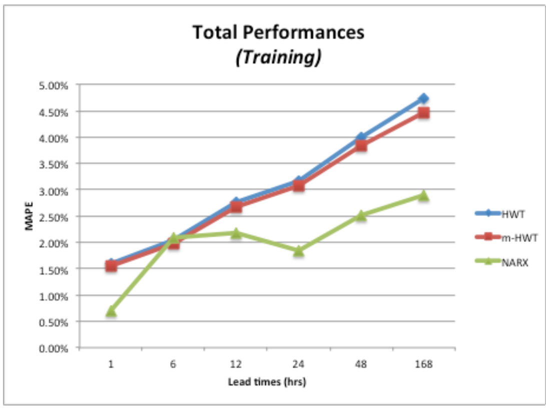
Table 4 Performances of NARX, HWT and m-HWT for different day types in terms of MAPE for Brabant data set.

	Lead times													
	1		6		12		24		48		168		Avg	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total														
HWT	1.59%	1.44%	2.05%	1.81%	2.76%	2.64%	3.18%	2.86%	4.01%	3.45%	4.74%	4.13%	3.05%	2.72%
m-HWT	1.55%	1.36%	1.99%	1.68%	2.68%	2.43%	3.08%	2.62%	3.84%	3.11%	4.47%	3.63%	2.93%	2.47%
NARX	0.71%	0.75%	2.09%	2.46%	2.17%	2.29%	1.84%	2.06%	2.52%	2.75%	2.89%	2.85%	2.04%	2.19%
Special														
HWT	2.01%	1.55%	2.94%	2.26%	4.36%	3.34%	5.30%	4.11%	6.85%	5.18%	7.87%	5.90%	4.89%	3.72%
m-HWT	1.82%	1.26%	2.65%	1.76%	4.00%	2.58%	4.75%	3.23%	5.97%	3.95%	6.58%	4.09%	4.30%	2.81%
NARX	0.80%	0.74%	2.16%	2.35%	2.44%	2.42%	2.18%	2.28%	3.01%	3.08%	3.57%	3.25%	2.36%	2.35%
Long Holidays														
HWT	1.58%	1.24%	2.01%	1.53%	2.96%	2.06%	3.32%	2.62%	4.53%	3.73%	5.57%	4.49%	3.33%	2.61%
m-HWT	1.52%	1.13%	2.00%	1.48%	2.99%	2.02%	3.22%	2.44%	4.13%	3.08%	4.71%	3.27%	3.10%	2.24%
NARX	1.33%	1.50%	2.57%	2.26%	2.97%	2.08%	2.51%	2.45%	3.05%	2.88%	4.78%	3.27%	2.87%	2.41%
Short Holidays														
HWT	4.10%	4.97%	7.49%	10.49%	11.23%	17.82%	14.98%	21.00%	18.19%	21.58%	19.09%	21.84%	12.51%	16.28%
m-HWT	3.25%	2.74%	5.81%	4.92%	8.95%	8.89%	12.25%	12.15%	14.95%	13.74%	15.70%	13.40%	10.15%	9.31%
NARX	0.76%	0.70%	2.17%	2.37%	2.46%	2.45%	2.19%	2.28%	3.04%	3.10%	3.56%	3.25%	2.36%	2.36%
Non-Spec														
HWT	1.48%	1.40%	1.80%	1.65%	2.29%	2.37%	2.58%	2.39%	3.22%	2.79%	3.86%	3.46%	2.54%	2.34%
m-HWT	1.48%	1.40%	1.80%	1.65%	2.29%	2.37%	2.58%	2.39%	3.22%	2.79%	3.86%	3.46%	2.54%	2.34%
NARX	0.67%	0.75%	2.08%	2.50%	2.07%	2.23%	1.72%	1.94%	2.35%	2.60%	2.65%	2.65%	1.92%	2.11%

In Table 4, both training and testing data performances are presented, because as the testing data refers to the period after 6th of March, it misses some of the special days, hence giving some idea to the reader about the performances on these days is aimed. It is remarkable that in general, training performance is smaller than testing performance for NARX, on the other hand, the opposite is the case for HWT and m-HWT. Observing larger errors in testing is an expected outcome for neural networks. For the case of HWT and m-HWT, larger errors in training data may be caused by two reasons. First, in HWT, as the forecasts are generated the state variables are updated with respect to the error terms. Therefore there occur larger errors in the beginning. Secondly, for application of adaptation, we need forecast of a full year, so the modification is applied after forecast generation of one year, forecast performance is worse for the first year. We believe evaluation of performances in terms of testing data set would reflect the real performances better.

Table 4 clearly shows that our m-HWT has improved the forecasting performance over Taylor's HWT. The improvement is around 8% on the average for five data sets, 9% in Brabant data set. As the modification is for special days, its effect is much more significant in terms of special days forecasting. In Brabant data set, Table 4 presents that the improvement varies between 16% and 30% for different lead times. When we consider the short holidays, the improvement is even more significant. It reaches to 53% in 6 hours ahead forecasting. Normal day forecast performances are identical as our modification applies only to the special days. The performance of our m-HWT even exceeds the performance of NARX in 6 hours ahead forecasting. For illustrative purposes we plot the performance of three methods for the region Brabant data set, for different day types in Figure 35 to Figure 44. In Figure 42 it is clear that the correction factor has clearly improved model performance of HWT on short holidays. The improvement is less on long holidays and Figure 36 shows that the improvement increases as the lead time gets larger.

Figure 36 presents that NARX performs better than m-HWT method in almost all lead times. We observe that especially for one-hour ahead forecasting NARX is quite effective and performs with a MAPE of as low as 0,71%. However NARX performance deteriorates in 6 hours ahead forecasting. Additionally, NARX is superior to m-HWT in short holidays' load forecasting, but not as much effective in forecasting load on long holidays. m-HWT is competitive with NARX in forecasting load on long holidays, even outperforms in 6 hours ahead forecasting, as graphed in Figure 40.



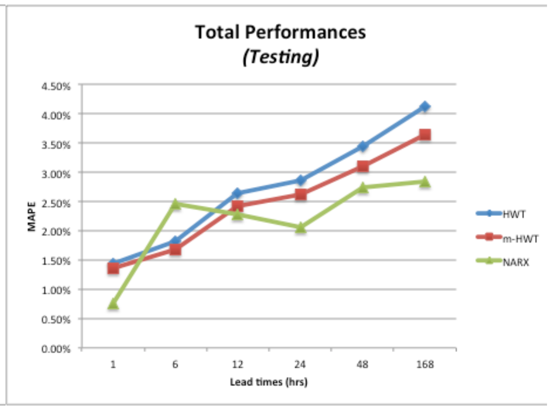
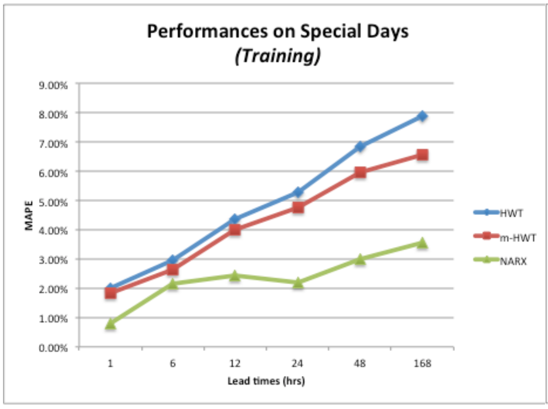


Figure 35. Training performances of three models for different forecast horizons in Brabant data set.

Figure 36. Testing performances of three models for different forecast horizons in Brabant data set.



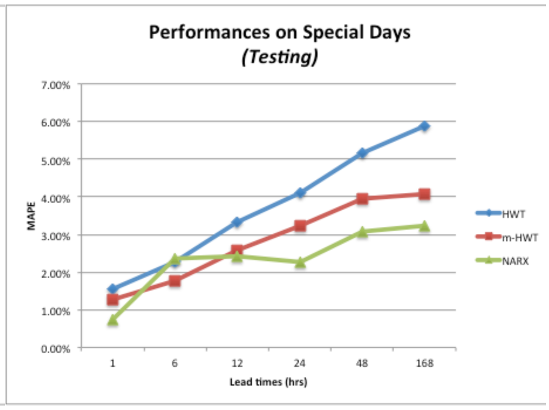
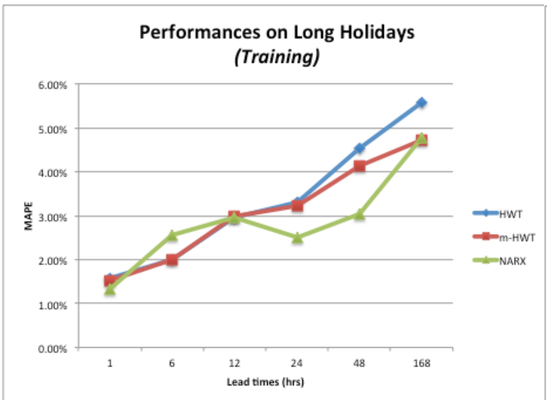


Figure 37. Training performances of three models on special days for different forecast horizons in Brabant data set.

Figure 38. Testing performances of three models on special days for different forecast horizons in Brabant data set.



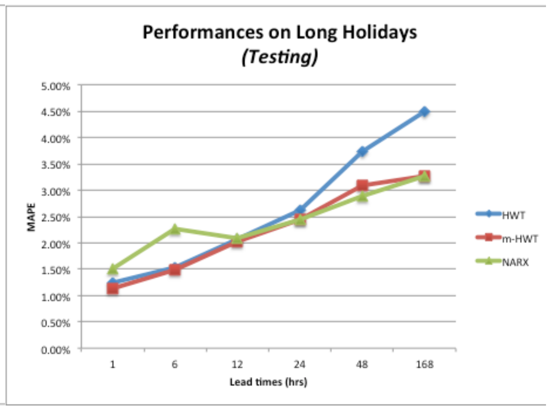


Figure 39. Training performances of three models on long holidays for different forecast horizons in Brabant data set.

Figure 40. Testing performances of three models on long holidays for different forecast horizons in Brabant data set.

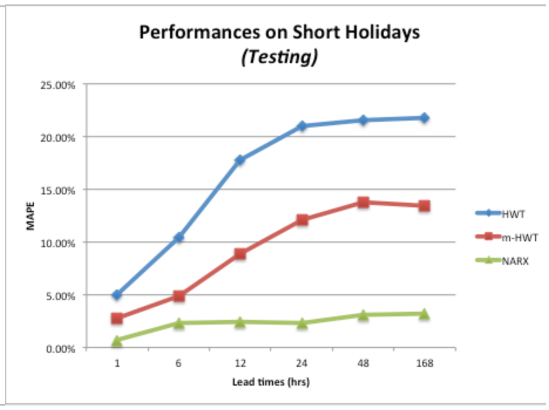


Figure 41. Training performances of three models on short holidays for different forecast horizons in Brabant data set.

Figure 42. Testing performances of three models on long holidays for different forecast horizons in Brabant data set.

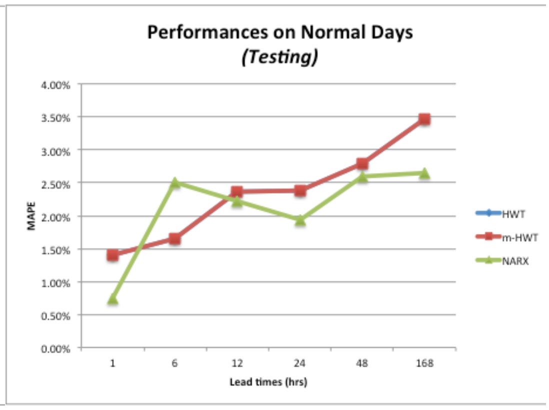


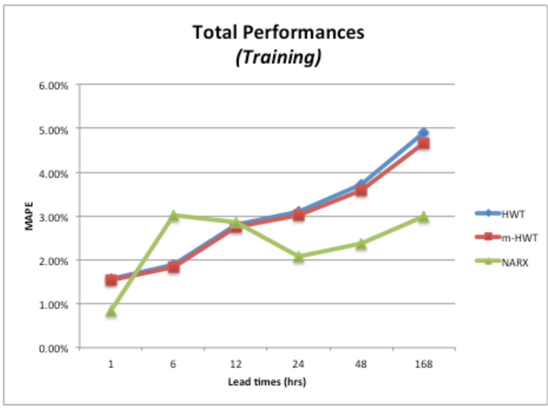
Figure 43. Training performances of three models on normal days for different forecast horizons in Brabant data set.

Figure 44. Testing performances of three models on normal days for different forecast horizons in Brabant data set.

In Table 5 performances of the methods on Noord data set is presented. Similar to results of Brabant data set, m-HWT outperforms HWT. NARX is more effective in forecasting short holidays, over all special days m-HWT and NARW are competitive but NARX outperforms, as the lead times get larger. As we claimed before, NARX performance worsens in forecasting with lead times 6 and 12 hours. m-HWT improved HWT 8% on the average. NARX improved HWT around 10%. Over forecasting short holidays, m-HWT performs 40% better than HWT on the average for all lead times, whereas for NARX the improvement is 83%. Similar to Brabant data set, we have provided figures for also Noord data set in Figure 45 to Figure 54. Results are materially similar for other forecasting horizons and regions. Therefore, we will not explain them in detail, but the MAPE values for these regions are presented in Table 6, Table 7 and Table 8 and the figures can be found in Appendix D.

Table 5 Performances of NARX, HWT and m-HWT for different day types in terms of MAPE for Noord data set.

	Lead times													
	1		6		12		24		48		168		Avg	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total														
HWT	1.55%	1.42%	1.89%	1.91%	2.82%	2.78%	3.10%	2.96%	3.71%	3.20%	4.90%	4.71%	3.00%	2.83%
m-HWT	1.53%	1.34%	1.84%	1.79%	2.75%	2.61%	3.01%	2.75%	3.59%	2.95%	4.67%	4.12%	2.90%	2.59%
NARX	0.85%	0.90%	3.02%	3.30%	2.87%	3.13%	2.07%	2.36%	2.38%	2.56%	2.99%	2.97%	2.36%	2.54%
Special														
HWT	2.03%	1.53%	2.99%	2.53%	4.86%	4.10%	5.49%	4.69%	6.63%	5.27%	7.82%	6.82%	4.97%	4.16%
m-HWT	1.90%	1.25%	2.77%	2.07%	4.55%	3.47%	5.11%	3.91%	6.04%	4.31%	6.73%	4.66%	4.52%	3.28%
NARX	0.97%	0.95%	3.29%	3.17%	3.18%	3.10%	2.53%	2.49%	2.89%	2.79%	3.78%	3.81%	2.77%	2.72%
Long Holidays														
HWT	1.74%	1.25%	2.36%	1.88%	3.70%	2.90%	4.17%	3.34%	5.20%	3.93%	6.35%	5.50%	3.92%	3.13%
m-HWT	1.71%	1.14%	2.36%	1.81%	3.72%	2.84%	4.16%	3.23%	5.04%	3.59%	5.84%	3.84%	3.81%	2.74%
NARX	1.85%	1.81%	4.17%	2.89%	4.13%	2.66%	3.76%	2.14%	3.78%	1.98%	4.98%	3.00%	3.78%	2.41%
Short Holidays														
HWT	4.28%	4.72%	7.98%	9.97%	14.00%	17.90%	15.86%	20.28%	17.90%	20.78%	19.35%	22.08%	13.23%	15.95%
m-HWT	3.35%	2.51%	6.01%	5.07%	11.07%	10.75%	12.55%	11.76%	13.98%	12.63%	13.81%	14.03%	10.13%	9.46%
NARX	0.93%	0.89%	3.29%	3.18%	3.18%	3.12%	2.54%	2.53%	2.91%	2.85%	3.78%	3.87%	2.77%	2.74%
Non-Spec														
HWT	1.44%	1.40%	1.60%	1.74%	2.26%	2.39%	2.45%	2.43%	2.91%	2.55%	4.09%	4.03%	2.46%	2.42%
m-HWT	1.44%	1.40%	1.60%	1.74%	2.26%	2.39%	2.45%	2.43%	2.91%	2.55%	4.09%	4.03%	2.46%	2.42%
NARX	0.81%	0.89%	2.92%	3.38%	2.76%	3.17%	1.90%	2.30%	2.19%	2.60%	2.72%	2.60%	2.22%	2.49%



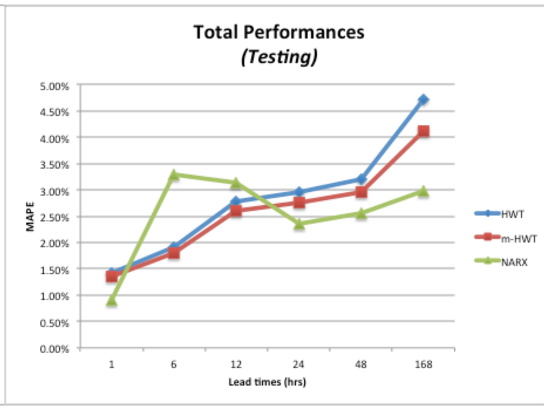


Figure 45. Training performances of three models for different forecast horizons in Noord data set.

Figure 46. Testing performances of three models for different forecast horizons in Noord data set.

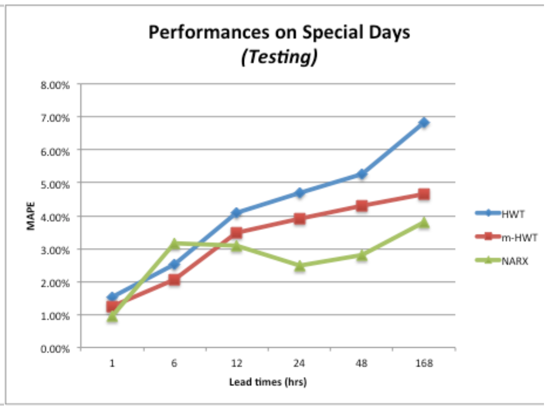
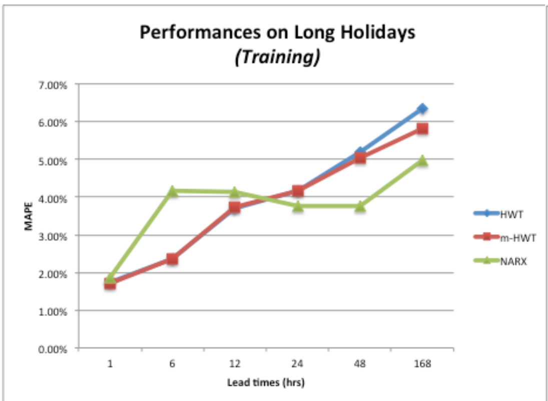


Figure 47. Training performances of three models on special days for different forecast horizons in Noord data set.

Figure 48. Testing performances of three models on special days for different forecast horizons in Noord data set.



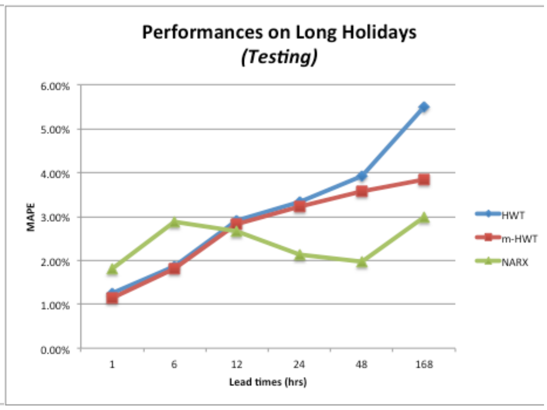


Figure 49. Training performances of three models on long holidays for different forecast horizons in Noord data set.

Figure 50. Testing performances of three models on long holidays for different forecast horizons in Noord data set.

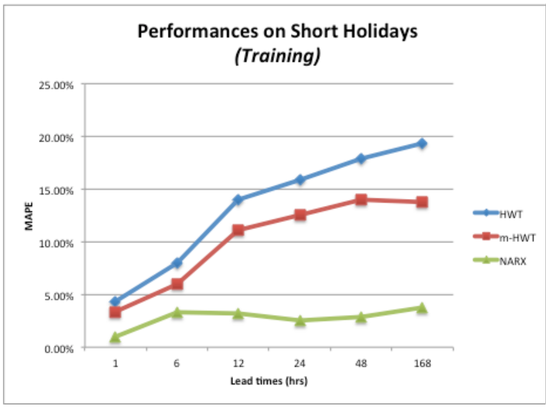


Figure 51. Training performances of three models on short holidays for different forecast horizons in Noord data set.

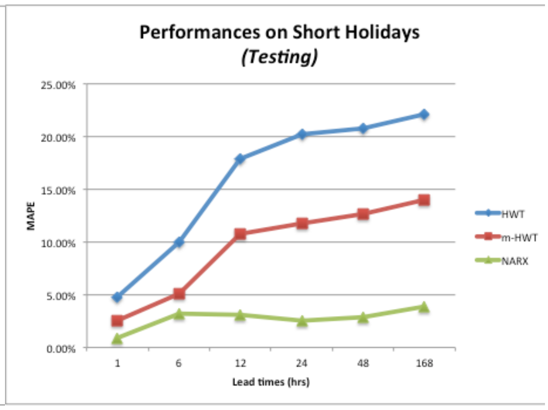


Figure 52. Testing performances of three models on long holidays for different forecast horizons in Noord data set.

Figure 53. Training performances of three models on normal days for different forecast horizons in Noord data set.

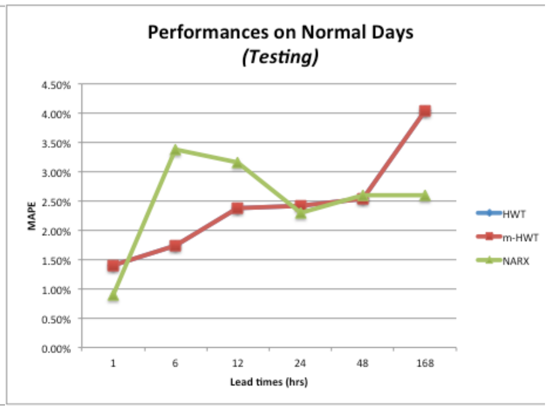


Figure 54. Testing performances of three models on normal days for different forecast horizons in Noord data set.

Table 6 Performances of NARX, HWT and m-HWT for different day types in terms of MAPE for Limburg data set.

	Lead times													
	1		6		12		24		48		168		Avg	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total														
HWT	2.12%	1.95%	2.57%	2.06%	3.71%	2.49%	3.03%	2.65%	4.64%	3.27%	4.75%	3.76%	3.47%	2.70%
m-HWT	2.06%	1.82%	2.57%	1.96%	3.67%	2.32%	2.92%	2.43%	4.53%	2.98%	4.48%	3.39%	3.37%	2.48%
NARX	0.79%	0.86%	2.62%	2.86%	2.25%	2.34%	1.84%	2.05%	2.97%	3.34%	2.85%	3.06%	2.22%	2.42%
Special														
HWT	3.14%	2.21%	3.38%	2.32%	5.30%	3.18%	5.25%	3.72%	7.52%	4.71%	8.55%	5.80%	5.52%	3.66%
m-HWT	2.78%	1.70%	3.31%	1.94%	5.04%	2.56%	4.61%	2.89%	6.75%	3.63%	7.18%	4.45%	4.95%	2.86%
NARX	0.86%	0.80%	2.81%	2.89%	2.50%	2.41%	2.22%	2.21%	3.56%	3.78%	3.46%	3.10%	2.57%	2.53%
Long Holidays														
HWT	2.25%	1.61%	2.65%	1.78%	3.91%	2.17%	3.49%	2.51%	5.29%	3.50%	6.28%	4.80%	3.98%	2.73%
m-HWT	2.17%	1.49%	2.73%	1.75%	3.89%	2.12%	3.36%	2.36%	5.04%	3.01%	5.54%	3.87%	3.79%	2.43%
NARX	1.26%	1.64%	3.57%	2.97%	3.29%	2.45%	2.74%	2.67%	3.68%	2.74%	4.20%	2.24%	3.12%	2.45%
Short Holidays														
HWT	7.52%	9.00%	6.95%	8.49%	12.08%	14.63%	13.86%	17.46%	18.38%	18.51%	19.62%	17.11%	13.07%	14.20%
m-HWT	5.72%	4.15%	6.16%	4.15%	10.66%	7.48%	10.74%	8.95%	15.10%	10.73%	15.19%	11.12%	10.59%	7.76%
NARX	0.85%	0.79%	2.81%	2.89%	2.51%	2.42%	2.23%	2.20%	3.59%	3.86%	3.47%	3.16%	2.58%	2.55%
Non-Spec														
HWT	1.87%	1.86%	2.37%	1.96%	3.30%	2.22%	2.45%	2.25%	3.92%	2.73%	3.76%	2.99%	2.95%	2.33%
m-HWT	1.87%	1.86%	2.37%	1.96%	3.30%	2.22%	2.45%	2.25%	3.92%	2.73%	3.76%	2.99%	2.95%	2.33%
NARX	0.77%	0.88%	2.55%	2.85%	2.17%	2.32%	1.70%	1.98%	2.75%	3.15%	2.64%	3.00%	2.10%	2.36%

Table 7 Performances of NARX, HWT and m-HWT for different day types in terms of MAPE for Maastricht data set.

	Lead times													
	1		6		12		24		48		168		Avg	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total														
HWT	1.86%	1.75%	2.36%	2.23%	3.52%	3.23%	3.90%	3.47%	4.67%	3.90%	5.41%	4.86%	3.62%	3.24%
m-HWT	1.82%	1.69%	2.30%	2.10%	3.43%	3.05%	3.76%	3.20%	4.47%	3.46%	5.11%	4.05%	3.48%	2.93%
NARX	1.06%	1.13%	2.89%	3.41%	2.89%	3.30%	2.37%	2.59%	3.17%	3.36%	3.33%	3.35%	2.62%	2.86%
Special														
HWT	2.35%	1.82%	3.51%	2.63%	5.69%	4.03%	6.93%	4.89%	8.56%	5.99%	9.68%	7.18%	6.12%	4.42%
m-HWT	2.14%	1.58%	3.17%	2.19%	5.16%	3.35%	6.08%	3.93%	7.26%	4.40%	8.01%	4.25%	5.30%	3.28%
NARX	1.13%	1.18%	3.19%	3.61%	3.20%	3.62%	2.84%	3.10%	3.79%	3.84%	4.21%	4.02%	3.06%	3.23%
Long Holidays														
HWT	1.89%	1.56%	2.57%	2.00%	4.02%	2.83%	4.75%	3.35%	5.89%	4.45%	6.94%	5.62%	4.34%	3.30%
m-HWT	1.83%	1.48%	2.54%	1.96%	3.96%	2.77%	4.54%	3.15%	5.40%	3.58%	6.02%	3.48%	4.05%	2.74%
NARX	1.46%	1.65%	3.85%	3.35%	3.64%	3.34%	3.34%	2.79%	4.15%	2.64%	4.80%	2.75%	3.54%	2.75%
Short Holidays														
HWT	4.60%	4.75%	8.07%	9.77%	13.82%	17.56%	17.60%	22.30%	21.62%	23.47%	23.09%	24.82%	14.80%	17.11%
m-HWT	3.66%	2.65%	6.27%	4.82%	11.03%	9.92%	13.58%	12.74%	16.36%	13.71%	17.70%	13.01%	11.43%	9.48%
NARX	1.11%	1.16%	3.21%	3.63%	3.22%	3.63%	2.86%	3.13%	3.82%	3.94%	4.22%	4.11%	3.07%	3.27%
Non-Spec														
HWT	1.74%	1.73%	2.07%	2.07%	2.97%	2.93%	3.13%	2.93%	3.70%	3.11%	4.32%	3.98%	2.99%	2.79%
m-HWT	1.74%	1.73%	2.07%	2.07%	2.97%	2.93%	3.13%	2.93%	3.70%	3.11%	4.32%	3.98%	2.99%	2.79%
NARX	1.04%	1.10%	2.79%	3.33%	2.78%	3.20%	2.20%	2.37%	2.94%	3.09%	3.02%	3.08%	2.46%	2.70%

Table 8 Performances of NARX, HWT and m-HWT for different day types in terms of MAPE for Friesland data set.

	Lead times													
	1		6		12		24		48		168		Avg	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total														
HWT	1.96%	1.81%	3.74%	2.64%	3.42%	3.15%	3.45%	3.15%	3.85%	3.45%	4.61%	3.86%	3.50%	3.01%
m-HWT	1.94%	1.75%	4.03%	2.52%	3.38%	2.95%	3.38%	2.94%	3.74%	3.19%	4.49%	3.53%	3.49%	2.81%
NARX	1.19%	1.25%	2.90%	3.46%	3.38%	3.81%	2.87%	2.85%	2.95%	3.09%	2.89%	3.30%	2.70%	2.96%
Special														
HWT	2.28%	1.85%	5.38%	3.44%	5.67%	4.74%	5.97%	4.95%	6.83%	5.36%	7.82%	6.16%	5.66%	4.42%
m-HWT	2.17%	1.62%	6.77%	2.99%	5.49%	3.96%	5.65%	4.13%	6.23%	4.39%	7.02%	4.93%	5.55%	3.67%
NARX	1.35%	1.40%	3.21%	3.26%	3.83%	3.71%	3.35%	3.00%	3.56%	3.56%	3.47%	3.47%	3.13%	3.07%
Long Holidays														
HWT	1.81%	1.53%	4.73%	2.63%	4.20%	3.19%	4.33%	3.24%	5.02%	3.60%	5.95%	4.71%	4.34%	3.15%
m-HWT	1.82%	1.44%	5.94%	2.60%	4.33%	3.08%	4.38%	3.15%	4.93%	3.32%	5.75%	3.85%	4.53%	2.91%
NARX	2.49%	2.50%	3.83%	3.18%	4.94%	3.71%	4.84%	2.37%	4.44%	3.07%	4.64%	3.29%	4.20%	3.02%
Short Holidays														
HWT	5.23%	5.47%	10.48%	12.71%	17.23%	22.62%	18.87%	24.63%	21.14%	25.58%	22.55%	22.83%	15.92%	18.97%
m-HWT	4.35%	3.63%	13.29%	7.45%	14.63%	14.02%	15.63%	15.49%	16.45%	16.77%	16.98%	17.29%	13.56%	12.44%
NARX	1.28%	1.34%	3.21%	3.26%	3.82%	3.71%	3.34%	3.05%	3.58%	3.60%	3.46%	3.52%	3.12%	3.08%
Non-Spec														
HWT	1.87%	1.80%	3.28%	2.42%	2.83%	2.70%	2.80%	2.63%	3.09%	2.89%	3.81%	3.15%	2.95%	2.60%
m-HWT	1.87%	1.80%	3.28%	2.42%	2.83%	2.70%	2.80%	2.63%	3.09%	2.89%	3.81%	3.15%	2.95%	2.60%
NARX	1.13%	1.18%	2.79%	3.55%	3.21%	3.84%	2.71%	2.76%	2.74%	2.90%	2.70%	3.16%	2.55%	2.90%

In summary, we observed NARX's superiority in forecasting load on short holidays over HWT and m-HWT. Also deterioration occurs in NARX's ability to forecast load for lead times that are not multiples of 24 hours (one day). m-HWT has improved performance of HWT significantly, however in general NARX is more effective in forecasting special days. However, NARX performs worse in load forecasting on long holidays compared to its performance on short holidays. m-HWT is competitive with NARX, even outperforms in some regions especially in 6 and 12 hours ahead forecasting.

In order to provide a deeper insight on superiority of methods over HWT, we have reported the relative percentages (RPs) of the models' MAPEs over all lead times. The value is obtained by subtracting respective method's MAPE value from HWT's MAPE and dividing it by HWT's MAPE. Therefore RP values larger than zero, refers to an improvement compared to HWT and negative values refer to deterioration.

We have also carried out stability analysis in order to evaluate if NARX's performance is stable in both training and testing data sets. The stability values for 5 data sets are presented in Table 9.

Table 9 Stability analysis for NARX in five data sets.

	Brabant	Limburg	Maastricht	Friesland	Noord
Lead times					
1	94.67%	91.86%	93.81%	95.20%	94.44%
6	84.96%	91.61%	84.75%	83.82%	91.52%
12	94.76%	96.15%	87.58%	88.71%	91.69%
24	89.32%	89.76%	91.51%	99.30%	87.71%
48	91.64%	88.92%	94.35%	95.47%	92.97%
168	98.62%	93.14%	99.40%	87.58%	99.33%

Table 10 Relative Percentages of m-HWT and NARX with respect to HWT in Brabant data set.

	Lead times											
	1		6		12		24		48		168	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total												
m-HWT	2.47%	5.37%	2.93%	7.60%	2.58%	7.90%	3.17%	8.41%	4.24%	9.83%	5.74%	12.01%
NARX	55.41%	47.81%	-1.97%	-35.64%	21.24%	13.10%	42.06%	28.05%	37.11%	20.21%	39.08%	31.00%
Special												
m-HWT	9.34%	18.20%	9.94%	22.21%	8.29%	22.71%	10.28%	21.35%	12.85%	23.81%	16.36%	30.64%
NARX	60.11%	52.11%	26.48%	-4.06%	44.06%	27.49%	58.86%	44.51%	56.07%	40.56%	54.64%	44.88%
Long Holidays												
m-HWT	3.26%	8.78%	0.41%	3.57%	-1.07%	1.79%	2.97%	6.62%	8.80%	17.41%	15.38%	27.20%
NARX	15.62%	-20.66%	-28.14%	-47.51%	-0.48%	-0.98%	24.35%	6.42%	32.68%	22.88%	14.20%	27.16%
Short Holidays												
m-HWT	20.75%	44.91%	22.39%	53.07%	20.32%	50.12%	18.19%	42.16%	17.78%	36.34%	17.75%	38.65%
NARX	81.48%	85.90%	71.03%	77.40%	78.09%	86.25%	85.38%	89.14%	83.28%	85.64%	81.36%	85.12%
Non-Spec												
m-HWT	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NARX	54.68%	46.28%	-15.75%	-51.96%	9.68%	5.87%	33.24%	18.88%	26.93%	6.78%	31.29%	23.45%

Table 11 Relative Percentages of m-HWT and NARX with respect to HWT in Limburg data set.

	Lead times											
	1		6		12		24		48		168	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total												
m-HWT	3.01%	7.06%	0.23%	5.09%	1.10%	6.88%	3.58%	8.55%	2.42%	9.07%	5.72%	9.81%
NARX	62.77%	55.99%	-1.76%	-38.79%	39.34%	5.88%	39.20%	22.71%	36.02%	-2.05%	40.05%	18.65%
Special												
m-HWT	11.65%	22.77%	1.93%	16.44%	4.89%	19.60%	12.21%	22.21%	10.15%	22.94%	16.02%	23.18%
NARX	72.64%	63.75%	16.80%	-24.40%	52.80%	24.18%	57.75%	40.57%	52.63%	19.81%	59.53%	46.53%
Long Holidays												
m-HWT	3.17%	7.40%	-3.17%	1.81%	0.56%	2.17%	3.79%	5.90%	4.71%	14.01%	11.83%	19.46%
NARX	43.91%	-1.99%	-34.95%	-66.96%	15.81%	-13.01%	21.52%	-6.54%	30.45%	21.63%	33.14%	53.32%
Short Holidays												
m-HWT	24.01%	53.89%	11.42%	51.17%	11.73%	48.86%	22.56%	48.74%	17.80%	42.04%	22.57%	35.00%
NARX	88.70%	91.22%	59.58%	65.97%	79.22%	83.46%	83.91%	87.40%	80.46%	79.15%	82.31%	81.54%
Non-Spec												
m-HWT	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NARX	58.72%	52.64%	-7.63%	-45.32%	34.21%	-4.32%	30.74%	11.94%	29.91%	-15.49%	29.82%	-0.32%

Table 12 Relative Percentages of m-HWT and NARX with respect to HWT in Maastricht data set.

	Lead times											
	1		6		12		24		48		168	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total												
m-HWT	2.10%	3.80%	2.45%	5.50%	2.42%	5.73%	3.41%	7.61%	4.47%	11.22%	5.71%	16.56%
NARX	43.04%	35.54%	-22.61%	-53.17%	17.84%	-2.07%	39.15%	25.31%	32.19%	13.80%	38.50%	31.06%
Special												
m-HWT	8.97%	13.34%	9.45%	16.92%	9.26%	16.74%	12.37%	19.66%	15.20%	26.57%	17.29%	40.81%
NARX	51.91%	35.15%	9.00%	-37.10%	43.71%	10.11%	59.04%	36.56%	55.74%	35.90%	56.52%	44.01%
Long Holidays												
m-HWT	3.27%	5.07%	1.25%	2.40%	1.54%	2.07%	4.39%	6.03%	8.36%	19.59%	13.17%	38.17%
NARX	22.75%	-5.71%	-49.72%	-67.23%	9.41%	-17.91%	29.65%	16.72%	29.56%	40.65%	30.80%	51.10%
Short Holidays												
m-HWT	20.42%	44.16%	22.22%	50.68%	20.22%	43.54%	22.88%	42.86%	24.30%	41.57%	23.35%	47.59%
NARX	75.85%	75.58%	60.20%	62.85%	76.71%	79.33%	83.75%	85.96%	82.33%	83.21%	81.72%	83.44%
Non-Spec												
m-HWT	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NARX	40.39%	36.34%	-34.86%	-60.69%	6.35%	-9.13%	29.69%	19.12%	20.60%	0.51%	30.06%	22.62%

Table 13 Relative Percentages of m-HWT and NARX with respect to HWT in Friesland data set.

	Lead times											
	1		6		12		24		48		168	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total												
m-HWT	1.33%	3.60%	-7.95%	4.55%	1.04%	6.27%	1.80%	6.84%	2.76%	7.62%	2.65%	8.54%
NARX	39.34%	31.01%	22.36%	-31.16%	1.03%	-21.04%	16.70%	9.56%	23.34%	10.46%	37.32%	14.48%
Special												
m-HWT	4.93%	12.83%	-25.85%	13.11%	3.22%	16.56%	5.39%	16.54%	8.89%	18.03%	10.27%	19.95%
NARX	40.72%	24.45%	40.30%	5.20%	32.47%	21.75%	43.90%	39.44%	47.90%	33.59%	55.64%	43.64%
Long Holidays												
m-HWT	-0.51%	6.28%	-25.58%	1.27%	-2.97%	3.32%	-1.13%	2.96%	1.78%	7.92%	3.33%	18.14%
NARX	-37.24%	-62.99%	19.00%	-20.76%	-17.54%	-16.40%	-11.70%	26.94%	11.48%	14.82%	22.03%	30.13%
Short Holidays												
m-HWT	16.98%	33.64%	-26.78%	41.36%	15.09%	38.03%	17.18%	37.12%	22.17%	34.43%	24.69%	24.27%
NARX	75.55%	75.51%	69.38%	74.34%	77.83%	83.60%	82.30%	87.62%	83.06%	85.92%	84.66%	84.58%
Non-Spec												
m-HWT	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NARX	39.58%	34.31%	14.95%	-46.64%	-13.24%	-42.11%	3.26%	-4.77%	11.21%	-0.29%	29.20%	-0.30%

Table 14 Relative Percentages of m-HWT and NARX with respect to HWT in Noord data set.

	Lead times											
	1		6		12		24		48		168	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Total												
m-HWT	1.69%	5.23%	2.71%	6.26%	2.54%	5.94%	2.87%	6.83%	3.38%	7.91%	4.74%	12.48%
NARX	45.34%	36.58%	-59.56%	-72.60%	-1.82%	-12.75%	33.27%	20.16%	35.88%	19.99%	38.97%	36.90%
Special												
m-HWT	6.45%	18.41%	7.25%	18.04%	6.32%	15.40%	6.93%	16.72%	8.81%	18.26%	13.83%	31.72%
NARX	52.14%	37.98%	-10.00%	-25.54%	34.56%	24.43%	53.90%	46.94%	56.40%	47.10%	51.64%	44.16%
Long Holidays												
m-HWT	1.70%	9.17%	-0.25%	3.67%	-0.71%	2.26%	0.22%	3.37%	3.09%	8.64%	8.11%	30.07%
NARX	-6.31%	-44.23%	-76.95%	-53.91%	-11.67%	8.38%	9.85%	35.90%	27.25%	49.58%	21.59%	45.43%
Short Holidays												
m-HWT	21.63%	46.69%	24.67%	49.18%	20.93%	39.93%	20.83%	42.01%	21.89%	39.19%	28.61%	36.47%
NARX	78.28%	81.13%	58.80%	68.12%	77.29%	82.57%	83.98%	87.52%	83.74%	86.28%	80.46%	82.47%
Non-Spec												
m-HWT	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
NARX	43.60%	36.58%	-82.90%	-93.90%	-21.97%	-32.62%	22.46%	5.23%	24.79%	-2.12%	33.47%	35.54%

Another performance measure considered in this study is MaxAPE. This measure is of significant managerial importance due to risk management and hedging reasons in practice. The maximum errors of the models for each region are presented in Table 15.

Table 15 Performances of three different models in terms of MaxAPE in five provinces.

Lead Times (hrs)	Brabant			Noord		
	NARX	HWT	m-HWT	NARX	HWT	m-HWT
1	9.04%	25.88%	15.90%	11.22%	24.63%	19.51%
6	24.03%	33.93%	24.54%	34.33%	28.64%	28.74%
12	17.94%	56.04%	46.78%	26.01%	53.32%	50.58%
24	15.26%	60.81%	48.67%	21.05%	55.22%	52.41%
48	33.23%	59.68%	48.88%	41.58%	55.69%	54.11%
168	28.83%	59.39%	51.19%	33.18%	53.92%	48.30%

Lead Times (hrs)	Limburg			Maastricht		
	NARX	HWT	m-HWT	NARX	HWT	m-HWT
1	9.67%	28.42%	21.41%	11.65%	27.18%	23.89%
6	29.88%	32.19%	22.78%	31.58%	35.51%	31.78%
12	29.22%	50.35%	36.24%	24.57%	58.87%	47.59%
24	35.61%	51.31%	37.04%	15.50%	64.05%	53.11%
48	16.87%	50.76%	39.51%	29.68%	62.45%	53.88%
168	18.80%	48.80%	37.03%	23.49%	63.30%	44.92%

Lead Times (hrs)	Friesland		
	NARX	HWT	m-HWT
1	16.61%	31.56%	18.96%
6	31.29%	41.64%	38.88%
12	31.72%	68.43%	57.58%
24	22.29%	69.55%	58.46%
48	31.11%	68.23%	65.40%
168	16.80%	67.60%	64.14%

In terms of MaxAPE, NARX outperforms both HWT methods for every forecast horizon and region. The only exceptions are the 6-hours ahead forecasts for Limburg and Noord data set, pointing out to the aforementioned loss of NARX accuracy for 6- and 12-hour ahead forecasting. For lead times except 6 hours, by improving MaxAPE performance of HWT from 40% to 105%; NARX is proved to be a good fit for the market parties who would like to avoid large risks. Furthermore, m-HWT has also decreased MaxAPE values up to 40% compared to Taylor's errors. In addition to the low MAPE values for one-hour ahead forecasts, NARX also gives very competitive MaxAPE values (below 17%) for one-hour ahead forecasts.

Overall we have observed that NARX is better at capturing complex effects of special days, except forecasting long holidays for 6 hours and 12 hours ahead. We believe the degeneration results from the overall decrease in NARX performance in forecasting for lead times that are not multiples of 24 hours (1 day). Even at 6 and 12-hours ahead forecasting, NARX is superior at capturing the effects of short holidays, which are the most complex periods. NARX's performance for longer lead times has led us to studying NARX neural networks for long-term load forecasting.

5.5. Methods for Long-term Load Forecasting

In this part of the study only NARX is adopted as the forecasting method. Results are compared to the outputs of a similar study, which utilized regression models [5] and to the performance of the existing time shifting model.

During the interviews with the company we were informed that long-term forecasts are derived in June for next year on hourly basis. Therefore we have developed a pilot model for Brabant data set to see whether NARX is effective in long-term forecasting.

In terms of implementation there are some differences in the NARX model compared to the one used in short-term load forecasting. The most important difference is in terms of recurrence. In short-term forecasting we used open loop networks, meaning that in the feedback loop actual load values were fed back to the network. For the lead times longer than one hour, we derived multi step forecasts. However in long-term forecasting lead times are too long, multi-step modeling would not work. Therefore we used closed loop networks, in which with the feedback loop not the actual values but past forecasts are fed back to the network.

In addition to the difference in terms of network structure, input variables are different than the ones used in short-term load forecasting. As mentioned in Section 5.2, in short-term forecasting effects of economical and meteorological variables can be negligible, however in long-term forecasting these variables can have significant effects, hence *'Temperature'* and *'Sun light'* variables are added to the data set.

Similar to short-term load forecasting, calendar variables are grouped. However, in training phase, despite the model generally captures special days successfully, the period between Christmas holiday and New Year's Eve performance was much lower compared to other special days. This is due to the fact that the Christmas period is the most difficult period of the year to forecast. During this period many companies and plants shut down, people take long holidays. Therefore, in order to clarify this unusual pattern during this period, another binary variable is defined, named *Christmas Period,* taking value one between 25th of December and 1st of January.

Analogous to short-term forecasting, in addition to the exact dates of special days, one-day prior and after these days and bridge days are defined as variables and included in the model. Furthermore, *day of the week, hour of the day* and *summer time* variables are defined as binary inputs, taking value one on respective days. Duration of sunlight is also included as binary input in the model, named as *Sun Down Percentage*; converts the information on percent of the respective hour the sun is down. For example, if sun rises at 6:30, then for 7th hour interval variable takes 0,5, meaning that half of the 7th hour will be with sunlight.

Lastly, as mentioned before, temperature is defined as another input. However, meteorological forecasts are not available for such a long period and hard to predict. Therefore, temperature data of last 5 years are fed into the network separately for the last five years. The model is run five times with every data set. And the results are averaged to obtain the final forecast. By following the mentioned methodology, extreme effects of too low or too high temperature values' effects are included into the model. However if we simply averaged the temperature values for five years and used as an input, such nonlinear effects of temperature on the infeed would be lost. Another advantage of the mentioned method is providing not only point forecasts but also intervals, which enable market parties to better adjust their risk.

Our second NARX model contains 18 exogenous inputs and uses the same feedback delays as short-term load forecasting NARX: *1, 2, 23, 24, 25, 168 (1 week)* and *169 hours*. In total input layer consists of 25 inputs and output layer contains 1 output node. Unlike short-term forecasting, the search for best architecture is limited to two hidden layers but the total number of hidden units is limited to a smaller number, 35. Every architecture is run 5 times with different initializations. In long-term load forecasting we have observed that performances deviate more with respect to the different architectures compared to the short-term load forecasting. As an example, we have provided the average performances of 5 initializations in terms of MAPE, for 10 to 20 hidden nodes in 1 hidden layer networks in Table 16. Similarly average performances of 5 initializations of a sample of 2 hidden layer networks tested are presented in Table 17. Architectures' performances may seem similar except some huge deviations; however in electricity load forecasting even the slightest improvement may bring high returns. Hence around 2500 models were run in architecture search.

Table 16 Average performances of 5 initializations of 1 hidden layer architectures in terms of MAPE in long-term load forecasting.

Number of hidden nodes	MAPE		
	Training	Validation	Testing
10	2.84%	3.18%	3.54%
11	2.42%	2.85%	3.12%
12	2.42%	2.73%	2.97%
13	2.72%	3.00%	3.26%
14	2.40%	2.85%	3.16%
15	2.34%	2.70%	3.02%
16	2.24%	2.70%	3.26%
17	61.67%	59.60%	64.26%
18	2.23%	2.62%	3.08%
19	68.96%	60.20%	72.90%
20	2.32%	2.71%	3.19%

Table 17 Average performances of 5 initializations of 2 hidden layer architectures in terms of MAPE in long-term load forecasting.

Number of hidden nodes		MAPE		
1st layer	2nd layer	Training	Validation	Testing
6	5	2.72%	3.00%	3.37%
6	6	2.54%	2.91%	3.06%
7	6	2.39%	2.71%	2.99%
7	7	2.62%	2.92%	3.10%
8	7	2.34%	2.70%	2.96%
8	8	2.36%	2.70%	2.85%
9	8	2.31%	2.71%	2.94%
9	9	2.41%	2.82%	3.24%
10	9	2.54%	2.90%	3.24%
10	10	2.37%	2.79%	3.46%
11	10	2.29%	2.69%	2.86%
11	11	2.31%	2.71%	2.94%
12	11	2.20%	2.66%	2.87%
12	12	25.86%	25.74%	27.34%
13	12	2.27%	2.78%	3.05%

The final architecture is shown in Figure 56 where sun_t is the sunlight variable, $temp_t$ is the temperature variable, HI_t stands for hour interval, WD_t is the weekday, $br_{i,t}$ stands for bridge day, day after special day and day before special day binary variables, are $s_{i,t}$special days binary variables and to $\hat{y}_t$ are the selected feedback loop variables.

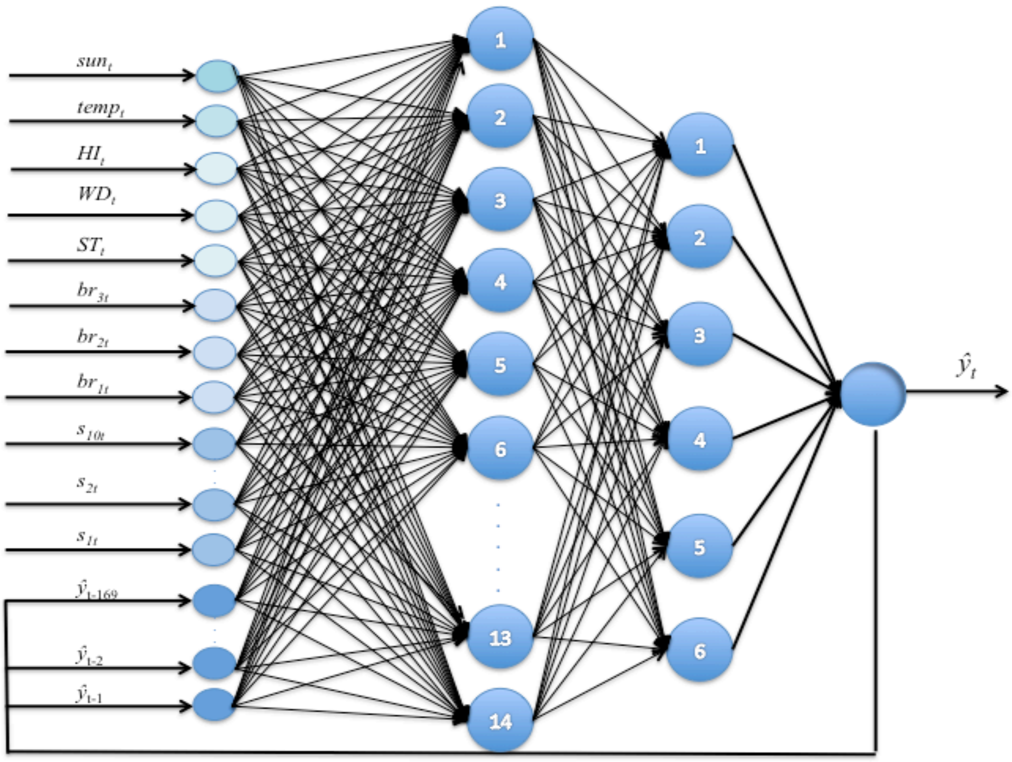


Figure 55. Final network architecture for long-term load forecasting in Brabant data set.

Parallel to our first NARX network training, we have used Levenberg-Marquardt' training algorithm and modeled the networks on MATLAB software, version 2012b. The training parameters are set to the default values, which are provided in Table 3.

5.6. Results of Long-term Load Forecasting

In this section performance of closed loop NARX network in long-term load forecasting is presented. Post-sample accuracies are measured in terms of MAPE and MaxAPE. Results are compared to another study implemented on the same data set by using regression based forecasting methods [5] and the existing methodology followed in the company, which is a time shifting method based on similar years.

In this model, we did not derive multi step forecasts; instead we derived forecasts starting from June till end of the next year. Therefore first, we will present model performance for the next one year and 5 months. Then we will carry out the comparison for one-year period. Because our data set contains data till the end of November 2012, we have included December 2011, to obtain a full year.

First of all, our model performance from June 2011 till November 2012 is provided in Table 18. In the table performances are listed for two different cases, actual temperature and unknown temperature. Actual temperature is the performance when actual temperature values are input for the respective forecasting period. Unknown temperature is the case when, as we mentioned above, past five years' temperature values are fed into the network and averaged. Despite the main indicator of the model performance is the second case, we provided the performance with the actual temperature values in order to give the reader an insight about temperature's effect on the performance. We observed that with the actual values model forecasts infeed 11% better in terms of MAPE. In Figure 56, actual and forecasted infeed values can be seen for the last three months on the same graph.

Table 18 Performance of closed loop NARX neural network in forecasting next 17 months' load in Brabant data set.

	MAPE	MaxAPE
Actual Temperature	2.48%	15.12%
Unknown Temperature	2.75%	18.61%

For an electricity distributor, accuracy in terms of monthly forecasting is also significant for the operations. Therefore, we have aggregated our forecasts monthly and checked the accuracy. In Table 19 aggregated forecasting performance is provided and a graph of 23 months actual infeed versus forecasted values are plotted in Figure 57. The figure proves very high accuracy values, forecasted infeed being very close to actual infeed.

Table 19 Performance of closed loop NARX neural network in forecasting 17 months' load in Brabant data set.

	MAPE	MaxAPE
Monthly forecasting perf.	1.69%	4.69%

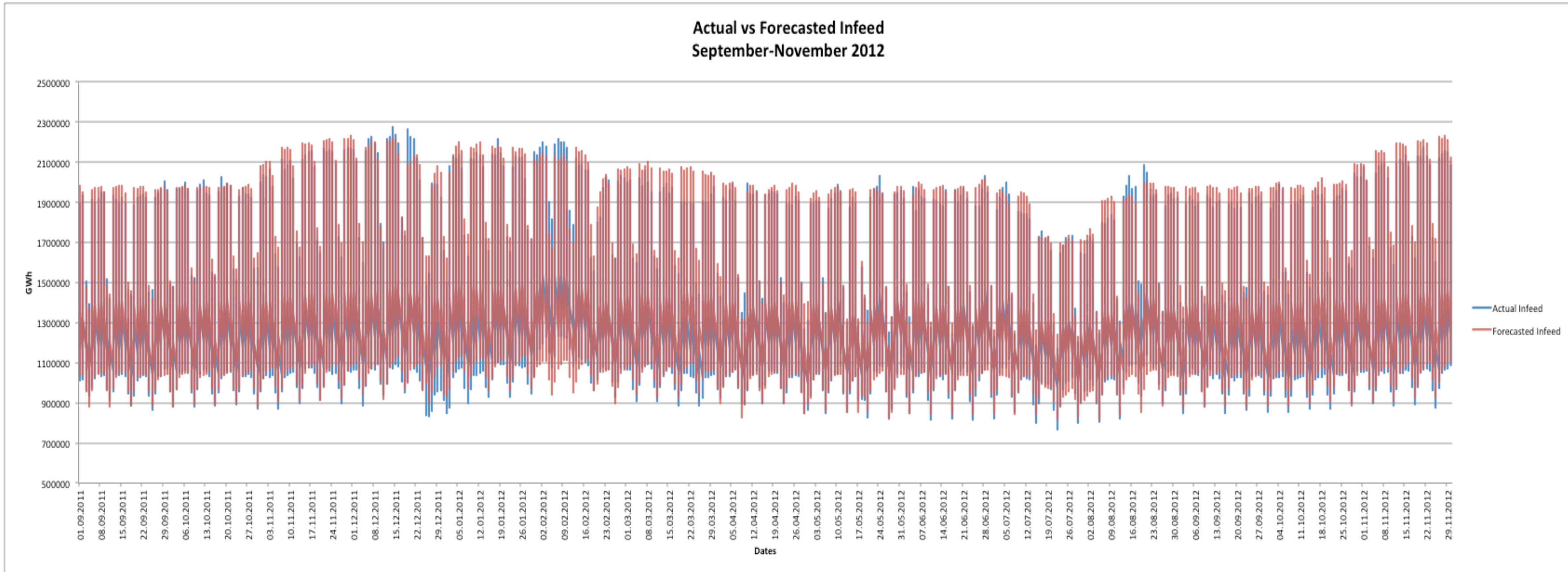


Figure 56. Actual Infeed vs Forecasted Infeed from September 2012 to November 2012 in Brabant data set.

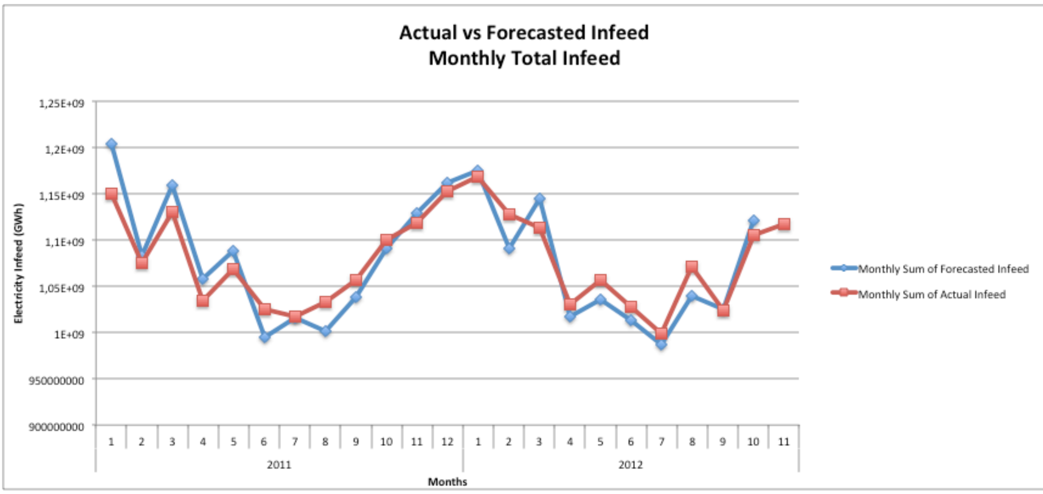


Figure 57. Monthly Actual Infeed vs Forecasted Infeed from January 2011 to November 2012 in Brabant data set.

In order to show how well the model performs at forecasting special days, we have provided some examples in Figure 58, presenting forecasting performance during School Holiday and Bouwvak period, highlighted with dashed circles. Figure 58 proves NARX meets the expectations of capturing special days. Clearly our model adjusts the forecasts during special days and converges very close to actual infeed. Average forecasting performance of our model throughout one year on Bouwvak and School Holidays is 3,23% in terms of MAPE. Figure 59 provides another example to special days forecasting. It includes Christmas and New Year period, which is particularly difficult for the electricity market parties to forecast. Figure 59 shows compared to its performance on other special days; our model also has some minor problems in forecasting the mentioned period. However in comparison with other methods we have received very positive feedbacks from the company about the performance during this period. Despite slight deviations from actual values, average percent error (APE) in December is 1,41%. Throughout one year, our models' MAPE on special days is 3,54%.

Thirdly, in order to also show the model's ability to capture bridge days we have provided an example from May. May 2012 was particularly difficult for the company to forecast, as it included two different types special days; Ascension and Whit Sunday & Monday. As Ascension Day is always on Thursdays, bridge day effect is expected to occur. Figure 60 shows the bridge day effect in May 2012 and our models ability to capture the effect.

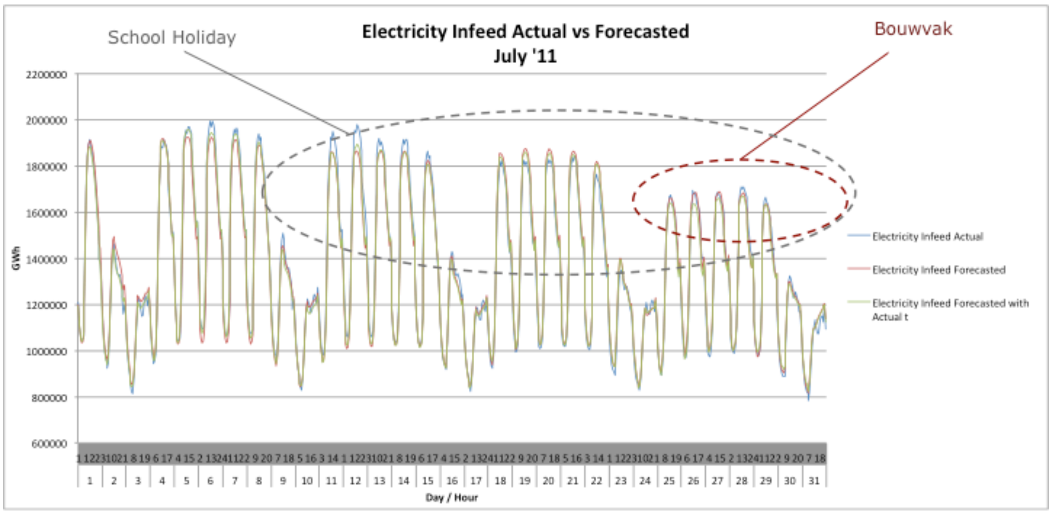


Figure 58. Actual and forecasted infeed values during school holiday and Bouwvak periods.

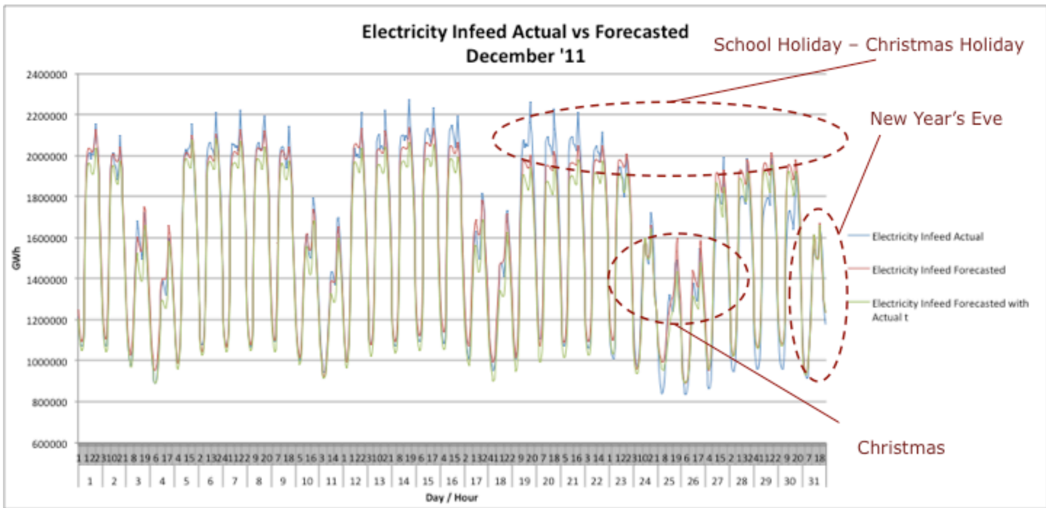


Figure 59. Actual and forecasted infeed values during Christmas and New Year periods.

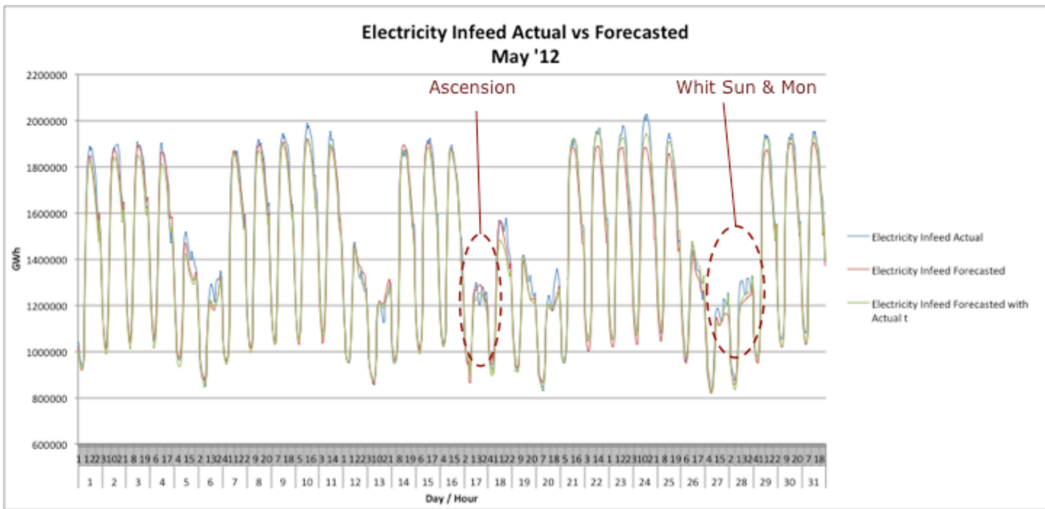


Figure 60. Actual and forecasted infeed values in May 2012.

Table 20 Comparison of performances of Regression, NARX and Time Shifting models in one year hourly forecasting

	Time Shifting	Regression	NARX
MAPE_hourly	3.93%	4.10%	2.85%
MAPE_monthly	2.70%	2.98%	1.66%
MaxAPE_hourly	35.11%	33.40%	18.61%

Lastly, we have provided a comparison of model performances for hourly forecasting of one year in Table 20. Tanrısever et al.'s [5] model is used as a reference regression model because they have used the same data set, but an older time period in training. Time shifting is the existing methodology being used in the company currently, which relies on using a past year as base year and assigning the load on the similar days on that year to the upcoming year.

Table 20 shows that our model has improved the model accuracy approximately 28% compared to current methodology. In addition, the maximum error is almost halved by our NARX model. Maximum error is important as it can be considered as a representation of risk the company takes. Therefore we can say that our model brings less risk to the company compared to other methods.

CHAPTER 6

CONCLUSION

After the deregulation in electricity markets, electricity load forecasting has become one of the main issues for market parties. Accurate electricity load forecasting is not an easy task, due to multiple seasonalities exiting in the data sets, complex and nonlinear relationships between the variables. In addition, different market parties need different types of load forecasting. Some need hourly forecasts, some need daily forecasts, some need peak load forecasts and for various lead times. Every need refers to different data sets; therefore there is no single correct methodology in load forecasting.

There are many studies on electricity load forecasting, dealing with aforementioned challenging issues on different data sets. However, a model performing well in dealing with one issue on one data set, usually fails in dealing with another issue or performing as good on a different data set. Additionally, despite the increase in number of studies, there is still dominancy of conventional and simpler methodologies in load forecasting literature. There is still the gap for applications on more recent methodologies on electricity load data.

Our scope in this thesis refers to mainly two issues. First, one of the main challenges in electricity load forecasting, incorporating special days into the models. Second, modeling of electricity load with a more recent methodology, NARX neural networks. Additionally, we provided a comprehensive computational study on the methods we proposed and compared the performances of those parametric and nonparametric models.

Special days are known as one of the most effective variables on the forecasting. As reported in Section 5.1.1, special days create significant deviations from regular load patterns. Additionally, the load levels on day before and after the special day are also affected. We provided a modification to HWT in order to enable inclusion of special days to the model. HWT is mostly studied and commonly used in the literature due to its simplicity and easiness of use in the real life. However, existing studies on HWT cover the data sets that do not include any special days or switched with smoothed values. Working with data sets without special days does not reflect the real life. Our modification has improved forecasting performance of HWT on special days up to 40%. In addition, we have proposed NARX as a promising way of capturing complex effects of special days. Our computational study on short-term load forecasting in five different regions for six different lead times, proved our proposition. We reported that NARX captures special days effect in also long-term prediction.

As a second contribution, we have reported that NARX performs very well in the electricity load forecasting. After our modification, HWT became very competitive with NARX for short-term load forecasting. But as the lead times get larger and in long-term forecasting NARX outperforms the parametric models. We suggest that NARX stands as a very promising method in short and long-term forecasting and more studies should be carried out on different data sets in the future.

In addition to NARX applications, we suggest the emphasis on hybrid models to be increased. Hybrid models include dividing load data into two: base data and nonlinear data. The first part is modeled by using more conventional methods, such as regression and/or time series models. And the nonlinear relationships existing in the data that simple methods cannot capture accounted for the errors. Therefore the errors are modeled by using promising methods in capturing complex and nonlinear relationships such as, fuzzy logic models. During the literature research we have seen that these models have very promising performances on load forecasting, therefore deeper analyses should be carried out on their effectiveness.

Another methodology whose effectiveness should be investigated is Multivariate Adaptive Regression Splines (MARS). There are no studies investigating effectiveness of MARS in electricity load forecasting - to our best of knowledge. MARS is a non-parametric methodology that includes an extension of recursive partitioning which uses linear functions for local fit. It is simple but effective in handling high dimensional data sets [54] Only in [55] MARS is applied to electricity price forecasting as a non-parametric regression method. In [55] results are compared to a NARX and a wavelet network performances and NARX is stated as the best performing methodology. However, MARS also produced promising results. Considering similarities between electricity load and price data and MARS' power in fitting nonlinear models in high dimensional cases, MARS stand as a promising method for electricity load forecasting. In addition, since least square regression methods are really popular in load forecasting, quantile regression stands as another promising methodology in load forecasting.

As we mentioned in Chapter 3, Turkish electricity market is not as mature as Dutch market yet, but developing in the same direction. There is need for more detailed, accurate models, which are able to capture the effects of both internal and external variables for better functioning of the Turkish market. Methodologies used in this paper can be utilized in also electricity price forecasting, considering price data set's patterns similar to electricity load data. We believe by developing more accurate models on Turkish market, market parties can obtain huge financial gains. Additionally, dependency on external models can be decreased and high savings can be obtained by the state.

REFERENCES

[1] Taylor, J W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8),799–805.

[2] Taylor, J. W. (2012). Short-Term Load Forecasting With Exponentially Weighted Methods. IEEE Transactions on Power Systems, 27(1), 458–464.

[3] Cancelo, J. R., Espasa, A., & Grafe, R. (2008). Forecasting the electricity load from one day to one week ahead for the Spanish system operator. International Journal of Forecasting, 24(4), 588–602.

[4] Kyriakides, E., & Polycarpou, M. (2007). Short-term electricity load forecasting: A tutorial. Trends in Neural Computation, Studies in Computational Intelligence, 34, 319-418.

[5] Tanrisever, F., Derinkuyu, K., Heeren, M., Forecasting electricity infeed for distribution system networks: An analysis of the Dutch case, Energy, Volume 58, Pages 247-257.

[6] Taylor, James W. (2010). Triple Seasonal Methods for Short-Term Electricity Demand Forecasting. European Journal of Operational Research, 44(0), 139–152.

[7] Ramanathan, R., Engle, R., Granger, C. W. J., Vahid-araghi, F., & Brace, C. (1997). Short-run forecasts of electricity loads and peaks. International Journal of Forecasting, 13, 161–174.

[8] Bianco, V., Manca, O., Nardini, S., & Minea, A. a. (2010). Analysis and forecasting of nonresidential electricity consumption in Romania. Applied Energy, 87(11), 3584–3590.

[9] Soliman, S. A. (2005). Long-term / mid-term electric load forecasting based on short-term correlation and annual growth. Electric Power Systems Research, 74, 353–361.

[10] Chen, H., Cañizares, C. A., & Singh, A. (2001). ANN-based Short-Term Load Forecasting in Electricity Markets. Power Engineering Society Winter Meeting. 2, pp. 411-415. Columbus,OH: IEEE.

[11] Myung Suk Kim, Modeling special-day effects for forecasting intraday electricity demand, European Journal of Operational Research, Volume 230, Issue 1, 1 October 2013, Pages 170-180, ISSN 0377-2217.

[12] Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Lalas, D., Moschovits, M., Karagiannis, F., & Papakonstantinou, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. Energy, 31(2-3), 208–227.

[13] Huang, S., & Shih, K. (2003). Short-Term Load Forecasting Via ARMA Model Identification Including Non-Gaussian. IEEE Transactions on Power Systems, 18(2), 673–679.

[14] Soares, Lacir J. and Medeiros, Marcelo C., (2005), Modelling and forecasting short-term electricity load: a two step methodology, No 495, Textos para discussão, Department of Economics PUC-Rio (Brazil), http://EconPapers.repec.org/RePEc:rio:texdis:495.

[15] Von Hirschhausen, C., & Andres, M. (2000). Long-term electricity demand in China — From quantitative to qualitative growth? *Energy Policy*, *28*(4), 231–241.

[16] Hahn, H., Meyer-Nieberg, S., & Pickl, S. (2009). Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research,* 199(3), 902–907.

[17] Al-Hamadi, H. M., & Soliman, S. a. (2005). Long-term/mid-term electric load forecasting based on short-term correlation and annual growth. *Electric Power Systems Research*, *74*(3), 353–361.

[18] Papalexopoulos, A., & Hesterberg, T. (1990). A regression-based approach to short-term system load forecasting. *Power Systems, IEEE Transactions on , 5* (4), 1535-1547.

[19] Bianco, V., Manca, O., & Nardini, S. (2009). Electricity consumption forecasting in Italy using linear regression models. *Energy*, 34, 1413–1421.

[20] Mohamed, Z., & Bodger, P. (2005). Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy*, *30*(10), 1833–1843.

[21] Hagan, M.T.; Behr, Suzanne M., "The Time Series Approach to Short Term Load Forecasting," *Power Systems, IEEE Transactions on* , vol.2, no.3, pp.785,791, Aug. 1987.

[22] Goh, T. N., & Choi, S. S. (1984). Short-term forecasting of electricity demand by decomposition analysis. *European Journal of Operational Research*, *17*, 79–84.

[23] Taylor, J W, Mcsharry, P. E., & Member, S. (2008). Short-Term Load Forecasting Methods : An Evaluation Based on European Data. *IEEE Transactions on Power Systems*, 2213–2219.

[24] Taylor, J. W. (2012). Short-Term Load Forecasting With Exponentially Weighted Methods. *IEEE Transactions on Power Systems*, 27(1), 458–464.

[25] Huang, C., Huang, C., & Wang, M. (2005). A Particle Swarm Optimization to Identifying the ARMAX Model for Short-Term Load Forecasting. *IEEE Transactions on Power Systems*, *20*(2), 1126–1133.

[26] Al-Saba, T., & El-Amin, I. (1999). Artificial neural networks as applied to long-term demand forecasting. *Artificial Intelligence in Engineering*, 13(April 1998), 189–197.

[27] Connor, J. T., Martin, R. D., & Atlas, L. E. (1994). Recurrent neural networks and robust time series prediction. *IEEE Transactions on Neural Networks*, 5(2), 240–54.

[28] Czernichow, T.; Germond, A.; Dorizzi, B.; Caire, P., "Improving recurrent network load forecasting," *Neural Networks, 1995. Proceedings., IEEE International Conference on* , vol.2, no., pp.899,904 vol.2, Nov/Dec 1995.

[29] Espinoza, M., Suykens, J. A. K., Belmans, R., & De Moor, B. (2007). Using kernel-based modeling for nonlinear system identification. *IEEE Control Systems Magazine*, (October 2007), 43–57.

[30] Varghese, S. S., & Ashok, S. (2012). Performance comparison of ANN models for short-term load forecasting. *International Conference on Electrical Engineering and Computer Science*, 97–102.

[31] Awan, S. M., Khan, A., Aslam, M., Mahmood, W., & Ahsan, A. (2012, May 28-31). Application of NARX based FFNN, SVR and ANN fitting models for long-term industrial load forecasting and their comparison. *Industrial Electronics (ISIE), 2012 IEEE International Symposium* , 803-807.

[32] Abraham, A., & Nath, B. (2001). A neuro-fuzzy approach for modelling electricity demand in Victoria. *Applied Soft Computing*, 1(2), 127–138.

[33] Taylor, James W., De Menezes, L. M., & McSharry, P. E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22(1), 1–16.

[34] Tzafestas, S., & Tzafestas, E. (2001). Computational Intelligence Techniques for Short-Term Electric Load Forecasting. *Journal of Intelligent and Robotic Systems*, *31*, 7–68.

[35] González-Romera, E., Jaramillo-Morán, M. a., & Carmona-Fernández, D. (2008). Monthly electric energy demand forecasting with neural networks and Fourier series. *Energy Conversion and Management*, *49*(11), 3135–3142.

[36] Amjady, N., & Keynia, F. (2008). Mid-term load forecasting of power systems by a new prediction method. *Energy Conversion and Management*, *49*(10), 2678–2687.

[37] Desouky, A. El, & Elkateb, M. (2000). Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA. *Generation, Transmission and Distribution, IEE Proceedings, 147*(4), 213–217.

[38] Azadeh, A., Ghaderi, S., & Sohrabkhani, S. (2007). Forecasting electrical consumption by integration of Neural Network, time series and ANOVA. *Applied Mathematics and Computation* , 186 (2), 1753-1761.

[39] Küçükbahar, D. (2008). *Modeling Monthly Electricity Demand in Turkey for 1990-2006.* M.S. Thesis. Middle East Technical University, Ankara.

[40] Gelper, S., Fried, R., & Croux, C. (2010). Robust Forecasting with Exponential and Holt – Winters Smoothing. Journal of Forecasting, 300(June 2009), 285–300.

[41] Chatfield, C., & Yar, M. (2012). Holt-Winters forecasting : some practical issues. *Journal of the Royal Statistical Society*, 37(2), 129–140.

[42] Smith, Michael S. "Short-term forecasting of New South Wales electricity system load". *Journal of Business and Economic* Statistics, 18 (2000), 465-478.

[43] Hippert, H.S., Bunn, D.W., Souza, R.C. Large neural networks for electricity load forecasting: Are they overfitted?, International Journal of Forecasting, Volume 21, Issue 3, July–September 2005, Pages 425-434.

[44] Taylor, James W. (2010). Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles. *International Journal of Forecasting*, 26(4), 627–646.

[45] Duda, R. O., Hart, P. E., & Stork, D. G. (1997). Pattern Classification. John Wiley and Sons. Nov 1, 2000.

[46] Catalão, J., Mariano, S., & Mendes, V. (2007). Short-term electricity prices forecasting in competitive market: A neural netowork approach. *Electric Power Systems Research*, 77 (10), 1297-1304.

[47] Bodén, M. (2002). A Guide to Recurrent Neural Networks and Backpropagation. In the Dallas Project, SICS Technical Report, 2002:03.

[48] Lin, T.-N., Giles, C., Horne, B., & Kung, S.-Y. (1997). A delay damage model selection algorithm for NARX neural networks. *Signal Processing, IEEE Transactions on*, 45 (11), 2719-2730.

[49] Tsungnan, L., Horne, B., Tino, P., & Giles, C. (1996). Learning long-term dependencies in NARX recurrent neural networks. *Neural Networks, IEEE Transactions on* , 7 (6), 1329-1338.

[50] Hyndman, R. J., Koehler, A.B., Snyder, R.D., A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, Volume 18, Issue 3, July–September 2002, Pages 439-454.

[51] Engle, R., & Manganelli, S. (1999). CAViaR: conditional value at risk by quantile regression. *National Bureau of Economic Research*.

[52] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks : The state of the art, *International Journal of Forecasting,* 14, 35–62.

[53] Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* , 5 (6), 989-983.

[54] Koc ,E. K., Iyigun, C. (2013). Restructuring forward step of MARS algorithm using a new knot selection procedure based on a mapping approach. *Journal of Global Optimization.* doi: 10.1007/s10898-013-0107-5.

[55] Arash Andalib, Farid Atry, Multi-step ahead forecasts for electricity prices using NARX: A new approach, a critical analysis of one-step ahead forecasts, *Energy Conversion and Management*, Volume 50, Issue 3, Pages 739-747.

APPENDIX A

EFFECTS OF SPECIAL DAYS

Ascension Day

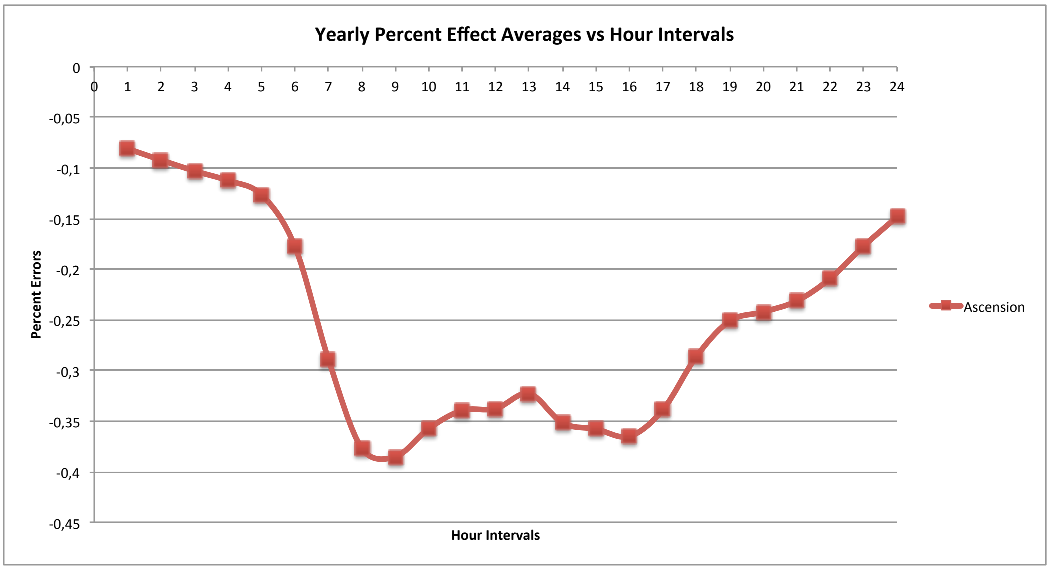


Figure 61. Effects of special days averaged for five years with respect to hours of day in Brabant data set (Ascension Day).

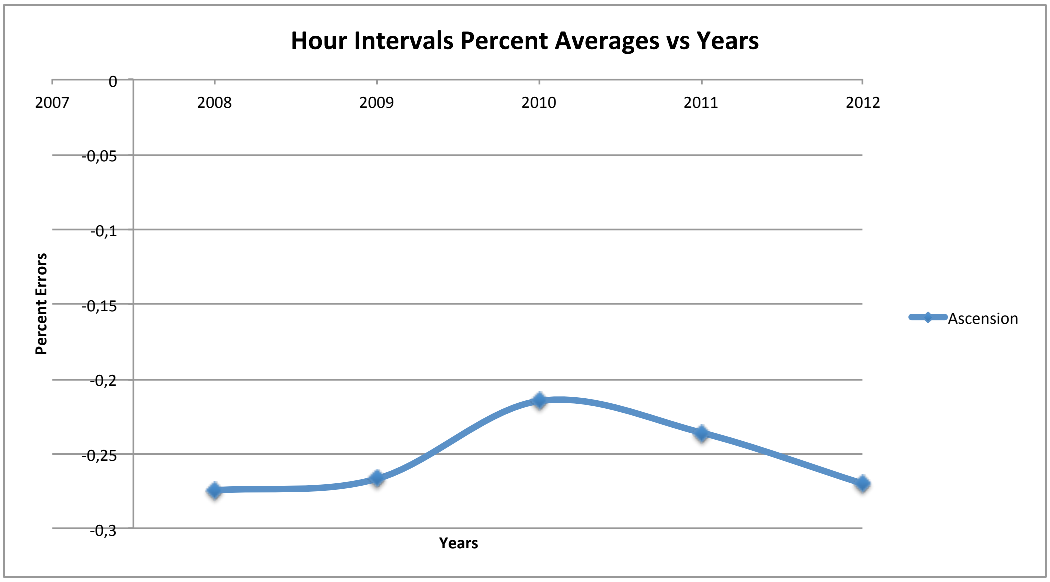


Figure 62. Average effects of special days in different years in Brabant data set (Ascension Day).

Queen’s Day

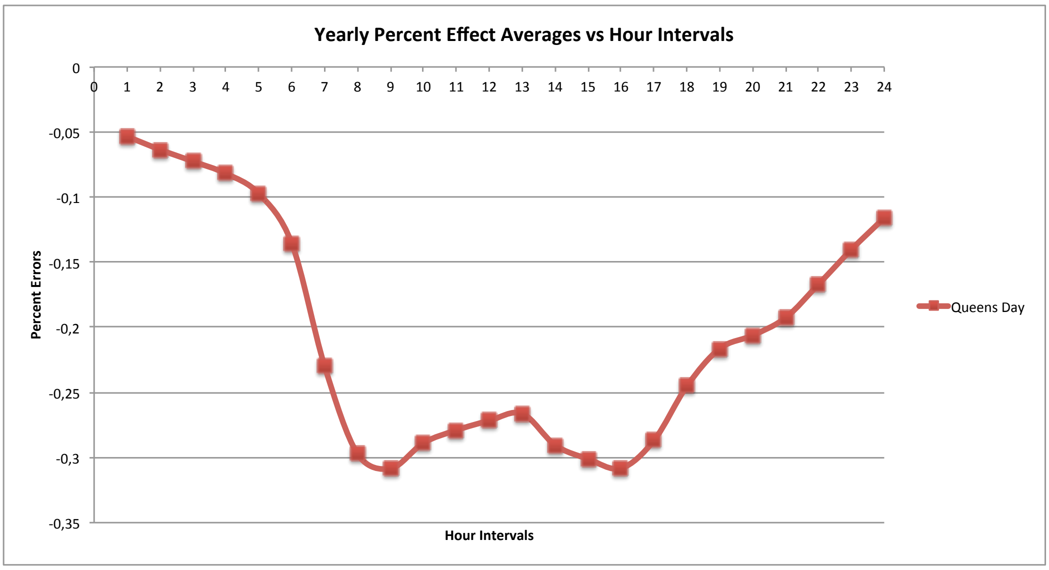


Figure 63. Effects of special days averaged for five years with respect to hours of day in Brabant data set (Queen’s Day).

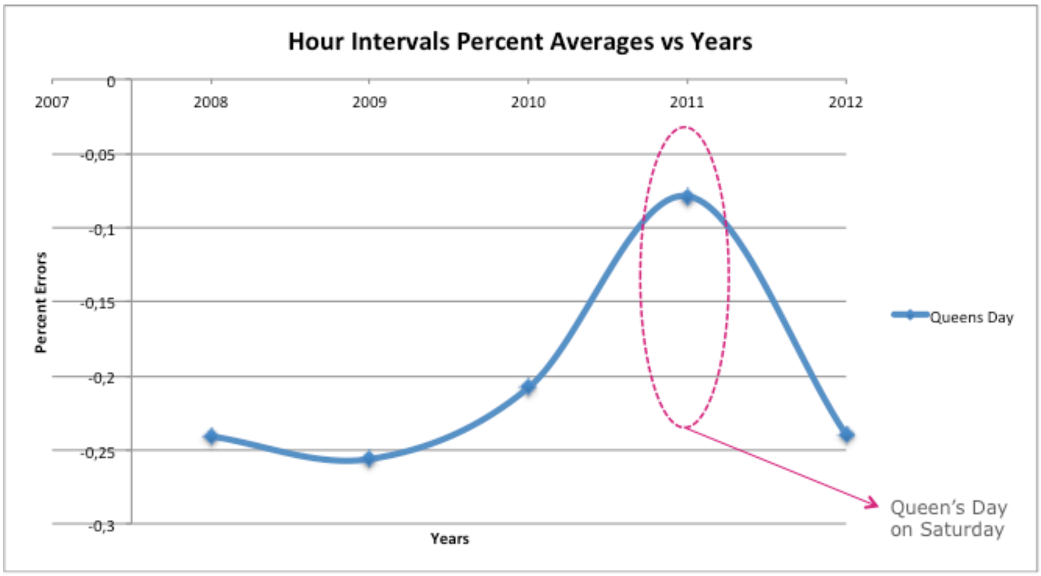


Figure 64. Average effects of special days in different years in Brabant data set (Queen’s Day).

New Year Holiday

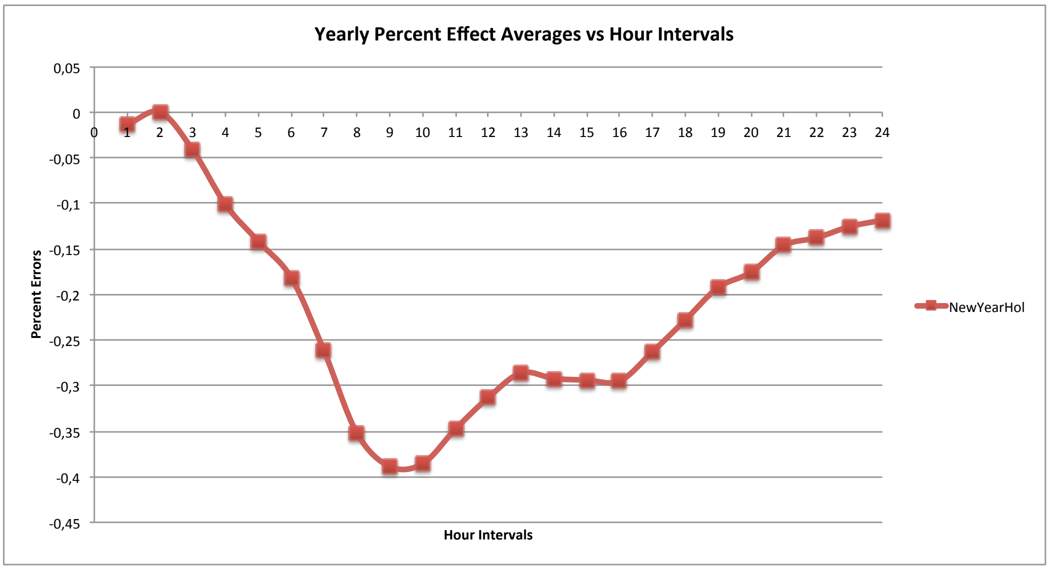


Figure 65. Effects of special days averaged for five years with respect to hours of day in Brabant data set (New Year Holiday).

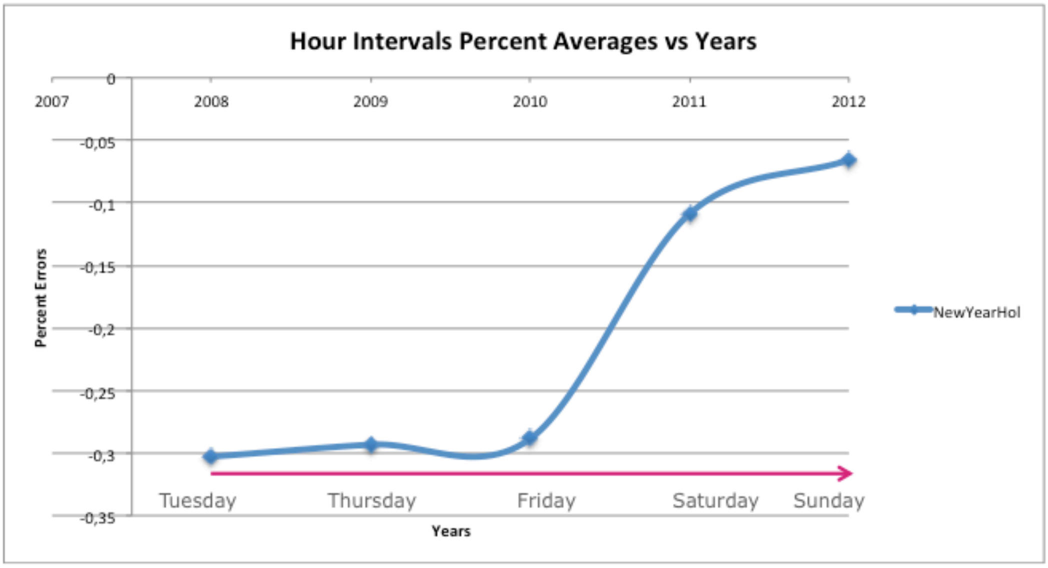


Figure 66. Average effects of special days in different years in Brabant data set (New Year Holiday).

Carnival

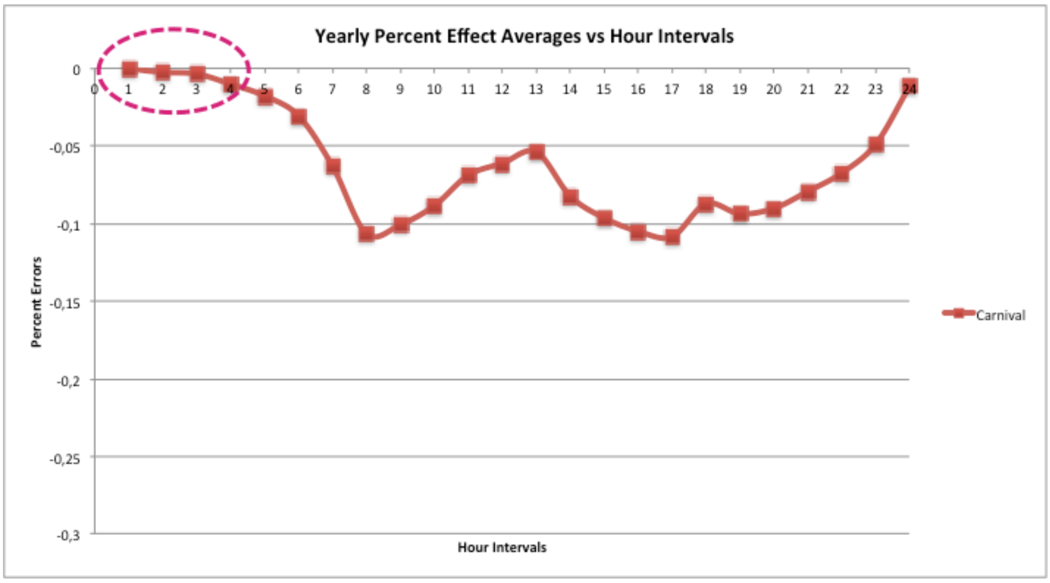


Figure 67. Effects of special days averaged for five years with respect to hours of day in Brabant data set (Carnival).

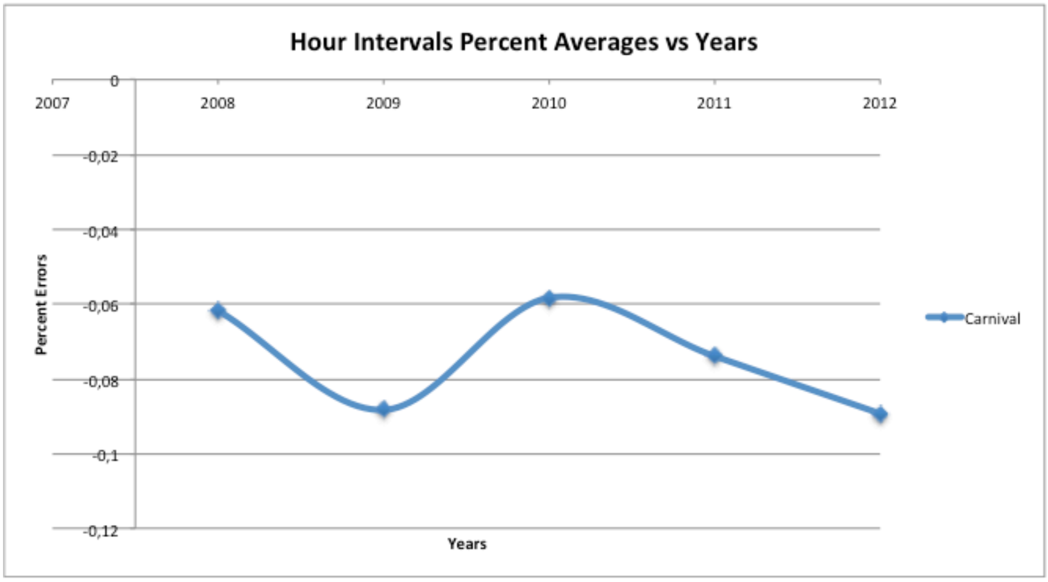


Figure 68. Average effects of special days in different years in Brabant data set (Carnival).

APPENDIX B

M-HWT MODEL PARAMETERS

Table 21 Parameters for m-HWT models of Maastricht data set.

Lead Times (hrs)	α_{best}	δ_{best}	ω_{best}	ϕ_{best}
1	0.6099	0.1877	0.2070	0.2839
6	0.0528	0.1888	0.1300	0.7814
12	0.0262	0.1800	0.0894	0.8017
24	0.0043	0.2087	0.0729	0.7990
48	0.0004	0.0905	0.0442	0.9183
168	0.0001	0.0068	0.2103	0.9753

Table 22 Parameters for m-HWT models of Limburg data set.

Lead Times (hrs)	α_{best}	δ_{best}	ω_{best}	ϕ_{best}
1	0.5623	0.1912	0.2168	0.3157
6	0.0433	0.1927	0.1305	0.7990
12	0.0203	0.2036	0.0735	0.8581
24	0.0023	0.2744	0.0943	0.9309
48	0.0035	0.0635	0.0719	0.9214
168	0.0009	0.0499	0.1648	0.8870

Table 23 Parameters for m-HWT models of Friesland data set.

Lead Times (hrs)	α_{best}	δ_{best}	ω_{best}	ϕ_{best}
1	0.6107	0.1966	0.2430	0.2738
6	0.0262	0.1614	0.1662	0.8236
12	0.0115	0.2000	0.1213	0.7941
24	0.0006	0.2378	0.1423	0.8906
48	0.0035	0.0635	0.0719	0.9214
168	0.0001	0.0947	0.2340	0.9477

Table 24 Parameters for m-HWT models of Noord data set.

Lead Times (hrs)	α_{best}	δ_{best}	ω_{best}	ϕ_{best}
1	0.5358	0.2287	0.2070	0.3493
6	0.0357	0.1884	0.1572	0.7887
12	0.0124	0.2078	0.1110	0.8236
24	0.0043	0.2087	0.0729	0.7990
48	0.0001	0.1435	0.0571	0.9396
168	0.0004	0.0068	0.2103	0.9753

APPENDIX C

NARX ARCHITECTURES

Table 25 Maastricht data set model architectures for different forecasting horizons.

Lead Times (hrs)	Hidden Layer	Number of Hidden Nodes
1	1	25
6	1	60
12	1	60
24	1	50
48	1	20
168	1	20

Table 26 Limburg data set model architectures for different forecasting horizons.

Lead Times (hrs)	Hidden Layer	Number of Hidden Nodes
1	1	20
6	1	30
12	1	70
24	1	55
48	1	20
168	1	30

Table 27 Friesland data set model architectures for different forecasting horizons.

Lead Times (hrs)	Hidden Layer	Number of Hidden Nodes
1	1	5
6	1	30
12	1	20
24	1	30
48	1	30
168	1	25

Table 28 Noord data set model architectures for different forecasting horizons.

Lead Times (hrs)	Hidden Layer	Number of Hidden Nodes
1	1	35
6	1	50
12	1	30
24	1	35
48	1	25
168	1	30

APPENDIX D

FIGURES OF SHORT-TERM LOAD FORECASTS' PERFORMANCES

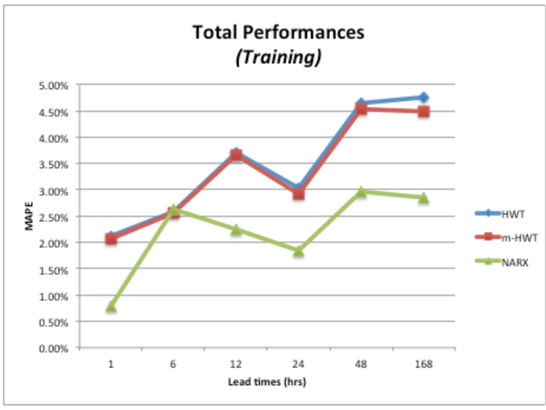


Figure 69. Training performances of three models for different forecast horizons in Limburg data set.

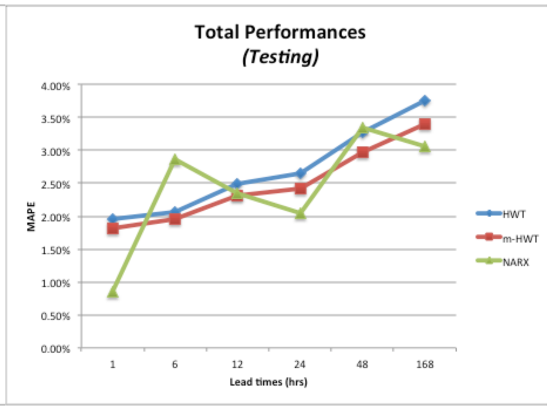


Figure 70. Testing performances of three models for different forecast horizons in Limburg data set.

Figure 71. Training performances of three models on special days for different forecast horizons in Limburg data set.

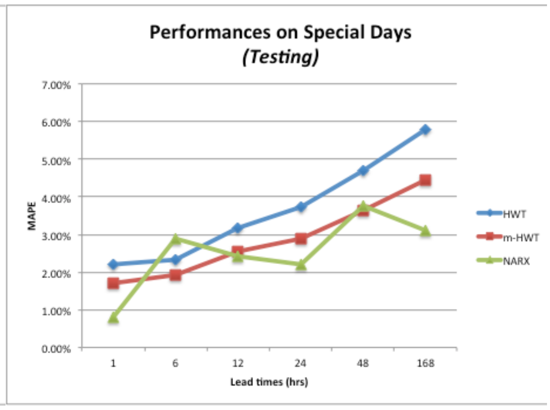


Figure 72. Testing performances of three models on special days for different forecast horizons in Limburg data set.

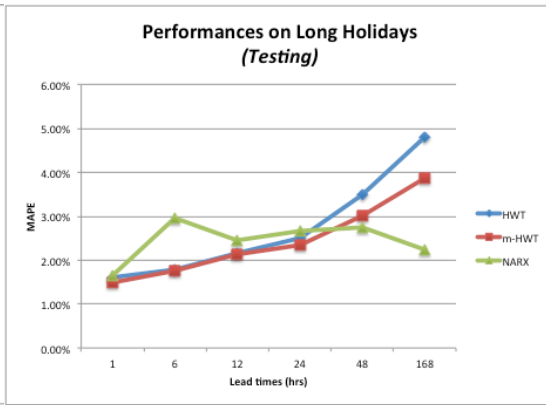
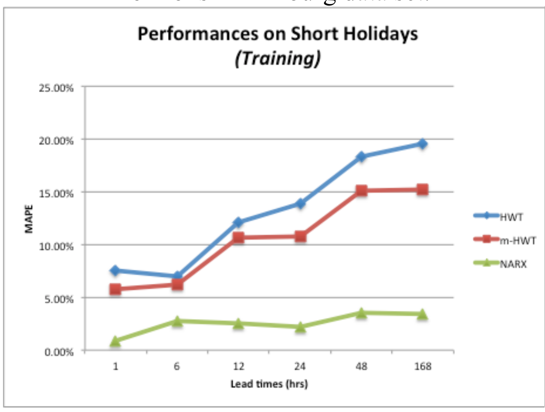


Figure 73. Training performances of three models on long holidays for different forecast horizons in Limburg data set.

Figure 74. Testing performances of three models on long holidays for different forecast horizons in Limburg data set.



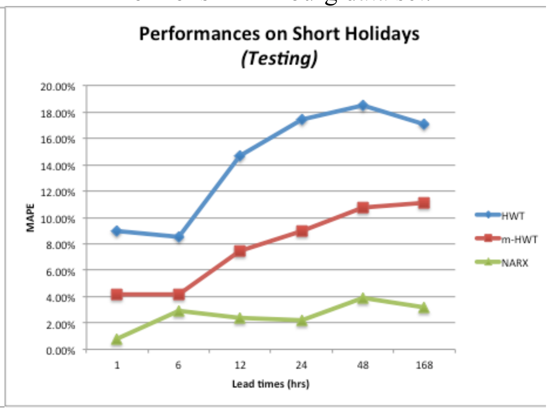
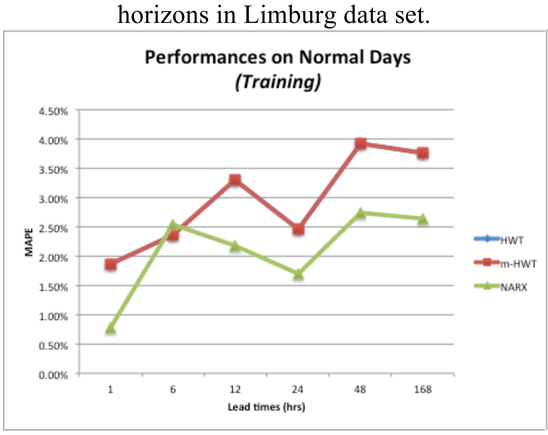


Figure 75. Training performances of three models on short holidays for different forecast horizons in Limburg data set.

Figure 76. Testing performances of three models on long holidays for different forecast horizons in Limburg data set.



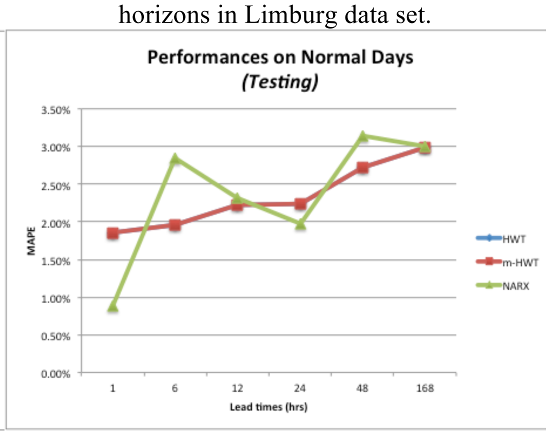


Figure 77. Training performances of three models on normal days for different forecast horizons in Limburg data set.

Figure 78. Testing performances of three models on normal days for different forecast horizons in Limburg data set.

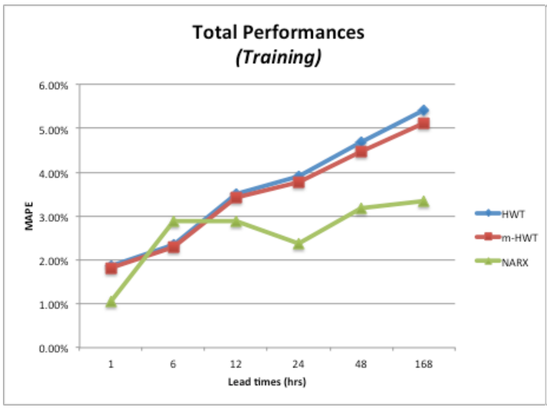


Figure 79. Training performances of three models for different forecast horizons in Maastricht data set.

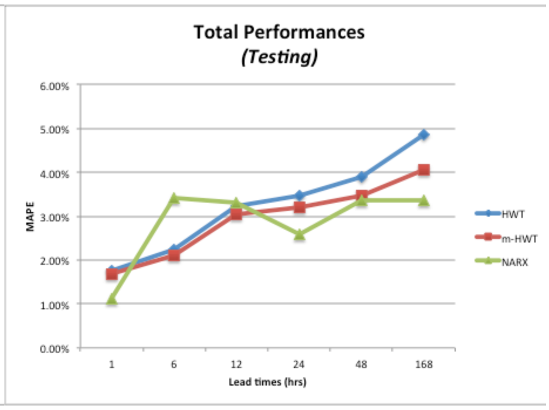


Figure 80. Testing performances of three models for different forecast horizons in Maastricht data set.

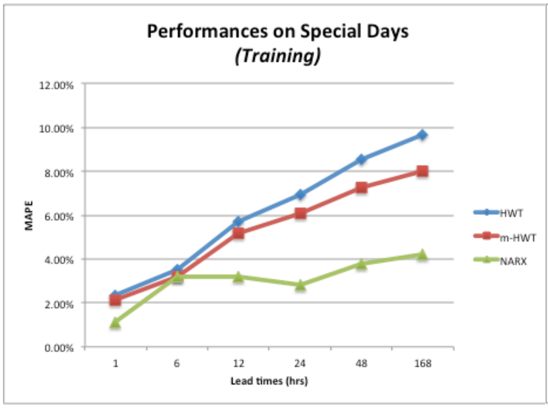


Figure 81. Training performances of three models on special days for different forecast horizons in Maastricht data set.

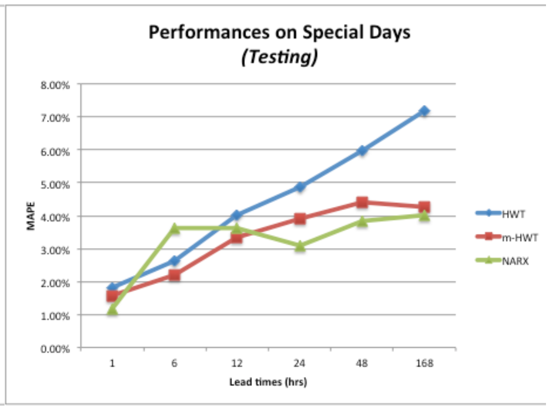


Figure 82. Testing performances of three models on special days for different forecast horizons in Maastricht data set.

Figure 83. Training performances of three models on long holidays for different forecast horizons in Maastricht data set.

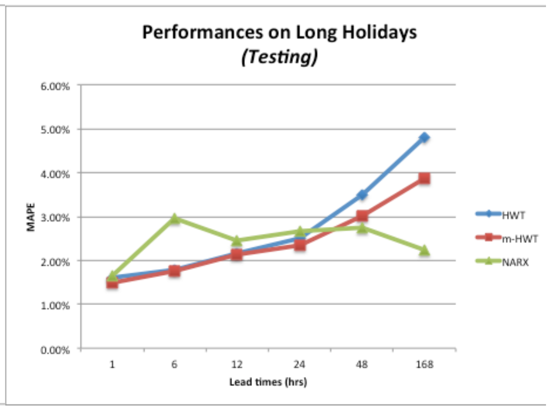
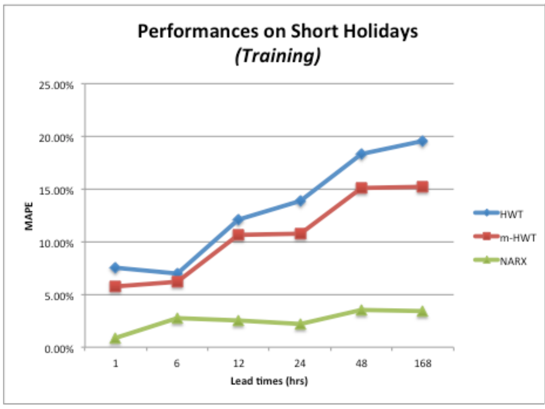


Figure 84. Testing performances of three models on long holidays for different forecast horizons in Maastricht data set.



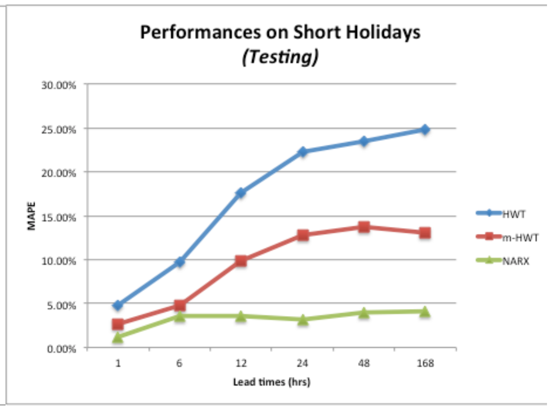


Figure 85. Training performances of three models on short holidays for different forecast horizons in Maastricht data set.

Figure 86. Testing performances of three models on long holidays for different forecast horizons in Maastricht data set.

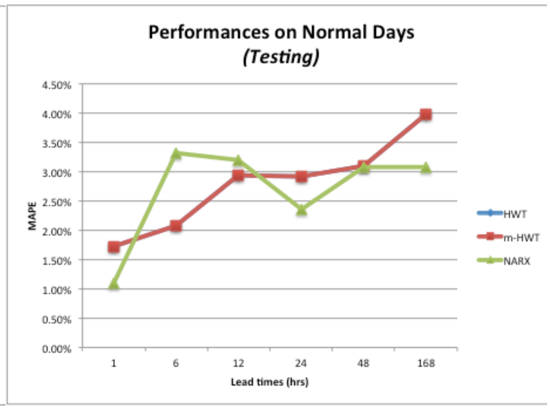
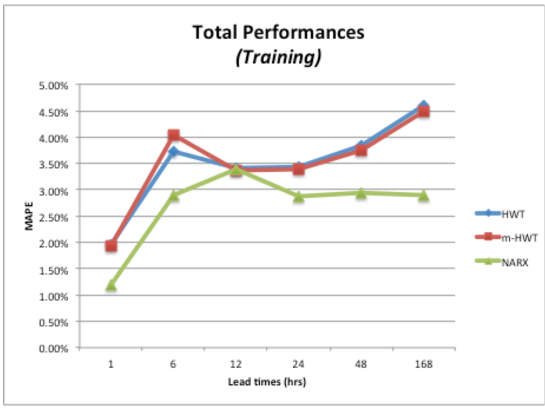


Figure 87. Training performances of three models on normal days for different forecast horizons in Maastricht data set.

Figure 88. Testing performances of three models on normal days for different forecast horizons in Maastricht data set.



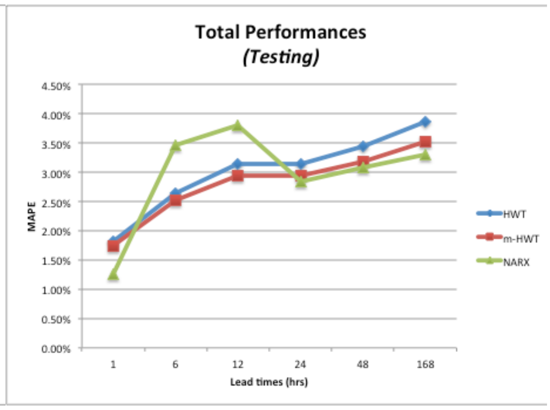


Figure 89. Training performances of three models for different forecast horizons in Friesland data set.

Figure 90. Testing performances of three models for different forecast horizons in Friesland data set.

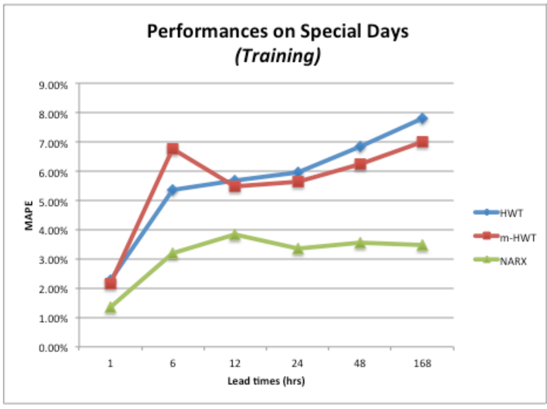


Figure 91. Training performances of three models on special days for different forecast horizons in Friesland data set.

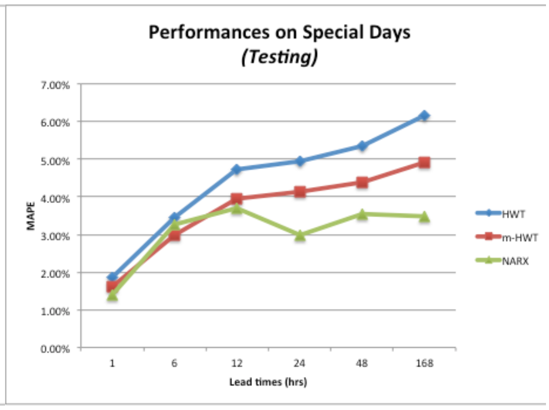


Figure 92. Testing performances of three models on special days for different forecast horizons in Friesland data set.

Figure 93. Training performances of three models on long holidays for different forecast horizons in Friesland data set.

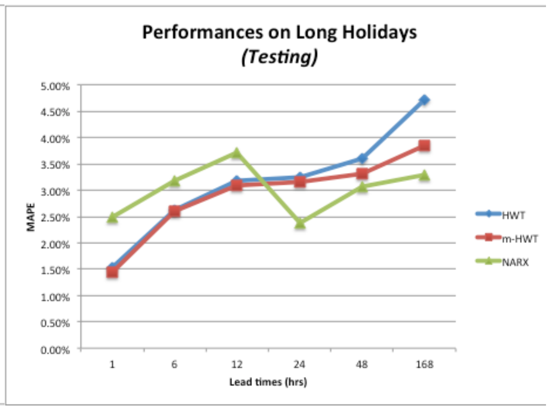


Figure 94. Testing performances of three models on long holidays for different forecast horizons in Friesland data set.

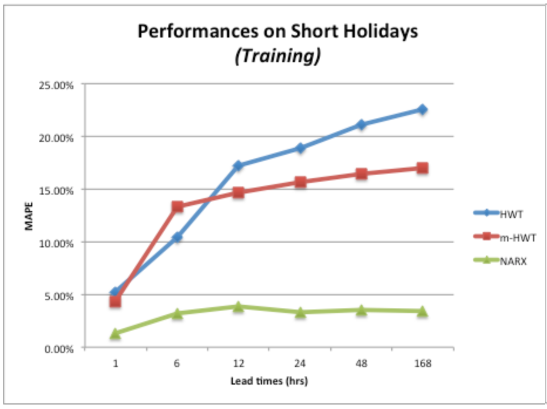


Figure 95. Training performances of three models on short holidays for different forecast horizons in Friesland data set.

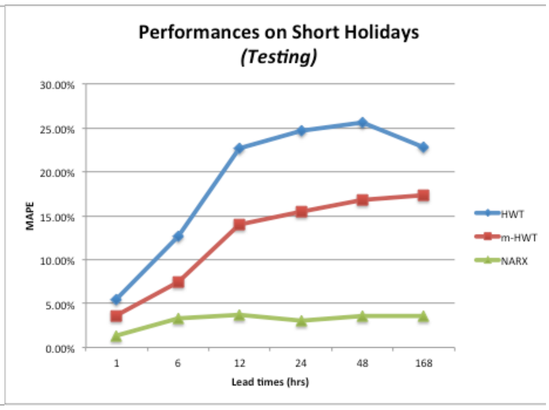


Figure 96. Testing performances of three models on long holidays for different forecast horizons in Friesland data set.

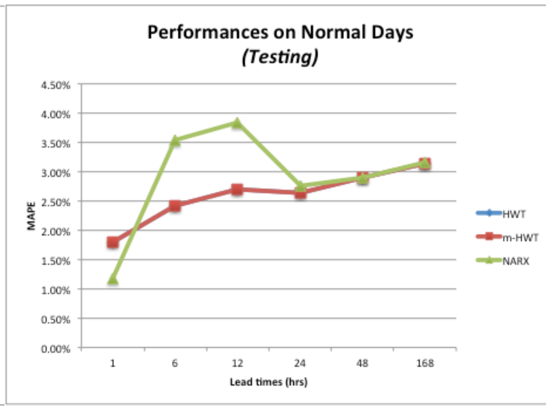


Figure 97. Training performances of three models on normal days for different forecast horizons in Friesland data set.

Figure 98. Testing performances of three models on normal days for different forecast horizons in Friesland data set.