ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

ESSAYS ON ELECTRICITY PRICE MODELING AND FORECASTING

Ph.D. THESIS

Umut UĞURLU

Department of Management Engineering

Management Engineering Programme

JANUARY 2019

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

ESSAYS ON ELECTRICITY PRICE MODELING AND FORECASTING

Ph.D. THESIS

Umut UĞURLU
(507132004)

Department of Management Engineering

Management Engineering Programme

Thesis Advisor: Prof. Dr. Oktay TAŞ

JANUARY 2019

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK FİYATLARININ MODELLENMESİ VE TAHMİNİ ÜZERİNE MAKALELER

DOKTORA TEZİ

Umut UĞURLU
(507132004)

İşletme Mühendisliği Anabilim Dalı

İşletme Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Oktay TAŞ

OCAK 2019

Umut Uğurlu, a **Ph.D.** student of ITU **Graduate School of Science, Engineering and Technology** student ID 507132004, successfully defended the **thesis** entitled "**ESSAYS ON ELECTRICITY PRICE MODELING AND FORECASTING**", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Prof. Dr. Oktay TAŞ** ....
Istanbul Technical University

Jury Members : **Prof. Dr. Mehmet Fuat BEYAZIT** ....
Bilgi University

Prof. Dr. Suat TEKER ....
Işık University

Assoc. Prof. Dr. Cumhur Enis EKİNCİ ....
Istanbul Technical University

Assoc. Prof. Dr. Kaya TOKMAKÇIOĞLU ....
Istanbul Technical University

Date of Submission : 7 December 2018
Date of Defense : 7 January 2019

FOREWORD

Although I had many obstacles and difficulties in the process of this thesis, I enjoyed this journey very much. The joy in this journey was mainly because of the people who I travelled with.

First, and foremost, I want to express my gratitude to Prof. Dr. Oktay Taş. It was such a fortune to work with him. First of all, he believed in me at the start of this journey. Then, he pushed me all the way to my limits. He listened to all my problems and tried to solve all of them patiently. He encouraged me in all the difficult moments and gave me the self-confidence, when it was lacked. I am indebted to him.

I want to thank Dr. Michael Coulon, who was my supervisor in my research visit to the University of Sussex. It was a pleasure for me to meet and work with this extraordinary person.

Furthermore, I would like to thank my co-authors in addition to Prof. Dr. Oktay Taş; Dr. Umut Gündüz, Dr. İlkay Öksüz and Res. Asst. Aycan Kaya. I have to acknowledge that all of them helped me very generously. Dr. Umut Gündüz, a great friend of mine, helped me a lot with the statistical methods and their application. For the last 17 years, I felt the privilage of being a friend of Dr. İlkay Öksüz. He helped me so generously in the second and third articles as he has done from my early high school days. Res. Asst. Aycan Kaya helped me in the optimization process of the third article. I admire her skills in the optimization area.

It was a pleasure for me having Prof. Dr. Burç Ülengin and Prof. Dr. Fuat Beyazıt in my thesis progress committee. I had the opportunity of taking courses from Prof. Dr. Burç Ülengin, which helped me a lot in the further steps of PhD. He also helped me with his comments and solutions in the early stages of my thesis. Prof. Dr. Fuat Beyazıt was one of the people who turned this journey into a joyful event. One of the most important gains of this PhD process was knowing him. His paper on the electricity price forecasting triggered me to work on this topic. I would also like to thank my thesis defense committee members, Prof. Dr. Suat Teker, Assoc. Prof. Dr. Cumhur Ekinci and Assoc. Prof. Dr. Kaya Tokmakçıoğlu. I am grateful to all the thesis defense committee members for their comments and suggestions on the draft of this thesis.

As this thesis consists of three papers, I have to mention some paper-based thanks as well. For the first paper, I would like to thank Dr. Michael Coulon and Prof. Dr. Burç Ülengin, for their comments and suggestions. In addition to them, I want to acknowledge all the participants of the 4th Young Finance Scholars' Conference and The Commodity and Energy Markets Annual Meeting 2017. I found the opportunity of having fruitful discussions with Prof. Dr. Derek Bunn and Prof. Dr. Rafal Weron, which I appreciate a lot. Moreover, I would like to thank the editor of the Emerging Markets Finance and Trade, Prof. Dr. Ali Kutan. For the second paper, I acknowledge

the Dr. İlkay Öksüz's EPSRC programme of King's College London, which paid the open access fee generously. I also want to thank Assoc. Prof. Dr. Tolga Kaya and Dr. Anirban Mukhopadhyay for their comments on the draft of the second paper. Lastly, I would like to express my gratitude to the anonymous referees of all the papers.

I have to acknowledge some of my colleagues specifically here. It was a pleasure for me to work in the Finance group of the ITU Management Engineering Department with Prof. Dr. Oktay Taş, Assoc. Prof. Dr. Cumhur Ekinci, Assoc. Prof. Dr. Kaya Tokmakçıoğlu and Res. Asst. Gülfem Levent Berkay. I am thankful to them for supporting me. In addition to them, I found the opportunity of working with some great academicians. On behalf of all of them, I have to mention Prof. Dr. Fahir Bilginoğlu's name. Although I couldn't find the opportunity of taking a course from him, I feel like a life-long student of him. It was such a big pleasure for me to meet and work with this special person. On behalf of all my colleagues, I would like to thank my fellow friend Res. Asst. Hakan Yorulmuş. He was always in the right place at the right time; he helped me a lot throughout this process.

I have to thank some institutions for supporting me in this journey. I am supported by The Scientific and Technological Research Council of Turkey, 2214/A Programme in my visit to the University of Sussex. Tinçel Cultural Foundation also supported me in this process. This thesis is supported by Research Fund of the Istanbul Technical University, project number: SDK-2018-41160. Moreover, NVIDIA Corporation generously donated a GPU for our machine learning computations.

I want to specifically mention the name of the beloved one, Tarryn Williams. She has changed me a lot and brought much more happiness to my life. I always felt her support and patience during this PhD process. It was great knowing that she appreciated what I have done.

Last, but not least, I want to thank my family; my mother, my father and my grandmother. They supported me patiently and raised me as an independent person. They allowed me to make my own decisions from an early age and always trusted in me. I appreciate everything they have done for me.

December 2018 Umut UĞURLU

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ESSAYS ON ELECTRICITY PRICE MODELING AND FORECASTING

SUMMARY

Electricity markets have been privatized since the beginning of 2000s. Due to this liberalization, electricity price forecasting has become an essential task for all the participants of the electricity markets. Especially day-ahead electricity price forecasts are used for many purposes. For example, generation companies can optimize their production schedule according to these electricity price forecasts and decrease the loss of profits caused by the inaccurate electricity price forecasts. Turkish electricity market is not exempt from this change all over the world. As an emerging market, Turkish day-ahead electricity market needs to be investigated.

This thesis consists of three essays on electricity price modeling and forecasting. First one suggests the usage of factorial ANOVA as a pre-treatment to the Turkish electricity price series before applying the forecast methods. Furthermore, it compares the statistical time series methods by using the lagged price series. Even though best performing methods vary from period to period, SARIMA is chosen as the best method among the statistical models, which include Markov regime-switching, SETAR, AR(24) and naïve method. In the second article, electricity prices are forecasted by various models. In addition to the lagged prices; temperature, forecast demand/supply, 24^{th} lag of realized demand/supply and the 24^{th} lag of the balancing market prices are utilized as the exogenous variables. The main contribution of this paper is that the success of the electricity price forecasts increase a lot by using the deep recurrent neural networks. This is the first work, which involves recurrent neural networks as well as deep learning methods into the electricity price forecasting literature. Third paper, which evaluates the financial effect of the inaccurate electricity price forecasts on a hydro-based generation company, optimizes the production schedule according to nine electricity price forecasts and compare the results with the "best" ex-post actual prices case. According to these inaccurate forecasts, generation companies might face significant loss of profits. The main finding of this paper is that the best method according to the performance evaluation criteria is not necessarily the best method according to the financial effect measures. In our example, ANN-LSTM method, which is not the best method among the nine methods in terms of the forecast performance measures is the best method in terms of the financial effect criteria. Another important point to mention is that the hybrid methods, which combine the forecasts of various methods, perform best in most of the examined periods.

It is the first study in the Turkish day-ahead electricity market, which looks at the electricity price forecasting from such a wide perspective. Moreover, it suggests using the factorial ANOVA method as a pre-treatment before forecasting the electricity prices. Furthermore, it proposes the deep recurrent neural networks as the best forecasting method compared to the shallow recurrent neural networks, artificial neural networks and the statistical time series methods. Lastly, it mentions the conflict

between the performance evaluation criteria and the financial effect measures. Generation companies must take this conflict into account, when they choose the forecast methods.

As an ever improving research area, electricity price forecasting needs further research in many branches. Lasso or principal component analysis type dimension reduction techniques could help to choose the variables. Hybrid methods, both combination of the electricity price forecasts from various methods and the simultaneous use of different methods as hybrid models could increase the forecast performance. Furthermore, energy derivatives pricing will be an important research area in the near future, especially due to the development of the intraday markets. Last, but not least, all the electricity markets have unique characteristics due to their location, development level, renewables share etc. Therefore, applying the models discussed in this thesis on other markets at different time periods will give us more robust results.

ELEKTRİK FİYATLARININ MODELLENMESİ VE TAHMİNİ ÜZERİNE MAKALELER

ÖZET

Elektrik piyasaları 2000'lerin başından itibaren ciddi bir değişim içine girmiştir. Daha önce devletlere ait olan ve tekel halinde bulunan elektrik emtiası, özelleştirilmiş ve rekabete açık bir ortam oluşmuştur. Bu vesileyle, daha önce bu tekeller tarafından belirlenen elektrik fiyatları, serbest piyasada belirlenir hale gelmiştir. Türkiye'de elektrik piyasası 2000'li yıllarda özelleşmiş ve 2011 Aralık ayında kurulan Gün Öncesi Elektrik Piyasası ile elektrik fiyatlarının oluşabileceği bir ortam meydana gelmiştir. Bu noktada, elektrik fiyatlarının gerçek fiyatlara en yakın tahmin edilmesi pek çok açıdan önem arz etmektedir. Öncelikli olarak bu piyasada teklif veren arz ve talep taraflarının mümkün olduğunca doğru ve tutarlı fiyat teklifleri vermeleri gerekmektedir. Örneğin, hidroelektrik santralleri üretimlerini elektrik fiyat tahminlerine göre optimize etmekte ve gerçekleşen fiyatlardan uzaklaşan elektrik tahminleri önemli miktarda zarar etmelerine neden olabilmektedir. Bununla beraber, daha doğru şekilde yapılan elektrik tahminleri üretici rantını ortadan kaldıracak olup, fiyat daha aşağı seviyede oluşacak ve buradaki farktan kazanç sağlayacak kişiler nihai tüketiciyi temsil eden vatandaşlar olacaktır. Elektrik fiyat tahmini ve modellenmesi konusu dünya çapında artan bir ilgiye mazhar olmakla beraber, Türkiye piyasasında bugüne kadar yapılmış olan çalışmalar hem sayıca hem de kapsam olarak kısıtlı kalmaktadır. Bu tezde, yayınlanmış üç makale ile elektrik fiyat modellenmesi ve tahmini konusu tartışılacaktır.

İlk makalede Türkiye Gün Öncesi Elektrik Piyasası'nda istatistiksel yöntemler kullanılmak suretiyle elektrik fiyatlarının saatlik olarak tahmin edilmesi konusu tartışılmaktadır. Bu bağlamda elektrik fiyatlarının önceki değerleri açıklayıcı değişkenler olarak kullanılmıştır. En önemli olarak görülen 1., 24., 48. ve 168. gecikmeli değerlere ek olarak 23., 72. ve 336. gecikmeli değerlerin kullanıldığı modeller de geliştirilmiştir. Kullanılan istatistiksel yöntemler SARIMA, Markov geçiş modelleri, SETAR, AR(24) ve naive metoddur. Türkiye piyasasında fiyatlarda bulunan 0'lardan dolayı, finansal enstrümanlarda her zaman kullanılan logaritmik getiriyi alma ve seriyi durağan hale getirme işlemi uygulanamamakta ve bu ciddi bir sorun yaratmaktadır. Benzer sorun çok sayıda sıfırın bulunduğu İspanya ve negatif değerlerin de bulunduğu Almanya piyasaları gibi pek çok piyasada gözlemlenmektedir. Bu makalenin getirdiği en önemli yenilik, bu sorunu ortadan kaldırmak için bir ön işlem olarak faktöriyel (factorial) ANOVA uygulaması ve seriyi bu yolla durağan hale getirmesidir. Buna ek olarak Türkiye piyasasında bu denli kapsamlı şekilde istatistiksel yöntemleri karşılaştıran ilk çalışma olan tezde, farklı zamanlarda farklı yöntemlerin başarısı görülmekle beraber, genel trend SARIMA modelinin Türkiye piyasasında, gözlemlenen zaman diliminde en başarılı model olduğudur.

İkinci makale ise aynı konuya çok daha kapsamlı bir perspektiften yaklaşmaktadır. İlk makalede eksik olan ekzojen değişkenler de makaleye eklenmiş; temel yenilik olarak ise istatistiksel modellere ek olarak, yapay sinir ağları ve derin öğrenme metodları da işin içine katılmıştır. Yine aynı şekilde Türkiye Gün Öncesi Elektrik Piyasası'nda elektrik fiyat tahmini yapmayı hedefleyen bu makalede bir önceki makaledeki istatistiksel yöntemler de kullanılmakla beraber, derin yapay sinir ağları, evrişimli sinir ağları ve devirli (recurrent) sinir ağları yöntemleri ile de tahminler yapılmaktadır. Bununla beraber, bir önceki makalede kullanılan gecikmeli fiyat değerlerine ek olarak; sıcaklık, tahmini talep/arz, gerçekleşen talep/arzın 24. gecikmeli değeri ve dengeleme piyasası fiyatının 24. gecikmeli değerleri kullanılmıştır. Türkiye piyasasında önceki üç seneyi kullanarak 2016 yılının her bir günü için elektrik fiyat tahmini yapan ve ortalamaları alarak berk sonuçlara ulaşan bu tezde, öncelikli olarak makina öğrenmesi yöntemlerinin istatistiki yöntemlerden anlamlı şekilde daha doğru tahminler yaptığı söylenmelidir. Buna ek olarak, zaman serisi problemlerinde daha başarılı olan LSTM ve GRU gibi devirli sinir ağları yöntemleri de en başarılı yöntemler olmuştur. Bir diğer önemli nokta ise daha fazla katman içeren derin sinir ağlarının, tek katman içeren sinir ağlarına göre daha başarılı sonuçlara ulaştığıdır. Bununla beraber, derin devirli sinir ağları içinde de GRU'nun LSTM'e göre daha iyi sonuçlar verdiği söylenebilir. Tüm bu sonuçların Diebold-Mariano testi ile istatistiksel anlamlılığı da saptanmıştır. Açıklayıcı değişkenlerin seçiminde, 1., 24., 48. ve 168. gecikmeli fiyat değerlerinin en önemli değişkenler olduğu gözlemlenmekte, ekzojen değişkenlerin ancak hepsi birden eklendiğinde endojen değişkenlere göre anlamlı bir tahmin performans başarısı üstünlüğüne sahip olduğu görülmektedir. Bu bağlamda, derin devirli sinir ağlarının, özellikle derin GRU'nun elektrik fiyat tahmininde kullanılmasını öneren bu çalışma, alanında ilk olma özelliğini taşımaktadır.

Üçüncü makalede ise elektrik fiyat tahminlerindeki yanlışlığın bir hidroelektrik santraline olan finansal etkileri tartışılmaktadır. Bu bağlamda, ikinci makalede kullanılan tahmin yöntemlerinin en başarılı beş tanesine ek olarak dört adet de bu tahminlerin kombinasyonundan oluşturulan hibrit modeller kullanılmıştır. Bu dokuz tahmin yöntemine göre hidroelektrik santralinde karışık tamsayılı doğrusal programlama yöntemi ile günlük üretim planlaması optimize edilmiştir. Buna göre oluşturulan üretim çizelgelerinin finansal etkileri, oluşan gerçek fiyatlara göre yapılan üretim çizelgelemesinden elde edilecek maksimum kar ile karşılaştırılmaktadır. Bu amaçla, kardan zarar gibi kimi finansal performans ölçütleri kullanılarak her bir tahmin modelinin ne kadarlık bir finansal etkiye sebep olduğu değerlendirilmektedir. Bu çalışmadaki en temel bulgu, literatür ile de uyumlu şekilde, tahmin performansı değerlendirme ölçütlerinden biri olan ortalama mutlak hataya göre en iyi model olmayan ANN-LSTM yönteminin, finansal performans ölçütlerine göre en iyi model olarak seçilmesi olmuştur. Bu da tahmin performansı ölçütleri ile finansal performans ölçütleri arasında bir çelişkiden söz etmeye sebep olmaktadır. Özellikle üretim çizelgeleme yapacak elektrik santrallerinin, kullanacakları elektrik fiyat tahmin modelini seçerken finansal performans ölçütlerine göre karar vermesi daha yerinde olacaktır. Bir diğer önemli bulgu ise hibrit modellerin başarısı olmuştur. Hibrit modellerden ANN-LSTM, diğer modellerin önünde birinci gelmekle beraber, diğer hibrit modeller de oldukça iyi sonuçlar vermektedir. Hibrit modellerin elektrik fiyat tahmininde kullanılmanın yanı sıra, elektrik fiyat tahminlerinin üretim tesisleri üzerine finansal etkileri hesaplanırken de değerlendirilmesi önerilmektedir.

Sonuç olarak, bu tez gelişmekte olan Türkiye Elektrik Piyasası'na geniş bir perspektiften bakma imkanı bulmaktadır. İstatistiksel yöntemlerin yanı sıra makina öğrenmesi temelli, yapay sinir ağları yöntemlerini ve çağımızın en önemli yeniliklerinden biri olan derin öğrenme yöntemlerini kullanan bu tez, Türkiye Gün Öncesi Elektrik Piyasası'nda elektrik fiyat tahmini yapmaktadır. Buna ek olarak, elektrik fiyat tahminlerinin önemini ve finansal etkilerini bir hidroelektrik santralini örnek alarak izlemekte ve üretimi elektrik fiyat tahminlerine göre optimize etmenin finansal etkilerini tartışmaktadır. Bu multidisipliner çalışma, Türkiye piyasasına ışık tutmanın ötesinde, global ölçekte de elektrik fiyatlarını durağan hale getirmek için bir ön yöntem olarak faktöriyel ANOVA'yı önermekte; buna ek olarak derin öğrenme yöntemlerini ve özellikle devirli sinir ağlarını da elektrik fiyat literatürüne kazandırmaktadır.

Sürekli gelişmekte olan bu alanda yeni çalışmalar pek çok daldan ilerleyebilir. Öncelikle değişken seçiminde de kullanılan Lasso regresyon ya da temel bileşenler analizi gibi yöntemler oldukça ilgi çekici sonuçlar vermektedir. Bizim de bulgularımız arasında olduğu gibi hem farklı modellerin sonuçlarını kombine etme ile oluşturulan hibrit sonuçlar, hem de farklı yöntemleri bir arada kullanma ile oluşturulan hibrit metodlar oldukça başarılı sonuçlara ulaşmaktadır. Buna ek olarak; petrol, doğalgaz fiyatları ya da döviz kurları gibi bağımsız değişkenler de özellikle gelişmekte olan piyasalarda önemli faktörler olabilir ve başka çalışmalarda değerlendirmeye katılmaları önerilir. Bir diğer nokta ise Gün İçi Piyasası'nın gelişmesi ile beraber, o alanda da çalışmalara ihtiyaç duyulmaya başlanmasıdır. Buna mukabil, enerjiye dayalı türev ürünlerin fiyatlandırılması da enerji finansıyla ilgili çalışılması gereken başka önemli bir konu olacaktır. Tüm bunların ötesinde, her bir piyasanın kendine özgü özellikler taşıdığı elektrik piyasaları ele alındığında, bu çalışmada önerilen yöntemlerin başka ülke piyasalarında, başka zaman dilimlerinde de incelenmesi çalışmanın sonuçlarının daha berk bir şekilde ortaya konabilmesine izin verecektir.

1. INTRODUCTION

In today's world, electricity has become almost as important as air or water. This paramount effect of the electricity is the result of the industrialized and the globalized system. If you assume a single-day electricity shut-down throughout a country, you can imagine the effects of such a big failure. All the systems are connected to power, which shows how essential energy is. Nowadays, the gap between the energy-importer and energy-exporter countries widen a lot, in terms of the current deficit. Turkey, as an energy-importer country, face with increasing current deficit amounts year by year. Therefore, research on the energy topic is especially important for the energy-importer and emerging countries such as Turkey.

Since the beginning of 2000s, state-owned energy companies are started to be liberalized all over the world. Turkey was not exempt from this process and most of the generation companies as well as the distribution networks are privatized. After this process, in December 2011, Energy Exchange Market is established for the day-ahead market. Balancing and intraday markets are followed this process. Although it is still an immature market, even energy derivatives are traded in Turkey nowadays. Today, about 75% of the electricity trade takes place in the bilateral contracts and approximately 20% of the trade occurs in the day-ahead market of Turkey. Most of the remaining is in the intraday market. At this point, it must be mentioned that the bilateral contracts, balancing and intraday markets, as well as energy derivatives take the day-ahead prices as a reference point. Therefore, modeling and forecasting the day-ahead electricity prices has a crucial effect.

Accurate electricity price forecasting is essential for many parties. First of all, buyers and sellers in the day-ahead market directly rely on the electricity price forecasts. They must submit their bids in terms of quantity and price for the each hour of the next day. Therefore, they need to give their bids precisely to avoid important losses. Secondly, in a similar way, balancing and intraday market participants should forecast the electricity prices. The advantage is that they have more information, most importantly day-ahead prices; and the forecast is for nearer future. Thirdly, bilateral contract parties and companies, which want to hedge by using the energy derivatives, must have medium or long-term electricity price forecasts. Fourthly, pumped storage units or some flexible companies or factories can schedule their buying behaviour according to the electricity price forecasts. Last, but not least, some types of the power plants, especially hydro ones, can make their procurement strategy according to the electricity price predicitons. With more accurate forecasts, they can decrease the loss of profits, which occur due to the self-scheduling by using the inaccurate electricity price forecasts.

Electricity differs from all the other assets, even commodities, due to its unique features such as sharp price spikes, high volatility, mean-reverting processes, seasonality in various frequencies, non-storability, demand inelasticity and the requirement of maintaining the constant balance between the demand and supply sides. Due to all these characteristics, forecasting the electricity prices has become not only an essential, but also a very challenging task. It takes the attention of the researchers from various fields to the electricity price forecasting topic, which makes it a multidisciplinary research area. According to state-of-the-art review of Weron (2014), there are mainly five ways of electricity price forecasting: Multi-agent models, fundamental models, reduced-form methods, statistical models and the computational intelligence ones. Additionally, there is an improving research area, which is called hybrid models. By hybrid models, two types of combinations are meant. First one is the combination of the electricity price forecasts from various models. Second one is the simultaneous use of the models in the electricity price forecasting process. Although, former one is a less developed research area, hybrid methods are applied widely int the latter way, and the results of these hybrid methods are relatively successful. Apart from the dimension reduction or seasonality removal type applications prior the forecast process, one of the newest topics is the application of the deep learning algorithms, especially recurrent neural networks, which are tailor-made for time-dependent data.

Although there are some research on the electricity price forecasting in the Turkish market, there is still a huge room for improvement. As Turkish market has many specific features and shares some common points with the other markets, it needs a special treatment. Turkish energy market has a very big share of hydro power by 34.3% and wind energy's share has increased to 7.4%. However, it is still fossil-fuels dominated energy market by 54.2%. Due to the high share of hydro and renewables, zero prices are frequently observed in the Turkish day-ahead market similar to the Spanish one (Diaz and Planas, 2016). It must be mentioned that if the Turkish day-ahead market electricity prices were not capped from 0 and 2000 TL/MWh, we could have observed negative prices similar to the German market (Keles et al., 2012). Turkish electricity market is an evolving market in many ways. First one is the establishment of the two nuclear power plants, which will provide the 5-7% of the installed capacity. Secondly, solar energy will be included in the grid in the near future. Thirdly, subsidies to the wind power companies will end in 2019, which will finish the 10 years wind energy subsidies era. Fourthly, transactions in the intraday market have more than doubled in the last two years. The share of the intraday market is predicted to be increased significantly in the near future. Fifthly, global warming has an upside effect on the electricity prices. As a country in the Mediterranean coast and very close to the Middle East; it affects Turkey more than most of the countries. Last, but not least, with the technological improvement and high growth rates, energy demand of Turkey follows an increasing trend. Therefore, it is difficult to generalize the results, not only to other markets, but also to the long-term future of Turkey. It is one of the reasons why continuing research is required to deal with the changes in the market. Although works of Hayfavi and Talasli (2014), Kolmek and Navruz (2015), Ozyildirim and Beyazit (2014) are highly acknowledged; today, they are a bit outdated due to the changes in the market. Furthermore, it is the first work, which provides such a wide perspective to the electricity price forecasting in the Turkish day-ahead market.

This thesis presents three published papers as independent, but strongly connected essays in Chapter 2,3 and 4. Chapter 2 has the topic of "Performance of Electricity Price Forecasting Models: Evidence from Turkey" (Ugurlu et al., 2018b). This chapter discusses the performance of the statistical methods such as SARIMA, SETAR, Markov regime-switching, and compares the forecast errors of these methods with the benchmark methods' , naive method (Nogales et al., 2002) and AR(24), forecast errors. Among the statistical methods, SARIMA model is chosen as the best performing one to forecast the electricity prices in the Turkish day-ahead market. The main contribution of this chapter is that the application of the factorial ANOVA as a method prior the forecast is a relatively good way of removing the non-stationarity in the data. Chapter 3, "Electricity Price Forecasting Using Recurrent Neural Networks" (Ugurlu et al., 2018a) proposes a novel approach to the electricity price forecasting. Although machine learning methods such as artificial neural networks, fuzzy logic or support vector machines are applied widely, it is the first application of the recurrent neural networks. Another novelty is that it applies them as deep recurrent neural networks by using 3-layers, which increases the accuracy of the forecasts significantly. This research compares machine learning methods LSTM, GRU, ANN and CNN with the statistical methods from Chapter 2. Although only endogenous variables are used in the previous chapter, Chapter 3 takes four exogenous variables into account, as well: Temperature, forecast demand/supply, 24^{th} lag of the realized demand/supply and the 24^{th} lag of the balancing market prices. This is the first research in the electricity price forecasting literature, which states the very successful performances of the recurrent neural networks and the deep learning applications. As Chapter 3 was connected to Chapter 2, Chapter 4 is strongly connected to the Chapter 3 as well. Chapter 4, "The Financial Effect of the Electricity Price Forecasts' Inaccuracy on a Hydro-based Generation Company" (Ugurlu et al., 2018c), uses the electricity price forecasts from Chapter 3. Furthermore, combination of the electricity price forecasts are also proposed as hybrid methods, in addition to the forecasts from the previous chapter. According to these forecasts and the ex-post actual prices, production schedule of a generation company is optimized by using mixed integer linear programming. This Price Based Unit Commitment problem is solved for a hydro-based power plant. Although electricity is a non-storable commodity, water can be hold in high capacities in dams. This allows the hydro power plants to organize their production strategy according to the electricity price forecasts for maximizing their profits. Reasonably, inaccurate electricity price forecasts could cause huge loss of profits, even to the relatively small generation companies. Therefore, accurate electricity price forecasting is essential for the hydro power plants. Our findings are in line with the literature (Mohammadi-Ivatloo et al., 2011; Mathaba et al., 2014) showing that there is a conflict between the electricity price forecasting evaluation methods and the financial effect measures. It means that the best method according to forecast method evaluations such as MAE is not necessarily the best method according to the financial effect measures such as loss of profit. For this reason, the evaluation of the financial effect measures and the selection of the methods according to them is essential for the power plants.

2. PERFORMANCE OF ELECTRICITY PRICE FORECASTING MODELS: EVIDENCE FROM TURKEY[1]

2.1 Introduction

Electricity differs from all other assets and even commodities due to its idiosyncratic properties such as non-storability, demand inelasticity, oligopolistic generation side and requirement of maintaining constant balance between demand and supply. Therefore, it needs a special effort and unique techniques to model and forecast the electricity prices.

Due to these features, electricity prices have some important characteristics: Seasonality, high volatility, sharp price spikes and mean reverting processes (Hayfavi and Talasli, 2014). Business activities, weather and industrial production cause intra-daily, weekly, monthly and annual seasonality, which affect the electricity prices dramatically. Non-storability and the requirement of having equilibrium between demand and supply sides, cause supply and, more importantly, demand shocks on the electricity prices. At low levels of demand, generators supply electricity at low marginal costs. However; at higher levels of demand, higher marginal cost generators also provide energy to the system, which is the main reason for the price shocks. Moreover, on the supply side, plant failures or maintenance and repair activities also activate high cost generation plants and are the reason for the price spikes to occur. When the reason for the price shock disappears, prices tend to revert back to the long term equilibrium level, which is formed mainly by the cost of production (Talasli, 2012).

Electricity price forecasting plays a major role in energy companies' decision making mechanisms (Bunn, 2004; Weron, 2006). For example, if the electric utilities

---

[1] This chapter is based on the paper "Performance of Electricity Price Forecasting Models: Evidence from Turkey". Ugurlu, U., Tas, O. & Gunduz, U. 2018. Performance of Electricity Price Forecasting Models: Evidence from Turkey. Emerging Markets Finance and Trade, 54 (8), 1720-1739.

over/under contract beforehand, and sell/buy the remaining amount in the balancing market; it might cause significant losses[2] to the companies, because it is difficult for them to pass the costs to the retail consumers swiftly. Due to the enormous volatility of electricity prices compared to other financial assets; companies need to hedge not only against volume, but against price movements as well. Therefore, price forecasting/modelling is important for all the parties; generators, utility companies or large industrial companies. Should they forecast the wholesale prices accurately, they can adjust their bidding strategy as well as their production or consumption schedule, which will decrease the risk and maximize the profit (Weron, 2014).

The Turkish Day Ahead Electricity Market is an emerging spot market, where generators and utilities submit their bids in hourly time frequency and the hourly market clearing price is determined by the Market Financial Settlement Centre. The establishment of the Turkish Day Ahead Electricity Market is due to several reasons: Firstly, it is an opportunity for the market participants to balance their portfolios in addition to bilateral contracts and providing the system operator with a balanced system. Secondly, it is used for power trading and balancing activities one day before the physical delivery of electricity. A breakthrough happened in September 2015 with the establishment of Power Exchange. Turkish Power Exchange operates the day ahead and intraday markets, and Borsa Istanbul has the operating right of derivatives market in the current situation (Avci-Surucu et al., 2016).

Turkish Day Ahead Electricity Market has unique features such as all the other markets; therefore it needs a special attention. According to Energy Exchange Istanbul Report (2016), 34.2% of the installed capacity is hydropower and 7.6% is wind power. Although fossil fuels still have the biggest share with 56.3%, the increasing trend is in favour of renewables. Governmental subsidies also support the instalment of the new wind turbines as well as hydropower plants. Turkish electricity prices are bounded

[2] It is surely possible that balancing market prices could be lower than the day-ahead market prices, however it is the last chance to buy/sell the electricity; and if prices occur in unexpectedly high/low levels, it might cause significant losses to the companies. Furthermore, in the Turkish Balancing Market, there is a 3% penalty for trading in the balancing market (EPDK, 2017).

from 0 and 2000 TL/MWh[3] and negative prices[4] are not allowed. Many zeros are monitored in the prices, especially in the early morning hours. This case is analysed in the Spanish market by Diaz and Planas (2016); they state that since 2010 many zeros have occurred in contrast to the other markets and they suggest a NATAF transformation for the electricity prices for this case, in which getting the logarithmic return series is impossible. Uniejewski et al. (2017) compare 16 variance stabilizing transformations in 12 markets[5] and find most of the transformations are better than the logarithmic returns transformation. Their best and most robust model is the probability integral transform-based N-PIT[6] in terms of Mean Absolute Error (MAE). It must also be mentioned that the prices are in a decreasing trend in the Turkish electricity market due to technological improvements in the fossil fuels plants, especially in the natural gas ones, which have the 22.5% of the installed capacity individually, in addition to the increase of the renewables' share. Another important point is that Turkey has many different climates, huge altitude differences, and great temperature variations, intra-daily and annually. According to the cooling requirement in the summer months and the heating requirement in the winter months, prices are quite high. On the other hand, price levels in the spring months are relatively low due to the activation of hydropower plants with the snowmelt effect.

Although there are some papers about electricity price forecasting in Turkish Electricity market (e.g. Talasli, 2012; Hayfavi and Talasli, 2014; Ozyildirim and Beyazit, 2014; Ozozen et al., 2016; Kolmek and Navruz, 2015) and some about the comparison of the forecast performances of various models in different markets (e.g. Aggarwal et al., 2009; Ziel and Weron, 2016) to our knowledge, this is the first work that intersects both areas. This paper will focus on the Turkish Day Ahead Electricity Market hourly prices and will use various models, such as Seasonal Autoregressive Integrated Moving Average (SARIMA), Threshold Autoregressive (TAR) and the variations of Markov regime switching; and will forecast the out-of-sample results by

---

[3] Electricity prices in the day-ahead markets are bounded in some markets for preventing big losses. See Negahdary and Ware (2016) for Alberta power prices.
[4] Negative prices is an important topic in the electricity prices, especially in the high renewables share markets such as Germany. Bublitz et al. (2017), Fanone et al. (2013) and Keles et al. (2012) can give more information about the negative prices.
[5] 11 of them are European markets and the other one is the GEFCom2014 competition data.
[6] Interestingly, this is the NATAF transformation of Diaz and Planas (2016). Uniejewski et al. (2017) mention that Diaz and Planas (2016) call NATAF transformation to the N-PIT transformation, misleadingly.

using the past prices as a univariate variable. These forecasts will be completed for the hours of every 15th day of 2016, for a week of winter and for a week of summer and the daily averages of MAE and Root Mean Square Error (RMSE) results will be compared for various methods, which will give the opportunity to choose the best model. More importantly, this paper applies a factorial Analysis of Variance (ANOVA) process to the price series as a pre-whitening method and works with the residuals series before transforming back to the price series. This method solves the non-stationarity[7] problem of the price series and also makes the series more linear, which gives a reasonable success chance to the linear forecasting methods.

Section 2.2 gives a comprehensive literature review, mostly about the time series methods in the electricity price forecasting. Section 2.3 is about the methodology, which also discusses the models used in the application. Section 2.4 illustrates the results in terms of MAE and RMSE for the forecasted days. Section 2.5 concludes and mentions some further research ideas.

2.2 Literature Review

There are many different ways of electricity price modelling and forecasting. Weron (2014) consolidates the methods under 5 main topics: Multi-agent, fundamental, reduced-form, statistical and computational intelligence. Multi-agent and fundamental models are difficult to apply due to various reasons. Both models require having all the data readily available. Multi-agent models are better to apply in small markets with relatively small number of players. Parameter rich fundamental models need the "all supply information" from generators and then intersect the supply curve with the forecasted demand curve to have the amount of demand and the price for the market. It is quite difficult to use this method in such a big market like Turkish Day Ahead Electricity Market with over 2000 generators and many more companies. Computational intelligence doesn't provide enough information about the working mechanisms or the coefficients of the parameters. These methods only use the inputs and give the final results. Other models, reduced-form and statistical, could be gathered under one topic, as well. However, reduced-form models, jump-diffusions and Markov regime-switching, focus more on the price spikes and the statistical

[7] It is impossible to remove the non-stationarity by taking log returns due to the zeros.

models are regression based and use past electricity prices in addition to exogenous variables.

The following papers can be categorized in two groups: Papers about electricity price modelling and forecasting in the Turkish electricity market and the papers that compare the performance of the models with each other for the specific markets. The first group starts with the following paper of Taysi et al. (2015), and the second group starts with the very comprehensive comparison of Ziel and Weron (2016).

Taysi et al. (2015) combine time series statistics with a neural network model in their paper for Turkish electricity market. The time series statistics method is SARIMA model, a seasonal approach of ARIMA, and the artificial intelligence model is named a feed forward neural network. Both methods use historical electricity price data and the performance of them is very close to each other with the average error rate of 8.5% for weekly frequency forecasting.

Ozyildirim and Beyazit (2014) forecast and model the electricity prices by radial basis function method, which implement a totally new approach to electricity price forecasting, in addition to a conventional linear regression technique. Regression method takes lagged series of price, time trend, hourly temperature degrees, square of temperature degrees and cosine function of time and all the dummy variables for seasonality, as independent variables; the dependent variable is the electricity price. Moreover, it takes into account the recurring structure of hourly prices and in this sense, proposes a radial basis function, which fits to the structure of data. Consequently, out of sample performance of radial basis function method slightly outperforms the regression method for a specific estimation period.

Ozozen et al. (2016) has a good starting point by combining seasonal ARIMA and artificial neural network (ANN) models and apply it to the Turkish hourly electricity prices, therefore it is in the interest area of this thesis. It gives a detailed discussion of the Turkish electricity market and the price formation process, which shows the authors have a strong understanding of the electricity price market in Turkey. This paper applies Box-Jenkins method to split the data into daily and hourly parts, then applies SARIMA to both methods and combines the forecasts. Moreover, it takes the error terms of this forecast and proposes an ANN model to these forecasts and combines both results, which outperforms the only SARIMA forecasts in in-sample and out-of-sample testing. So far, however, they choose the most stable periods for in-sample and out-of-sample testing, which give them the opportunity to report the best Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) results.

Talasli (2012) models the electricity markets stochastically in her thesis and analyses the Turkish electricity market in great detail. The thesis of Talasli (2012) tries to capture all the characteristics of the electricity prices such as mean reversion, seasonality and the spiky behaviour. Prices are modelled as a summation of deterministic function, which captures the seasonality, and multi-factor stochastic process, which takes care of spiky behaviour and mean reversion. However, the method is mainly based on jump diffusion models; and Markov regime switching models have not been used in the thesis. Another important point to mention is that the thesis focuses on the daily prices and tries to forecast at a daily frequency. Hayfavi and Talasli (2014)'s proposed stochastic multifactor model uses daily spot market electricity prices and composes three jump processes which take into account the spiky behaviour and mean reversion, and an iterative threshold function, constructed by GARCH (1, 1), is used to separate the jumps. The main addition to the thesis (Talasli, 2012) is that the results are compared with the mean-reverting jump diffusion model of Cartea and Figueroa (2005) and Markov regime switching model of Janczura and Weron (2010). Although the authors compare the four moments of the three methods and evaluate that the multifactor model performs better than the others, their method needs further performance evaluation measures.

Ziel and Weron (2016) conduct an extensive empirical study on electricity price forecasting models and compare their performance in different electricity markets. They take expert models as benchmarks and combine these 32 multi-parameter regression models and compare their performance in 12 power markets. They have an important finding that the multivariate modelling approach does not uniformly outperform univariate models across all datasets. This paper also strengthens our point of view by saying that combining advanced structures or the corresponding forecasts from both modelling classes may bring further improvement in the forecast accuracy. They also try to analyse variable selection of the best performing models and give an opinion about variable selection in their paper. It must be mentioned that they use mean absolute error (MAE) as the performance measure and Diebold and Mariano (1995) test to compare the performances of the models statistically.

Nowotarski et al. (2014) work on the performance of the forecast averaging in electricity price forecasting literature. They use seven averaging and one selection scheme and perform back-testing in the day-ahead electricity prices of three major European and US markets. They indicate that combining the forecasts of individual methods helps to make more accurate predictions. However, it is not uniform in all the markets and periods. Moreover, not all the averaging methods are as successful as the others.

Janczura and Weron (2010) stress one of the most important features of the electricity prices: spikes. It is clear that one of the most convenient methods for modelling the spikes is Markov regime-switching models. They try to calibrate and test Markov regime-switching models, which are successful in forecast accuracy as well as statistical correctness. They find the best models as independent spike 3-regime model with time-varying transition probabilities. It allows seasonal spike intensity throughout the year as well as consecutive spikes or price drops, which is consistent with electricity prices.

An overview literature review is given in Table 2.1. Some important information about the papers, such as methodology, regional scope, time frequency, training data, predicted period and time horizon are given in addition to level of forecast accuracy and key findings.

Table 2.1: An overview literature table.

Study	Methodology	Regional scope	Training Data	Time frequency	Predicted period	Time horizon	Level of accuracy	Key findings
Keles et al. (2016)	ANN	Germany	2008-2012	Hourly	First 9 months of 2013	1 hour-ahead	MAE 4.67 €/MWh, RMSE 6.58 €/MWh	The selected ANN model is compared to other models such as SARIMA and forecast error are smaller in ANN. It also mentions that ANN requires relatively less observations.
Mandal et al. (2005)	ANN	Australia	2000-2002	Half-hourly	2003	1 and 6 hour-ahead	MAPE 1 hour-ahead 9.75%, 6 hour-ahead 20.03%	MAPE results confirm that the proposed ANN models are good tools compared to the other simple methods in terms of accuracy and convenience.
Gonzalez et al. (2005)	Input/Output Hidden Markov	Spain	January-June 2001	Hourly	July-September 2001	1 hour-ahead	MAPE 15.83%	Non-stationary model categorizes the spot prices into three fundamental regimes. Different probability density functions can be used to evaluate diferent risk measures and for the generation of scenario trees. This model is like a piecewise dynamic regression and the state is chosen by the underlying Markov chain automatically.
Conejo et al. (2005a)	Wavelet transformation, then ARIMA	Spain	2002 (previous 48 days)	Hourly	4 weeks of 2002 (different seasons)	24 hour-ahead	MAPE; Winter 4.78%, Spring 5.69%, Summer 10.70%, Fall 11.27%	Wavelet-ARIMA, which uses the wavelet transformation as pre-processing method outperforms ARIMA and Naive method.
Shafie-Khah et al. (2011)	Wavelet transformation, then Radial basis function neural network	Spain	Previous 50 days	Hourly	4 weeks corresponding to four seasons of 2002	24 hour-ahead	MAPE, ARIMA-RBFN; February 4.27%, May 4.58%, August 6.76%, November 7.35%	Wavelet-ARIMA is used to analyze the linear part and RBFN network worked on the residuals of the wavelet-ARIMA. It needs lower number of input data and the results of this hybrid method are more accurate compared to other methods.
Ziel et al. (2015)	Time series models	12 European Markets	Previous 2 years (rolling window)	Hourly	13.08.2012 to 12.08.2014	24 hour-ahead	i.e. EPEX.DE&AT 2-dimensional AR(p) MAE 4.59 €/MWh, RMSE 7.96 €/MWh	This study investigates the EXAA day-ahead prices' effect on the other electricity day-ahead markets. Various time series methods are applied on 12 different markets; MAE and RMSE results vary from market to market. The main finding is that using the EXAA prices, which announced before the price submission of other markets, help to the forecast process.

Table 2.1 (continued) : An overview literature table.

Study	Methodology	Regional scope	Training Data	Time frequency	Predicted period	Time horizon	Level of accuracy	Key findings
Ziel and Steinert (2016)	Fundamental model	Germany and Austria	Previous 2 years (rolling window)	Hourly	01.11.2014 to 19.04.2015	24 hour-ahead	MAE 4.35 €/MWh, RMSE 6.46 €/MWh	The model estimates the price as the intersection of the sale and purchase curves. This model allows to predict extreme and rare price events. In addition, it also outperforms most of the consisting models.
Contreras et al. (2003)	ARIMA	Spain and California	Previous months (2-5 months) in 2000	Hourly	3 weeks Spanish and 1 week Californian market.	24 hour-ahead	MAPE average for 3 week, 10% in Spanish market and 11% in California	These are very good forecast results compared to previous studies. Explanatory vairables are only needed when the correlation is high between hydro production and price.
Cuaresma et al. (2004)	Time series models	Germany	16.06.2000 to 31.08.2001	Hourly	01.09.2001 to 15.10.2001	168 hour-ahead	Best model is ARMA model with jumps RMSE 3.99 €/MWh and MAE 2.57 €/MWh	Hour-by-hour modelling strategy improves the forecasting abilities of linear univariate time series models and inclusion of the simple probabilistic approaches could improve the forecast ability.
Weron and Misiorek (2008)	12 Time series methods	California and Nordic	05.07.1999-02.04.2000 and 07.04.2003-05.12.2004 (expanding window)	Hourly	10 weeks in 2000 and 4 weeks in 2004	24 hour-ahead	One of the best models is AR model with a smoothed nonparametric ML estimator; MAE 13.87 for California, 4.04 for Nordpool and 3.22.	In the California market system load as an exogenous variable has a positive effect on the forecast accuracy. On the other hand, air temperature in the Nordpool data doesn't have the same effect. Semiparametric models perform quite well, especielly under different market conditions.
Karakatsani and Bunn (2008)	Time-varying parameter models	UK	06.06.2011 to 01.04.2002 (expanding window)	Half-hourly	17.01.2002 to 01.04.2002	24 hour-ahead	Various results; i.e. For time-varying parameters regression of 35th period MAE 1.14, RMSE 1.48	Price models with fundamentals and their time-varying effects outperform the various alternatives including autoregressive models with similar coefficients.
Bordignon et al. (2013)	Time-varying parameter models	UK	01.04.2005 to 31.12.2005	Half-hourly	01.01.2006 to 30.09.2006	24 hour-ahead	Various results; i.e. For period 18, best individual model, MAE 3.82, MAPE 9.05 and best combination of models, MAE 3.79, MAPE 8.86	The paper forecasts with 5 individual models and then combine with the simple average combinations. In comparisons combined forecasts outperform individual models in 76% of the cases, but the finding is not significant in most of the cases.

2.3 Methodology

2.3.1 Data

Hourly price series are obtained from EPIAS[8] from 01.01.2013 to 31.12.2016, which contains 35064 observations[9]. It is observed that there are many zeros in the price series as well as sharp price spikes, which cause outliers.

2.3.2 Descriptive statistics

Electricity prices have different levels of seasonality such as intra-daily, weekly, monthly and annually. The most challenging and the most effective impact is due to intra-daily seasonality, which causes substantial differences in the prices of various hours. Therefore, descriptive statistics according to hours of the day are given in Table A.1. It should be mentioned that price levels are relatively low in the night hours[10] from 2-7 and the highest standard deviation also occurs in these hours. Prices for most of the hours (21 out of 24) are left skewed and the kurtosis levels are relatively low all the hours, however it is shown that there are some zeros in the hours 1-7.

Before mentioning this problem, Figure 2.1 indicates the intra-daily and weekly seasonality. Left panel of the Figure 2.1 shows the distribution of the hourly prices (TL/MWh) in 24-hours point of view. Prices tend to be relatively low, mainly due to the consumption levels during the night; then increase sharply to 11 and have a global maximum at 181.14 TL/MWh. Hereafter, they decrease slightly during lunch break and have the second maximum at 14. Although prices decrease smoothly by 21, they have another peak at 22 due to the high level of consumption in the households. Right panel of the Figure 2.1 illustrates the weekly seasonality in terms of electricity prices based on the hours of the week. It shows the change of the prices according to the 168-hours of the week. Prices on the weekdays[11] and Saturday follow a similar pattern up until the afternoon of Saturday[12], which follows a decreasing trend and have a

[8] (Url-1)
[9] 8760 hours for 2012-2015 and 8784 for 2016.
[10] Numbers given under the hours represent the following hour. For example; 1 represents, 01:00:00-01:59:59.
[11] "Weekdays" is used for Monday, Tuesday, Wednesday, Thursday, and Friday in this case.
[12] In Turkey, most of the companies work half-day on Saturday and factories work all-day.

minimum on Saturday night, which is followed by low price levels on Sunday and another minimum on Sunday night.

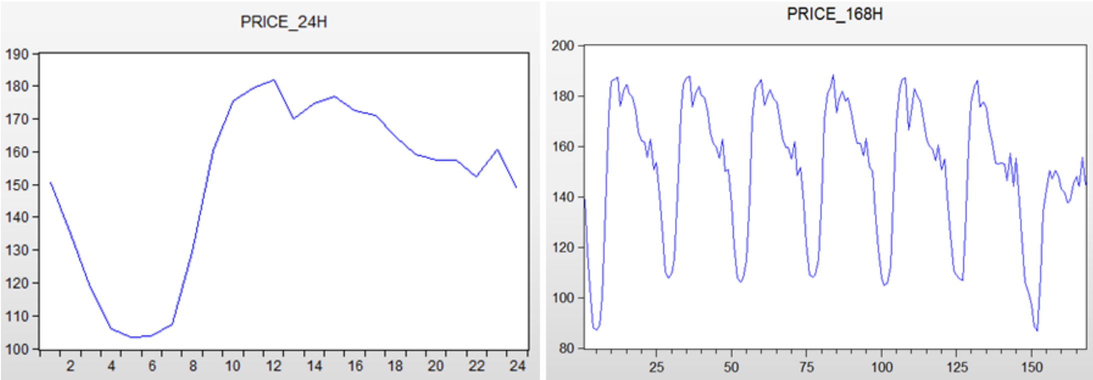


Figure 2.1 : Left panel:Price distribution of hourly prices (TL/MWh) according to the hours of the day. Right panel: Price distribution of hourly prices (TL/MWh) according to the hours of the week (based on 168 hours).

2.3.3 Factorial ANOVA as a pre-whitening method

Prior to applying the statistical methods on the price series, the Augmented Dickey-Fuller test is implemented to check the stationarity of the series. The null hypothesis of the price series has a unit root, couldn't be rejected even at 10% significance level[13], which means that the price series is non-stationary.

Therefore, logarithmic returns of the series should be taken. However, due to many zeros in the series, it is impossible to take the log returns. Even though price series is not stationary, estimation and forecast of the models are tried on the level price series, however estimations had singular covariance problem. Therefore coefficients were not unique and we couldn't obtain the standard errors, which do not allow us to use these coefficients.

For this reason, a factorial Analysis of Variance (ANOVA) process is applied as a pre-whitening method to eliminate the deterministic part and have stationary series. Factorial ANOVA equation, which is applied to the price series as a pre-whitening method could be found below in equation 2.1.

$$\mathrm{Y_i} = a_t + \sum_{j=1}^{31} \mathrm{b_j}(\mathrm{Day})_\mathrm{j} + \sum_{k=1}^{12} \mathrm{c_k}(\mathrm{Month})_\mathrm{k} + \sum_{l=2013}^{2015} \mathrm{d_l}(\mathrm{Year})_\mathrm{l} + \sum_{m=1}^{24} \mathrm{e_m}(\mathrm{Time})_\mathrm{m} + \sum_{n=1}^{7} \mathrm{f_n}(\mathrm{Weekday})_\mathrm{n} + \sum_{o=0}^{1} \mathrm{g_o}(\mathrm{Holiday})_\mathrm{o} + \varepsilon_t \quad (2.1)$$

[13] P-value of the Augmented Dickey-Fuller test is 0.1175.

ANOVA equation uses the 3 years data in an expanding window basis and splits the price series into deterministic and stochastic part by using the day of the month, month of the year, year, hours of the day, day of the week and holiday effects.

Although factorial ANOVA has some specific features, it is still a special case of the multiple linear regression. Firstly, it is a simple and appealing approach of understanding the analysis of covariance, which is a quite difficult to understand technique in the traditional way. Secondly, it is a relatively easy and straight-forward way of handling the unequal sample sizes. According to Howell (2013), last and most important one, which is also the reason of our choice, is that the ease of the application with the improved performance of the computers. It is the easiest way of applying a dummy-type multivariate linear regression. Having said these advantages, the novelty in our approach is still present. Most of the papers in literature, such as Ziel and Weron (2016), use that type of multivariate linear models; however, to our knowledge, all of them use them as direct forecast methods. In our paper, factorial ANOVA (a special case of multivariate linear regression) is used as a pre-processing technique before performing the forecast models. It also takes all the important variables; such as hours of the day, days of the week, months of the year, years and holiday effects; together into account[14].

A dependent variable is taken as the price series and the independent variables are hours of the week, days of the month, days of the week, months of the year, holiday or normal day and the years. Only main effects are used and the interaction effects[15] are not taken into account because of the model's size. Effects of between-subjects and the parameter estimates for 01.01.2013-14.01.2016 can be found in the Table A.2 and A.3, respectively. The rationale behind applying factorial ANOVA is that we know the independent variables, which we put in the ANOVA process. For example, we know that 15.01.2016 is a Friday, normal day, 15th day of the month, 2016, January. It is also known, which hour will be forecasted as well. Therefore, it is not needed to forecast this deterministic part. Calculations for finding the value of this deterministic

---

[14] ANOVA has the assumptions of the independence of the cases, normal distribution of the residuals, homoscedasticity and no multicollinearity. It is very difficult to fulfill them, which was the case for us even after attempts with various transformations from Uniejewski et al. (2017). It must be kept in mind that ANOVA assumptions couldn't be fulfilled.

[15] Interaction effects are the effects between the independent variables. For example, if day of the week have an additional effect via month of the year, it is named as interaction effect. It is a kind of combined effect of two or more independent variables.

part and the remaining from the price, which is called residuals, are given in Table A.4 for two different example hours.

This process is applied for all the values in SPSS 17 for the 26280 observations between 01.01.2013-31.12.2015 and the price series transformed to the residuals series. A demonstration of the price and the residuals series can be found in Figure 2.2. Application of the Augmented Dickey-Fuller (ADF) test proves that the residuals series is stationary[16] [17].

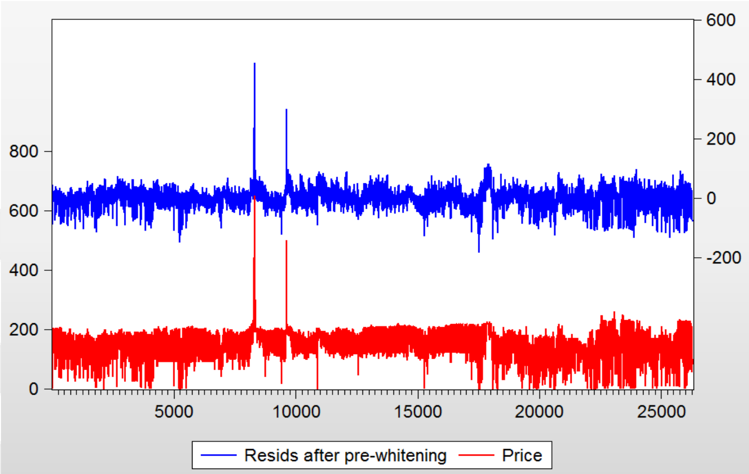


Figure 2.2 : Prices and resids after pre-whitening for 2013-2015.

2.3.4 Models

Naive method and AR(24) models are used as benchmark models in the paper. SARIMA, TAR[18] with less lags (TAR-less), TAR with more lags (TAR-more), Markov-2-regime switching with less lags (Markov-2-less), Markov-2-regime switching with more lags (Markov-2-more) and Markov-3-regime switching models (Markov-3-less and Markov-3-more) are applied to the residuals series. In this section, these models will be discussed.

[16] P-value of the Augmented Dickey-Fuller test is 0.0000.
[17] Although autocorrelation partial autocorrelation functions of the residuals show that the effect of the autocorrelations and partial autocorrelations are decreased in the residuals, according to TBATS model (De Livera et al., 2011) seasonality is still in the residuals (stochastic part) for the 24th and 168th lags.
[18] It is actually a SETAR model, which will be discussed in this section.

2.3.4.1 Naïve method

Naive method as a similar day technique, which is invented by Nogales et al. (2002) mentions that the hourly prices of Monday, Saturday and Sunday are the same with the previous week's corresponding hour and day; and Tuesday, Wednesday, Thursday and Friday take the previous day's hourly prices.

$$P_{d,h} = \begin{cases} P_{d-7,h} + \varepsilon_{d,h}, & \text{Monday,Saturday,Sunday} \\ P_{d-1,h} + \varepsilon_{d,h}, & \text{Tuesday,Wednesday,Thursday,Friday} \end{cases} \tag{2.2}$$

According to Nogales et al. (2002) and Conejo et al. (2005a), forecasting procedures, which are not calibrated well enough, can't outperform the naive method reasonably often.

2.3.4.2 AR(24) model

AR(24) process is included in the application as a second benchmark. It assumes that the price series is affected only by the previous day's same hour prices. AR(24) process can be defined as follow:

$$P_t = \beta P_{t-24} + \varepsilon_t \tag{2.3}$$

As it is examined in many papers, such as Uniejewski et al. (2016), 24th lag of the price series is one of the most important variables in the electricity price forecasting literature. Therefore, this method is also selected as a benchmark in addition to the naive method.

2.3.4.3 SARIMA model

ARMA (p,q) models try to forecast the spot prices by its p past values (autoregressive part) and q previous values of the noise (moving average part). ARMA modelling approach requires that the time series are stationary and ARMA model, which includes the differencing in the formulation is called ARIMA (p,d,q). d is the number of the differences of the series to obtain a weak form stationarity. There are many types of the ARIMA-type models, but the most important one in electricity pricing due to the nature of the prices is seasonal ARIMA, named SARIMA[19]. In this paper a comprehensive triple SARIMA model, which deals with the previous hour's, previous

[19] For detailed information about the SARIMA model, see Box and Jenkins (1976).

day's and previous week's same hours, is used. Below in equation 2.4, single seasonal ARMA model is given.

$$\phi_p(L)\Omega_{P_2}(L^{s_2})(y_t - a - bt) = \theta_q(L)\Psi_{Q_2}(L^{s_2})\varepsilon_t \qquad (2.4)$$

yt is the price in period t; a is a constant; b is the coefficient of the linear deterministic trend term, εt is a white noise error term; L is the lag operator; and ϕp, Ω_{P_2}, θ_q, and Ψ_{Q_2} are polynomial functions of orders p, P_2,q and Q_2, respectively (Taylor 2010).

Single SARIMA model could be used for the intraweek cycle, if it is enough for the estimation, but in the high seasonality environment of the electricity prices, generally an improved model is required. Therefore, in our case triple SARIMA model is used. Triple SARIMA model can be found below in equation 2.5 (Taylor, 2010).

$$\phi_p(L)\Phi_{P_1}(L^{s_1})\Omega_{P_2}(L^{s_2})\Gamma_{P_3}(L^{s_3})(y_t - a - bt) = \theta_q(L)\Theta_{Q_1}(L^{s_1})\Psi_{Q_2}(L^{s_2})\Lambda_{Q_3}(L^{s_3})\varepsilon_t \qquad (2.5)$$

The main difference from the single SARIMA is that there are four more terms, Φ_{P_1}, Γ_{P_3}, Θ_{Q_1}, and Λ_{Q_3} are added to the triple SARIMA model. These are the polynomial functions of P_1, P_3, Q_1, and Q_3, respectively. These functions enable ARMA modeling of the intraday and intrayear cycle. SARIMA model is performed by maximum likelihood assuming Gauss-Newton optimization.

There are double and triple seasonal ARIMA applications in the literature. Some double SARIMA applications are Gould et al. (2008), Cancelo et al. (2008), Sunaryo et al. (2011) and triple SARIMA application is Taylor (2010). Our model can be shown as a triple SARIMA (23,0,23), $(1,0,1)_1$ x $(1,0,1)_{24}$ x $(1,0,1)_{168}$.

2.3.4.4 Threshold autoregressive (TAR) model

Threshold autoregressive models are the models, which have different regimes according to a threshold variable. The difference between these models and the Markov-regime switching models is that they have an observable threshold variable. On the other hand, the threshold variable is latent in the Markov-regime switching models. When the threshold variable is found by the model itself, it is called Self Exciting Threshold Auto-Regressive (SETAR) model.

A time series x_t is said to follow a k-regime self exciting TAR (SETAR) model with threshold variable x_{t-d} if it satisfies the equation 2.6.

$$x_t = \phi_0^{(j)} + \phi_1^{(j)}x_{t-1} + \cdots + \phi_p^{(j)}x_{t-p} + \alpha_t^{(j)}, \quad \text{if } \gamma_{j-1} \le x_{t-d} \le \gamma_j \qquad (2.6)$$

Where k and d are positive integers, j= 1, …, k, γi are real numbers such that $-\infty = \gamma_0 < \gamma_1 < \cdots < \gamma_{k-1} < \gamma_k = \infty$, the superscript (j) is used to signify the regime, and $\{a_t^{(j)}\}$ are iid sequences with mean 0 and variance σ_j^2 and are mutually independent for different j. The parameter d is referred to as the delay parameter for different regimes. As it can be seen from the equation 2.6, SETAR model is a piecewise linear AR model. However, SETAR model is nonlinear provided that k > 1 (Tsay, 2005). Thresholds are determined according to sequential Bai-Perron L+1 breaks vs. L test.

For SETAR model, it is impossible to use a simple iterative scheme to generate multiperiod forecasts. Therefore, bootstrap simulation method is used to forecast 24-step ahead prices. The threshold variable is determined according to first lag values of the series (d=1). 1st, 24th, 48th, and 168th lags are used in the estimation of TAR-more model. 24th and 168th lags are used for the TAR-less model, delay parameter stayed same.

2.3.4.5 Markov regime switching model

In this paper, 3 different Markov regime switching models are used. Actually, these models are Markov-regime switching AR (MS-AR) models. First one is a 2 state Markov chain with relatively less lags than the second 2 state model. 1st, 24th, 48th, and 168th lags are used in the estimation. Second one is the 2 state MS-AR model with the addition of 23rd, and 72nd. Third one is the 3 state MS-AR-less model with the same parameters of the first one, 1st, 24th, 48th, and 168th; and the last one is the 3 state MS-AR-more model with the same parameters of the second one. This is the general representation of MS-AR model:

$$y_t = \alpha_s + \sum_{i=1}^{p} \phi_{s,i} y_{t-i} + \epsilon_t \tag{2.7}$$

where st is a two-state discrete Markov chain with S= {1, 2}, or three-state with S{1, 2, 3} for the last model, and $\varepsilon t \sim$ i.i.d. $N(0, \sigma^2)$. The estimation of MS-AR models are performed by maximum likelihood algorithm expectation-maximization.

Point forecasting is less complicated compared to the other models such as TAR-type models. The h-step forecasts from the MS-AR model are

$$\hat{y}_{t+h|t} = P\,(s_{t+h} = 1|\, y_t, \dots, y_0)\big(\alpha_{s=1} + \textstyle\sum_{i=1}^{p} \hat{\phi}_{s=1,i}\, y_{t+h-i}\big) + P\,(s_{t+h} = 2|\, y_t, \dots, y_0)\big(\alpha_{s=2} + \textstyle\sum_{i=1}^{p} \hat{\phi}_{s=2,i} y_{\,t+h-i}\big) + P\,(s_{t+h} = 3|\, y_t, \dots, y_0)\big(\alpha_{s=3} + \textstyle\sum_{i=1}^{p} \hat{\phi}_{s=3,i}\, y_{t+h-i}\big) \tag{2.8}$$

where $P\,(s_{t+h} = i|\, y_t, \dots, y_0)$ is the ith element of the column vector $P^h \xi_{t|t}$. In addition, $\xi_{t|t}$ represents the filtered probabilities vector, and Ph is the constant transition probabilites matrix[20] (Ozkan and Yazgan, 2015). Coefficient covariance matrix is Hessian, optimization method is BFGS and step method is Marquardt in the estimation.

2.3.5 Application

Naive method and AR(24) models are used as benchmark models in the paper. SARIMA, TAR with less lags (TAR-less), TAR with more lags (TAR-more), Markov-2-regime switching with less lags (Markov-2-less), Markov-2-regime switching with more lags (Markov-2-more) and Markov-3-regime switching models (Markov-3-less and Markov-3-more) are applied to the residuals series. In this section, these models will be discussed.

Hourly forecasts of the 15th days of each month in 2016 and the forecast for all the hours of the days in a winter week (18-24 January 2016) and in a summer week (4-10 July 2016) are performed by using the following steps:

1. Residuals series are obtained by using the hourly data from 01.01.2013 to the previous day of the forecast[21] day in SPSS 17.
2. Estimation of the equations[22] are done for AR(24), SARIMA, TAR with less variables, TAR with more variables, Markov-2-regime switching with less variables, Markov-2-regime switching with more variables, Markov-3-regime switching with less and Markov-3-regime switching with more variables by using these residuals in Eviews 9.
3. 24-step-ahead out-of-sample forecasts are calculated for the forecast days by using the estimations in Eviews 9.
4. Deterministic part is computed for the forecast day by using the data from 01.01.2013 to the forecast day in SPSS 17.

[20] See Hamilton (1994).
[21] Expanding window scheme has been used. The estimation period was from 01.01.2013 to the previous day of the forecast day.
[22] Lag selections are done according to autocorrelation – partial autocorrelation functions and all the estimations are available upon request.

5. These forecast components from part 3 and deterministic components from part 4 are added for the forecast day to have the price forecast for the model.
6. Naive model values are computed for the forecast day by using the price series.
7. Actual prices are taken from the system.
8. Absolute errors and square errors are calculated for the 24 hours of the forecast day.
9. Mean absolute errors (MAE) and root mean square errors (RMSE) are calculated.
10. First 9 steps are repeated for 6 models and for all the forecast days.
11. Averages of MAE and RMSE are calculated for all models.

2.4 Results

Electricity differs from all other assets and even commodities due to its idiosyncratic properties such as non-storability, demand inelasticity, oligopolistic generation side and requirement of maintaining constant balance between demand and supply. Therefore, it needs a special effort and unique techniques to model and forecast the electricity prices. Forecasts are done on the 15th day of each month and on the all days of a winter week and all days of a summer week by using SARIMA, TAR (less), TAR (more) Markov-2 (less), Markov-2 (more) and Markov-3 (less) and Markov-3 (more) models in addition to benchmark naive and AR(24) models, which were discussed in Section 2.2.3. As an example, comparison of the forecasted values for each model and the actual values for November 15 are given in Table A.5.

The main performance evaluation criteria are Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) in the electricity price forecasting literature. Although MAPE measures give the opportunity to compare the electricity price forecasts from various markets, it is most of the time misleading when the actual prices are closer to zero. In these times, MAPE values become very large and affect the performance measurement remarkably. On the other hand, when electricity prices spike and have very high values, then MAPE values become very small. Furthermore, in the markets, which allow negative prices, MAPE values get negative numbers and it is difficult to interpret these values (Weron, 2014).

Therefore, only MAE and RMSE are used in the performance evaluation of the forecasts. Both of them based on the absolute errors (AE), which is $\mathrm{AE_h} = \left|\mathrm{P_h} - \widehat{\mathrm{P}}_\mathrm{h}\right|$. This is the difference between the forecasted value and the actual value. MAE is computed as the mean of T=24 as a daily MAE.

$$\mathrm{MAE} = \frac{1}{\mathrm{T}} \sum_{\mathrm{h}=1}^{\mathrm{T}} \left|\mathrm{P_h} - \widehat{\mathrm{P}}_\mathrm{h}\right| \tag{2.9}$$

RMSE methods uses the squares of the absolute errors. Second performance evaluation method used in this paper is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{T}} \sum_{\mathrm{h}=1}^{\mathrm{T}} \left(\mathrm{P_h} - \widehat{\mathrm{P}}_\mathrm{h}\right)^2} \tag{2.10}$$

For T=24, RMSE value is obtained.

MAE and RMSE results are calculated for each of the 15th days of the months. MAE results are given in Table 2.2 and RMSE results are given in Table 2.3 in the form of heat maps, respectively. Green shows the most successful method and red shows the least successful one in the daily basis (line wise). Our results state the difference between the methods. There are relatively big differences between the methods for each of the days.

As an example, we can analyse 15th of February; MAE of SARIMA is at least 5.61 TL/MWh lower than the competitors. It is even less than the half of naive method. We can also observe the effect of the day; for example, SARIMA method has MAE of 7.83 for the 15th of November, in comparison to 38.48 for the 15th of December. Although, it is seen that in some days, especially in May, July and August, naive method couldn't be outperformed, in general SARIMA method seems the most successful one. The averages of 12 days also support our finding; SARIMA model is better than the closest rival by 3.12 TL/MWh. However, second best method is the naive method with close differences. A clear success of the SARIMA model can be mentioned in 7 of the 12 observations. In other 2 days, it outperforms all the models, except naive method.

Table 2.2 : MAE results for the models according to the days of each month (green→ most successful, red→ most unsuccessful).

MAE	Naive	AR(24)	SARIMA	TAR(less)	TAR (more)	Markov-2 (less)	Markov-2 (more)	Markov-3 (less)	Markov-3 (more)
January	28.15	28.76	30.19	27.57	27.51	26.76	27.90	26.32	26.73
February	48.43	33.88	22.56	29.03	29.79	30.51	28.17	32.27	32.77
March	29.50	27.11	25.87	21.08	22.05	20.83	22.00	21.52	22.55
April	18.79	16.13	15.55	18.29	18.12	18.94	20.62	17.72	17.74
May	19.03	38.37	29.66	34.42	33.13	38.12	33.08	39.05	38.13
June	37.30	31.88	30.60	37.66	36.28	36.62	35.16	38.03	36.26
July	5.61	19.61	27.90	29.38	27.75	28.21	26.09	25.25	23.77
August	19.83	47.85	30.18	42.21	43.84	44.23	44.08	47.53	49.00
September	39.02	21.07	20.95	24.01	24.01	23.90	22.55	24.39	24.69
October	26.70	13.79	13.14	15.23	15.75	15.67	14.28	14.68	15.16
November	11.50	10.89	7.83	11.72	10.70	12.18	11.07	11.05	9.98
December	46.49	53.48	38.48	44.39	42.77	46.57	46.57	42.48	42.03
Average	**27.53**	**28.57**	**24.41**	**27.92**	**27.64**	**28.54**	**27.63**	**28.36**	**28.23**

When it is checked according to RMSE, results are very similar (Table 2.3). SARIMA outperforms all the other methods in 5 observations. According to the averages, RMSE of SARIMA is lower than the closest rival by 3.22 TL/MWh. The most important difference between MAE and RMSE is that the benchmarks are outperformed by all the models in terms of RMSE averages. In terms of MAE averages, benchmark naive model was the second best after SARIMA model (Table 2.2).

Table 2.3 : RMSE results for the models according to the days of each month (green→ most successful, red→ most unsuccessful).

RMSE	Naive	AR(24)	SARIMA	TAR (less)	TAR (more)	Markov-2 (less)	Markov-2 (more)	Markov-3 (less)	Markov-3 (more)
January	44.79	40.01	40.07	33.87	33.93	32.97	33.44	36.70	36.51
February	57.83	42.47	30.60	36.73	37.24	38.63	35.69	39.91	39.99
March	38.39	35.97	37.87	30.48	31.13	30.73	31.20	30.49	31.09
April	24.79	21.81	21.97	24.08	24.20	24.31	25.61	24.03	24.20
May	26.43	48.21	36.90	44.64	43.46	49.16	42.25	49.56	48.80
June	46.06	36.54	35.60	43.65	42.15	42.64	40.90	43.99	42.32
July	7.55	21.63	30.19	34.18	32.38	32.87	31.23	29.49	28.01
August	32.65	57.92	40.16	54.67	55.22	55.92	54.99	58.92	59.61
September	49.14	26.51	28.96	29.73	29.32	30.40	29.35	28.77	28.65
October	31.52	17.84	15.96	18.79	19.13	18.88	17.33	17.96	18.59
November	15.43	14.11	10.63	15.42	15.32	15.85	14.18	14.39	14.38
December	63.27	61.91	49.38	57.12	54.20	58.09	60.76	55.31	53.35
Average	**36.49**	**35.41**	**31.52**	**35.28**	**34.81**	**35.87**	**34.74**	**35.79**	**35.46**

In the second part of the application one winter week (18-24 January 2016) and one summer week (4-10 July 2016) are examined. Results can be found in Table 2.4 and Table 2.5 for the winter week. According to both MAE and RMSE results, SARIMA is the best model in the averages; however in the 3 of the 7 days, it is not the best performing model. It is observed that benchmark AR(24) model is one of the best performing ones, even in the averages. Another interesting finding is that the Markov-2 less lags model is better than the Markov-2 more lags and Markov-3 models.

Table 2.4 : MAE results of the models for 18-24 January 2016 (green→ most successful, red→ most unsuccessful).

MAE	Naive	AR(24)	SARIMA	TAR (less)	TAR (more)	Markov-2 (less)	Markov-2 (more)	Markov-3 (less)	Markov-3 (more)
18 Jan	32.08	26.69	17.07	28.72	24.54	26.35	21.69	25.41	22.29
19 Jan	40.12	30.96	34.02	36.30	37.92	30.76	34.19	29.62	32.03
20 Jan	21.19	17.53	28.58	26.23	33.16	20.48	26.88	23.89	26.71
21 Jan	21.77	25.83	29.17	25.33	30.74	25.57	27.32	27.93	28.48
22 Jan	26.07	17.05	15.71	24.52	23.44	18.68	15.79	20.03	17.55
23 Jan	28.55	18.73	20.63	22.60	24.57	16.20	18.03	20.13	17.69
24 Jan	52.24	48.68	32.78	49.21	40.53	46.60	44.21	47.57	43.61
Average	**31.72**	**26.50**	**25.42**	**30.42**	**30.70**	**26.38**	**26.87**	**27.80**	**26.91**

Table 2.5 : RMSE results of the models for 18-24 January 2016 (green→ most successful, red→ most unsuccessful).

RMSE	Naive	AR(24)	SARIMA	TAR (less)	TAR (more)	Markov-2 (less)	Markov-2 (more)	Markov-3 (less)	Markov-3 (more)
18 Jan	45.47	33.56	21.40	37.40	32.24	34.25	29.21	34.83	29.68
19 Jan	52.83	37.25	40.88	42.99	43.95	38.03	40.72	37.51	38.22
20 Jan	35.44	26.09	37.92	31.91	39.13	31.48	36.59	32.27	36.34
21 Jan	30.22	32.23	34.42	31.03	34.92	33.83	33.43	36.47	35.01
22 Jan	36.14	23.45	22.97	30.17	29.58	25.44	24.04	26.79	25.90
23 Jan	40.85	22.67	25.39	24.36	27.69	19.47	23.17	22.77	21.15
24 Jan	61.86	60.04	46.30	61.56	51.92	59.69	56.10	59.98	55.40
Average	**43.26**	**33.61**	**32.75**	**37.06**	**37.06**	**34.60**	**34.75**	**35.80**	**34.53**

MAE and RMSE results for 4-10 July 2016 can be found in Table 2.6 and Table 2.7. According to the results, SARIMA model is the best in the averages for both models. This is consistent with the winter week and the 15th days of the year results. In the summer week, second best model is the Markov-2 (more) model, but it is difficult to conclude that there is much difference between the Markov models. It is possible to mention that TAR model is the second worst after Naive method; and AR(24) benchmarks give relatively good results.

Table 2.6 : MAE results of the models for 4-10 July 2016 (green→ most successful, red→ most unsuccessful).

MAE	Naive	AR(24)	SARIMA	TAR (less)	TAR (more)	Markov-2 (less)	Markov-2 (more)	Markov-3 (less)	Markov-3 (more)
04 Jul	66.62	32.34	34.48	36.48	33.17	28.82	25.74	26.81	26.25
05 Jul	58.49	58.90	58.05	58.22	57.92	58.88	58.53	60.12	60.00
06 Jul	65.68	55.25	55.58	69.11	68.38	57.83	59.11	60.23	60.85
07 Jul	31.07	36.16	35.33	39.89	42.67	36.44	37.26	38.93	38.88
08 Jul	60.13	46.73	46.80	52.10	52.21	46.31	48.33	45.34	47.74
09 Jul	33.51	25.49	18.39	31.03	30.48	23.37	20.50	22.19	19.00
10 Jul	36.10	31.90	22.01	34.93	33.23	25.58	23.53	25.54	22.54
Average	**50.23**	**40.97**	**38.66**	**45.97**	**45.44**	**39.60**	**39.00**	**39.88**	**39.33**

Table 2.7 : RMSE results of the models for 4-10 July 2016 (green→ most successful, red→ most unsuccessful).

RMSE	Naive	AR(24)	SARIMA	TAR (less)	TAR (more)	Markov-2 (less)	Markov-2 (more)	Markov-3 (less)	Markov-3 (more)
04 Jul	85.60	40.12	44.13	47.20	43.57	39.42	36.45	38.02	37.14
05 Jul	69.51	64.32	64.57	66.64	66.00	67.36	67.48	69.09	69.17
06 Jul	82.75	66.06	68.77	81.62	81.18	67.58	68.76	69.82	70.97
07 Jul	50.28	47.18	40.83	46.58	50.33	43.81	43.96	46.21	46.88
08 Jul	80.39	57.81	56.37	59.59	60.81	57.19	58.74	56.03	57.89
09 Jul	52.24	29.00	21.94	34.55	33.49	26.97	24.37	25.79	22.66
10 Jul	49.10	38.61	26.55	38.62	37.30	30.59	28.32	31.01	27.66
Average	67.13	49.01	46.17	53.54	53.24	47.56	46.87	48.00	47.48

Comparing the results with different models is always a difficult issue because of some reasons. Mainly, all the markets have their specific features and different to compare the results from various markets. Moreover, MAE and RMSE aren't comparable for different markets; and MAPE is very high when prices are close to zero, non-calculable when the price is zero, and meaningless when prices are negative. Another problem is that all the research are done for different time periods and comparing different time periods is impossible due to high level of seasonality. Our paper uses only the lagged price series and including exogenous variable could help the situation. Noting all these problems, our model will still be compared with the other papers from the Turkish Day-Ahead Market. Ozozen et al. (2016) prefer to give their out-of-sample results in a calmer period of 02.10.2015 – 06.10.2015. Their SARIMA, MAPE results are 13.8% and the hybrid model of SARIMA and ANN achieves the MAPE of 10.2%. When our SARIMA model is run for the same period, it has the MAPE of 13.02%, which outperforms the results of Ozozen et al. (2016). Ozyildirim and Beyazit (2014) have a regression model by adding the temperature as an exogenous variable and has the MAE of 8.44 Turkish Lira. In addition, they perform a radial basis function forecast and the results are very close to the linear model, has the MAE of 8.39 Turkish Liras. Out-of-sample forecast is done for the period of 01.07.2013 to 30.09.2013. Our SARIMA model is used to compare the results; however our MAE results are 13.74 Turkish Liras in the same period. One of the reasons is that our estimation period decreased to 01.01.2013 to 31.06.2016 and the forecast is done for entire out-of-sample period. It is assumed that Ozyildirim and Beyazit (2014) forecast hourly by adding the available data. In addition, authors have the temperature variable in their model. Taysi et al. (2015) have the average MAPE for SARIMA at 9.38% and for ANN at 8.24%, for 12 weeks in 12 months from 2013 to 2014. Another work is Kolmek and Navruz (2015), which has the ARIMA, MAPE of 15.60% and ANN, MAPE of 14.15% for 2010. As mentioned, it is difficult to compare with these models. MAPE results are not calculated for our application due to zeros and Kolmek and Navruz (2015)'s forecast period is even not in our sample. It must be stressed again that our main objective is comparing various models in the Turkish Day-Ahead market and comparing the success of these models with each other, and the second aim is using the well-known and used in different-style approach factorial ANOVA in a different aspect such as a pre-whitening method instead of forecasting scheme. Keeping in mind that our model doesn't include any exogenous variable, it was not expected that it would outperform the other models.

2.5 Conclusions

This paper is important because of two main reasons. Firstly, it is the first study that uses the factorial ANOVA as a pre-whitening methodology to the price series. It is known, that electricity price series have many features such as seasonality, high volatility, sharp price spikes, and mean reverting processes; which make the forecasting very difficult. In addition to these features; it is impossible to take logarithmic returns to make the series stationary, because of the zeros and the negatives[23] in the price series. Therefore, a factorial ANOVA process is applied to the prices series, which removed the deterministic part. This method removes the deterministic part caused by the day of the month, hour, weekday, month, year, holiday

[23] Turkish Day-Ahead Market doesn't allow prices to exceed 0 and 2000 TL/MWh, downside and upside, respectively. In other markets, such as German EEX, market-makers also allow negative prices, which cause forecasting problems as well.

effects and makes the data stationary and relatively linear. Although ANOVA is a special case of multivariate linear regression and it is used as a forecasting method in many papers, this is the first application of factorial ANOVA as a pre-whitening method to split the price series into deterministic and stochastic parts.

According to both MAE and RMSE performance measures, SARIMA model outperforms the non-linear TAR and Markov models' variations. It reveals the second advantage of the factorial ANOVA methods' application. Linear models like SARIMA require relatively less mathematical background and computational time in comparison to non-linear models. Therefore, applying such a pre-whitening method affects the forecast procedure efficiently.

Secondly; it is the first paper, which compares so many forecast methods by taking benchmark naive method (Nogales et al., 2002) and AR(24) into account in the Turkish Day-Ahead Electricity Market in hourly basis. Works of Ozyildirim and Beyazit (2014), which compares the regression with radial basis function; Ozozen et al. (2016), which applies a SARIMA model to hourly electricity prices and an ANN model to the residuals; Hayfavi and Talasli (2014), which compares the generated daily jump diffusion forecasts with the mean-reverting jump diffusion model of Cartea and Figueroa (2005) and the Markov regime switching model of Janczura and Weron (2010); Taysi et al. (2015), which compares the weekly forecast performance of ARIMA and ANN models; and Kolmek and Navruz (2015), which compares the hourly forecast performance of ARIMA and ANN; should not be undervalued. However, this is the first paper that comprehensively discusses the performance of econometric models such as SARIMA, TAR and Markov regime-switching variations in addition to benchmark models, naive method and AR(24). It is hoped that this paper would encourage the researchers and practitioners to have various transformations to the price series and to compare the forecast performance of different methods and work on the emerging Turkish Day-Ahead Market. Regulated Turkish electricity market needs more attention with the increasing share of the electricity trade in the Day-Ahead Market as well as balancing and intraday markets, in addition to currently improving derivatives market. General price levels tend to decrease due to the technological improvements and the increasing share of the renewables. Especially dam-type hydro power plants were one of the main electricity providers in the Turkish electricity market for years, but in the last years there is a huge increase in the wind electricity supply share. Therefore, works on the effect of the renewables in the Turkish electricity market would be very interesting.

It should be mentioned that Turkish day-ahead electricity market is a quite young market and the price bidding behaviour is debateable due to the lack of the knowledgeable participants. Due to the lack of knowledge, most of the players in this market don't have a better method other than bidding the same prices of the previous day[24]. It can be observed that the same prices occur for the same hours consecutively. Especially, in the relatively difficult times with high prices as well as high volatility such as summer months[25] most of the market players tend to be cautious and give the previous day's prices[26]. The immaturity of the Turkish day-ahead market causes advanced problems in addition to the known difficulties of electricity price forecasting. By keeping in mind that Nogales et al. (2002) and Conejo et al. (2005a) mention that the models, which are not calibrated well, cannot outperform the naive method; this special issue of the Turkish day-ahead should also be covered as a limitation.

This application is performed in the specific Turkish Day-Ahead Market, for the specific timeframe and data frequency, specific estimation and forecast periods, specific models and lag selection. Especially for a commodity like electricity, it mustn't be forgotten that the results are affected by many different factors. Therefore, results and outperformance of the SARIMA method must be evaluated under these circumstances.

It must be stressed that the analysis has been applied on the 15th days of the year, a randomly selected week in summer and a week in winter, which might cause biases. Although, similarly, forecast periods are only a few weeks in most of the papers; forecasting all the days of the year and taking the average errors is a better approach. Using this approach in further research would give much more robust results. It is also a univariate model approach, which does not include any exogenous variable and outperforming the other models is not among the aims of the paper. It is just a comparison of various methods in the same dataset and the novel usage of the well-

[24] 2.88% of the prices are exactly same with the previous day's prices and 11.01% have less than 1 TL/MWh difference in 2016, which are quite high levels compared to the other methods' forecasts. For further information, bid structure must be examined.
[25] Due to high levels of air-conditioning usage.
[26] As an example, prices of 15 July, which is analysed in this paper, is given in the Table A.6 with the previous day's prices and absolute errors.

known ANOVA approach. This kind of analysis could be performed in mature markets to see the difference between mature and immature environments. Additionally; in this analysis, a single series of the prices are used to forecast the electricity prices. It can be analysed by using 24 hourly series and forecast each and every hour separately. Another point is that this study is focused on a single market like most of the other studies in the literature. However, a comparative analysis of emerging countries' energy markets could be a nice approach.

3. ELECTRICITY PRICE FORECASTING USING RECURRENT NEURAL NETWORKS[27]

3.1 Introduction

Since the liberalization of the electricity markets, electricity price forecasting has become an essential task for all the players of the electricity markets for several reasons. Energy supply companies, especially dam-type hydroelectric, natural gas, and fuel oil power plants could optimize their procurement strategies according to the electricity price forecasts. As the share of the regulated electricity markets, such as day-ahead and balancing markets, increase day by day, bilateral contracts also take the market prices as a benchmark (Keles et al., 2016). Moreover, prices of the energy derivatives are also based on electricity price forecasts (Carmona et al., 2013). From the demand side, some companies can schedule their operations according to the low-price zones and operate in these hours or months. Zareipour et al. (2010) stressed the importance of the short-term electricity forecasting accuracy. A 1% improvement in the mean absolute percentage error (MAPE) would result in about 0.1–0.35% cost reductions from short term electricity price forecasting (Uniejewski et al., 2016), which results to circa $1.5 million per year for a medium-size utility with a 5 GW peak load (Hong, 2015).

Electricity prices differ from all other assets and even commodities due to its unique features such as requirement of having constant balance between the supply and demand sides, demand inelasticity, oligopolistic generation side, and non-storability (Ugurlu et al., 2018b). These features cause some important characteristics of the electricity prices: high volatility, sharp price spikes, mean reverting process, and seasonality in different frequencies (Hayfavi and Talasli, 2014). Because of all these

---

[27] This chapter is based on the paper "Electricity Price Forecasting Using Recurrent Neural Networks". Ugurlu, U., Oksuz, I. & Tas, O. 2018. Electricity Price Forecasting Using Recurrent Neural Networks. *Energies*, *11* (5), 1255.

idiosyncratic features and characteristics, forecasting the electricity prices accurately becomes a very challenging task.

Machine learning models are able to solve very complicated classification and regression problems with great success. Recently, deep learning models have become the state-of-the-art in speech recognition (Greff et al., 2017), handwriting recognition (LeCun et al., 1998) and image classification (Krizhevsky et al., 2012).

This paper presents a Gated Recurrent Unit (GRU) based method for electricity price estimation with the goal of using the valuable time series information fully in a neural network architecture. Neural network based methods showed great promise in computer vision, speech recognition and natural language processing (Greff et al., 2017). In particular, Recurrent Neural Networks are capable of faithfully preserving the key time-dependent patterns for natural language processing type problems. This motivated us to propose a thorough analysis of multiple features for the electricity prices estimation using Recurrent Neural Networks (RNNs). In particular, the main contributions of this paper are:

- A multi-layer GRU Recurrent Neural Network setup for estimating electricity prices is used.
- A wide analysis of multiple feature settings for neural networks, Convolutional Neural Networks (CNN), Long Short Term Networks (LSTM) and state-of-the-art statistical methods is performed.
- Extensive electricity price estimation performance analysis with both daily andmonthly comparisons is made.
- Detailed analysis between the state-of-the-art statistical models and the neural network based methods is made.

3.1.1 Literature review

Electricity price forecasting literature started to develop in the beginning of the 2000s (Szkuta et al., 1999; Bunn, 2000; Contreras et al., 2003; Nogales et al., 2002; Shahidehpour et al., 2002; Cuaresma et al., 2004; Bunn, 2004). Following the review by Weron (2014), we partition the main methods of electricity price forecasting into five groups: multi-agent, fundamental, reduced-form, statistical, and computational intelligence models.

Multi-agent models simulate the operation of the system and build the price process by matching the demand and the supply. The papers by Shafie-Khah et al. (2015) and Ziel and Steinert (2016) are very good and recent examples of these type of papers. Shafie-Khah et al. (2015) modelled wind power producers, plug-in electricity vehicle owners and customers, who participated into demand response programs, as independent agents in a small Spanish market. Furthermore, Ziel and Steinert (2016) proposed a model for the German European Power Exchange (EPEX) market, which considers all the supply and demand information of the system and discusses the effects of the changes in supply and demand.

Fundamental or structural methods discuss the effects of the physical and economic factors on the electricity prices. In this part of the literature, variables are modelled and predicted independently, often via other methods such as reduced-form, statistical or machine learning methods. For example, Howison and Coulon (2009) developed a model for electricity spot prices using the stochastic processes of the independent variables. Their method also takes the bid stack function of the price drivers and the electricity prices into account. In another study, Carmona and Coulon (2014) focused on the role of the energy prices and effect of the fundamental factors on the electricity prices in a survey about the structural methods. Carmona and Coulon (2014) also discussed the superiority of the fundamental models to the reduced-form models. Both Carmona et al. (2013) and Füss et al. (2015) constructed fundamental models to achieve the final aim of electricity derivatives pricing.

Reduced-form models mainly consist of two methods: Markov regime-switching and jump diffusion. These models are relatively better than structural and statistical models in terms of handling spikes. Geman and Roncoroni (2006) used mean-reverting jump diffusion (MRJD) model. Their approach captures both trajectory and statistical components of the electricity prices. Cartea and Figueroa (2005) and Janczura et al. (2013) used more hybrid methods. First, theyed filter out the jumps using a jump diffusion model and then they proposed more statistical methods to model the remaining, stationary part of the series. Hayfavi and Talasli (2014) applied a hybrid-jump diffusion model to the Turkish market and compared the results with Carte and Figueroa (2005) and Janczura and Weron (2010). Janczura and Weron (2010) compared some of the examples in the literature with their own three-regime-switching Markov model, which captures both positive and negative spikes, in addition to exhibiting the inverse leverage effect of the electricity spot prices. Furthermore, Eichler and Türk (2013) proposed a semi-parametric Markov regime-switching model. In their method, model parameters are employed by robust statistical techniques. Moreover, it is easier to estimate, and needs less computational time and distributional assumptions. Keles et al. (2012) and Bordignon et al. (2013) used jump diffusion and Markov regime-switching, respectively, in hybrid works.

Statistical and computational intelligence are the most common models in the electricity price forecasting literature. Statistical models are in great variety from basic naive method (Nogales et al., 2002) to very developed methods (Ziel and Weron, 2018). As Ziel and Weron (2018) discussed, there are univariate and multivariate frameworks in the electricity price forecasting. In day-ahead electricity price forecasting, players bid the prices and the quantities for the 24 h of the next day. In this sense, the first way is to predict all the prices in a univariate framework from a single price series as a 24-step-ahead forecast. Forecasting the prices from 24 different time series as one-step-ahead forecasts is another option, which is called multivariate framework. Weron and Misiorek (2008) applied the univariate framework to the Nordic data. Kristiansen (2012) utilized the multivariate framework on the same dataset in a follow-up study and argued that using univariate framework increases the prediction accuracy. However, it contradicts with the findings of Cuaresma et al. (2004), who mentioned that using the multivariate framework presents better forecasting results than univariate method. In the same Nordpool market, Raviv et al. (2015) have a different point of view. It compares the one-step-ahead daily average price forecasts in a univariate framework with the aggregated 24-step-ahead forecasts of the hourly prices. From empirical evidence, Raviv et al. (2015) stated that multivariate framework has lower out-of-sample errors than the univariate one. Nogales et al. (2002), Contreras et al. (2003), and Conejo et al. (2005b) presented some substantial examples of the auto-regressive models. Nogales et al. (2002) proposed the naive method and, as mentioned by Contreras et al. (2003), Nogales et al. (2002) and Conejo et al. (2005b), poorly-calibrated forecasting methods cannot outperform the naive method. Although Conejo et al. (2005b) found that Auto-regressive Integrated Moving Average (ARIMA) model is worse than the model with exogenous variables in the American PJM market, Contreras et al. (2003) stated that adding an exogenous variable does not necessarily increase the prediction accuracy.

Many types of computational intelligence models are applied in the electricity price forecasting literature. Some of the early stage papers were presented by Mandal et al. (2006), Catalão et al. (2007) and Zhang and Cheng (2008). Mandal et al. (2006) forecasted the electricity loads and prices in the Australian market by applying Artificial Neural Network (ANN) model for 1–6 h ahead. MAPE increased from 9.75% to 20.03% when one-step ahead forecast increased to six-step ahead forecast. In another study, Catalao et al. (2007) utilized a three-layered feed-forward neural network, which is trained by Levenberg–Marquardt method, and forecasted 168-step-ahead in the Spanish and Californian markets. Although they gave the results for all the seasons of the Spanish market, in the Californian market, results are available only for the Spring term. Therefore, it is difficult to compare the results of both markets. Differently, Zhang and Cheng (2008) forecasted the daily average prices and required only one-step-ahead forecast. In the Nordpool market, a standard error back-propagation method is used, which is improved by self-adaptive learning rate and momentum coefficient algorithms. Results indicate that ANN model outperforms the standard ARIMA method. Recent studies by Keles et al. (2016) and Panapakidis and Dagoumas (2016) apply mainly ANN methods. Keles et al. (2016) proposed ANN models with different variables by utilizing the clustering methods. Their ANN based method outperforms the benchmark naive-type models and the Seasonal Auto-regressive Integrated Moving Average (SARIMA) model. An important contribution of this work is the thorough analysis of the forecast accuracy according to the months, extreme price levels, and small and extreme price changes. Panapakidis and Dagoumas (2016) compared the forecast performances of different ANN models with various numbers of variables, layers and neurons. The main approach they applied is the clustering of the groups. According to their results, clustering gives 20% better results. Amjady et al. (2006) applied fuzzy neural network, Zhao et al. (2008) performed support vector machines, Alamaniotis et al. (2015a) used kernel machines and Pindoriya et al. (2008) utilized adaptive wavelet-neural network.

3.1.2 Turkish market

Electricity markets differ from country to country for several reasons. The main difference is the supply share of different production methods. When share of renewables, i.e., wind and solar, as well as hydro power plants increase, prices tend to decrease. As Diaz and Planas (2016) mentioned, Spanish market has many zeros, which is the minimum price allowed, as well as in the Canadian market (Filipovic et al., 2017). Turkish market has the same price floor of 0 and the price cap of 2000 Turkish Liras/MWh (about 598 Euros/MWh, by the 2016 average exchange rate). Furthermore, as Fanone et al. (2013) and Keles et al. (2012) mentioned, many negative prices occur due to increased wind share in the German market and it needs special attention. Ugurlu et al. (2018b) mentioned some information about the shares of the installed capacity in the Turkish market: 34.2% for hydro and 7.6% for wind. In addition to the improved technology in the other supply methods, increasing shares of hydro and wind trigger the decrease in the Turkish day-ahead market electricity prices, which causes many zeros in the price series. These zeros require a special treatment and transformation prior the forecasting procedure (Ugurlu et al., 2018b; Diaz and Planas, 2016; Uniejewski et al., 2017). Avci-Surucu et al. (2016) and Ozozen et al. (2016) gave some information about the working mechanism of the Turkish day-ahead market. Day-ahead market is used to balance the electricity requirement one day before the physical delivery of the electricity (Ugurlu et al., 2018b). As in many other markets, market participants give their bids in terms of quantity and price until 11:00, and the price for each hour of the next day is determined by the market maker until 14:00 according to the intersection of the supply and demand curves. It is aimed to meet the required demand with the lowest possible price.

Turkish day-ahead electricity market has an improving literature. Hayfavi and Talasli (2014) reported one of the first works, which proposes a multifactor model and compares the model with Cartea and Figueroa (2005) and Janczura and Weron (2010). The stochastic model composed of three jump processes outperforms Cartea and Figueroa (2005) and Janczura and Weron (2010) according to the comparison of the empirical moments and model moments in the daily Turkish data. Kolmek and Navruz (2015) compared an artificial neural network (ANN) model with the ARIMA model. According to their results, performance of the models differ widely in respect to the selected evaluation period. However, overall, ANN model is a little better than the ARIMA model. In another work, Ozguner et al. (2017) proposed an ANN model to forecast the hourly electricity prices and loads in the Turkish market and compared the results with multiple linear regression. Findings of this paper is very similar to Kolmek and Navruz (2015); in both papers, ANN model outperforms ARIMA model with a small difference. Ozyildirim and Beyazit (2014) compared another machine learning method, radial basis function, with the multiple linear regression. In their work, difference between the prediction performance of themodels are negligible. Ozozen et al. (2016) adapted a method from the literature to Turkish electricity prices and takes the residuals of the SARIMA forecast and puts it into ANN procedure. However; the simple model of Ugurlu et al. (2018b), which even does not include an exogenous variable, outperforms Ozozen et al. (2016). In our opinion, the reason for the better performance is the factorial Analysis of Variance (ANOVA) application of Ugurlu et al. (2018b) on the electricity price series prior to forecasting. Although the best model varies from period to period, SARIMA is chosen as the best statistical model for the Turkish day-ahead market in (Ugurlu et al., 2018b).

3.1.3 Deep learning

Neural networks transform into deep neural networks (deep learning) with the addition of more layers into the neural network mechanisms. Besides, recurrent neural networks such as LSTM and GRU have started to give better results in the time series data, which triggered the application of these methods in the electricity price forecasting and related literature. RNNs have shown great success in speech recognition, handwriting recognition and polyphonic music modelling (Greff et al., 2017). In the electricity load forecasting literature, Zheng et al. (2017) applied similar days selection and empirical mode decomposition methods in addition to LSTM, and their method outperforms many state-of-the-art methods such as support vector regression, ARIMA or ANN. Xiaoyun et al. (2016) made wind power forecast by combining principal component analysis (PCA) with LSTM. In a solar power forecast research, Gensler et al. (2016) applied LSTM method with AutoEncoder and the results show that LSTM usage gives much better results than ANN. In another work, Bao et al. (2017) applied very similar method to the stock price forecasting and used wavelet transformation, stacked AutoEncoders and LSTM. Hosein et al. (2017) made similar findings as the superiority of the deep neural networks (various deep neural networks including LSTM ones are used) in the power load forecasting, but mentioned the computational complexity as a drawback. The only deep neural networks (deep learning) application in the day-ahead electricity price forecasting literature was by Lago et al. (2018b), who only used a simple multi-layer perceptron with more than single layer and did not propose a RNN algorithm such as LSTM or GRU. Another point is that the paper’s main research question is the effect of the market integration on the electricity price forecasting in Europe and deep neural network is only used as the forecast model and is not compared with any other method. We want to acknowledge two simultaneous works that are published after our submission on the same topic (Lago et al., 2018a; Kuo and Huang, 2018). Lago et al. (2018a) proposed a framework for deep learning applications in the electricity price forecasting and also suggested a benchmark by comparing various price forecasting models. Results are threefold: First, machine learning models outperform the statistical methods. Second, moving average terms do not improve the success of the predictions. Third, hybrid models do not perform better than the individual ones. An important point to discuss is that they applied recurrent neural networks, LSTM and GRU as well as deep neural networks (DNN). Surprisingly, they found that DNN has a better predictive accuracy compared to LSTM and GRU. Although the authors had two hypotheses about these results, which are low amount of data and different structure of the models, they suggested further research about the same topic. Our work differs with these work in the number of features we utilized and by proposing deep RNNs in comparison to DNNs. In another very recent paper, Kuo and Huang (2018) also proposed CNN and LSTM as deep network structures. According to their results, combining CNN and LSTM gives lower errors than the individual forecasts, in addition to the state-of-the-art machine learning methods. Lago et al. (2018a) used EPEX Belgium hourly data from 2010 to 2016 and, Kuo and Huang (2018) utilized U.S. PJM half-hourly data of 2017.

In this paper, we propose to use RNNs for the time-dependent problem of electricity price estimation. To the best of our knowledge, our paper is the first in the electricity price forecasting literature to apply deep RNNs, LSTM and GRU. Furthermore, these models are compared with simple deep neural networks (multi-layer ANN), single layer neural networks and the statistical time series methods. In addition to the lagged values of the price series, forecast Demand/Supply (D/S), temperature, realized D/S and balancing market prices are used as the exogenous variables. Various combinations of these features are selected to measure the effects of the variables. Moreover, Diebold–Mariano (DM) test (Diebold and Mariano, 1995) is applied to evaluate the statistical significance of the performance difference achieved with all different architectures and features.

The remainder of the paper is structured as follows. Section 3.2 gives information about the data. The neural networks based methods are described in Section 3.3 with a particular interest in RNNs. Experimental setup, methods of comparison and corresponding results are shared in Section 3.4. We conclude the paper with a detailed discussion on the results in Section 3.5.

3.2 Data

Turkish Day-ahead Market electricity prices are effected by various types of seasonality. Early morning hours (2:00–7:00) have relatively low prices, even some zeros. Moreover, there are double peaks in the day, one before and one after the lunch time, 11:00 and 14:00, respectively, as visualized in Figure 3.1. In weekly terms, Saturday morning prices are as high as the other weekdays, which shows the working pattern on Saturday mornings. Furthermore, there are two minimums on Saturday night and Sunday night. From a seasonal point of view, both heating and cooling requirements cause high prices in winter and summer, respectively. However, due to the high share of hydro power plants in the electricity production, prices tend to decrease in spring time. An example from the data for each season of 2016 is visualized in Figure 3.2. The detailed statistics of the test data from 2016 are illustrated in the Appendix B.1.

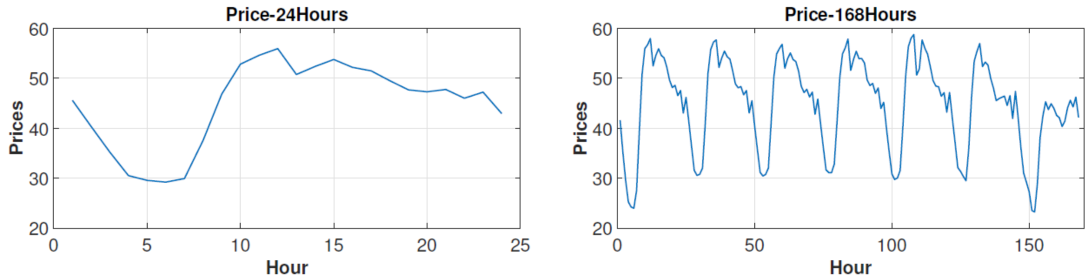


Figure 3.1 : Left panel: Price distribution of hourly prices (Euro/MWh) according to the hours of the day (based on 24 h). Right panel: Price distribution of hourly prices (Euro/MWh) according to the hours of the week (based on 168 h).

Hourly day-ahead electricity prices of the Turkish Day-Ahead Market are obtained from 1 January 2013 to 21 December 2016 (EPIAS, 2018). The Turkish Day-Ahead Market was established on 1 December 2011. The first 13 months was excluded due to the learning-by-doing process, which limited us to start our data from 1 January 2013.

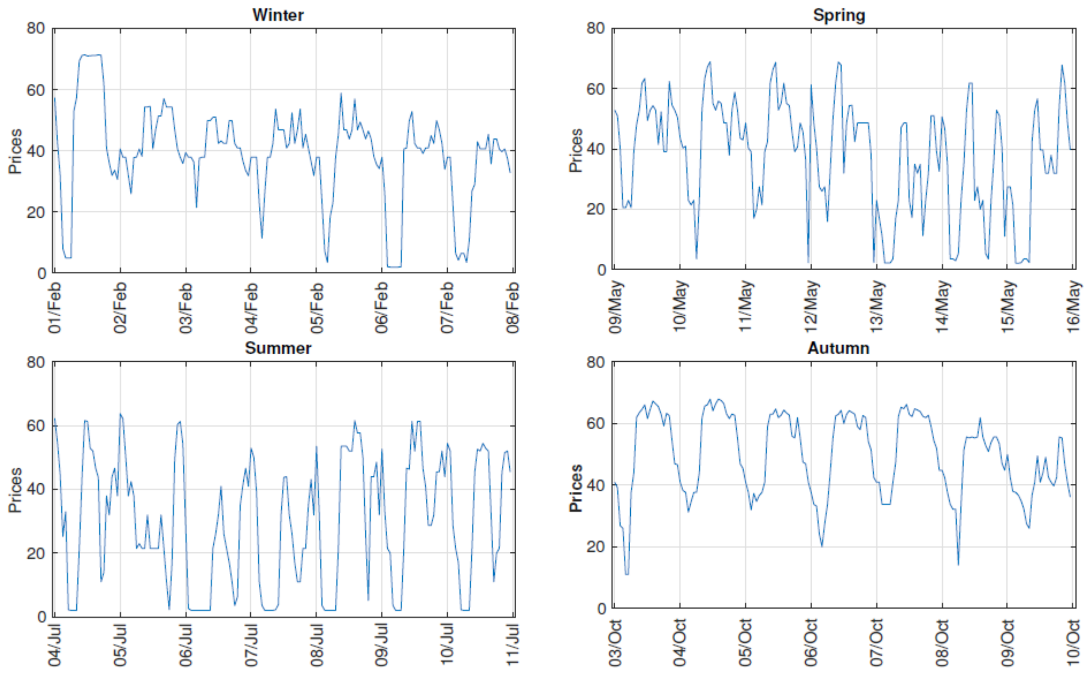


Figure 3.2 : Price time series of sample weeks from each season.

In neural network applications, the first three years (1 January 2013–31 December 2015) are used for training and each hour of the next day (1 January 2016) is predicted using the 24-step-ahead forecast scheme. This process is repeated using rolling window method by moving the window 24 h in every forecast. Training period remained as three years and the forecast period as 24-h of the following day. This process is repeated for 356 days of 2016. The reason forf not including the last 10 days of 2016 in the forecast procedure is the very high prices, which occurred in this term due to the natural gas shortage and inactivity of the natural gas power plants. Prices increased up to 515 Euro/MWh on 23 December at 14:00, which is approximately 14 times higher than the average price level.

In the statistical time series methods, such as Markov, Threshold Auto Regressive (TAR) and SARIMA, due to non-stationary nature of the price series and zeros, factorial ANOVA (Ugurlu et al., 2018b) transformation was applied and the series split into deterministic and stochastic parts. Then, stationary stochastic part was forecasted and added to the deterministic part values, which include the hour, weekday, month, holiday and year components. This process was repeated in the rolling window scheme for 356 days as in the neural network methods.

Variable selection is a very important topic in the electricity price forecasting. In our paper, we have chosen the lagged price values as variables according to autocorrelation and partial auto-correlation functions. The chosen lags are also coherent with the lagged price series used in the literature. Furthermore, exogenous variables are also selected according to the electricity price literature (Uniejewski et al., 2016; Bordignon et al., 2013). Due to the high correlation between them and the independent variable, forecast D/S, temperature and the 24th lags of realized D/S and balancing market price are selected as exogenous variables. One advantage is that the market maker (EPIAS) provides forecast D/S before the bids are given into the system for the next day. Another variable is temperature, which was taken from the Turkish State Meteorological Service as 81 city-based hourly temperatures. Then, annual energy consumption for all the cities was taken from Republic of Turkey Energy Market Regulatory (EPDK, 2018) and energy consumption-weighted hourly temperatures (T) were calculated for every hour. Furthermore, we took the 24th lags of realized D/S and balancing market prices into account because both have very high correlation with the price series and also used as variables in the literature. In addition to the above mentioned exogenous variables, 1, 23, 24, 48, 72, 168 and 336 h lagged prices were also utilized as features to estimate the day-ahead prices for the upcoming 24 h. To report the results with aforementioned features, we use the symbols stated in Table 3.1.

Table 3.1 : Utilized features for electricity price estimation.

Symbol	Feature
F1	24-h lagged price
F2	168-h lagged price
F3	1-h lagged price
F4	48-h lagged price
F5	23-h lagged price
F6	72-h lagged price
F7	336-h lagged price
F8	Forecast demand/supply
F9	Temperature
F10	Realized demand/supply with 24 h lag
F11	Balancing market price with 24 h lag

3.3 Methods

In this section, we describe the Neural Network architectures we used for electricity price estimation. A simple neural network with three input neurons is visualized in Figure 3.3. The guiding equation of a neuron can be described as:

$$Y = f\left(\sum_i^{Inputs}(x_i w_i + b_i)\right) \quad (3.1)$$

where w is the weight on each connection to the neuron, b is the bias and x is the input of the neuron. f can be described as the activation function to introduce non-linearity and, in our experiments, we used Rectified Linear Units (ReLU) (Glorot et al., 2011).

In Section 3.1, basic neural network structure, Artificial Neural Networks, is defined. In Section 3.2, we give a brief definition of Convolutional Neural Networks and their application on the time series data for electricity price estimation. Then, we move to RNNs in Section 3.3, which is the focal point of our work. In Section 3.3.1, we define the LSTM networks and their benefits for time series prediction tasks. Finally, in Section 3.3.2, we define the GRUs and their fundamental differences from LSTMs.

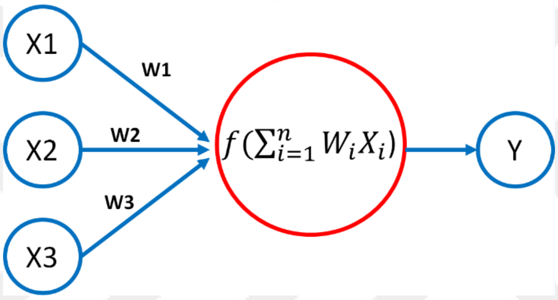


Figure 3.3 : Simple neural network.

3.3.1 Artificial neural networks

ANN is a basic architecture of a neural network, which consists of layers of neurons connected densely (Wasserman, 1988). This type of networks is also known as Multi-layer Perceptrons (MLP) and they are early examples of the neural networks. We used a shallow network with a single layer with 10 neurons and a deeper three-layer network, each consisting of 10 neurons, for our experiments. We added a final layer to estimate the target values.

3.3.2 Convolutional neural networks

Convolutional Neural Networks have been successfully applied to many problems in computer vision (Krizhevsky et al., 2012) and medical image analysis (Oksuz et al., 2018). In our application, the convolutional layers were constructed using one-dimensional kernels that move through the sequence (unlike images where 2D convolutions are used). These kernels act as filters which are being learned during training. As in many CNN architectures, the deeper the layers get, the higher the number of filters become. We used two convolutional layers and a final fully connected layer for prediction. Each convolution is followed by pooling layers to reduce the sequence length.

3.3.3 Recurrent neural networks

RNNs are networks with loops in them, allowing information to persist. They are used to model time-dependent data (Dorffner, 1996). The information is fed to the network one by one and the nodes in the network store their state at one time step and use it to inform the next time step. Unlike MLP, RNNs use temporal information of the input data, which make them more appropriate for time series data. An RNN realizes this ability by recurrent connections between the neurons. A general equation for RNN hidden state ht given an input sequence x= (x_1, x_2, . . . , x_T) is the following:

$$h_t = \begin{cases} 0, \; if \; (t = 0) \\ \varphi(h_{t-1}, x_t), \; otherwise \end{cases} \tag{3.2}$$

where φ is a non-linear function. The update of recurrent hidden state is realized as:

$$h_t = g(Wx_t + Uh_{t-1}) \tag{3.3}$$

where g is a hyperbolic tangent function.

In general, this generic setting of RNN without memory cells suffers from vanishing gradient problems. In this study, we investigated the performance of two RNNs with memory cells for electricity price forecasting, namely, LSTMs and GRUs.

3.3.3.1 Long short-term memory networks

LSTM (Hochreiter and Shmidhuber, 1997) is a special type of RNN that is able to deal with remembering information for much longer time. In LSTM, each node is used as a memory cell that can store other information in contrast to simple neural networks, where each node is a single activation function. Specifically, LSTMs have their own cell state. Normal RNNs take in their previous hidden state and the current input, and output a new hidden state. An LSTM does the same, except it also takes in its old cell state and outputs its new cell state c_t^j (Vanishing Gradients & LSTMs, 2018). This property helps LSTMs to address the vanishing gradients problem from the previous time-steps.

We visualize the LSTM structure in Figure 3.4 (left panel) to define the guiding equations of LSTM. LSTM has three gates: input gate it, forget gate ft and output gate ot, as visualized in Figure 3.4 (left panel). Sigmoid function is applied to the inputs st and the previous hidden state h_{t-1}. The goal of the LSTM is to generate the current hidden state at time t. The hidden state hjt of LSTM unit is defined as:

$$h_t^j = o_t^j \tanh(c_t^j) \tag{3.4}$$

where o_t^j modulates the memory influence on the hidden state. The output gate is computed as:

$$o_t^j = \sigma(W_0 x_t + U_0 h_{t-1} + V_0 c_t)^j \quad , \tag{3.5}$$

where σ is the logistic sigmoid function and V_0 is a diagonal matrix. The memory cell c_t^j is updated partially following the equation

$$c_t^j = f_t^j c_{t-1}^j + i_t^j \tilde{c}_t^j \quad , \tag{3.6}$$

where the memory content is defined by a hyperbolic tangent function:

$$\tilde{c}_t^j = \tanh(W_c x_t + U_c h_{t-1})^j \tag{3.7}$$

Forget gate f_t^j controls the amount of old memory loss. Instead, input gate i_t^j controls new memory content that is added to the memory cell. Gates are computed by:

$$f_t^j = \sigma(W_f x_t + U_f h_{t-1} + V_f c_{t-1})^j \tag{3.8}$$

$$i_t^j = \sigma(W_i x_t + U_i h_{t-1} + V_i c_{t-1})^j \tag{3.9}$$

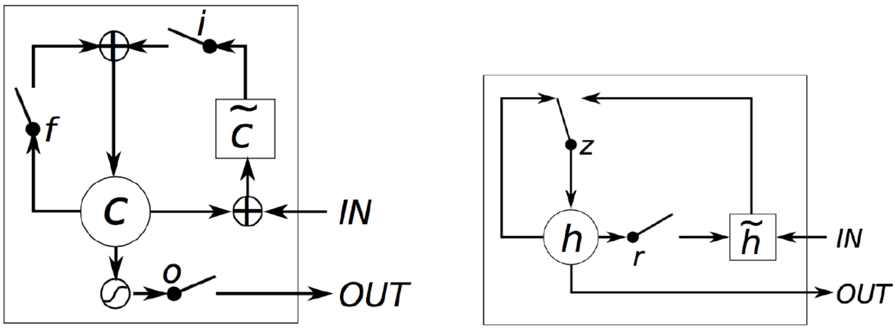


Figure 3.4 : Illustration of: Left panel: LSTM; and right panel: GRU. Left panel: i, f and o are the input, forget and output gates, respectively. c and $\tilde{c}$ denote the memory cell and the new memory cell content. Right panel: r and z are the reset and update gates, and h and $\tilde{h}$ are the activation and the candidate activation. (Figure adapted from Chung et al. (2014)).

LSTM unit is robust compared to traditional RNN, thanks to the control over the existing memory via the introduced gates. LSTM is can pass information that is captured in early stages and easily keeps memory of this information for long term, which enables the opportunity to generate potential long-distance dependencies as underlined by Chung et al. (2014).

3.3.3.2 Gated recurrent units

A GRU (Cho et al., 2014) has two gates, a reset gate r and an update gate z, as visualized in Figure 3.4 (right panel). The update gate defines how much of the previous memory to be kept and the reset gate determines how to combine the new input with the previous memory. GRUs become equivalent to RNNs, if the reset gates are all 1 and update gates all 0.

Following Chung et al. (2014), we formulated the guiding equations. The activation h_t^j of the GRU at time t is a linear interpolation between the previous activation h_{t-1}^j and the candidate activation h_t^j:

$$h_t^j = (1 - z_t^j)h_{t-1}^j + z_t^j \tilde{h}_t^j \tag{3.10}$$

where an update gate z_t^j is in charge of the content update. The update gate is computed by:

$$z_t^j = \sigma(W_z x_t + U_z h_{t-1})^j \tag{3.11}$$

This procedure of taking a linear sum between the existing state and the newly computed state is similar to the LSTM unit. Unlike LSTM, GRU does not have any control on the state that is exposed, but exposes the whole state each time.

The candidate activation $\tilde{h}_t^j$ is computed similarly to RNN:

$$\tilde{h}_t^j = \tanh(W x_t + U(r_t \odot h_{t-1}))^j \tag{3.12}$$

Where r_t is a set of reset gates and $\odot$ is an element-wise multiplication.The reset gate r_t^j is computed similarly to the update gate:

$$r_t^j = (W_r x_t + U_r h_{t-1})^j \tag{3.13}$$

GRUs have the same fundamental idea of gating mechanism to learn long-term dependencies compared to LSTM, but there are couple of significant differences. First, GRU has two gates and fewer parameters compared to LSTM. The input and forget gates are coupled by an update gate z and the reset gate r is applied directly to the previous hidden state in GRUs. In other words, the responsibility of the reset gate in an LSTM is divided into both reset gate r and the update gate z. GRUs do not possess any internal memory that is different from the exposed hidden state. LSTMs have output gates and GRUs do not possess output gates. In addition, in LSTMs, there is a second non-linearity applied when computing the output, which is not present in GRUs (Implementing a GRU/LSTM RNN with Python and Theano, 2018).

3.4 Results

This section offers a qualitative and quantitative analysis of the proposed method, as well as comparison of RNNs with respect to state-of-the-art methods, to demonstrate its robustness for electricity price estimation.

Our quantitative analysis consists of comparing our method with others and also looking into monthly and weekly performance. In Section 3.4.1, we describe the evaluation metrics and then explain the state-of-the-art statistical methods in Section 3.4.2. We report the quantitative results achieved by all network types with a different combination of layers in Section 3.4.3 and evaluate the statistical significance in Section 3.4.4. Finally, we mention some implementation details about the neural network training and hyper-parameters in Section 3.4.5.

3.4.1 Evaluation metrics

In the performance evaluation of the forecasting techniques, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are the most used metrics. Although MAPE gives opportunity to compare the electricity price forecasts' performances from various markets, for the prices around zero, it does not give interpretable results. For zeros, MAPE can not be calculated; for negative prices, there are negative values, which are meaningless; and for small positive prices, MAPE values are very high. In the comparisons, there is not an important difference between the MAE and RMSE values, because both are based on the absolute errors (Ugurlu et al., 2018b). Therefore, MAE method is used as the performance evaluation criterion in this paper. Equation 3.14 shows the MAE formula.

$$MAE = \frac{1}{T}\sum_{i=1}^{T}\left|P_i - \hat{P}_i\right| \quad (3.14)$$

3.4.2 State-of-the-art statistical methods

Traditionally, Naive method, SARIMA, Markov regime-switching and Self exciting threshold auto-regressive regression (SETAR) have been used with great success for time series estimation in the electricity price forecasting literature (Ugurlu et al., 2018b). We compared the robustness of these techniques with the neural network architectures.

3.4.2.1 Naïve method

One of the most important benchmark techniques in the electricity price forecasting literature, naïve method (Nogales et al., 2002), can be found below in equation 3.15. According to Nogales et al. (2002) and Conejo et al. (2005b), forecasting methods that are poorly calibrated cannot outperform the naive method (Ugurlu et al., 2018b).

$$P_{d,h} = \begin{cases} P_{d-7,h} + \varepsilon_{d,h}, & Monday, Saturday, Sunday \\ P_{d-1,h} + \varepsilon_{d,h}, & Tuesday, Wednesday, Thursday, Friday \end{cases} \quad (3.15)$$

$P_{d,h}$ states the price of the selected day and hour. ed,h stands for the noise term.

3.4.2.2 Markov regime-switching auto-regressive (MS-AR) model

As another benchmark method, two-state Markov regime-switching auto regressive model (Hamilton, 1989) with the 1st, 24th, 48th and 168th lags of the price series are used in the estimation. This method allows the observations to be distributed into different states by a latent variable. Equation 3.16 relates the Markov Regime-Switching Auto Regressive (MS-AR) model.

$$y_t = \alpha_s + \sum_{i=1}^{p} \phi_{s,i} y_{t-i} + \epsilon_t \quad (3.16)$$

where st is a two-state discrete Markov-chain with S = 1,2 and $\epsilon_t \sim$ i.i.d. N(0, σ^2). The estimation of the MS-AR model is performed by maximum likelihood algorithm (Ugurlu et al., 2018b; Ozkan and Yazgan, 2015).

3.4.2.3 Self-exciting threshold auto-regressive (SETAR) model

Threshold auto-regressive (TAR) models are similar to Markov regime-switching models in terms of placing the observations into different groups. The main difference of the TAR models is that the threshold variable is observable compared to the latent one in the Markov models. TAR models allow to choose the threshold according to an exogenous variable. If the threshold variable is selected according to a lagged value of the dependent variable, then it is called SETAR model. In equation 3.17, SETAR model is given.

$$x_t = \phi_0^{(j)} + \phi_1^{(j)} x_{t-1} + \cdots + \phi_p^{(j)} x_{t-p} + \alpha_t^{(j)}, \quad \text{if } \gamma_{j-1} \le x_{t-d} \le \gamma_j \qquad (3.17)$$

Where k and d are positive integers, j= 1, …, k; γ_i are real numbers such that $-\infty = \gamma_0 < \gamma_1 < \cdots < \gamma_{k-1} < \gamma_k = \infty$, the superscript (j) is used to signify the regime, and $\alpha_t^{(j)}$ are i.i.d sequences with mean 0 and variance σ_j^2 and are mutually independent for different j. The parameter d is referred to as the delay parameter for different regimes (Ugurlu et al., 2018b; Tsay, 2005).

As in Markov model, 1st, 24th, 48th and 168th lags of the price series are used in the estimation, in addition to the delay parameter, d = 1.

3.4.2.4 Seasonal auto-regressive integrated moving average (SARIMA) model

ARIMA is a special kind of regression, which takes the past prices (AR), previous values of the noise (MA) and the integration level (I) of the price series into account. In SARIMA, seasonal component (S) are also involved in the estimation process. Generally, only intra-weekly nature of the series is incorporated as a seasonal component, but, in the electricity price series, it is required to deal with the intra-daily and intra-yearly seasonality as well. Therefore, triple SARIMA model of Taylor (2010) is performed by maximum likelihood assuming Gauss–Newton optimization. Equation 3.18 refers to the triple SARIMA model.

$$\phi_p(Ł)\Phi_{P_1}(Ł^{s_1})\Omega_{P_2}(Ł^{s_2})\Gamma_{P_3}(Ł^{s_3})(y_t - a - b_t) = \theta_q(Ł)\Theta_{Q_1}(Ł^{s_1})\Psi_{Q_2}(Ł^{s_2})\Lambda_{Q_3}(Ł^{s_3})\varepsilon_t \qquad (3.18)$$

y_t is the load in period t, a is a constant term, b_t is the coefficient of linear deterministic trend term; ε_t is a white noise error term; Ł is the lag operator; and ϕ_p, Φ_{P_1} , Ω_{P_2} , Γ_{P_3} , θ_q, Θ_{Q_1} , Ψ_{Q_2} and Λ_{Q_3} are the polynomial functions of orders p, P_1, P_2, P_3, q, Q_1, Q_2 and Q_3, respectively (Ugurlu et al., 2018b; Taylor, 2010).

Our triple SARIMA model can be stated as $(1, 0, 1)_1 x(1, 0, 1)_{24} x(1, 0, 1)_{168}$. To comply with the other statistical methods, ARMA(48,48) component is also added to this model.

3.4.3 Quantitative analysis

In this section, we report the performance analysis of neural networks in comparison with the state-of-the-art methods. We also use a different combination of features for shallow and deep networks to analyze the prediction accuracy. Finally, we report the monthly average results and illustrate the price estimation accuracy of GRU on a graph.

3.4.3.1 Comparison with the state of the art methods

In our first experimental setup, we use key features of lagged price values 1, 24, 48 and 168 on all described algorithms to compare the one-layered neural network algorithm performance with the state-of-the-art methods. Results in Table 3.2 indicate the neural network models' success compared to the statistical ones. Recurrent neural networks, LSTM and GRU are the best methods in this comparison. As a note, naive method outperforms two other methods, which is in line with the findings of Contreras et al. (2003), Nogales et al. (2002) and Conejo et al. (2005b), mentioning the relatively good performance of naive method.

Table 3.2 : Single-layer 24-step-ahead prediction MAE results comparison of neural network based methods with state-of-the-art techniques.

Features	Markov	Naive	SETAR	SARIMA	CNN	ANN	LSTM	GRU
F1-4	8.04	7.95	7.89	7.29	9.82	6.37	5.91	5.71

3.4.3.2 Shallow network comparisons

Our first comparison is on shallow network architectures to see the performance of each neural network method. We experiment different network architectures using the many different combinations of features in Table 3.1 following the findings of the literature. Table 3.3 demonstrates the addition of new variables into the single-layer neural networks. It should be stated that the addition of 1st and 48th lagged values of the price series to the 24th and 168th lags decrease the MAE values, but addition of the exogenous variables do have a very little or even negative effect.

Table 3.3 : Single-layer 24-step-ahead prediction MAE results. Each network of one layer and a final fully connected layer for prediction. CNNs have been implemented two convolutional layers stacked together.

Features	CNN	ANN	LSTM	GRU
F1-2	9.82	8.51	7.79	7.70
F1-4	8.57	6.37	5.91	5.71
F1-7	9.47	6.65	6.01	5.64
F1-8	10.05	8.05	6.22	5.83
F1-9	10.51	9.27	6.16	5.83
F1-10	10.64	9.85	6.02	5.58
F1-11	10.58	9.48	5.93	5.55

3.4.3.3 Deep network comparisons

To showcase the performance of deeper networks we concatenate three layers for simple ANNs, LSTMs and GRUs. It is evident in Table 3.4 that the GRU still performs the best compared to other techniques. The multiple layer structure comes up with an additional computational cost and, to find the optimal number of layers, we do a test on the algorithms.

In this deep neural networks comparison, CNN is excluded due to the low performance. Addition of the new layers increased the performance in every neural networks mechanism. However, the positive effects of the additional variables are still very small, which is in line with our findings in the shallow network comparison section.

Table 3.4 : Multi-layer 24-step-ahead prediction MAE results. Each network of stacked three layers and a final fully connected layer for prediction.

Features	ANN	LSTM	GRU
F1-2	7.63	7.66	5.86
F1-4	5.66	5.66	5.68
F1-7	5.59	5.58	5.57
F1-8	5.84	5.62	5.56
F1-9	6.08	5.70	5.57
F1-10	6.29	5.51	5.41
F1-11	6.20	5.47	5.36

3.4.3.4 Monthly comparisons

We also evaluated the monthly performance of each technique, as shown in Figure 3.5. The results for each month are generally consistent with the overall average performance with some exceptional cases. Results demonstrate the relatively good performance of the LSTM and GRU models. Although there are some months that single-layer is better than the multi-layer neural networks, in most of the months, deep neural networks give much better results. With the exception of Naive method in

August and three-layer ANN in October, recurrent neural networks, LSTM and GRU, have the best results in every month.

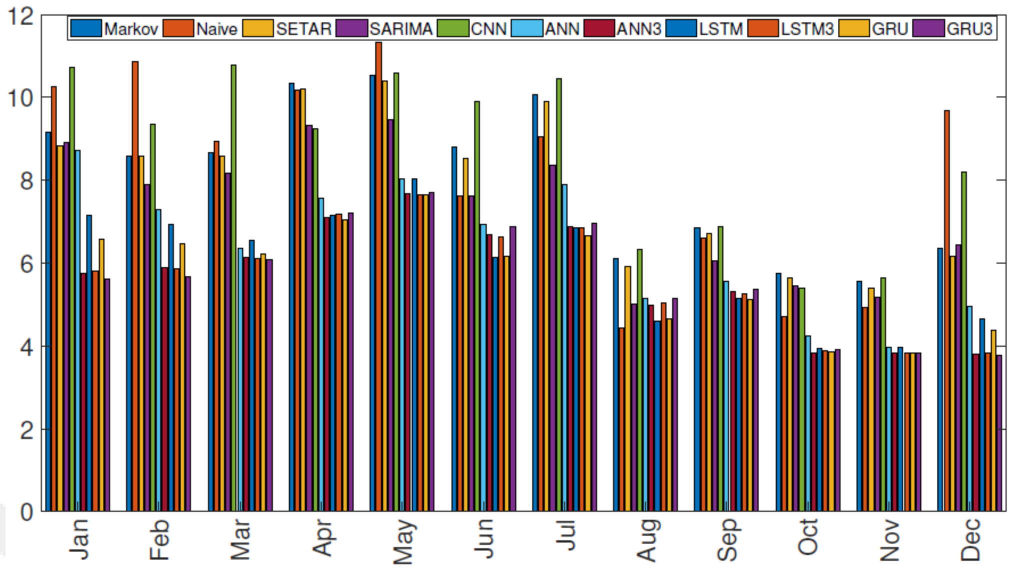


Figure 3.5 : Monthly MAE comparison of all the price estimation methods

3.4.3.5 Seasonal prediction results

We illustrate the prediction results of GRU for the sample weeks from each season we defined in Section 3.2. Figure 3.6 shows the successful performance of GRU with a good match to the original prices. We observe the ability of capturing the spikes, as well as the good performance in relatively calmer periods. It is clear that the performance of the GRU model is great in the relatively calmer autumn week. Moreover, the performance in the summer week, which has a high volatility, gives evidence about the spike detection of the model.

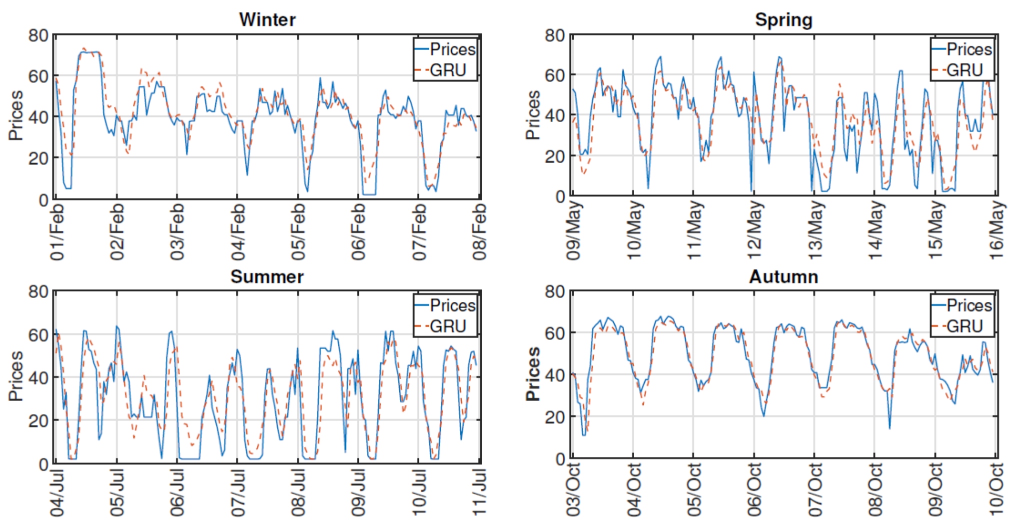


Figure 3.6 : Prediction results of GRU for a sample week from each season.

3.4.4 Diebold-Mariano tests

Tables 3.2 – 3.4 provide a ranking of the various methods, but not statistically significant conclusions on the performance of the forecasts of one method compared to others. To showcase the statistical significance of the performance difference between all model variations and features combinations, we use a Diebold–Mariano test (Diebold and Mariano, 1995), which takes the correlation structure into account. In Figure 3.7, we show the p-values for the Diebold–Mariano tests between neural network based methods and the state-of-the-art statistical methods. In Figure 3.8, we repeat the same tests for shallow and deep networks using different number of features. It tests the forecasts of each pair of transformations against each other and uses a colour map to show p-values. The low p-values show statistically significant better performance of the methods in X-axis. For example, F1-11 GRU model outperforms all the other models significantly in the three-layer networks comparison (Figure 3.8, right panel).

Figure 3.7 demonstrates the successful performance of the neural networks models, except CNN, compared to the statistical methods. Especially, good performance of the recurrent neural network models, GRU and LSTM, is statistically proven by Diebold–Mariano test.

In Figure 3.8 (left panel), single layer networks are compared with each other. F1-10 GRU and F1-11 GRU are significantly better than all the other models. Performance of F1-7 GRU and F1-4 LSTM, which do not include any exogenous variables, should also be mentioned. In Figure 3.8 (right panel), in three-layer networks, addition of new features has a much more significant effect than the single layer network. F1-11 GRU, F1-10 GRU, F1-11 LSTM, and F1-10 LSTM are the best methods among three-layer networks.

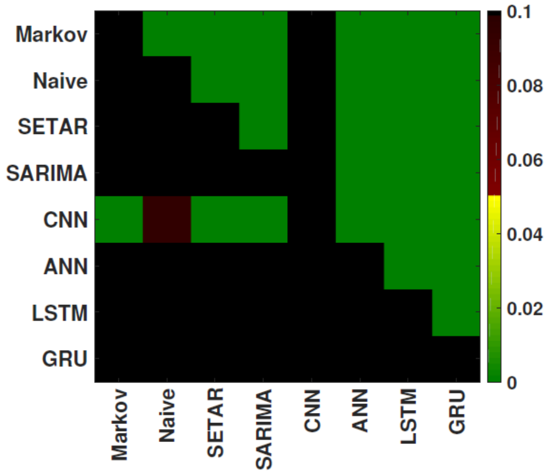


Figure 3.7 : Results of the Diebold-Mariano tests defined by the loss differential series as absolute errors in between all investigated parameters for F1-4. The figure indicates the statistical significance (green) for which the forecasts of a model on the X-axis are significantly better than those of a model on the Y-axis.

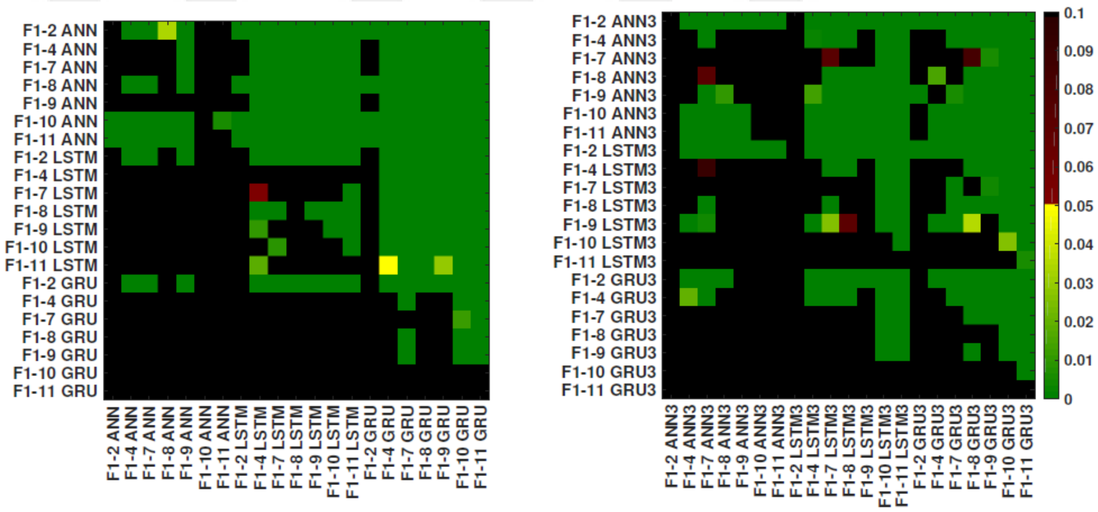


Figure 3.8 : Left panel: Single layer networks. Right panel: Three-layer networks. Results of the Diebold-Mariano tests defined by the loss differential series as absolute errors in between all investigated parameters and used features for different number of layers. The figure indicates the statistical significance (green) for which the forecasts of a model on the X-axis are significantly better than those of a model on the Y-axis.

3.4.5 Implementation details

The training of a neural network can be viewed as a combination of two components, a loss function or training objective, and an optimization algorithm that minimizes this function. In this study, we used the Adam optimizer to minimize the mean absolute error loss function. The training ends when the network does not significantly improve for a predefined number of epochs (300).

During training, a batch-size of three years was used. The momentum of the optimizer was set to 0.90 and the learning rate was 0.001. The parameters of the fully-connected, convolutional, and recurrent layers were initialized randomly from a zero-mean Gaussian distribution. The training continued until no substantial progress was observed in the training loss.

We performed multiple tests to see the performance of different numbers of layers in ANN, LSTM and GRU architectures for selecting the optimal number of layers. Figure 3.9 shows that the optimal results can be achieved using three layers. Additional layers increase in the total number of parameters and add to the computational cost without achieving a significant gain in the performance.

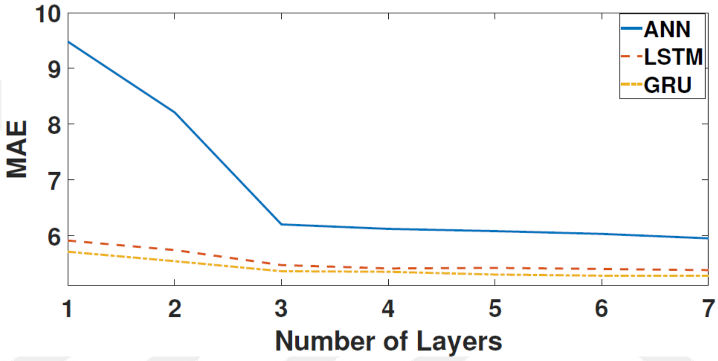


Figure 3.9 : Performance change when applying different number of layers to ANN, LSTM and GRU algorithms.

3.5 Discussion

In this paper, we investigate the application of various neural network architectures on electricity price forecasting. Our experiments in Table 3.2 highlight that neural network based methods produce better results compared to the state-of-the-art statistical forecasting methods in the literature such as SARIMA and Markov models. We use simple artificial neural networks (ANNs), CNNs, LSTMs and GRUs to estimate the electricity prices in the Turkish market. We see that the RNN models, namely LSTM and GRU, are able to separate themselves in terms of performance compared to CNNs and simple ANNs in Table 3.3. This is because RNN models have memory about the previous time steps, which makes them the method of choice for time series type problems. They keep a memory of the previous instances effectively, which is crucial for estimating electricity prices of the day-ahead market.

The deep learning paradigm of stacking multiple layers increases the performance for ANNs, LSTM and GRUs, as highlighted in Table 3.3 in comparison with Table 3.4. GRUs still give the best performance among all available techniques and we reached the best results of 5.36 Euros/MWh MAE using three-layered GRUs. The results show good alignment with the prices as illustrated in Figure 3.6.

Neural networks are data-driven models and their performance heavily depends on the availability of the large training data. The limited data are a deteriorating factor for all training based methods, but in particular for neural network based methods. We show in Figure 3.9 that the performance does not improve after three layers for any of the networks due to the limited data. With the availability of further data, we believe the overall performance of LSTM and GRU methods will be better.

Another significant observation is the fact that GRUs perform better than the LSTM models. This can explained by the fewer number of parameters that are needed to be learned by GRUs. In the literature, Yin et al. (2017) and Chung et al. (2014) compared the two models for polyphonic music modelling and speech signal modelling task. They showed the better performance of GRU for these tasks. Moreover, GRUs train faster due to the fact that they require fewer parameters.

We see that the key features are lagged price values for estimating the electricity prices, which is in line with the findings of Uniejewski et al. (2016). In terms of single layer, addition of 1st and 48th lagged values to the 24th and 168th lagged values have an important effect. Especially for LSTM single layer using the 1st, 24th, 48th and 168th lagged values is as good as using all the variables. For GRU, adding 23rd, 72nd and 336th lagged values give better results. Addition of exogenous variables have a very small effect in LSTM. Although addition of forecast D/S and temperature do not have a significant effect in GRU, further addition of 24th lags of realized D/S and balancing market price have significant effects. In three-layer networks, results are similar, but addition of features help much more to have better results. If we do not use any exogenous variables, F1-7 gives better results than F1-4. In three-layer GRU networks, addition of all the variables, except temperature, change the performance significantly. On the other hand, LSTM F1-7 is only worse than LSTM F1-10 and F1-11, which is similar to the single layer results. To conclude, endogenous variables are the most important ones and using the 1st, 24th, 48th and 168th lagged prices give relatively good results. In most cases, adding one or two exogenous variables does not improve the results, but if we use the lagged values of the other exogenous variables, in addition to forecast D/S and temperature, then these models with all the variables significantly outperform the models with fewer variables.

One additional comparison we made was grouping the results in terms of months. It is possible to say that the general error levels are lower in autumn and wintermonths compared to spring and summer months. In relatively mild weather months of Turkey -October, November and December- three-layer GRU networks' MAE values are lower than 4 Euros/MWh. On the other hand, relatively hot weather months of Turkey -May, June, and July- have MAE values around 7 Euros/MWh, which is almost double of the mild weather months. It must be mentioned that, in most countries, prices during summer months are not high compared to the other months, but, as mentioned in Section 3.1.2 on the Turkish market, due to the requirement of air conditioning, prices during summer months are very close to the winter months prices. We can conclude that the MAE values show a similar pattern with the price levels, which demonstrate the effect of the seasonality.

Our results are in line with the main findings of Lago et al. (2018a), Kuo and Huang (2018), which is that machine learning models, especially deep neural networks, outperform the state-of-the-art statistical models and shallow neural networks. On the other hand, in our experiment, deep recurrent neural networks, LSTM and GRU, which are tailor-made for time-dependent problems, give lower errors than DNN, which contradicts with the results of Lago et al. (2018a). Lago et al. (2018a) made two hypotheses about the unexpected superiority of DNN in their paper: first, low amount of data; and, second, different structure of the models. Moreover, they underlined the necessity of further research. In our opinion, having deep LSTM and GRU, instead of shallow LSTM and GRU, causes the conflict between the results. Lago et al. (2018a) applied single-layer LSTM and GRU, or apply LSTM and GRU as one layer of the hybrid deep neural networks. In our case, there are three layers of LSTM and GRU in the experiments. Another possible explanation is the market specifics. Turkish market has an increasing share of hydro and renewables in the energy production and the market is similar to the Spanish (Diaz and Planas, 2016) and German (Keles et al., 2016) markets in some aspects. However, as we know that all the markets have unique characteristics, generalizability to other markets needs further research. Incredibly fast changing nature of the energy markets, especially in the emerging economies, must also be mentioned. Establishment of two nuclear plants in the next five years, inclusion of the solar energy into system in near future and expiration of the subsidies for the wind power plants in two years will change the dynamics of the Turkish market as well. Therefore, further research in Turkish market and in the emerging economies, such as Southeast Europe markets (Hryshchuk and Lessmann, 2018) is also required.

Generalization capability of machine learning models is promising for applying our model fordifferent market data. The GRU network architecture can accurately predict the electricity prices in the Turkish market. With the availability of the multiple feature data for each market, the model can be applied to various markets using domain adaptation. However, Aggarwal et al. (2009) underlined the superiority of different methods in different markets and combination of multiple methods might be promising in these type of problems. We would like to investigate possibility of using hybrid models to merge benefits of multiple methods. Zhang (2003) proposed combining ARIMA and ANN models to forecast the linear and non-linear components of price separately. Chaabanae (2014) developed the Zhang (2003) method and combined auto-regressive fractionally integrated moving average (ARFIMA) with neural networks model. Guo and Zhao (2017) also utilized decomposition, optimization and support vector machine techniques in a hybrid work. In another example, Shrivastava and Panigrahi (2014) applied a hybrid wavelet extreme learning machine. Moreover, Alamaniotis et al. (2015b) combined relevance vector machines and linear regression ensemble optimization. These types of hybrid approaches can aid the performance of RNNs.

The uncertainty of the predictions made by the neural network models can be of great value to assess their utility. Currently Bayesian based neural networks are used to predict the uncertainty of the neural network based predictions (Iwata and Ghahramani, 2017). With the developments in machine learning literature, we would like to estimate the uncertainty values of GRUs and LSTMs to increase the reliability of both methods. Recent work by Hwang et al. (2018) opens the path for fast and accurate uncertainty estimations of GRUs (Hwang et al., 2018).

One avenue of improvement for our method is to investigate the decomposition techniques. Related to the hybrid models, Neupane et al. (2017) proposed an ensemble prediction method by choosing the algorithm and features among a set of them, which give much better forecast results than state-of-the-art techniques. In another work, Hong and Wu (2012) applied principal component analysis (PCA) as a dimension reduction method. Ziel (2016) and Ludwig et al. (2015) used Lasso shrinkage method for variable selection. Zheng et al. (2017) proposed using empirical mode decomposition for decomposing the signal to several intrinsic mode functions (IMFs) and residuals. They used these IMFs to train LSTM to forecast short-term load. In the future, we would like to include dimension reduction algorithms and investigate their contribution to seasonality of the data, in particular in RNN setting.

In conclusion, this study instigated the utility of neural networks for electricity price estimation. Development of new conditions in electricity markets across the world brings new challenges. Accurate price estimation is a crucial task for adapting to the new market conditions, and machine learning methods are capable of addressing these issues with high accuracy. Recurrent Neural Networks set the state-of-the-art in addressing time-dependent problems. With this work, we show a detailed analysis on RNNs for electricity price forecasting and highlight the superior performance of GRUs in comparison to various neural network based methods and state-of-the-art statistical techniques.

4. THE FINANCIAL EFFECT OF THE ELECTRICITY PRICE FORECASTS' INACCURACY ON A HYDRO-BASED GENERATION COMPANY[28]

4.1 Introduction

Electricity price forecasting has become an essential task since the liberalization of the electricity markets. It is integral for all the players in the energy markets, due to several reasons. Firstly, both supply and demand sides present their bids in the regulated markets according to electricity price forecasts. Secondly, bilateral contracts and energy derivatives also use longer term electricity price forecasts as reference points. Thirdly, large-in-scale demand side bidders, such as distribution companies, large industrial companies or pumped storage units can manage their purchasing behavior according to electricity price predictions. Last, but not least, generation companies (GenCos), such as hydro, natural gas and fuel oil can schedule their generation and bidding behavior according to the day-ahead price forecasts to maximize their profits.

This paper presents the influence of electricity price forecast accuracy on the profit maximization of GenCos. In particular, we use Mixed Integer Linear Programming (MILP) to schedule production strategies of a hydro-based power plant to minimize the profit loss of the companies. We make use of five individual and four hybrid forecast models to schedule the electricity production of the hydro-power plant. The main contributions of this paper are particularly:

- Extensive analysis of the financial influence of electricity price estimation inaccuracy;

[28] This chapter is based on the paper "The Financial Effect of the Electricity Price Forecasts' Inaccuracy on a Hydro-based Generation Company". Ugurlu, U., Tas, O., Kaya, A. & Oksuz, I. 2018. The Financial Effect of the Electricity Price Forecasts' Inaccuracy on a Hydro-based Generation Company. *Energies*, *11* (8), 2093.

- Analysis of statistical methods, Artificial Neural Networks (ANN), Long Short Term Memory (LSTM), Gated Recurrent Units (GRU) and hybrid methods for electricity price estimation;
- Use of a hybrid ANN–LSTM method for estimating electricity prices to maximize the profit;
- Detailed statistical analysis between electricity price estimation and profit maximization of GenCos.

4.1.1 Electricity price forecasting

Electricity price forecasting is an ever-improving research area and many different methods are implemented to forecast electricity prices. The review of Weron (2014) categorizes electricity price forecasts into five main groups: (1) Multi-agent models, (2) fundamental models, (3) reduced-form methods, (4) statistical time-series methods, and (5) machine learning models. Generally, the first two groups are more applicable to the smaller markets. The first group includes game-theory type models for the smaller markets with less numbers of participants. Fundamental models require all supply and demand information to intersect both curves to obtain the price. Reduced-form methods are mainly successful in the price spikes, which are one of the important characteristics of electricity prices, and statistical methods include regression-type methods from relatively easy naïve methods (Nogales et al., 2002) to complex models (Ziel and Weron, 2018). Machine learning methods include several different sub-categories, such as neural networks, fuzzy logic, support vector machines, etc. A time-dependent type of neural networks, recurrent neural networks, provide notably impressive results nowadays, especially with the addition of more than one layer, which is then called a deep neural network (Lago et al., 2018a; Ugurlu et al., 2018a). According to Ugurlu et al. (2018a), deep neural networks, especially deep recurrent neural networks, such as Long-short Term Memory (LSTM) and Gated Recurrent Units (GRU), outperform the statistical time series methods like Seasonal Auto Regressive Integrated Moving Average (SARIMA), as well as shallow and deep Artificial Neural Networks (ANN). These findings are mainly in line with the results of Lago et al. (2018a). Although LSTM and GRU, which are tailor-made for time series, are expected to perform better than deep ANN, Lago et al. (2018a) find out that deep ANNs are better than deep recurrent neural networks. On the other hand, a superiority of Lago et al. (2018a) to Ugurlu et al. (2018a) is that it proposes 27 models and Lago et al. (2018a) can be used as a benchmark in the electricity price forecasting literature. Models of Lago et al. (2018a) also contain some deep hybrid methods, which motivated us to use the deep hybrid methods. Although current research has had promising results in favor of machine learning methods, Lasso regression applications (Ziel, 2016; Ludwig et al., 2015), ensemble predictions (Neupane et al., 2017; Alamaniotis et al., 2015a; Shrivastava and Panigrahi, 2014), and hybrid works (Chaabane, 2014; Hong and Wu, 2012; Keles et al., 2012; Bordignon et al., 2013) also have successful results. One important example is the work of Chaabane (2014), which combines SARIMA with Auto-Regressive Fractionally Integrated Moving Average (ARFIMA). However, as Aggarwal et al. (2009) mentioned, still, none of the methods outperform the others regularly and continuously.

4.1.2 Generator companies' profit maximization

There are mainly two problems generation companies need to solve related to electricity price forecasting. The first problem is the self-scheduling of the power plants, which means the optimization of the production quantities for each hour of the next day according to the price forecasts. The second problem is presenting the correct price bids related to these quantities. This article will focus on the first problem. The purpose of this paper is to propose a Price Based Unit Commitment (PBUC) to a hydro-based GenCo according to the price forecasts of various methods, both from statistical and machine-learning methods, applied in Ugurlu et al. (2018a). According to the price forecasts, a GenCo will procure the production of electricity. It should be mentioned here that electricity is a non-storable commodity, whereas water can be stored. Thus, hydro GenCos have the opportunity of storing electricity in the shape of water, which eases most of the production costs and constraints compared to the other types of GenCos, i.e., thermal, wind or solar. In this sense, a self-scheduling optimization problem must be solved by using a technique, such as Mixed Integer Linear Programming (MILP), Lagrangian relaxation, dynamic programming or genetic algorithms (Li and Shahidehpour, 2005). There are two assumptions when optimizing the hydro-based GenCo's self-scheduling (Delarue et al., 2010). The first assumption is that the GenCo is a price taker, which means that the price bids the GenCo presents do not have a significant effect on the determined market price. In a relatively big market, this assumption is easily justified, related to the capacity of the GenCo. The second assumption is that all the quantities offered will be accepted and sold. This assumption is also reasonable, considering that most of hydro GenCos, as well as solar and wind power plants, give their bids with zero prices and accept the clearing price.

In a pioneer work, Zareipour et al. (2010) measure the economic impact of inaccurate electricity price forecasts from the demand side. Two different typical industrial loads, process-industry load and municipal water-plant load, are investigated, and the actual prices and inaccurate electricity price forecasts for the next 24-h are compared. One of the main findings of Zareipour et al. (2010) is that Mean Absolute Percentage Error (MAPE) cannot reflect the economic value of improving the forecast accuracy. Therefore, other financial indicators should be checked to evaluate the financial effect of inaccurate electricity price forecasts. Another point is that the expectation from the forecast varies according to the type of customer. For example, the process industry needs accurate forecasts with respect to an exact threshold. On the other hand, for the water-plant, knowing the general trend of electricity prices over the planning period is quite helpful. Another important study, which was published right after Zareipour et al. (2010), is Delarue et al. (2010). The main difference of Delarue et al. (2010) from Zareipour et al. (2010) is that its point of view is mainly from the supplier side. Another distinction is that it uses MILP to solve the PBUC problem. In Delarue et al. (2010), four different power-plant types are examined. Combined cycle power-plant and pumped storage power plants are affected more by inaccurate forecasts than hydro and coal-fired classical thermal power plants, in terms of profit loss. Another interesting finding is that if inaccurate forecasts have an upside or downside bias, then profit loss gets affected by this bias as well. Downside bias, which means predicting the prices lower than the actual ones, cause higher profit losses. Mohammadi-Ivatloo et al. (2011) also examine the economic impact of four different price forecasts compared to the actual prices for GenCos. Mohammadi-Ivatloo et al. (2011) take a hydro power-plant and a thermal power-plant into account. This research proposes two indices to evaluate the effect of inaccurate forecasts: The first one is the Economic Loss Index (ELI), which is the profit loss of the electricity price forecasting model, in terms of percentage, from the actual price profit; the second one is the Price Forecast Disadvantage Index (PFDI), which shows the profit loss per energy sold. According to the results of Mohammadi-Ivatloo et al. (2011), traditional error measures do not always cause significantly high economic losses. This means that a model with lower forecast performance errors could cause higher economic losses than a model with higher forecast performance errors, and vice versa. Thus, according to Mohammadi-Ivatloo et al. (2011), using ELI or PFDI instead of Mean Absolute Error (MAE) or MAPE as a financial effect evaluation method would yield less profit loss. In a related paper, Mathaba et al. (2014) work on the same topic and propose a method which could be used in choosing the best forecast mechanism. According to Mathaba et al. (2014), using the Rank Correlation (RC) method instead of Root Mean Square Error (RMSE) or MAPE would cause less profit loss. Mathaba et al. (2014) use a coal-conveying system with storage, which allows the combining of the supply and demand sides. Three forecasting methods in the U.S. Pennsylvania-New Jersey-Maryland (PJM) market are performed. There are mainly three findings of Mathaba et al. (2014): Firstly, RC is a better indicator than RMSE and MAPE in terms of having less profit loss; secondly, price volatility, rather than mean price, has a higher effect on the profit loss. Therefore, models which take volatility into account could have less profit loss. Thirdly, profit loss is very dependent on the responsiveness of the load to electricity price changes. Research of Doostmohammadi et al. (2017) proposes a completely different evaluation method for the same problem. First of all, it prepares a financial loss/gain (FLG) time series by using the real conditions of the electricity market. Secondly, they quantize this signal by using the Silhouette criterion and k-means clustering technique to simplify the problem. Then, the most informative variables are chosen from a feature selection problem by a combined technique. Lastly, by using all these methods and extreme learning machine, FLG predictions can be made which help the GenCos to optimize their scheduling. This is an interesting work because it combines the forecasting procedure with the profit loss calculation mechanism and the results show the positive effect of this evaluation.

To solve the self-scheduling optimization problem of the supplier company, different optimization methods are used in the literature. The most common technique is the MILP proposed in a pioneer work by Conejo et al. (2002), which introduces a self-scheduling plan for a hydro producer in a pool-based electricity market. There are eight cascaded hydro power plants along a river basin in this system. It is a relatively big power plant, in terms of production capacity. Thus, it is difficult to fulfill the assumption that the supplier is a price taker and its bids do not change the market price. The objective of the optimization is maximizing the profit by selling electricity in the day-ahead market. For each plant, it takes nonlinear and non-concave three-dimensional relationships between the power produced, water discharged and the head of the related reservoir. Start-up costs are definitely in the calculation of profit. Similar to common techniques in the literature (Esmaeily et al., 2017; Ahmadi et al., 2012; Karami et al., 2013), Conejo et al. (2002) also utilize the IEEE 118-bus test system (IEEE, 2018) as the hydro power plant. Applied to the Spanish day-ahead market prices in 2001, daily profits are around $600,000, which shows the great capacity of the hydro power-plant system. Esmaeily et al. (2017) add some practical constraints, such as valve-loading cost, dynamic ramp rate and prohibited operation zones. As in Conejo et al. (2002), they take the price forecast errors to understand the price uncertainty. They also performs a Lattice Monte Carlo Simulation for the effects of spinning and non-spinning reserve prices. Working hours of the hydro power-plant correspond to the high price hours in this model, which cause relatively high profits. Bisanovic et al. (2008) is another study about the hydro-thermal self-scheduling problem in the day-ahead electricity market. As in the previous papers, Bisanovic et al. (2008) use MILP, because this optimization method allows one to have non-convex and non-linear items as constraints. The difference for Bisanovic et al. (2008) is that it takes the long-term bilateral contracts, in addition to the hourly day-ahead electricity price forecasts, into account. As another difference, the system in this research is the combination of the thermal and hydro power plants. On the other hand, a common point with the other papers (Li and Shahidehpour, 2005; Conejo et al., 2002; Esmaeily et al., 2017) is that this model also utilizes the piece-wise linear model to represent the non-linear functions. Bisanovic et al. (2008) solve the PBUC problem for four different types of plants: Thermal, combined cycle, cascaded hydro and pumped storage. This paper also utilizes the IEEE 118-bus system and applies mixed integer programming (MIP), which is compared with the Lagrangian relaxation (LR) method. Yamin and Shahidehpour (2004) utilized transmission congestion and locational marginal prices as well as fuel and emission constraints in their model. Shahidepour et al. (2002) give a broad overview in the forecasting, scheduling and risk management of power systems. With the availability of high memory and greater computational power, MIP and MILP type optimization techniques have become state-of-the-art methods for PBUC problems.

4.1.3 Turkish market

Although there are some works (Ugurlu et al., 2018a; Hayfavi and Talasli, 2014; Ozyildirim and Beyazit, 2014; Ugurlu et al., 2018a) on electricity price forecasting in the Turkish electricity market, the financial effect of electricity price forecasts' inaccuracy has not been investigated. Due to the nature of the Turkish market, with many zeros similar to the Spanish market (Diaz and Planas, 2016), and an increasing renewable share similar to the German market (Keles et al., 2012), the Turkish market needs investigation, and can give some insight about the other Southeast Europe markets (Hryshchuk and Lessmann, 2018) in addition to Spanish and German markets. The Turkish market is an emerging market with very specific features as discussed in Ugurlu et al. (2018a), Avci-Surucu et al. (2016) and Ozozen et al. (2016). Mainly due to the country's climate, in the summer months, electricity consumption and the prices are almost as high as the winter months. Turkey has an inter-connected grid with Greece and Bulgaria, but the share of the import and export electricity never exceeds 1% of the daily consumption. One of the most important features of the Turkish market is the high share of hydro energy of 34.2% and the increasing share of the wind energy of 7.6% at the end of 2016, in terms of the installed capacity (EPDK, 2018). Due to the snow-melt effect in the spring months and the wind effect in spring and autumn months especially, prices tend to decrease in these seasons. Turkish electricity prices are limited from 0 to 2000 TL/MWh, which is approximately 562 $/MWh by the 2016 average exchange rate (Ugurlu et al., 2018a). This rule does not allow the Turkish market to have negative prices, which is similar to the Canadian market (Filipovic et al., 2017) and the opposite of the German market (Keles et al., 2012; Fanone et al., 2013). Taking the number of zeros in the Turkish market into account, we can mention that it has a negative effect on the market efficiency. However, on the other hand, the price cap of approximately 562 $/MWh is very high and has never been reached in the short history of the market since December 2011. However, prices tend to have very high values, especially in the lack of natural gas for the power plants. In the near future, two nuclear power plants will be established as the first attempts of Turkey; solar power will be integrated into the grid; wind share still has an increasing trend, but the subsidies on the wind power plants will stop in two years. As another point of interest, intra-day market and energy derivatives are also developing and in need of research. To sum up, there is a rapidly changing and improving environment in this major emerging market, which motivates our work to investigate the specific characteristics of this market. Evaluating the effects of electricity price forecasts' inaccuracy in the Turkish Day-Ahead Market and comparing it with the results of the other markets could help us to understand the nature of the market better. In addition to this, our paper is the first one which takes deep neural networks, especially deep recurrent neural networks, as forecast methods to investigate the relationship between the forecasts' inaccuracy and the financial effect caused by this inaccuracy. Moreover, it also combines the predicted electricity prices of these models as hybrid methods and compares them with the individual counterparts, in terms of profit loss. The remainder of the paper will be structured as follows: Section 4.2 describes the dataset and the method. In particular, electricity price forecasting, Price Based Unit Commitment (PBUC) and the financial effect measures are detailed. In Section 4.3, results are given and discussed. Section 4.4 concludes the paper and investigates some further research ideas.

4.2 Data and Methods

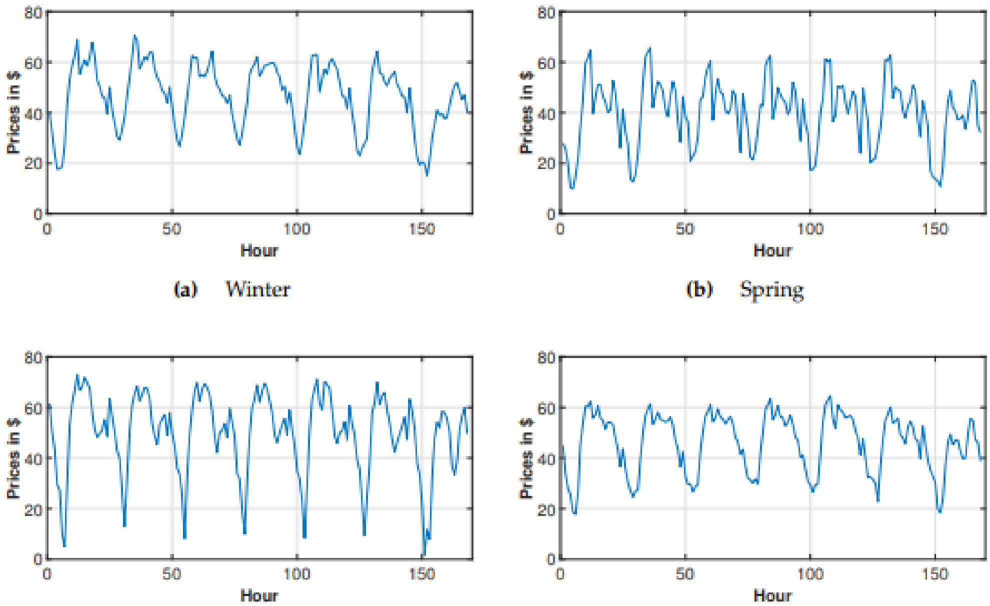
We use data from mainly two sources in our simulations. Firstly, Turkish Day-ahead Market electricity prices are taken from EPIAS (2018) between 2013 and 2016. Secondly, IEEE 118-bus-system test data (IEEE, 2018) are used in the hydro-based GenCo's self-scheduling. The seasonal average of each hour of the week from the year 2016 is visualized in Figure 4.1. We have listed the descriptive statistics of the test data for each hour of the data in the year 2016 in Table 4.1.

Firstly, we need to mention the intra-day seasonality of the data. The price average at 6 am is nearly one-third of the price average of 11 am. Early morning hours have the lowest prices with the highest variation. It is especially difficult to forecast these early morning prices. Secondly, there are many zeros in the prices, which make the preliminary studies of the data difficult. In the statistical methods, to make the data stationary, there is a need for transformation. Due to these low prices, it is impossible to take the logarithmic returns. Moreover, prices around zero cause biased results in the MAPE numbers. Thirdly, the highest price of 2016 is 132.36 \$/MWh, which is beyond $\mu + 5\sigma$ for 10 am. Figure 4.1 mainly visualizes the intra-year seasonality. Although the range is very small in the spring and autumn months, in the summer the average daily range of the prices are as high as 70 \$/MWh. Figure 4.1c illustrates the average summer prices of 2016. Due to religious holidays, which are not coherent with the Europe, prices do not show co-movement in Europe's and Turkey's holidays. The prices show a sharper decrease on Fridays at lunchtime due to the Friday prayer, which can be seen especially in Figure 4.1a. Another difference to the European market is the half-day working habit on Saturdays. In Figure 4.1, relatively high prices can be observed in the morning hours on Saturdays. Due to the usage of the air-conditioning because of the hot climate in the summer months, prices are very high in the day-time. On the other hand, prices in the early morning hours are very low. This causes lots of spikes, which makes electricity price forecasting especially difficult in the summer months. In Figure 4.1b, due to the snow-melt effect, hydro power-plants work in high levels and produce relatively low-priced electricity.

Table 4.1 : Descriptive statistics of the Turkish day-ahead electricity prices ($/MWh) according to hours of the day.

Hours	Mean	Std. Deviation	Lower Bound	Upper Bound	Median
0	49.27	13.47	0.28	78.42	47.98
1	41.80	14.99	0.00	78.05	42.84
2	35.89	15.98	0.00	76.69	37.85
3	27.40	16.31	0.00	76.13	27.57
4	25.42	16.55	0.00	76.12	26.27
5	23.77	15.36	0.00	77.86	25.12
6	22.65	17.62	0.00	77.94	24.55
7	34.52	17.54	0.00	78.08	40.27
8	44.76	17.91	0.00	79.20	48.88
9	55.74	15.32	0.00	99.83	58.98
10	59.53	13.93	0.00	132.36	60.40
11	62.66	13.20	0.32	99.26	51.11
12	51.64	15.30	0.32	99.26	51.11
13	54.27	14.18	1.71	99.26	55.14
14	57.26	14.51	0.36	113.16	59.41
15	55.13	14.27	0.36	96.14	57.41
16	54.12	14.34	0.34	96.14	54.23
17	50.80	15.56	1.70	124.03	50.69
18	48.85	13.46	0.27	90.88	49.25
19	49.24	12.15	3.60	81.36	50.66
20	51.56	9.75	20.52	78.66	52.30
21	49.16	9.78	17.88	78.61	49.33
22	46.30	12.58	1.59	78.74	45.26
23	39.17	14.34	0.00	78.42	40.41

Firstly, we need to mention the intra-day seasonality of the data. The price average at 6 am is nearly one-third of the price average of 11 am. Early morning hours have the lowest prices with the highest variation. It is especially difficult to forecast these early morning prices. Secondly, there are many zeros in the prices, which make the preliminary studies of the data difficult. In the statistical methods, to make the data stationary, there is a need for transformation. Due to these low prices, it is impossible to take the logarithmic returns. Moreover, prices around zero cause biased results in the MAPE numbers. Thirdly, the highest price of 2016 is 132.36 \$/MWh, which is beyond $\mu + 5\sigma$ for 10 am. Figure 4.1 mainly visualizes the intra-year seasonality. Although the range is very small in the spring and autumn months, in the summer the average daily range of the prices are as high as 70 \$/MWh. Figure 4.1c illustrates the average summer prices of 2016. Due to religious holidays, which are not coherent with the Europe, prices do not show co-movement in Europe's and Turkey's holidays. The prices show a sharper decrease on Fridays at lunchtime due to the Friday prayer, which can be seen especially in Figure 4.1a. Another difference to the European market is the half-day working habit on Saturdays. In Figure 4.1, relatively high prices can be observed in the morning hours on Saturdays. Due to the usage of the air-conditioning because of the hot climate in the summer months, prices are very high in the day-time. On the other hand, prices in the early morning hours are very low. This causes lots of spikes, which makes electricity price forecasting especially difficult in the summer months. In Figure 4.1b, due to the snow-melt effect, hydro power-plants work in high levels and produce relatively low-priced electricity.



(a) Winter **(b)** Spring

(c) Summer **(d)** Autumn

Figure 4.1 : Hourly averaged electricity prices for each season of 2016.

4.2.1 Electricity price forecasts

Electricity prices are forecasted for 6 weeks of every season in 2016 by using various methods following Ugurlu et al. (2018a). A 3-year rolling window scheme is used for the estimation and 24-step ahead hourly forecasts are done by using the endogenous variables, which are the 1st, 24th, 48th and 168th lags of the price series. In this way, forecasts are done for 2 weeks from all the months of 2016. Five different forecast methods, namely the Naïve method, SARIMA, ANN, LSTM, and GRU are utilized in this paper. In addition to these models, four different combinations of the best performing ANN, LSTM and GRU models' forecasts are also used to evaluate the financial effect of the forecast inaccuracies compared to the "best" ex-post actual prices case. This paper uses the models from Ugurlu et al. (2018a); brief descriptions of all the models are given below in the related sections.

4.2.1.1 Naïve method

The naïve method is a benchmark in the electricity price forecasting literature, which takes the previous day's or previous week's same hour as the forecast price (Nogales et al., 2002). According to Nogales et al. (2002) and Conejo et al. (2005b), unsuccessful forecasts cannot outperform this benchmark method. The naïve method is described as:

$$\mathrm{P_{d,h}} = \begin{cases} \mathrm{P_{d-7,h}},\ \varepsilon_{d,h} & \text{Monday, Saturday, Sunday} \\ \mathrm{P_{d-1,h}},\ \varepsilon_{d,h} & \text{Tuesday, Wednesday, Thursday, Friday} \end{cases} \tag{4.1}$$

$\mathrm{P_{d,h}}$ states the price of the selected day and hour. $\varepsilon_{d,h}$ stands for the noise term.

4.2.1.2 Seasonal auto-regressive integrated moving average model

ARIMA is a special kind of regression, which takes the past prices (AR), previous values of the noise (MA) and the integration level (I) of the price series into account. In SARIMA, a seasonal component (S) is also involved in the estimation process. Generally only the intra-weekly nature of the series is incorporated as a seasonal component, but in the electricity price series it is required to deal with the intra-day and intra-year seasonality as well. Therefore, the triple SARIMA model (Taylor, 2010) is performed by maximum likelihood assuming Gauss-Newton optimization. Equation 4.2 refers to the triple SARIMA model.

$$\phi_{\mathrm{p}}(\text{Ł})\Phi_{\mathrm{P_1}}(\text{Ł}^{s_1})\Omega_{\mathrm{P_2}}(\text{Ł}^{s_2})\Gamma_{\mathrm{P_3}}(\text{Ł}^{s_3})(\mathrm{y_t} - \mathrm{a} - \mathrm{b}_t) = \theta_{\mathrm{q}}(\text{Ł})\Theta_{\mathrm{Q_1}}(\text{Ł}^{s_1})\Psi_{\mathrm{Q_2}}(\text{Ł}^{s_2})\Lambda_{\mathrm{Q_3}}(\text{Ł}^{s_3})\varepsilon_{\mathrm{t}} \tag{4.2}$$

y_t is the load in period t, a is a constant term, b_t is the coefficient of linear deterministic trend term; ε_t is a white noise error term; Ł is the lag operator; and ϕ_p, Φ_{P_1} , Ω_{P_2} , Γ_{P_3} , θ_q, Θ_{Q_1} , Ψ_{Q_2} and Λ_{Q_3} are the polynomial functions of orders p, P_1, P_2, P_3, q, Q_1, Q_2 and Q_3, respectively (Ugurlu et al., 2018b; Taylor, 2010).

Our triple SARIMA model can be stated as $(1, 0, 1)_1$ x $(1, 0, 1)_{24}$ x $(1, 0, 1)_{168}$. To comply with the other statistical methods, ARMA (48,48) component is also added to this model.

4.2.1.3 Artificial neural networks

There is a growing interest in Artificial Neural Networks (ANN) in the electricity price forecasting literature (Keles et al., 2016; Mandal et al., 2006; Kolmek and Navruz, 2015) as well as many other areas. ANN consists of layers of neurons, which are connected densely. They are also called Multi-layer Perceptrons (MLP). In this paper, we use three-layer ANN, where each layer has 10 neurons and a final layer estimates the forecast values. The batch-size is 3-years during training, the learning rate is 0.001, the momentum of the optimizer is 0.90 and 300 epochs are used (Ugurlu et al., 2018a).

4.2.1.4 Long short term memory

Long short term memory (LSTM) is a type of the recurrent neural networks (RNNs). RNNs are the best fit for the time-dependent problems, because they allow the information to persist, with their loops-allowing architecture. Due to their nature, which allows using the temporal information as the input, RNNs are the best models for the time series data. In a unique type of the recurrent neural network, LSTM, each node can be used as a memory cell, which can store the information from the other cells as well. Therefore, LSTM addresses the vanishing gradients problem of the previous time steps. The input, forget and output gates of LSTM control the existing memory and take the information from the first moments of the learning process and use it much later. This feature gives the opportunity to modeling long-term dependencies. The same batch-size, learning rate, momentum of the optimizer and epochs with the ANN model are used for the 3-layer LSTM model (Ugurlu et al., 2018a).

4.2.1.5 Gated recurrent units

Gated recurrent units (GRU) are another type of RNN, which is utilized in time-dependent problems with considerable success. GRU consists of two gates, namely the reset gate and update gate. The update gate determines how much of the previous memory will be used and the reset gate decides how to combine the previous memory and the new input. The main aim of the GRU is very similar to LSTM, which is taking long-term dependencies into account. However, in GRU there are only two gates and fewer parameters than LSTM. Instead of having only a reset gate, as in LSTM, in GRU there is both a reset gate and update gate. Another difference is that LSTM has output gates, but GRU does not have any. In our experiment, we use a 3-layer GRU model with the same features of the ANN and the LSTM model (Ugurlu et al., 2018a).

4.2.1.6 Hybrid models

We also form 4 different hybrid models by combining the forecasts:

1. 50% LSTM – 50% GRU
2. 50% ANN – 50% GRU
3. 50% ANN – 50% LSTM
4. 33% ANN – 33% LSTM – 33% GRU

The combinations are selected according to the best performing models in Ugurlu et al. (2018a). We examine the performance of the hybrid models, consisting of neural networks, and compare them with the individual model counterparts in terms of the financial effect of the forecast inaccuracy. This is the first work, which investigates the hybrid models from this point of view.

4.2.2 Hydro-based power plant

In order to model the behavior of GenCos, we used IEEE 118-bus-system test data (IEEE, 2018). These data are from a hydro-based power plant with eight cascaded units, which is a state of the art data set in the literature (Conejo et al., 2002; Esmaeily et al., 2017). Although this data give much information about the units, topology, start-up costs, reservoir levels etc., they are for a massive system, which has the generation capacity of approximately double of the biggest hydro power plant in Turkey. Therefore, it is impossible to assume that this hydro power plant will work as a price taker in the Turkish market without affecting the market prices. For this reason, a modified version of these data are used and only the first two cascaded units are taken into account. The goal function, costs and the constraints are given to the model in General Algebraic Modeling System (GAMS) software. The PBUC problem is solved for the production amounts of each hour of the forecast days by MILP, which allows the hydro-power plant to self-schedule. Production amounts for each hour of the day are calculated for all the mentioned models: Actual prices, five different models and four hybrid models. This process is repeated for the predicted 168 days of 24 weeks. Then, the production amounts are multiplied by the ex-post actual prices to calculate the revenue. Lastly, costs are subtracted from the revenue and the profits are obtained.

4.2.2.1 Price based unit commitment according to mixed integer linear programming

In this study, we used a Mixed Integer Linear Programming (MILP) model adapted from Conejo et al. (2002) and Esmaeily et al. (2017) to solve the self-scheduling problem of the hydro GenCo. Conejo et al. (2002) represented a set of non-concave and non-linear performance curves showing the relationship between the reservoir head, the water discharged and the power output. They used piece-wise linearization to deal with these non-concavities and non-linearities of the performance curves, and proposed a mixed integer linear programming model. In this study, we use the same mathematical model by adding only one constraint related to maximum spillage amount adapted from Esmaeily et al. (2017). The formulation of the mathematical model can be found in Appendix C.

4.2.3 Financial effect of the forecast inaccuracy measures

Having these production amounts and the related costs gives us the opportunity of calculating the actual profits for each model. The accurate prediction of the actual prices should translate to maximum profit during the sale of electricity. The difference between the profit of the forecast model and profit of the ex-post actual price model is called profit loss (Delarue et al., 2010). For choosing the best performing model, Mean Absolute Error (MAE) or Mean Absolute Percentage Error (MAPE) are the most common methods. We prefer to use MAE instead of MAPE because of the reason that MAPE values are biased with the actual electricity price values, which are around zero. Previous literature (Mohammadi-Ivatloo et al., 2011; Mathaba et al., 2014) suggest that there is a discrepancy between the general forecast model decision methods and the profit loss models. Therefore, Economic Loss Index (ELI) and Price Forecast Disadvantage Index (PFDI) (Mohammadi-Ivatloo et al., 2011) are also used to measure the financial effect of forecast inaccuracy. These measures are described as follows:

$$\mathrm{ELI} = \frac{\mathrm{Profit_{actual}} - \mathrm{Profit_{forecast}}}{|\mathrm{Profit_{actual}}|} \tag{4.3}$$

$$\mathrm{PFDI} = \frac{\mathrm{Profit_{actual}} - \mathrm{Profit_{forecast}}}{\sum_{t=1}^{T} \mathrm{E_t}} \tag{4.4}$$

where, $\sum_{t=1}^{T} \mathrm{E_t}$ is the total energy sold in the market, in terms of MWh.

ELI demonstrates the profit loss as a percentage due to the inaccuracy of the forecast. Although it is not the case in our experiment, actual profit can be negative due to the ramp-up prices, constraints and limits. Thus, an absolute value of the actual profit is used in the denominator. Another point is that negative ELI is also possible due to unexpected higher profits of the forecast model than the ex-post actual prices case. PFDI is another financial effect measure, which calculates the profit loss per energy sold in terms of $/MWh. It must be mentioned that these models do not show the accuracy of the forecast, but the financial effect of forecast inaccuracy (Mohammadi-Ivatloo et al., 2011).

4.3 Results and Discussion

In this section, we report the profit loss comparisons in relation to electricity price forecast accuracy. In Section 4.3.1 we report profits obtained from all the methods, including ex-post actual prices: Profit losses, Economic Loss Index (ELI), Price Forecast Disadvantage Index (PFDI) and Mean Absolute Error (MAE) of the forecast methods for 24 weeks, two weeks from each month. Then, in Section 4.3.2, we compare the seasonal performance of the methods in terms of profit loss. In Section 4.3.3, we visually show the relationship between MAE and ELI for each hour or the day. We also illustrate the energy price profile and production schedule of the power plant for an exemplary day. This allows us to measure the inaccurate forecasts' financial effect on the hydro-based power plant. Finally, we evaluate the statistical significance in Section 4.3.4 and discuss the impact of the results.

4.3.1 Profit loss comparison

This section demonstrates profit, profit loss, ELI, PFDI and MAE results for the hydro-based GenCo's self-scheduling scheme according to various forecast methods, in addition to the ex-post day-ahead electricity prices. Table 4.2 gives the results as the total of the 24 weeks, six weeks from each season and best results are highlighted in the table. According to our results, ANN–LSTM is the best method in terms of financial effect measures. Scheduling the GenCo according to ANN–LSTM method would cause a profit loss of \$216410, 2.20% ELI and 1.1856 PFDI, compared to the ex-post actual prices scheduling. On the other hand, ANN is the best-performer by a small margin in terms of forecast performance measure MAE. This shows us that the forecast performance measures and the financial effect measures are not necessarily coherent with each other. Moreover, other hybrid methods ANN–LSTM–GRU and LSTM–GRU are the second and third best performing models, respectively, according to profit loss.

Table 4.2 : Results of the hydro-based Genco's self-scheduling according to various forecast methods for 24 weeks.

24 Weeks	Profit	Profit loss	ELI	PFDI	MAE
Actual	9815726	-	-	-	-
Naïve	9513169	302557	0.0308	1.6576	9.3066
SARIMA	9576815	238911	0.0243	1.3089	8.3289
ANN	9594006	221721	0.0226	1.2147	6.3774
LSTM	9589966	225760	0.0230	1.2369	6.5489
GRU	9584191	231536	0.0236	1.2685	6.4586
LSTM-GRU	9596195	219532	0.0224	1.2027	6.4472
ANN-GRU	9591269	224457	0.0229	1.2297	6.3929
ANN-LSTM	9599316	216410	0.0220	1.1856	6.3851
ANN-LSTM-GRU	9597754	217972	0.0222	1.194	6.4018

4.3.2 Seasonal performance comparison

Figure 4.2 demonstrates the profit loss of the hydro-based GenCo's self-scheduling scheme according to various forecast methods divided into seasons. We report the average results for six weeks of each season. First of all, we investigate the variation of the profit loss levels according to the seasons of the year. In the examined period, profit loss levels of winter and autumn are relatively small. On the other hand, profit losses are very high, especially spring, at the level of \$100,000. Although ANN–LSTM is the best model only in winter, we observe the stable performance of the hybrid models. On the contrary, the performance of the individual models are not very stable.

Relatively good performance of LSTM is shadowed by the poor performance in winter.

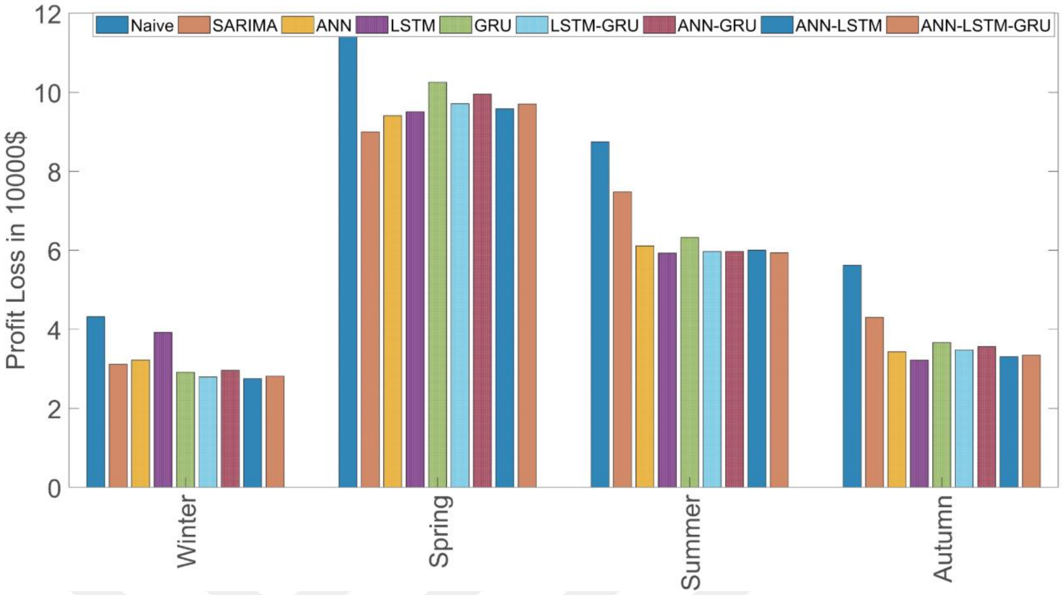


Figure 4.2 : Seasonal profit loss results.

4.3.3 Energy price profile and production scheduling

Figure 4.2 demonstrates the profit loss of the hydro-based GenCo's self-scheduling scheme according to various forecast methods divided into seasons. We report the average results for six weeks of each season. First of all, we investigate the variation of the profit loss levels according to the seasons of the year. In the examined period, profit loss levels of winter and autumn are relatively small. On the other hand, profit losses are very high, especially spring, at the level of $100,000. Although ANN–LSTM is the best model only in winter, we observe the stable performance of the hybrid models. On the contrary, the performance of the individual models are not very stable. Relatively good performance of LSTM is shadowed by the poor performance in winter.

Figure 4.3 reveals the relationship between MAE and ELI according to the LSTM model's forecasts and the related scheduling of the GenCo. It must be mentioned that these values are from the 24 weeks of 2016, and, therefore, they are not continuous values. However, they give information about the co-movement of MAE and ELI. Although we observe the co-movement of MAE and ELI in general, on the last days, MAE levels do not follow the decreasing trend in the ELI numbers. This is further evidence for a difference between the forecast evaluation measures, such as MAE, MAPE, RMSE and the financial effect measures, such as profit loss, ELI and PFDI.

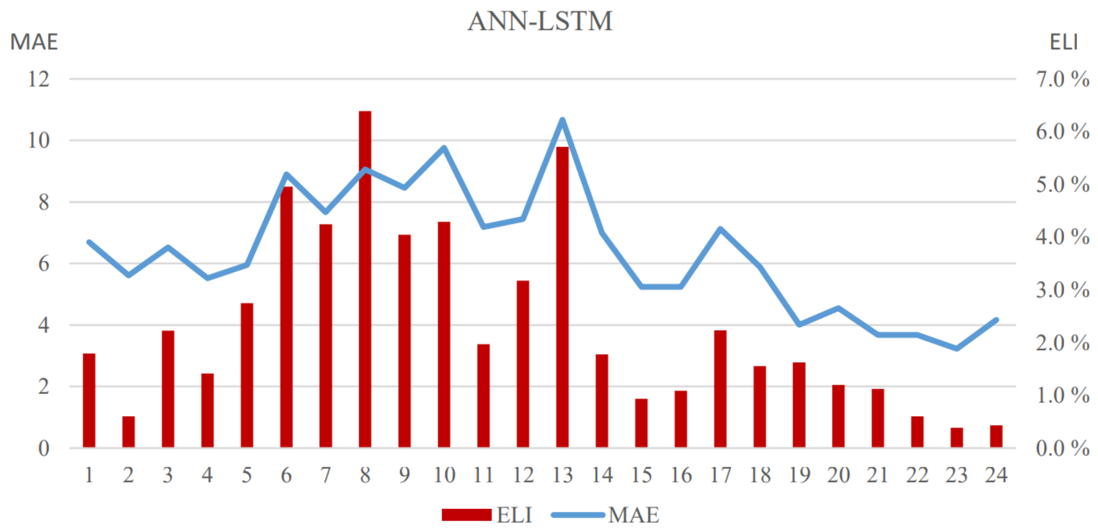


Figure 4.3 : MAE and ELI results of the ANN-LSTM model for 24 forecast weeks.

Figures 4.4 and 4.5 are the graphs for the scheduling and profits of the GenCo on a randomly chosen day, 14 November. Figure 4.4 demonstrates the relative success of the ANN– LSTM model in the scheduling of the hydro power plant by showing the energy prices and power output of both units. It is observed that the power plant does not work in the lower price zone during early morning hours, and both units work in the maximum capacity in the peak-price hours. Due to the ramp-up costs, first, the greater unit, Plant 2, starts to operate, and in the higher demand moments Plant 1 gets activated as well. As a negative point, we can mention the 1 h delay in the production at 9. Figure 4.5, which shows the profits according to ex-post day-ahead scheduling and ANN-LSTM scheduling comparatively, supports this point by showing the profit of the ex-post day-ahead scheduling compared to almost no profit of ANN-LSTM scheduling at 9. Furthermore, ANN–LSTM scheduling produces more electricity and makes a profit in relatively low price levels at 21 and 22. On this day, scheduling according to ex-post day-ahead price forecasts resulted in $57223.38 profit compared to the LSTM model scheduling, which caused $56097.34 profit. It means $1126.04 profit loss for the ANN–LSTM model in one day for a relatively small power plant, which has only approximately 1089 MW production on this day.

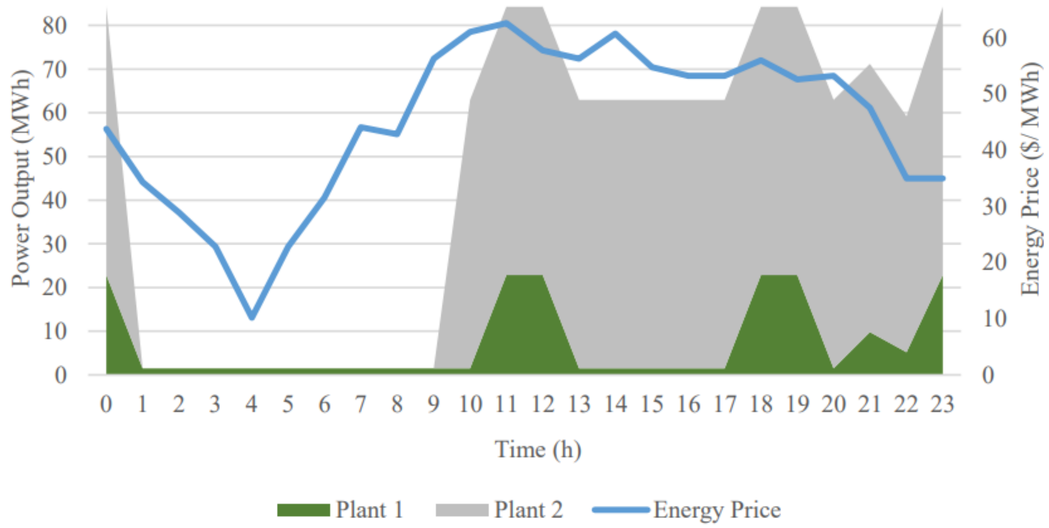
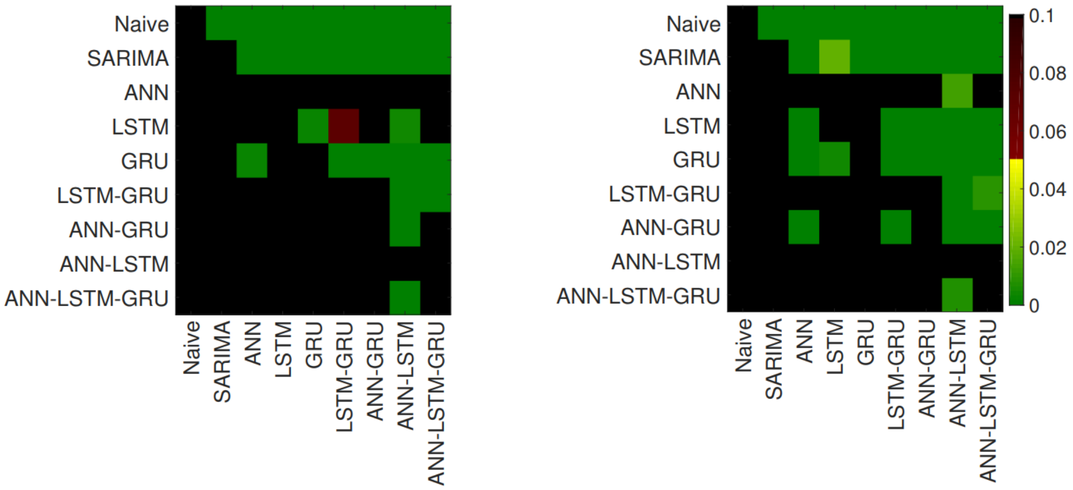


Figure 4.4 : Energy price profile and production schedule of Plant 1 and 2 based on ANN-LSTM method on 14 November.

Figure 4.5 : Profit of ANN-LSTM model and ex-post day-ahead prices according to the hours of the day on 14 November.

4.3.4 Diebold-Mariano tests

The Table 4.2 can be used to provide a ranking of the various methods, however no statistically significant conclusions can be drawn on the performance of the forecasts. To showcase the statistical significance of the performance difference between all model variations, we use a Diebold-Mariano test (Diebold and Mariano, 1995), which takes the correlation structure into account. In Figure 4.6a, we show the p-values for the Diebold-Mariano tests between hybrid methods, neural networks-based methods and the statistical methods for 4032 h of the 24 weeks we investigated. In Figure 4.6b, we show the statistical significance test in between the MAE values and profit losses for each method to illustrate the difference between the MAE and profit loss.



(a) Diebold-Mariano for Electricity Price Forecast MAE **(b)** Diebold-Mariano for Profit loss

Figure 4.6 : Results of the Diebold-Mariano tests defined by the MAE values and profit loss differential series in between different models. The figure indicates the statistical significance (green) for which the forecasts of a model on the X-axis are significiantly better than those of a model on the Y-axis. The statistical significance on the difference between the MAE values does not translate fully into a difference in profit loss.

Figure 4.6a demonstrates the successful performance of the ANN model and the ANN–LSTM model, in terms of MAE. However, there is not a significant difference between both models and it is not possible to choose one model over another. However, according to the profit loss values, ANN-LSTM model significantly outperforms all the other models, including ANN. It is evident that forecast performance measures, such as MAE and the financial effect measures, such as profit loss give different results, in statistically significant terms.

4.4 Conclusions

In this paper, we propose self-scheduling schemes according to nine different forecast methods by using Mixed Integer Linear Programming for a relatively small hydro-based GenCo with approximately 1 GWh production per day. Nine forecast schemes include a benchmark naive method, a statistical triple SARIMA model, machine learning ANN, LSTM and GRU models, in addition to the hybrid methods by the combination of the machine learning models. Additionally, we utilized ex-post actual prices as the perfect prices to schedule the power plant for optimum performance. This allowed us to calculate the profit loss, ELI, and PFDI as the financial effect measures. We also compare the relationship between the financial effect measures and the forecast performance evaluation measure, namely MAE. This is the first paper which explores the use of hybrid methods from the financial effect of the forecast inaccuracies point of view. According to our results, ANN–LSTM model is the best performing one in statistically significant terms. Moreover, other hybrid methods ANN–LSTM–GRU and LSTM–GRU are the second and third best models, respectively. On the other hand, as individual models, ANN performs relatively well, but especially naive method, SARIMA and GRU would cause higher losses to the generation company. As the literature (Chaabane, 2014; Hong and Wu, 2012) suggest, the usage of hybrid works and the combination of the price forecasts (Bordignon et al., 2013; Nowotarski et al., 2016), we also found out the statistically significant superiority of the hybrid ANN–LSTM method. Our findings are also in line with the works of Lago et al. (2018a) and Kuo and Huang (2018), which advocate the use of hybrid methods in deep learning electricity price forecasting applications. Another finding of this paper is that it supports the conflict between the forecast performance evaluation measures, such as MAE, and the inaccurate electricity price forecasts' financial effect measures, such as profit loss, ELI, and PFDI. Even though the ANN is the best method in terms of MAE, and there is not a statistically significant difference between the ANN and ANN–LSTM methods according to the forecast errors; in terms of financial effect measures, ANN–LSTM is better than all the other methods, including ANN. Although the general trend is the same for MAE and financial effect measures, ELI in Figure 4.3, there are some conflicting weeks, which cause this discrepancy. Our results support the findings of Zareipour et al. (2010), Mohammadi-Ivatloo et al. (2011) and Mathaba et al. (2014) in the Canada and U.S. markets on this conflict; in a smaller power plant, in a different market and with various new forecast methods from Ugurlu et al. (2018a) and the combinations of these forecasts. As Figure 4.2 illustrates, the seasonality influences financial effect measures. Even though LSTM is successful in spring, summer and autumn; the poor performance in winter affects the success and reliability of this method. On the other hand, hybrid methods, especially ANN-LSTM, give stable and reliable results. For instance, SARIMA model shows a very good performance in spring. It could be helpful to change the models according to time periods. This is an avenue of improvement for our work.

In this paper, we focus on the relation between electricity price forecasting accuracy and profit maximization. For a fair comparison between different methods, we do not focus on the variable selection and just compare the methods with the same input variables. We compared five forecast techniques and four combinations of these techniques. In all the models we used the same endogenous variables; 1st, 24th, 48th, and 168th lags of the prices to compare the effects of the various models. In Ugurlu et al. (2018a), it is observed that machine learning models outperformed the triple SARIMA and the other statistical models. In our current experimental setup, statistical methods, such as SARIMA, and naive method do not perform well. However, by using more complex statistical methods (Ziel and Weron, 2018), different results could be obtained. We appreciate the hybrid models (Chaabane, 2014; Hong and Wu, 2012; Zhang, 2003), dimension reduction techniques (Ziel, 2016; Ludwig et al., 2015) and the automated variable selection works (Uniejewski et al., 2016), which utilize a variety of variables and choose the best ones. An obvious further research topic is using the available input variables to the the fullest and evaluate the relation between electricity price forecast accuracy and profit maximization. This paper opens the path for further research on the relation between electricity price prediction accuracy and profit maximization. Firstly, hybrid models show impressive potential and can be instrumental techniques for profit maximization. Secondly, the best performing models vary according to the test period. It could give the opportunity of using different models in different periods. Further research on this issue has the potential to minimize the profit loss of the suppliers. Thirdly, applying a non-linear programming technique for the PBUC could increase the scheduling performance. Lastly, solving the same problem for other types of suppliers, more markets and different time periods would check the generalizability, robustness and accuracy of these results.

5. CONCLUSIONS

With the liberalization of the electricity markets, electricity price forecasting has become an essential task for all the participants of the electricity markets. Due to this requirement, many electricity price forecasting models are developed by using various methods from different areas. This thesis focuses on choosing the best performing model among the statistical and machine learning methods in the first two papers (Ugurlu et al., 2018a, Ugurlu et al., 2018b). The third paper (Ugurlu et al., 2018c), which is included as the fourth chapter mainly discusses the financial effect of the inaccurate electricity price forecasts on a generation company by using the applied models in the previous papers (Ugurlu et al., 2018a, Ugurlu et al., 2018b). This thesis takes the Turkish day-ahead market as the research field due to the lack of research in this emerging electricity market. Due to its location, high growth rate and young population; Turkey has an increasing electricity demand, which is met by the installed capacity of over 80 GW per hour in 2017.

Electricity price forecasting has a paramount effect on the market participants from the marketmaker to the end users. Although short-term electricity price forecasts seem to affect only the supply side generation companies and the demand side companies, which give bids in the day-ahead market; in reality, accurate electricity price forecasts give the opportunity of deciding the prices with less consumer or producer surplus to the marketmaker. As we take the electricity market as an oligopolistic medium, most of the time higher prices occur due to producer surplus. Therefore, having the accurate electricity price forecasts would decrease the consumer prices and let the citizens pay less for the electricity in longer term.

There are many different ways of electricity price forecasting and there is still a challenge between the electricity price models. Hybrid models, which combine various methods, are frequently used to forecast the electricity prices. As Aggarwal et al. (2009) suggested, there is still a competition between different methods and we can not mention that one model is better than the other one consistently in different time periods and markets. However, by the improvements in the technology and the domination of the artificial intelligence and the machine learning in various fields nowadays, the usage of machine learning related models (Lago et al., 2018a; Kuo and Huang, 2018) started to give better results than the conventional ones. Another important topic is the usage of the pre-methods in the electricity price forecasting. Due to the non-stationary nature of the electricity prices and the seasonality in many different frequencies, electricity prices require pre-treatments before the analysis, which increase the success of the electricity price forecasting models significantly (Uniejewski et al., 2017; Ziel, 2016). As mentioned, electricity price forecasts affect all the participants of the electricity markets. But, the most effected one, by the short-term electricity price forecasts' accuracy, is the hydro-based generation companies. The main reason of it is that they can make their production strategy according to the forecasts, which would cause them the smallest loss of profit.

This thesis comprises three papers, which are independent, but strongly connected to each other, on the electricity price forecasting. The first paper proposes the usage of the factorial ANOVA model as a pre-whitening method to the electricity price series and suggests that it will turn the non-stationary series to the stationary ones, which is a requirement for applying the statistical methods. As discussed in the literature (Diaz and Planas, 2016, Keles et al., 2012), electricity prices can take zeros and negative values. Thus, taking the log returns, which is the state-of-the-art process to have the stationary series, is not applicable for the electricity prices. Therefore, factorial ANOVA is suggested as a transformation method to the electricity price series. Moreover, this first paper of the thesis compares the statistical electricity price forecasting methods such as SARIMA, Markov regime-switching and SETAR, in addition to the benchmark naïve method and the AR(24) model. Although best performing models vary according to the chosen time of the year, in a relatively robust way, SARIMA is chosen as the best performing statistical method in the Turkish day-ahead market. The main drawback of this paper is that it only uses the lagged price series as independent variables. In the second paper, temperature, forecast demand/supply, 24^{th} lag of the realized demand/supply and the 24^{th} lag of the balancing market prices are added as the exogenous variables. But the main addition in this paper is the neural networks models, especially deep recurrent neural networks, LSTM and GRU. This paper proposes a new framework to the electricity price forecasting. Deep learning methods are used in speech recognition, image processing and natural language processing successfully (Greff et al., 2017), but this is the first paper[29] which performs the deep learning methods in the electricity price forecasting literature. In a similar sense, it is also the first work utilizing recurrent neural networks, which are tailor-made for the time series data. Deep recurrent neural networks outperform the shallow recurrent neural networks, artificial neural networks, convolutional neural networks and statistical time series methods significantly in the Turkish day-ahead market data. GRU is the better performing one compared to the LSTM among the recurrent neural networks models. Another contribution of this paper is that adding the meaningful exogenous variables also increase the performance of the electricity price forecasting models. However, adding less important variables do not cause a significant improvement. In the meantime, endogenous variables, 1st, 24th, 48th and 168th, which are the ones used in the first paper, are the most explanatory variables. Third paper discusses the financial effect of the inaccurate electricity price forecasts on the hydro-based generation company. As, a hydro power plant can organize it's generation schedule according to the electricity price forecasts, accurate electricity price forecasts have a major financial effect on the hydro-based generation companies. In this sense, the main problem is to decide the best electricity price forecasting model for the generation company. In general, the best performing models are chosen according to the performance evaluation criteria such as MAE, MAPE or RMSE. However, best model, in terms of the financial effect, is not necessarily consistent with the performance evaluation criteria (Mohammadi-Ivatloo et al., 2011; Mathaba et al., 2014). In the Turkish day-ahead market, ANN-LSTM model, which is not the best according to MAE, is chosen as the best model in terms of loss of profit. From the generation company point of view, forecast models must be evaluated according to the financial effect performance measures. This paper also evaluates the combinations of the electricity price forecasts as hybrid models, in terms of financial effect. In most of the periods, hybrid models are the best performing ones. Therefore, it would be a good idea to use different models in different periods by giving chance to the hybrid models, as well.

This thesis is the first study in the Turkish electricity market with such a broad view to the electricity price forecasting. Although works of Ozyildirim and Beyazit (2014),

[29] Lago et al. (2018a) and Kuo and Huang (2018) are the simultanous works on the same topic.

Kolmek and Navruz (2015), Hayfavi and Talasli (2014) are acknowledged, this thesis gives a wider perspective to the reader about the Turkish electricity market and the electricity price forecasting models, which can be used in the Turkish electricity market. Moreover, it mentions the financial effect of the inaccurate electricity price forecasts on a generation company. Furthermore, it proposes factorial ANOVA as a pre-treatment method and deep recurrent neural networks as the electricity price forecasting models. According to the results, deep GRU is the best method in the forecasting of the Turkish day-ahead electricity prices. Another contribution of the thesis is that the generation companies must select the electricity price forecasting method according to the financial effect measures for the electricity generation scheduling.

Electricity price forecasting is an ever improving research area. There are many topics, which can be discussed in terms of further research. First and foremost, hybrid models, both combination of the different models' forecasts and the combined usage of some forecast methods give relatively good results in the literature (Chaabane, 2014; Bordignon et al., 2013). Therefore, especially combination of statistical and machine learning methods would be tried. In this thesis we used limited number of variables; gas and oil prices and exchange rates could help to forecast more accurately. Accessing the hourly data for a long term could be challenging in this aspect. It is also much better automatizing the variable selection process by using the dimension reduction techniques such as Lasso or Principal component analysis (Uniejewski et al., 2016).

As mentioned, electricity markets differ from country to country according to many features such as location, technology usage, temperature, share of renewables etc. Therefore, all the markets have different conditions and unique features. A model successful in a market is not necessarily successful in another one. Even a model, which performs well now, could be outdated in a short while, especially in the emerging economies. Furthermore, day-ahead markets, balancing markets and the intraday markets have different characteristics and variables. Therefore, working on different markets in different time periods is necessary to forecast the electricity prices accurately.

In Turkey, besides balancing market, intraday market also has an increasing volume and transaction numbers. High frequency research can be done in the Turkish intraday electricity market by using the relevant variables. In the intraday market, 1 hour ahead forecasts are required and more recent data can be used. Another advantage for deep learning models is that there are approximately 500000 observations in the last two years in the Turkish intraday market. As it is known, higher number of observations allow the machine learning models to learn better and forecast more accurately. Additionally, financial instruments such as forwards, futures and options can be priced according to the intraday and day-ahead prices. This will allow the companies to hedge their positions and protect themselves from the sharp price increases and decreases in the electricity markets. The developing energy derivatives market of Turkey will also need pricing models. Pioneer work of Talasli (2012) was faced the lack of data in the Turkish intraday market. Nowadays, with the increasing number of observations, energy derivatives pricing would be a very intriguing topic.

REFERENCES

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

Ahmadi, A., Jamshid A., Heidar A. S. & Rabiee A. (2012). Mixed integer programming of multiobjective hydro-thermal self scheduling. *Applied Soft Computing*, *12*, 2137-2146.

Alamaniotis, M., Bargiotas, D., Bourbakis, N. G. & Tsoukalas, L. H. (2015a). Genetic Optimal Regression of Relevance Vector Machines for Electricity Pricing Signal Forecasting in Smart Grids. *IEEE Transactions on Smart Grid*, *6* (6), 2997–3005.

Alamaniotis, M., Bourbakis, N. & Tsoukalas, L. H. (2015b). Very-short term forecasting of electricity price signals using a Pareto composition of kernel machines in smart power systems. *In Proceedings of the 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, FL, USA, 14–16 December 2015.

Amjady, N. (2006). Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Transactions on Power Systems*, *21*, 887–896.

Avci-Surucu, E., Aydogan A. K. & Akgul, D. (2016). Bidding structure, market efficiency and persistence in a multi-time tariff setting. *Energy Economics*, *54*, 77–87.

Bao, W., Yue, J. & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE*, *12*, e0180944.

Bisanovic, S., Mensur H. & Dlakic M. (2008). Hydrothermal self-scheduling problem in a day-ahead electricity market. *Electric Power Systems Research*, *78*, 1579-1596.

Bordignon, S., Bunn, D. W., Lisi, F. & Nan, F. (2013). Combining day-ahead forecasts for British electricity prices. *Energy Economics*, *35*, 88–103.

Box, G. E. P. & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco, CA: Holdan-Day.

Bublitz, A., Keles, D. & Fichtner, F. (2017). An analysis of the decline of electricity spot prices in Europe: Who is to blame? *Energy Policy*, *107*, 323–336.

Bunn, D.W. (2000). Forecasting loads and prices in competitive power markets. *Proceedings of IEEE*, *88*, 163–169.

Bunn, D. W. (2004). *Modelling prices in competitive electricity markets*. England: John Wiley&Sons.

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.

Carmona, R. & Coulon, M. (2014). A survey of commodity markets and structural models for electricity prices. *In Quantitative Energy Finance*, Springer: Berlin, Germany, 41–83.

Carmona, R., Coulon, M. & Schwarz, D. (2013). Electricity price modeling and asset valuation: A multi-fuel structural approach. *Mathematics and Financial Economics*, *7* (2), 167–202.

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

Catalão, J. P. D. S., Mariano, S. J. P. S., Mendes, V. & Ferreira, L. (2007). Short-term electricity prices forecasting in a competitive market: A neural network approach. *Electric Power Systems Research*, *77* (10), 1297–1304.

Chaâbane, N. (2014). A hybrid ARFIMA and neural network model for electricity price prediction. *International Journal of Electrical Power & Energy Systems*, *55*, 187–194.

Cho, K., Van Merriënboer, B., Bahdanau, D. & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. *arXiv*, arXiv:1409.1259.

Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv*, arXiv:1412.3555.

Conejo, A. J., Contreras, J., Espínola, R. & Plazas, M. A. (2005a). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting 21* (3), 435–462.

Conejo, A. J., Nogales, F. J. & Arroyo, J. M. (2002). Price-taker bidding strategy under price uncertainty. *IEEE Transactions on Power Systems*, *17*, 1081-1088.

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

Contreras, J., Espinola, R., Nogales, F. J. & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, *18* (3), 1014–1020.

Cuaresma, J. C., Hlouskova, J., Kossmeier, S. & Obersteiner, M. (2004). Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*, *77* (1), 87–106.

Delarue, E., Bosch, P. V. D. & D'haeseleer, W. (2010). Effect of the accuracy of price forecasting on profit in a price based unit commitment. *Electric Power Systems Research*, *80*, 1306-1313.

de Livera, A. M., Hyndman, R. J. & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, *106* (496), 1513–15227.

Díaz, G. & Planas, E. (2016). A note on the normalization of Spanish electricity spot prices. *IEEE Transactions on Power Systems*, *31*, 2499–2500.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–263.

Doostmohammadi, A., Amjady, N. & Zareipour, H. (2017). Day-Ahead Financial Loss/Gain Modeling and Prediction for a Generation Company. *IEEE Transactions on Power Systems*, *32*, 3360-3372.

Dorffner, G. (1996). Neural Networks for Time Series Processing. *Neural Network World*, *6*, 447–468.

Eichler, M. & Tuerk, D. (2013). Fitting semiparametric Markov regime-switching models to electricity spot prices. *Energy Economics*, *36*, 614–624.

Energy Exchange Istanbul. (2016). Annual Report. 01 January 2016 – 31 December 2016, 2016.

EPDK. (2017). Republic of Turkey Energy Market Regulatory. Electricity Market Balancing and Settlement Regulation. Official Gazette No. 27200, Turkey.

EPDK. (2018). Republic of Turkey Energy Market Regulatory. Retrieved from: http://www.epdk.org.tr/TR/Dokumanlar/Elektrik/YayinlarRaporlar/ElektrikPiyasasiGelisimRaporu. Last access on 23 February 2018.

EPIAS. (2018). EXIST Transparency Platform. Retrieved from: https://seffaflik.epias.com.tr/transparency. Last access on 21 May 2018.

Esmaeily, A., Ahmadi, A., Raeisi F., Ahmadi, M. R., Nezhad, A. E. & Janghorbani, M. (2017). Evaluating the effectiveness of mixed-integer linear programming for day-ahead hydro-thermal self-scheduling considering price uncertainty and forced outage rate. *Energy*, *122*, 182-193.

Fanone, E., Gamba, A. & Prokopczuk, M. (2013). The case of negative day-ahead electricity prices. *Energy* Economics, *35*, 22–34.

Filipovic, D., Larsson, M. & Ware, T. (2017). Polynomial processes for power prices. *arXiv*, arXiv:1710.10293.

Füss, R., Mahringer, S. & Prokopczuk, M. (2015). Electricity derivatives pricing with forward-looking information. *Journal of Economic Dynamics and Control*, *58*, 34–57.

Geman, H. & Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *Journal of Business*, *79*, 1225–1261.

Gensler, A., Henze, J., Sick, B. & Raabe, N. (2016). Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks. *In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Budapest, Hungary, 9–12 October 2016, 2858–2865.

Glorot, X., Bordes, A. & Bengio, Y. (2011). Deep sparse rectifier neural networks. *In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA, 11–13 April 2011, 315–323.

González, A. M., Roque, A. S. & García-González, J. (2005). Modeling and forecasting electricity prices with input/output hidden Markov models. *IEEE Transactions on Power Systems*, *20* (1), 13–24.

Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J. & Vahid-Araghi, F. (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, *191* (1), 207–222.

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R. & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks & Learning Systems*, *28*, 2222–2232.

Guo,W. & Zhao, Z. (2017). A Novel Hybrid BND-FOA-LSSVM Model for Electricity Price Forecasting. *Information*, *8*, 120.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, *57*, 357–384.

Hamilton, J. D. (1994). *Time series analysis*. Vol. 2, Princeton, NJ: Princeton University Press.

Hayfavi, A. & Talasli, I. (2014). Stochastic multifactor modeling of spot electricity prices. *Journal of Computational and Applied Mathematics*, *259* (10), 434–442.

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780.

Hong, T. (2015). Crystal ball lessons in predictive analytics. *EnergyBiz Mag*, 35–37.

Hong, Y.Y. & Wu, C.P. (2012). Day-ahead electricity price forecasting using a hybrid principal component analysis network. *Energies*, *5*, 4711–4725.

Hosein, S. & Hosein, P. (2017). Load forecasting using deep neural networks. *In Proceedings of the Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*,Washington, DC, USA, 23–26 April 2017, 1–5.

Howell, D. C. (2013). *Statictical methods for psychology*, 8th ed, Belmont, CA: Wadsworth/Cengage Learning.

Howison, S. & Coulon, M. (2009). Stochastic behavior of the electricity bid stack: From fundamental drivers to power prices. *Journal of Energy Markets*, *2*, 29–69.

Hryshchuk, A. & Lessmann, S. (2018). Deregulated Day-Ahead Electricity Markets in Southeast Europe: Price Forecasting and Comparative Structural Analysis. *SSRN*.

Hwang, S.J., Mehta, R. & Singh, V. (2018). Sampling-free Uncertainty Estimation in Gated Recurrent Units with Exponential Families. *arXiv*, arXiv:1804.07351.

IEEE. (2018). IEEE 118-bus test system. Retrieved from: http://motor.ece.iit.edu/data. Last access on 21 May 2018.

Implementing a GRU/LSTM RNN with Python and Theano. (2018). Retrieved from: http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4-implementing-a-grulstm-rnn-with-python-and-theano/. Last access on 6 April 2018.

Iwata, T. & Ghahramani, Z. (2017). Improving Output Uncertainty Estimation and Generalization in Deep Learning via Neural Network Gaussian Processes. *arXiv*, arXiv:1707.05922.

Janczura, J., Trück, S., Weron, R. & Wolff, R.C. (2013). Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics*, *38*, 96–110.

Janczura, J. & Weron, R. (2010). An empirical comparison of alternate regime-switching models for electricity spot prices. *Energy Economics*, *32* (5), 1059–1073.

Karakatsani, N. V. & Bunn, D. W. (2008). Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, *24* (4), 764–785.

Karami, M., Shayanfar, H. A., Aghaei, J. & Ahmadi, A. (2013). Mixed integer programming of security-constrained daily hydrothermal generation scheduling (SCDHGS). *Scientia Iranica. Transaction D, Computer Science & Engineering, Electrical Engineering*, *20*, 2036-2050.

Keles, D., Genoese, M., Möst, D. & Fichtner, W. (2012). Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy Economics*, *34* (4), 1012–1032.

Keles, D., Scelle, J., Paraschiv, F. & Fichtner, W. (2016). Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Applied Energy*, *162*, 218–230.

Kolmek, M. A. & Navruz, I. (2015). Forecasting the day-ahead price in electricity balancing and settlement market of Turkey by using artificial neural networks. *Turkish Journal of Electrical Engineering & Computer Sciences*, *23* (3), 841–852.

Kristiansen, T. (2012). Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy*, *49*, 328–332.

Krizhevsky, A., Sutskever, I. & Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105.

Kuo, P.H. & Huang, C.J. (2018). A High Precision Artificial Neural Networks Model for Short-Term Energy Load Forecasting. *Sustainability*, *11*, 213.

Lago, J., Ridder, F. D. & Schutter, B. D. (2018a). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, *221*, 386–405.

Lago, J., Ridder, F. D., Vrancx, P. & Schutter, B. D. (2018b). Forecasting day-ahead electricity prices in Europe: The importance of considering market integration. *Applied Energy, 211*, 890–903.

LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceeding of IEEE, *86* (11), 2278–2324.

Li, T. & Shahidehpour, M. (2005). Price-based unit commitment: A case of Lagrangian relaxation versus mixed integer programming. *IEEE Transactions on Power Systems*, *20*, 2015-2025.

Ludwig, N., Feuerriegel, S. & Neumann, D. (2015). Putting Big Data analytics to work: Feature selection for forecasting electricity prices using the LASSO and random forests. *Journal of Decision Systems, 24*, 19–36.

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

Mandal, P., Senjyu, T., Uezato, K. & Funabashi, T. (2005). Several-hours-ahead electricity price and load forecasting using neural networks. *In Power Engineering Society General Meeting*, 2146–53. Piscataway, NJ: IEEE.

Mathaba, T., Xia, X. & Zhang, J. (2014). Analysing the economic benefit of electricity price forecast in industrial load scheduling. *Electric Power Systems Research*, *116*, 158-165.

Mohammadi-Ivatloo, B., Zareipour, H., Ehsan, M. & Amjady, N. (2011). Economic impact of price forecasting inaccuracies on self-scheduling of generation companies. *Electric Power Systems Research*, *81*, 617-624.

Negahdary, E. & Ware, T. (2016). Modeling Alberta power prices through fundamentals. *Journal of Energy Markets*, *9* (4), 49–70.

Neupane, B., Woon, W. & Aung, Z. (2017). Ensemble Prediction Model with Expert Selection for Electricity Price Forecasting. *Energies*, *10*, 77.

Nogales, F. J., Contreras, J., Conejo, A. J. & Espinola, R. (2002). Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems*, *17* (2), 342–348.

Nowotarski, J., Raviv, E., Trück, S. & Weron, R. (2014). An empirical comparison of alternative schemes for combining electricity spot price forecasts. *Energy Economics*, *46*, 395–412.

Oksuz, I., Ruijsink, B., Anton, E., Sinclair, M., Rueckert, D., Schnabel, J. & King, A. (2018). Automatic Left Ventricular Outflow Tract Classification For Accurate CardiacMR Planning. *In Proceedings of the 2018 IEEE International Symposium on Biomedical Imaging (ISBI),*Washington, DC, USA, 4–7 April 2018.

Ozguner, E., Tor, O. B. & Guven, A. N. (2017). Probabilistic day-ahead system marginal price forecasting with ANN for the Turkish electricity market. *Turkish Journal of Electrical Engineering & Computer Sciences*, *25*, 4923–4935.

Ozkan, H. & Yazgan, M. E. (2015). Is forecasting inflation easier under inflation targeting? *Empirical Economics*, *48* (2), 609–626.

Ozozen, A., Kayakutlu, G., Ketterer, M. & Kayalica, O. (2016). A combined seasonal ARIMA and ANN model for improved results in electricity spot price forecasting: Case study in Turkey. *In Management of Engineering and Technology (PICMET), 2016 Portland International Conference*, 2681–2690, IEEE.

Ozyildirim, C. & Beyazit, M. F. (2014). Forecasting and Modelling of Electricity Prices by Radial Basis Functions: Turkish Electricity Market Experiment. *Iktisat Isletme ve Finans*, *29* (344), 31–54.

Panapakidis, I. P. & Dagoumas, A. S. (2016). Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, *172*, 132–151.

Pindoriya, N., Singh, S. & Singh, S. (2008). An Adaptive Wavelet Neural Network-Based Energy Price Forecasting in Electricity Markets. *IEEE Transactions on Power Systems*, *23*, 1423–1432.

Raviv, E., Bouwman, K. E., van Dijk, D. (2013). Forecasting day-ahead electricity prices: Utilizing hourly prices. *Energy Economics*, *50*, 227–239.

Shafie-khah, M. & Catalão, J.P. (2015). A stochastic multi-layer agent-based model to study electricity market participants behavior. *IEEE Transactions on Power Systems*, *30*, 867–881.

Shafie-Khah, M., Moghaddam, M. P. & Sheikh-El-Eslami, M. K. (2011). Price forecasting of day-ahead electricity markets using a hybrid forecast method. *Energy Conversion and Management*, *52* (5), 2165–69.

Shahidehpour, M. Yamin, H. & Li, Z. (2002). *Market Operations in Electric Power Systems*. IEEE: New York, NY, USA.

Shrivastava, N. A. & Panigrahi, B. K. (2014). A hybrid wavelet-ELM based short term price forecasting for electricity markets. *International Journal of Electrical Power & Energy Systems*, *55*, 41–50.

Sunaryo, S., Suhartono, S. & Alfonsus, J. E. (2011). Double seasonal recurrent neural networks for forecasting short term electricity load demand in Indonesia. *In Recurrent neural networks for temporal data processing*, ed. H. Cardot. InTech.

Szkuta, B., Sanabria, L.A. & Dillon, T.S. (1999). Electricity price short-term forecasting using artificial neural networks. *IEEE Transactions on Power Systems*, *14*, 851–857.

Talasli, I. (2012). Stochastic Modeling of Electricity Market. Doctoral Dissertation Thesis. Middle East Technical University.

Taylor, J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, *204*, 139–152.

Taysi, Z. C., Biricik, G. & Bozkurt, O. (2015). Evaluation of price forecast systems for Turkish Electric Market. *In Signal Processing and Communications Applications Conference (SIU)*, 620–623. 23th IEEE.

Tsay, R. S. (2005). *Analysis of financial time series*. Vol. 543, Hoboken, NJ: John Wiley & Sons.

Ugurlu, U., Oksuz, I. & Tas, O. (2018a). Electricity Price Forecasting Using Recurrent Neural Networks. *Energies*, *11* (5), 1255.

Ugurlu, U., Tas, O. & Gunduz, U. (2018b). Performance of Electricity Price Forecasting Models: Evidence from Turkey. *Emerging Markets Finance and Trade*, *54* (8), 1720-1739.

Ugurlu, U., Tas, O., Kaya, A. & Oksuz, I. (2018c). The Financial Effect of the Electricity Price Forecasts' Inaccuracy on a Hydro-based Generation Company. *Energies*, *11* (8), 2093.

Uniejewski, B., Nowotarski, J. & Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, *9* (8), 621.

Uniejewski, B., Weron, R. & Ziel, F. (2017). Variance stabilizing transformations for electricity spot price forecasting. *IEEE Transactions on Power Systems*, *33*(2), 2219-2229.

Url-1 <https://seffaflik.epias.com.tr/transparency/>, date retrieved 09.02.2017.

Vanishing Gradients & LSTMs. (2018). Retrieved from: http://harinisuresh.com/2016/10/09/lstms/. Last access on 10 April 2018.

Wasserman, P. D. & Schwartz, T. (1988). Neural networks. II. What are they and why is everybody so interested in them now? *IEEE Expert*, *3*, 10–15.

Weron, R. (2006). *Modeling and forecasting electricity loads and prices: A statistical approach*. Vol. 403, Chichester, UK: John Wiley & Sons.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, *30* (4), 1030–1081.

Weron, R. & Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, *24* (4), 744–763.

Xiaoyun, Q., Xiaoning, K., Chao, Z., Shuai, J. & Xiuda, M. (2016). Short-Term Prediction of Wind Power Based on Deep Long Short-Term Memory. *In Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Conference*, Xi'an, China, 25–28 October 2016.

Yamin, H. Y. & Shahidehpour, S. M. (2004). Self-scheduling and energy bidding in competitive electricity markets. *Electric Power Systems Research*, *71*, 203-209.

Yin,W., Kann, K., Yu, M. & Schütze, H. (2017). Comparative study of cnn and rnn for natural language processing. *arXiv*, arXiv:1702.01923.

Zareipour, H., Canizares, C. A. & Bhattacharya, K. (2010). Economic impact of electricity market price forecasting errors: a demand-side analysis. *IEEE Transactions on Power Systems*, *25*, 254-262.

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

Zhang, J. & Cheng, C. (2008). Day-ahead electricity price forecasting using artificial intelligence. *In Proceedings of the Electric Power Conference*, Vancouver, BC, Canada, 6–7 October 2008, 1–5.

Zhao, J. H., Dong, Z. Y., Xu, Z. & Wong, K. P. (2008). A statistical approach for interval forecasting of the electricity price. *IEEE Transactions on Power Systems*, *23*, 267–276.

Zheng, H., Yuan, J. & Chen, L. (2017). Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation. *Energies*, *10*, 1168.

Ziel, F. (2016). Forecasting electricity spot prices using lasso: On capturing the autoregressive intraday structure. *IEEE Transactions on Power Systems*, *31*, 4977–4987.

Ziel, F. & Steinert, R. (2016). Electricity price forecasting using sale and purchase curves: The X-Model. *Energy Economics*, *59*, 435–54.

Ziel, F., Steinert, R. & Husmann, S. (2015). Forecasting day ahead electricity spot prices: The impact of the EXAA to other European electricity markets. *Energy Economics*, *51*, 430–444.

Ziel, F. & Weron, R. (2016). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate models, *(No. HSC/16/08). Hugo Steinhaus Center*, Wroclaw University of Technology.

Ziel, F. & Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, *70*, 396–420.

APPENDICES

APPENDIX A: Performance of electricity price forecasting models: Evidence from Turkey
APPENDIX B: Electricity price forecasting using recurrent neural networks
APPENDIX C: The financial effect of the electricity price forecasts' inaccuracy on a hydro-based generation company

APPENDIX A : Performance of electricity price forecasting models: Evidence from Turkey

Table A.1 : Descriptive statistics of the Turkish day-ahead electricity prices (TL/MWh) according to hours of the day (2013-2015).

Hours	Mean	Median	Std. Deviation	Minimum	Maximum	Range	Skewness	Excess Kurtosis
0	151.08	149.74	30.36	10.00	228.08	218.08	-0.42	1.28
1	135.61	134.99	33.15	0.00	216.60	216.60	-0.70	2.15
2	118.79	122.00	37.94	0.00	216.26	216.26	-0.57	1.25
3	106.22	115.01	39.48	0.00	216.24	216.24	-0.56	0.62
4	103.27	114.91	38.58	0.00	210.77	210.77	-0.68	0.68
5	103.38	110.49	37.94	0.00	212.99	212.99	-0.54	0.82
6	107.26	118.99	41.91	0.00	216.64	216.64	-0.64	0.65
7	129.75	130.01	38.55	0.00	227.79	227.79	-0.96	1.99
8	161.03	167.65	35.89	10.02	230.96	220.94	-0.86	0.93
9	175.76	180.00	31.37	31.71	232.95	201.24	-1.06	1.73
10	178.90	181.91	29.93	30.00	270.00	240.00	-0.96	1.49
11	181.14	185.00	29.37	44.19	299.76	255.57	-0.89	1.38
12	170.04	171.99	30.89	40.68	233.92	193.24	-0.43	0.06
13	174.52	179.84	30.13	39.01	234.70	195.69	-0.60	0.35
14	176.64	180.00	30.82	38.15	276.00	237.85	-0.66	0.80
15	172.33	177.99	32.00	33.95	256.01	222.06	-0.55	0.39
16	170.46	174.85	33.71	24.27	250.00	225.73	-0.58	0.46
17	164.46	168.33	36.65	16.92	274.93	258.01	-0.55	0.26
18	159.26	160.00	33.57	19.67	231.58	211.91	-0.38	0.25
19	157.42	155.02	30.50	29.71	231.24	201.53	-0.15	-0.07
20	157.39	153.00	28.79	47.69	230.38	182.69	0.03	-0.40
21	152.51	149.99	29.61	54.78	230.00	175.22	0.11	-0.41
22	160.57	161.54	28.29	56.51	230.35	173.84	-0.25	-0.31
23	148.76	148.00	29.66	50.37	230.00	179.63	0.02	-0.12

Table A.2 : Tests of between-subjects effects for 01.01.2013-14.01.2016 as a result of factorial ANOVA.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2.793E7	74	377474.066	462.398	.000	.563
Intercept	3.556E7	1	3.556E7	43565.770	.000	.621
DAY	125320.803	30	4177.360	5.117	.000	.006
MONTH	4605340.497	11	418667.318	512.859	.000	.175
YEAR	3213615.910	3	1071205.303	1312.206	.000	.129
TIME	1.747E7	23	759359.934	930.201	.000	.446
WEEKDAY	1650976.197	6	275162.699	337.069	.000	.071
HOLIDAY	1082763.351	1	1082763.351	1326.364	.000	.048
Error	2.167E7	26541	816.340			
Total	6.521E8	26616				
Corrected Total	4.960E7	26615				

Table A.3 : Parameter estimates for 01.01.2013-14.01.2016 as a result of factorial ANOVA[30].

Parameter	B	Parameter	B	Parameter	B
Intercept	83.892				
[DAY=1]	7.653	[MONTH=1]	-11.294	[TIME=0:00]	2.124
[DAY=2]	5.976	[MONTH=2]	-27.052	[TIME=1:00]	-13.645
[DAY=3]	6.421	[MONTH=3]	-44.988	[TIME=2:00]	-30.309
[DAY=4]	4.901	[MONTH=4]	-39.616	[TIME=3:00]	-42.912
[DAY=5]	6.692	[MONTH=5]	-39.847	[TIME=4:00]	-45.792
[DAY=6]	7.398	[MONTH=6]	-34.250	[TIME=5:00]	-45.068
[DAY=7]	9.562	[MONTH=7]	-18.316	[TIME=6:00]	-41.042
[DAY=8]	7.709	[MONTH=8]	-11.816	[TIME=7:00]	-19.277
[DAY=9]	4.917	[MONTH=9]	-14.029	[TIME=8:00]	11.520
[DAY=10]	7.653	[MONTH=10]	-25.592	[TIME=9:00]	26.751
[DAY=11]	9.838	[MONTH=11]	-20.651	[TIME=10:00]	30.638
[DAY=12]	10.144	[MONTH=12]	0[a]	[TIME=11:00]	33.438
[DAY=13]	10.549	[YEAR=2013]	29.242	[TIME=12:00]	21.544
[DAY=14]	8.456	[YEAR=2014]	43.137	[TIME=13:00]	26.223
[DAY=15]	9.880	[YEAR=2015]	17.175	[TIME=14:00]	28.525
[DAY=16]	8.089	[YEAR=2016]	0[a]	[TIME=15:00]	24.132
[DAY=17]	4.159	[WEEKDAY=1]	18.150	[TIME=16:00]	23.075
[DAY=18]	6.301	[WEEKDAY=2]	24.497	[TIME=17:00]	17.012
[DAY=19]	7.661	[WEEKDAY=3]	22.610	[TIME=18:00]	11.084
[DAY=20]	6.815	[WEEKDAY=4]	23.801	[TIME=19:00]	9.340
[DAY=21]	10.059	[WEEKDAY=5]	22.145	[TIME=20:00]	9.137
[DAY=22]	7.747	[WEEKDAY=6]	18.051	[TIME=21:00]	4.052
[DAY=23]	7.684	[WEEKDAY=7]	0[a]	[TIME=22:00]	12.072
[DAY=24]	5.751	[HOLIDAY=0]	34.263	[TIME=23:00]	0[a]
[DAY=25]	3.733	[HOLIDAY=1]	0[a]		
[DAY=26]	3.023				
[DAY=27]	7.652				
[DAY=28]	8.706				
[DAY=29]	9.620				
[DAY=30]	7.641				
[DAY=31]	0[a]				

[30] a. This parameter is set to zero because it is redundant.

Table A.4 : Calculation of the deterministic part and the residuals part of the prices.

Time	Price	Day of the month	Month	Year	Hour	Weekday	Holiday	Deterministic	Residuals
01012013 00:00	145	1	1	2013	00:00	2	1	135.89	9.11
02012013 14:00	184.79	2	1	2013	14:00	3	0	193.17	-8.38
01012013 00:00	Unstd= Intercept + First Day of the Month + January + 2013 + 00:00 + Tuesday + Holiday Unstd= 101.700 +7.076 -11.290 +12.076 +1.811 +24.514 +0 = 135.887 Residuals= Price-Deterministic Residuals= 145.00 -135.89 = 9.11								
02012013 14:00	Unstd= Intercept + Second Day of the Month + January + 2013 + 14:00 + Wednesday + Unstd= 101.700 + 5.159 -11.290 +12.076 + 28.321 + 22.886 +34.318 Residuals= Price-Deterministic Residuals= 184.79 -193.17 = -8.38								

Table A.5 : Comparison of the models forecasted prices with each other and the actual prices for November 15.

Hour	Price	Naive	AR(24)	SARIMA	TAR	Markov-2 (less)	Markov-2 (more)	Markov-3
0	144.99	144.01	135.23	140.91	135.65	134.01	132.73	133.50
1	101.99	112.97	120.96	119.97	118.14	116.39	114.06	115.87
2	91.78	94.99	106.18	106.83	101.95	100.74	98.20	100.29
3	91.77	75	92.13	88.76	87.92	86.55	80.65	85.96
4	75	33.49	89.97	79.21	79.47	78.29	76.09	75.68
5	75	74.99	90.40	85.69	78.78	77.77	75.37	76.20
6	102	103.92	92.76	106.48	91.70	88.93	88.96	89.59
7	140.9	145.04	118.78	132.94	122.75	120.48	118.95	123.39
8	159	140.99	150.29	143.25	141.14	140.92	143.55	142.82
9	175.11	185.1	170.02	174.65	165.76	164.20	167.11	167.82
10	185.1	200.77	175.80	182.00	176.08	174.26	177.73	179.59
11	192.28	205.93	180.22	187.47	183.31	181.63	185.36	187.84
12	170.13	190.01	163.72	154.62	164.30	164.01	164.81	170.80
13	158.1	185.1	169.45	165.67	166.23	166.93	167.85	173.29
14	170.93	199.77	173.47	175.82	171.61	172.48	170.45	179.69
15	167.45	180.11	168.66	164.77	163.06	165.16	160.77	171.50
16	158.06	175.11	166.79	164.38	158.86	160.52	154.66	166.30
17	175.1	175.1	159.41	170.54	155.78	156.17	152.60	162.32
18	180.1	184.25	153.58	179.66	161.89	160.84	159.09	168.01
19	175.11	172.95	152.73	178.02	167.10	165.69	164.38	171.62
20	168.06	175.1	154.62	167.22	170.85	169.71	166.05	174.52
21	148.99	156.38	149.23	150.25	161.88	161.29	153.37	164.10
22	114.99	114.99	153.67	142.35	158.51	158.16	150.50	155.57
23	101.99	114.99	139.44	124.11	137.23	138.69	129.92	134.89

Table A.6 : Hourly electricity prices for 14 July 2016, 15 July 2016 and absolute errors for the hours of the days.

Hour	14-Jul-16	15-Jul-16	Absolute error
0	169.99	188.01	18.02
1	150.58	159.99	9.41
2	127.99	145.24	17.25
3	127.98	127.98	0
4	127.99	127.98	0.01
5	89	89	0
6	0.09	0	0.09
7	122.4	123.96	1.56
8	199.67	198.89	0.78
9	201.82	203.63	1.81
10	208.88	215.72	6.84
11	211.26	221.91	10.65
12	209.59	214.16	4.57
13	211.58	214.07	2.49
14	216.45	220.67	4.22
15	214.81	220.98	6.17
16	210.98	218.98	8
17	204.99	202.88	2.11
18	166.9	169.98	3.08
19	159.99	156.74	3.25
20	171.61	179.89	8.28
21	166.98	177.12	10.14
22	185	196.01	11.01
23	159.99	164.99	5
Average	167.36	172.45	5.61

APPENDIX B : Electricity price forecasting using recurrent neural networks

Table B.1 : Descriptive statistics of the Turkish day-ahead electricity prices (Euro/MWh) according to hours of the day (2016).

Hours	Mean	Standard Deviation	Lower Bound	Upper Bound	Median
0	45.61	10.34	0.00	70.53	45.38
1	40.38	11.44	0.00	69.90	40.89
2	35.25	12.70	0.00	69.73	36.50
3	30.53	13.53	0.00	69.38	33.33
4	29.57	13.47	0.00	69.38	33.03
5	29.22	13.24	0.00	69.72	30.91
6	29.93	15.00	0.00	69.82	33.34
7	37.57	13.64	0.00	70.02	39.39
8	46.85	13.40	0.00	71.54	48.49
9	52.85	12.08	0.00	211.87	54.55
10	54.62	12.96	0.00	303.03	55.32
11	55.96	13.56	0.29	351.27	57.27
12	50.78	13.02	0.28	303.02	51.51
13	52.39	12.25	1.55	242.42	53.93
14	53.79	17.73	0.33	575.75	54.55
15	52.22	15.60	0.32	454.56	53.03
16	51.52	13.38	0.31	242.42	51.54
17	49.54	16.18	1.53	354.41	50.00
18	47.71	12.69	0.25	235.45	47.27
19	47.31	10.87	3.15	151.52	46.97
20	47.78	9.75	14.45	139.41	46.94
21	46.03	9.46	9.09	90.00	45.45
22	47.24	10.08	1.35	72.12	47.75
23	42.95	11.28	0.00	72.12	43.32

APPENDIX C : The financial effect of the electricity price forecasts' inaccuracy on a hydro-based generation company

Set values	
I	Plants of the hydro generating company
T	Time periods (hour) {1,.., T}
K	Performance curves {1,..,K}
L	Set of blocks relating to the performance curve {1,.., L}
Ω_i	Upstream reservoirs of plant i
Parameters	
M	Conversion factor (3.6 x 10-3 Hm3s/m3h)
λ_t	Forecasted price of energy in period t ($/MWh)
P_i	Capacity of plant i (MW)
$P0_{ki}$	Minimum power output of plant i for performance curve k (MW).
SU_i	Start-up cost of plant i
U_i^{min}	Minimum water discharge of plant i (m3/s)
U_{li}	Maximum water discharge of block l of plant i (m3/s)
W_{it}	Forecasted natural water inflow of the reservoir associated to plant i in period t (Hm3/h)
$X0_i$	Initial water content of the reservoir associated to plant i (Hm3)
XF_i	Final water content of the reservoir associated to plant i (Hm3)
XL_i	Lower bound of the water content pertaining to the reservoir of plant i (Hm3)
XU_{ki}	Upper bound of the water content to the k^{th} performance curve of plant i (Hm3)
β_{lki}	The slope of the l^{th} block of the k^{th} performance curve of plant i (MW/m3/s)
Υ_{ij}	Time delay between reservoir of plant i and plant j (h)
s_i^{max}	Maximum spillage of the reservoir associated to plant i (m3/s)
Decision variables	
d_{kit}	0/1 variable used for the discretization of the performance curve k
v_{it}	0/1 variable which is equal to 1 if plant i is on-line in period t
y_{it}	0/1 variable which is equal to 1 if plant i is started-up at the beginning of period t
z_{it}	0/1 variable which is equal to 1 if plant i is shut-down at the beginning of period t
w_{lit}	0/1 variable which is equal to 1 if water discharged by plant i has exceeded block l in period t
p_{it}	Power output of plant i in period t (MW)
s_{it}	Spillage of the reservoir associated to plant i in period t (m3/s)
u_{it}	Water discharge of plant i in period t (m3/s)
μ_{lit}	Water discharge of block l of plant i in period t (m3/s)
x_{it}	Water content of the reservoir associated to plant i in period t (Hm3)

The objective function (C.1) maximizes the total profit of the hydro GenCo. In this equation, total profit equals to total revenue coming from the sales of the produced energy minus total start-up costs of the plants. Constraint sets (C.2)–(C.5) determine water volume of the plants according to the performance curves. Each performance curve is active at the predetermined intervals of the water volume based on the discretization of the non-linear functions. Constraints (C.6) and (C.7) calculate power generation of a plant according to the minimum power output associated with active performance curve, the total discharged water of the blocks and the power output capacity of the plant. Constraint set (C.8) is the water balance equation. The total amount of the water content, spillage and discharged water from a plant in a period is equal to the total amount of the previous water content, natural water inflow, spillage and the discharged water amounts of the upstream reservoirs associated with the plant. Constraint set (C.9) determines the discharged water amount of a plant based on water discharge of the reservoir blocks and minimum water discharge. Constraint sets (C.10)–(C.13) determine the discharged water by the reservoir blocks of a plant. Constraint set (C.14) ensures the spillage of a plant does not exceed the maximum spillage amount. Constraint sets (C.15) and (C.16) ensure the initial and final amount of the water content equals to the predetermined amounts. Constraint set (C.17) is the logical statements to arrange the start-up and shut down status of the plants. Constraint sets (C.18)–(C.21) show the type of the variables and sign restrictions.

Mathematical Model

$$Maximize \sum_t \sum_i \lambda_t \, po_{it} - \sum_t \sum_i \mathrm{SU}_i \, y_{it} \tag{C.1}$$

s.t.

$$x_{it} \leq XU_{ki} d_{k-1,it} + \sum_{k=2}^{K} XU_{k-1,i} \left[d_{k-2,it} - d_{k-1,it} \right] \qquad \forall\, i \in I, \forall\, k \in K, \forall t \in T \tag{C.2}$$

$$x_{it} \geq XU_{k-1,i} d_{k-1,it} + \sum_{k=3}^{K} XU_{k-2,i} \left[d_{k-2,it} - d_{k-1,it} \right] \qquad \forall\, i \in I, \forall\, k \in K, \forall t \in T \tag{C.3}$$

$$x_{it} \geq XL_i \qquad \forall\, i \in I, \forall\, k \in K, \forall t \in T \tag{C.4}$$

$$d_{1it} \geq d_{2it} \geq \cdots \geq d_{kit} \qquad \forall\, i \in I, \forall\, k \in K, \forall t \in T \tag{C.5}$$

$$p_{it} - P0_{ki} v_{it} - \sum_l \mu_{lit} \, \beta_{lki} - P_i \Big[(k-1) - \sum_{n=1}^{k-1} d_{kit} + \sum_{n=k}^{K-1} d_{kit} \Big] \leq 0 \qquad \forall\, i \in I, \forall\, k \in K, \forall t \in T \tag{C.6}$$

$$p_{it} - P0_{ki} v_{it} - \sum_l \mu_{lit} \, \beta_{lki} + P_i \Big[(k-1) - \sum_{n=1}^{k-1} d_{kit} + \sum_{n=k}^{K-1} d_{kit} \Big] \geq 0 \qquad \forall\, i \in I, \forall\, k \in K, \forall t \in T \tag{C.7}$$

$$x_{it} = x_{it-1} + W_{it} - M[u_{it} + s_{it}] + M \sum_{i \in \Omega_i} \left[u_{i,t-\Upsilon_{ij}} + s_{i,t-\Upsilon_{ij}} \right] \qquad \forall\, i \in I, \forall\, t \in T \tag{C.8}$$

$$u_{it} = \sum_l \mu_{lit} + U_i^{min} v_{it} \qquad \forall\, i \in I, \forall\, t \in T \tag{C.9}$$

$$\mu_{1it} \leq U_{1i} v_{it} \qquad \forall\, i \in I, \forall\, t \in T \tag{C.10}$$

$$\mu_{1it} \geq U_{1i} w_{1it} \qquad \forall\, i \in I, \forall\, t \in T \tag{C.11}$$

$$\mu_{lit} \leq U_{li} w_{l-1,it} \qquad \forall\, i \in I, \forall\, t \in T,, \forall\, l \in L \tag{C.12}$$

$$\mu_{lit} \geq U_{li} w_{lit} \qquad \forall\, i \in I, \forall\, t \in T, \forall\, l \in L \qquad \text{(C.13)}$$

$$s_{it} \leq s_i^{max} \qquad \forall\, i \in I, \forall t \in T \qquad \text{(C.14)}$$

$$x_{i0} = X0_i \qquad \forall\, i \in I \qquad \text{(C.15)}$$

$$x_{iT} = XF_i \qquad \forall\, i \in I \qquad \text{(C.16)}$$

$$y_{it} - z_{it} = v_{it} - v_{i,t-1} \qquad \forall\, i \in I, \forall t \in T \qquad \text{(C.17)}$$

$$d_{kit}, v_{it}, y_{it}, z_{it} \in \{0,1\} \qquad \forall\, i \in I, \forall t \in T \qquad \text{(C.18)}$$

$$w_{lit} \in \{0,1\} \qquad \forall\, i \in I, \forall\, t \in T, \forall\, l \in L \qquad \text{(C.19)}$$

$$p_{it}, s_{it}, u_{it}, x_{it} \geq 0 \qquad \forall\, i \in I, \forall t \in T \qquad \text{(C.20)}$$

$$\mu_{lit} \geq 0 \qquad \forall\, i \in I, \forall\, t \in T, \forall\, l \in L \qquad \text{(C.21)}$$

CURRICULUM VITAE

Name Surname : Umut Uğurlu

Place and Date of Birth : Istanbul, 25.04.1987

E-Mail : umut.ugurlu@itu.edu.tr

EDUCATION :

- **B.Sc.** : 2010, Istanbul Technical University, Civil Engineering Faculty, Environmental Engineering
- **M.A.** : 2013, Istanbul Technical University, Management Faculty, Management

PROFESSIONAL EXPERIENCE AND REWARDS:

- Research Assistant at the Department of Management Engineering since 2013.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Ugurlu, U.**, Tas, O., Gunduz, U. 2018. Performance of Electricity Price Forecasting Models: Evidence from Turkey, *Emerging Markets Finance & Trade*, 54(8), 1720-1739. (SSCI)

- **Ugurlu, U.**, Oksuz, I., Tas, O. 2018. Electricity Price Forecasting Using Recurrent Neural Networks, Energies, 11(5), 1255. (SCI-Expanded)

- **Ugurlu, U.**, Tas, O., Kaya, A., Oksuz, I. 2018. The Financial Effect of the Electricity Price Forecasts' Inaccuracy on a Hydro-based Generation Company, *Energies*, 11(8), 2093. (SCI-Expanded)

- **Ugurlu U.**, Ulengin B., Tas, O. 2016: Basic Time Series Patterns of Turkish Electricity Prices. The 2016 WEI International Academic Conference Proceedings, Rome, Italy.

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- **Ugurlu, U.**, Tas, O., Guran, C. B., Guran, A. 2018. SSD Efficiency at Multiple Data Frequencies: Application on the OECD Countries, *Prague Economic Papers*, 27(2), 169-195. (SSCI)

- Tas, O., Tokmakcioglu, K., **Ugurlu, U.**, Isiker, M. 2016. Comparison of Ethical and Conventional Portfolios with Second-order Stochastic Dominance Efficiency

Test. *International Journal of Islamic and Middle Eastern Finance and Management*, 9(4), 492-511. (SSCI)

- Tas, O., Barjiough, F. M., **Ugurlu, U.** 2015. A Test of Second-order Stochastic Dominance with Different Weighting Methods, *Journal of Business Economics and Finance*, 4(4), 723-731. (Econlit)

- Tokmakcioglu, K., Ozcelebi, O., **Ugurlu, U.** 2015. Revisiting the Interactions between the Balance Sheet Accounts of US Telecommunications and Transportation Sectors: Empirical Evidence from Panel-VAR Models, *Journal of Business Economics and Finance*, 4(4), 770-795. (Econlit)

- Yorulmus H., **Ugurlu U.**, Tas O. 2018: A Long Short Term Memory Application on the Turkish Intraday Electricity Price Forecasting, *Pressacademia Procedia*, 7(1), 126-130.

- Tas O., Yakak E., **Ugurlu U.** 2018: Using Artificial Neural Network and a Statistical Method for the Estimation of Euro/Turkish lira Exchange Rate, *Pressacademia Procedia*, 7(1), 414-417.

- **Ugurlu U.** 2017: Comparison of the Electricity Price Models' Forecast Performances in the Turkish Day-Ahead Market. 4th Young Finance Scholars Conference, University of Sussex, UK.

- Tokmakcioglu K., Ozcelebi O., **Ugurlu U.** 2016: Examining the Relationship between the Balance Sheet Accounts of US Biotechnology, Telecommunicaitons and Transportation Sectors, *Procedia Economics and Finance*, 38, 492-498.

- **Ugurlu U.**, Tas O., Guran C. B., Guran, A. 2015: An Application of Second-order Stochastic Dominance Efficiency Test on Global Fossil Fuels Energy Stocks. ISER 15th International Conference on Economics and Business Research, Lisbon, Portugal.

- Tas O., Guran C. B., **Ugurlu, U.** 2014: Second Order Stochastic Dominnace Portfolio Efficiency by Changing Data Frequency: An application in ten different country indexes. 13th Eurasia Business and Economics Society (EBES) Conference, Istanbul Bilgi University, Istanbul, Turkey.

- **Ugurlu U.**, Tas O. 2014: Selection of Turkish Mutual Funds by Second-order Stochastic Dominanace Efficiency Test and Comparison with Conventional Portfolio Performance Measures. International Journal of Arts&Sciences (IJAS) Conference for Academic Disciplines, Freiburg, Germany.