KADIR HAS UNIVERSITY

SCHOOL OF GRADUATE STUDIES

PROGRAM OF INDUSTRIAL ENGINEERING

MODELS FOR ELECTRICITY DEMAND FORECASTING, CLASSIFICATION, AND IMBALANCE REDUCTION IN COMPETITIVE MARKETS

Ergün YÜKSELTAN

DOCTOR OF PHILOSOPHY THESIS

ISTANBUL, JANUARY, 2023

ERGÜN YÜKSELTAN Ph.D. Thesis 2023

MODELS FOR ELECTRICITY DEMAND FORECASTING, CLASSIFICATION, AND IMBALANCE REDUCTION IN COMPETITIVE MARKETS

ERGÜN YÜKSELTAN

ADVISOR: PROF.DR. AHMET DENIZ YÜCEKAYA

A thesis submitted to

the School of Graduate Studies of Kadir Has University

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in

Industrial Engineering

Istanbul, January, 2023

APPROVAL

This thesis titled **Models For Electricity Demand Forecasting, Classification, and Imbalance Reduction in Competitive Markets** submitted by **ERGÜN YÜKSELTAN**, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering and Natural Sciences is approved by

Prof. Dr. Ahmet Deniz Yücekaya (Advisor) ....

Kadir Has University

Prof. Dr. Ayşe Bilge (Co- Advisor) ....

Kadir Has University

Prof. Dr. Funda Samanlıoğlu ....

Kadir Has University

Assoc. Prof. Dr. Mehmet Aydın ....

Kadir Has University

Dr. Emre Çelebi ....

Yeditepe University

Assoc. Prof. Dr. Hatice Tekiner Moğulkoç ....

Marmara University

I confirm that the signatures above belong to the aforementioned faculty members

Prof. Dr. Mehmet Timur AYDEMİR

Director of the School of Graduate Studies

Approval Date: 09/01/2023

DECLARATION ON RESEARCH ETHICS AND

PUBLISHING METHODS

I, ERGÜN YÜKSELTAN ; hereby declare

- that this Ph.D. Thesis that I have submitted is entirely my own work and I have cited and referenced all material and results that are not my own in accordance with the rules;
- that this Ph.D. Thesis does not contain any material from any research submitted or accepted to obtain a degree or diploma at another educational institution;
- and that I commit and undertake to follow the "Kadir Has University Academic Codes of Conduct" prepared in accordance with the "Higher Education Council Codes of Conduct".

In addition, I acknowledge that any claim of irregularity that may arise in relation to this work will result in a disciplinary action in accordance with the university legislation.

Ergün Yükseltan

____

09/01/2023

ACKNOWLEDGEMENT

I would like to extend my heartfelt gratitude to my advisor, Prof. Dr. Ahmet Deniz Yücekaya, for his invaluable guidance, support and mentorship throughout my doctoral journey. His unwavering support and insightful advice helped me navigate through the challenges of the thesis process and brought me to where I am today.

I am also deeply grateful to my co-advisor, Prof. Dr. Ayşe Bilge, for her dedication, motivation, and encouragement. Her constant guidance and support has been instrumental in enhancing my skills and overcoming the obstacles in my academic journey.

I extend my thanks to the faculty at Kadir Has University and the Department of Industrial Engineering for providing me with a rich educational experience and ample resources. Their support has been instrumental in helping me achieve my academic goals and complete my thesis with excellence.

Last but not the least, I would like to extend my sincerest gratitude to my family for their unwavering love, support, and encouragement. They have been my rock throughout this journey and I am forever grateful for their sacrifice and understanding.

MODELS FOR ELECTRICITY DEMAND FORECASTING, CLASSIFICATION, AND IMBALANCE REDUCTION FOR COMPETITIVE MARKETS

ABSTRACT

In liberalized energy markets, hourly forecasts of consumers and producers are crucial for efficiently using energy resources and reducing environmental impacts. In this study, the countries' consumption in the ENTSO-E common network between 2006 and 2018 was analyzed using the time series method. With the created model, short, medium, and long-term demand forecasts are made using Fourier Series Expansion. In order to improve the error rate of short-term forecasts, a hybrid model was created with alternatively created feedback and autoregressive methods. While annual forecasts are made with an average error rate of 6%, the error rate in daily forecasts is around 4.5%. With the hybrid models created, hourly estimates can be made with approximately 1.5% and 1% error rates. Accurate estimations are of great importance in terms of the efficiency of energy markets, and the emergence of energy storage opportunities with the developing technology increases this importance. For this reason, the amount of imbalance was estimated by using the forecast result of the hybrid model in the Turkish Energy Market, and a strategy was developed to reduce the imbalance cost accordingly. With this strategy, simulations have been made for situations with and without storage, and the results have been shared.

Keywords: Energy Storage, Free Electricity Markets, Time Series Analysis, Demand Forecasting, Market Strategies

ELEKTRİK TALEP TAHMİNİ, SINIFLANDIRILMASI VE REKABETÇİ PİYASALAR İÇİN DENGESİZLİK AZALTMA MODELLERİ

ÖZET

Liberalleşen enerji piyasalarında, tüketici ve üreticilerin saatlik tahminleri, enerji kaynaklarının verimli kullanılması ve çevresel etkilerin azaltılması açısından büyük önem taşımaktadır. Bu çalışmada ENTSO-E ortak ağında yer alan ülkelerin 2006-2018 yılları arasındaki tüketimleri zaman serisi yöntemi kullanılarak analiz edilmiştir. Oluşturulan model ile Fourier Serisi Genişletmesi kullanılarak kısa, orta ve uzun vadeli talep tahminleri yapılmaktadır. Kısa vadeli tahminlerin hata oranını iyileştirmek için alternatif olarak oluşturulan geri besleme ve otoregresif yöntemlerle hibrit bir model oluşturulmuştur. Yıllık tahminler ortalama %6 hata oranı ile yapılırken, günlük tahminlerde hata oranı %4,5 civarındadır. Oluşturulan hibrit modeller ile yaklaşık %1,5 ve %1 hata oranları ile saatlik tahminler yapılabilmektedir. Doğru tahminler enerji piyasalarının etkinliği açısından büyük önem taşımakta ve gelişen teknoloji ile birlikte enerji depolama imkanlarının ortaya çıkması bu önemi artırmaktadır. Bu nedenle hibrit modelin tahmin sonucu kullanılarak Türkiye Enerji Piyasasındaki dengesizlik miktarı tahmin edilmiş ve buna göre dengesizlik maliyetinin düşürülmesine yönelik bir strateji geliştirilmiştir. Bu strateji ile depolamanın mümkün olduğu ve olmadığı durumlar için simülasyonlar yapılmış ve sonuçlar paylaşılmıştır.

Anahtar Kelimeler: Büyük Ölçekli Enerji Depolama, Serbest Enerji Pazarları, Zaman Serisi Analizi, Talep Tahmini, Piyasa Stratejileri

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Brief History of Electricity

The common belief about the discovery of electricity is that it was found suddenly. However, it is the result of a series of discoveries, from simple to complex. The magnetic and electrical properties of lodestone and amber have been known by humans very long time. When the lodestone is placed to a woodblock on water, it behaves as a compass and aligns in north-south directions. If certain materials are rubbed with amber, they attract light objects like paper. William Gilbert is the first person who systematically examines these first samples of electricity and magnetism (Gilbert, 1958).

The foundations of electric machines were laid in the 17th and 18th centuries by studies of Otto von Guericke and Francis Hauksbee. Guericke invented the primitive electrostatic generator in 1663 with a rotating and rubbing sulfur globe. This design inspired future simple generators, and many scientists devised different ideas to improve them. Isaac Newton offered to use glass material instead of sulfur (Newton et al., 1951). At the beginning of the 18th century, Francis Hauksbee improved the design and made the frictional electrical machine that rotates and glows a glass ball when it closes to woolen cloth (Pumfrey, 2004). These experiments led to many studies on electricity, and Stephen Gray and Charles Du explained how electricity moves on a stick. Stephen Gray discovered the differences between conductors and insulators and proved that the amount of electricity on any object is independent of its mass (Hellström, 1998). Charles Du completed these studies by discovering two electricity charges, positive and negative (Guarnieri, 2012).

In the mid of 18th century, Benjamin Franklin proved that all materials are in an electrically natural state and can be charged positively or negatively with friction. If materials charge positively or negatively, the electric discharge and materials return to their natural state. In the following century, Italian scientist Alessandro Volta created the first electric cell that allowed scientists to use electricity from a reliable source (Guarnieri, 2012).

Finally, Michael Faraday was the first person that realize the relationship between magnetism and electricity. An electric current may be produced by slipping a magnet through a copper wire (“V. Experimental Researches in Electricity,” 1832). This discovery converts electrical energy to motion energy and motion energy to electrical energy.

1.2. Generation, Transmission, and Distribution

1.2.1. Electricity Generation

Energy became one of the industrialization’s primary inputs, especially with the start of the 20th century. The race between developed countries in industrialization caused an increase in energy consumption.

Energy resources can be classified under two main titles: non-conventional and conventional resources. Conventional resources like fossil fuels and nuclear power are not renewed by any natural process; however, non-conventional resources are continuously recycled by natural processes. These resources are alternative resources to conventional ones. The most common non-conventional resources were hydropower and biomass previously. With the advance in technology, wind and solar have become significant energy resources, and their portion in consumption will continue to increase in the future.

All conventional energy resources have commonly been used in the last decades, and the number of alternative energy resources increased in recent years. Hence, non-conventional energy resources are still developing, and their share in total energy consumption is minimal.

Figure 1.1 below shows the yearly energy resource shares in total global consumption. The usage of fossil fuels, especially coal and gas, rapidly increased in the 18th and 19th centuries. With the oil crisis in 1972, the share of oil consumption started to decrease in total consumption, but oil is still the most common and primary energy resource.

The overall picture is slightly different from total consumption when it comes to electricity generation. The electricity generation from alternative resources has considerably increased in the last ten years. The main reason for the increase is the limitation of fossil fuels. Moreover, the rapid increase in fossil fuel consumption, especially after the 1980s, causes climate problems and questions about fossil fuel reliability.

These concerns will expedite the development and usage of alternative energy resources. Figure 1.2 presents the share of energy resources in electricity generation. Coal and natural gas are the primary sources of electricity generation, while oil share has decreased since the 1972 oil crisis. Wind and solar have become very popular and will play an essential role in transitioning from fossil fuel to renewable energy resources.

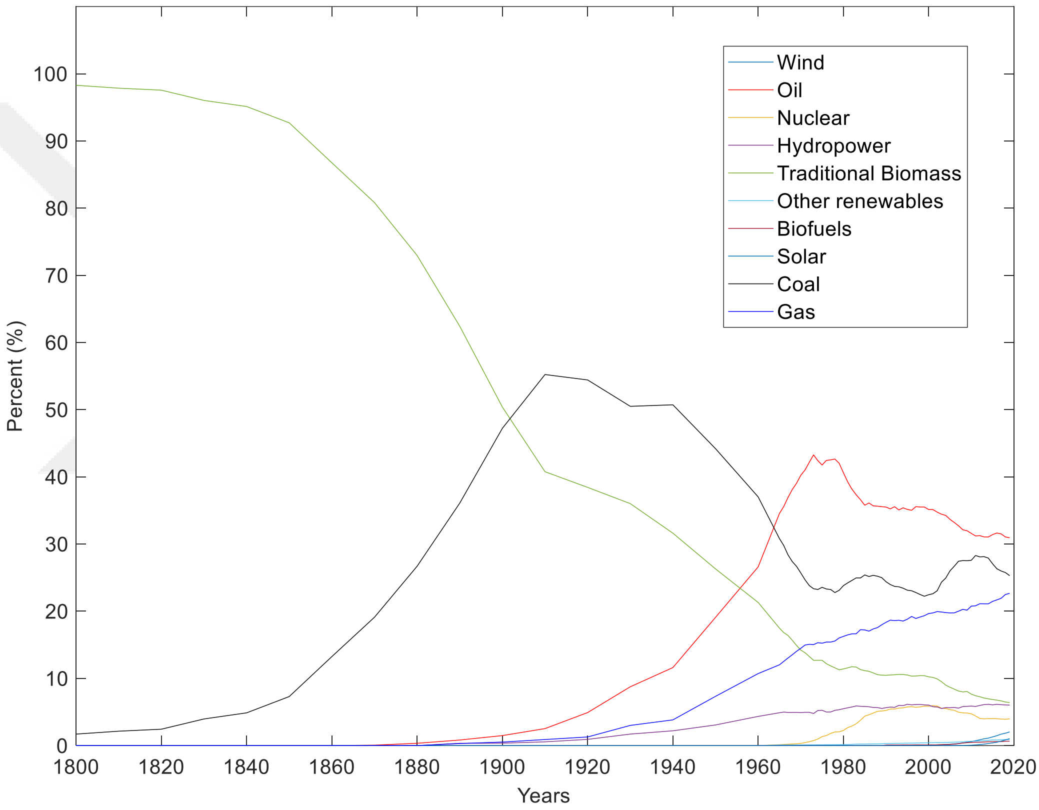


Figure 1.1: Yearly energy resources shares in total world consumption

The industrialization race between countries, especially after World War II, makes energy important strategically. The distribution of energy resources in the world is unequal, and some countries have a geographical advantage in reaching energy resources. As a result, countries with energy resources or sovereignty over energy resources rapidly reach the higher technological and industrializations level. Hence, energy gained importance strategically for modernization, and developed countries started to use their sovereignty on energy resources to obtain a political privilege over other countries.

To summarize, energy became one of the main drivers of industrialization in the 19th century. It shaped the political strategies of developed countries to secure and reach energy resources for sustainable growth.

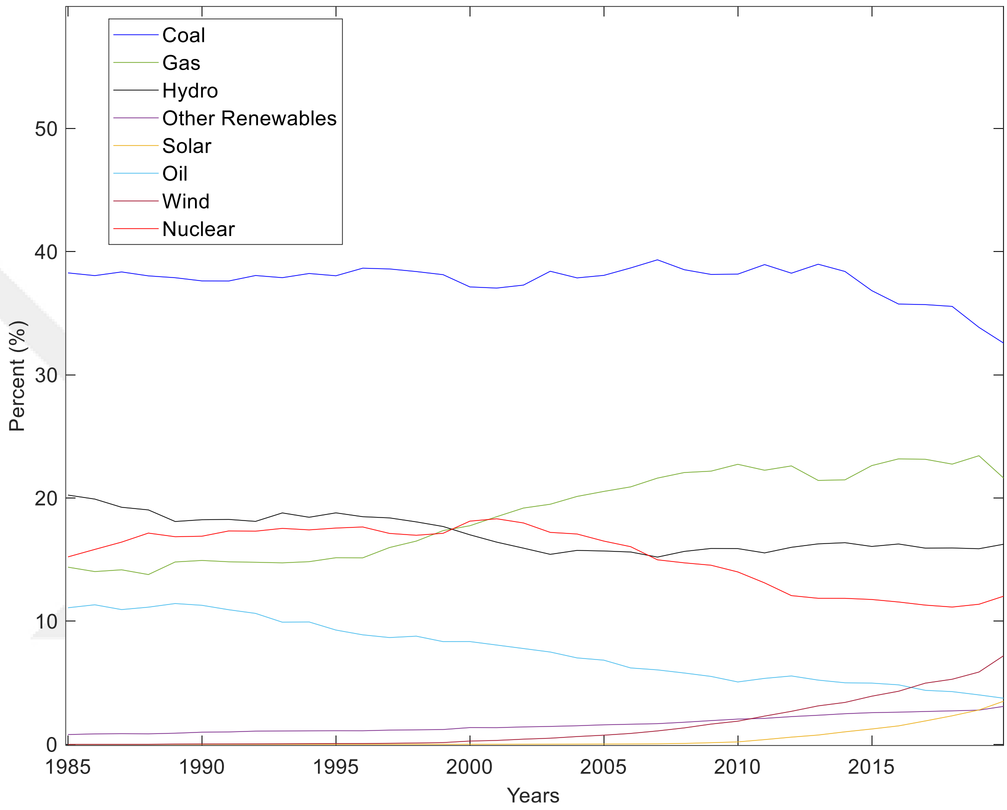


Figure 1.2: Energy resources share in electricity generation by years

1.2.2. Transmission and Distribution

In the beginning, electricity usage was mainly for industry. However, with the invention of the incandescent lamp by Thomas Edison in the 19th century, it started to be used in public areas and homes for lighting. The end of the 19th century is called the "War of the Current" in literature because of the competition between alternating and direct currents.

Direct current (DC) transmissions were initially used but were inefficient for long-range electricity transfers because of the different voltage requirements of different electrical items. Direct current transmission is based on high current and low voltage, which is unsuitable for a device that needs a high voltage. That is why transmission lines with different specializations were needed for different devices. For this reason, transmission lines cannot expand rapidly, and the number of distributed local generation systems increased (Brown & Sedano, 2004).

Transmission with alternative current (AC) became possible in the following years with the efforts of White Westinghouse and Nicola Tesla (Allerhand, 2017). The AC transmission system needs a more complex design than the dc transmission lines. However, it allows electricity transmission with minimum loss and for an extended range. Moreover, it is cheaper than DC transmission lines. Even though it is more complicated than DC transmission, it meets different voltage requirements of different equipment on the end customer level. DC lines need a more complex converter system to meet different voltage requirements.

The electricity supply is divided into two steps by centralization and capacity increase of power plants: transmission and distribution. In early times, the transmission was used only. However, with the increasing and expansion of grid lines and end-users, it was separated as distribution for low-voltage electricity transfer between substations and consumer's end and transmission for high-voltage electricity transfer for long-range between power plants and substations. These terms formed because of the need to separate bulk and small electricity transfers. The connection of consumers directly to large power plants is not practical and is much more expensive. Transmission line capacity is for extra high voltage, 330-1100 kV (500-2000 MWA), and high voltage, 60-220 kV (50-300 MWA), while distribution line capacity is for mid voltage 1-66 kV (1-50 MWA) and low voltage 0.4-1 kV (0.01- 1 MWA)(Majstrović, 2020).

With the increase of renewable energy resources in electricity generation, new technologies are emerging for efficient transmission because of the difference from conventional generation methods. The different characteristics of renewable energy resources are listed in five categories by International Energy Agency (International Energy Agency, n.d.):

- Low short-run marginal cost:

Bids from renewable energy generators should be at the short-run marginal cost level. Additional payments like incentives should not affect market prices drastically.

- Variability:

Due to variability in renewable resources, prices are valid for the short term, so large differences in prices are imported for the market.

- Uncertainty:

Uncertainty about renewable energy resources increases the importance of short-term price signals because it causes variability in generation. Prices are formed close to real-time prices or based on the current grid status.

- Location and Modularity:

Renewable energy generation does not have to be centralized and can be done with a small capacity and at many locations. The less centralization makes prices differ from place to place.

- Non-synchronous technology:

Wind and solar power cannot connect to the system synchronously, and penetration of large amounts of non-synchronous generation to the grid is challenging, so high-voltage direct current (HVDC) transmission lines have developed and gained popularity in recent years. They will be essential between renewable energy resources and substations for transmissions. However, AC distribution lines keep their importance for end customers, and existing substations need more investment or renewal to convert DC to AC for distribution. We can list the advantages and disadvantages of HVDC as follows;

- Interconnection flexibility
- Lower power loss due to not requiring capacitive-reactive power like in AC
- Able to connect through a sea (Submarine DC connection)
- Able to transfer non-synchronous power generation
- High cost and complexity due to conversion requirements for end customers

1.3. Evaluating the energy status of Türkiye

Reaching sustainable and cheap energy resources is one of the critical inputs for industrialization and development. There is a direct relationship between energy consumption and economic growth; they are positively correlated, and economic growth is highly dependent on energy consumption (Zhixin & Xin, 2011). We can say that higher economic standards also support higher energy consumption by people for their well-being. That is why energy consumption per person is one of the significant indicators of a country's economic and development level. Figure 1.3 below shows energy consumption per person for Türkiye, high-income, upper-middle-income, low-middle, and low-income countries. The World Bank economic income classification is used (*World Bank Country and Lending Groups – World Bank Data Help Desk*, n.d.), and details are in Appendix A.

On the other hand, reaching higher economic and social levels cause efficient energy usage, especially in high-income countries; even if GDP increases, their energy consumption level does not change significantly because of productivity. For this reason, energy intensity has become popular in measuring countries' capacity to turn consumed energy into US dollars. The definition of "Energy Intensity" according to "Our World in Data" is; "Energy intensity level of primary energy is the ratio between energy supply and gross domestic product measured at purchasing power parity. Energy intensity is an indication of how much energy is used to produce one unit of economic output. A lower ratio indicates that less energy is used to produce one unit of output". Figure 1.4 shows calculated energy intensity according to the 2011 US dollar index.

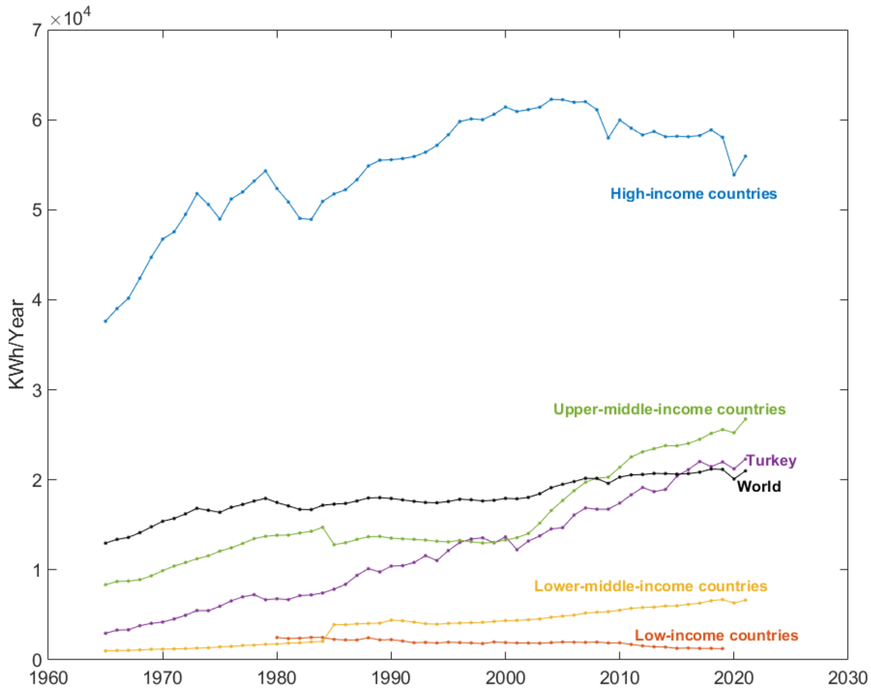


Figure 1.3: Energy consumption per person

Both data indicate how countries consume primary energy resources. The energy intensity of Türkiye is slightly smoother than the other class. Comparing it with "Upper-middle income countries," we see that Türkiye's economic growth is underperformed. Even though the lower energy intensity is better and shows a higher productivity level, historical data of Türkiye is a result of less industrialization according to the upper-middle income group.

Moreover, even though the countries continue to increase their GDP, their energy intensity is not increasing due to technological levels. According to a study of 21 developed countries, a 1% increase in GDP results in a 0.62-0.78% reduction in energy intensity (Zhou et al., 2021).

These indicators are significant for the countries, especially those lacking energy resources, and Türkiye is in that group. Increasing population and industrialization cause higher consumption, but limited energy resources create difficulties.

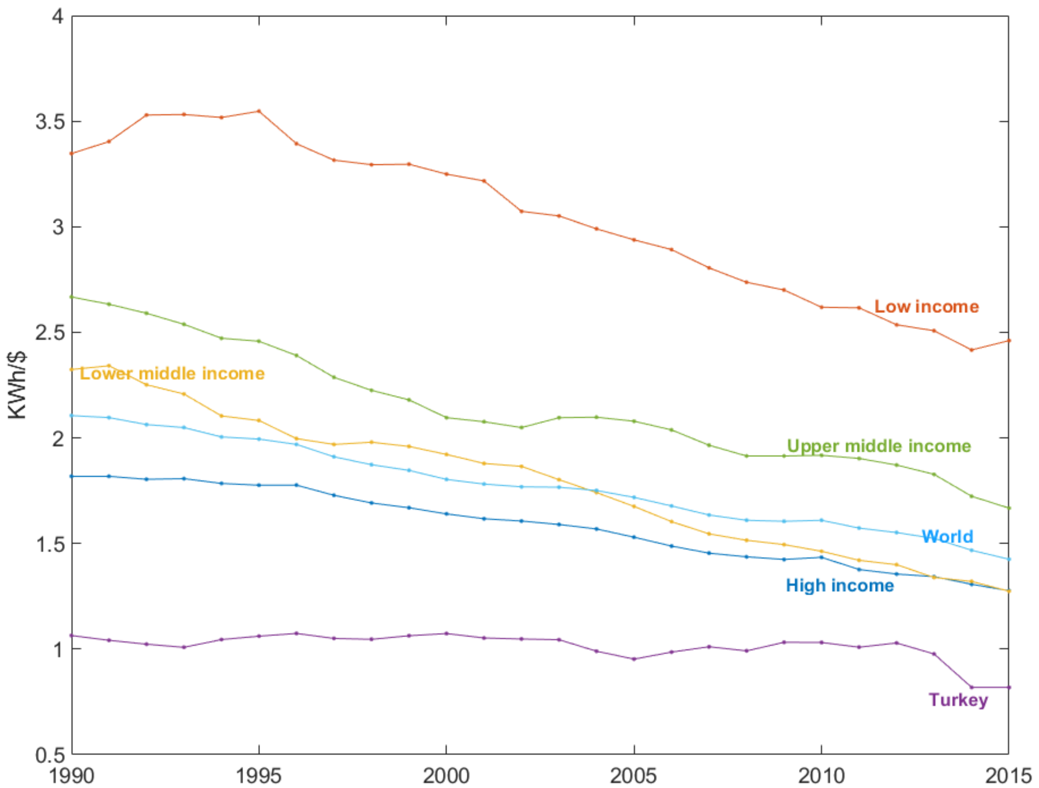


Figure 1.4: Energy intensity changes between 1990-2015 (2011 US dollar)

Türkiye is a net energy importer country. Energy import costs are the main component of the national account deficit. Figure 1.5 shows that Türkiye differs from other developing countries in energy imports. The portion of imported energy in total consumption increases due to high population and industrialization. It is too much more when compared to other countries. This picture proves the strategic importance of energy resources and how the 1970s oil crisis affected energy imports.

Türkiye must follow strict policies to overcome energy problems because of the increase in population in the last decade and economic growth causing high import dependency. Türkiye follows European energy policies mostly and shows efforts to reach decarbonization targets. Electricity market reforms and the pace of transition from fossil fuels to renewable energy resources are decisive steps to diversify the energy mix. Moreover, diversification of import resources of natural gas is also favorable for energy security—the dominance of Russia in supplying natural gas decreases with importing natural gas from Iran and Azerbaijan. This prevents of using energy supply as a political pressure by the exporters.

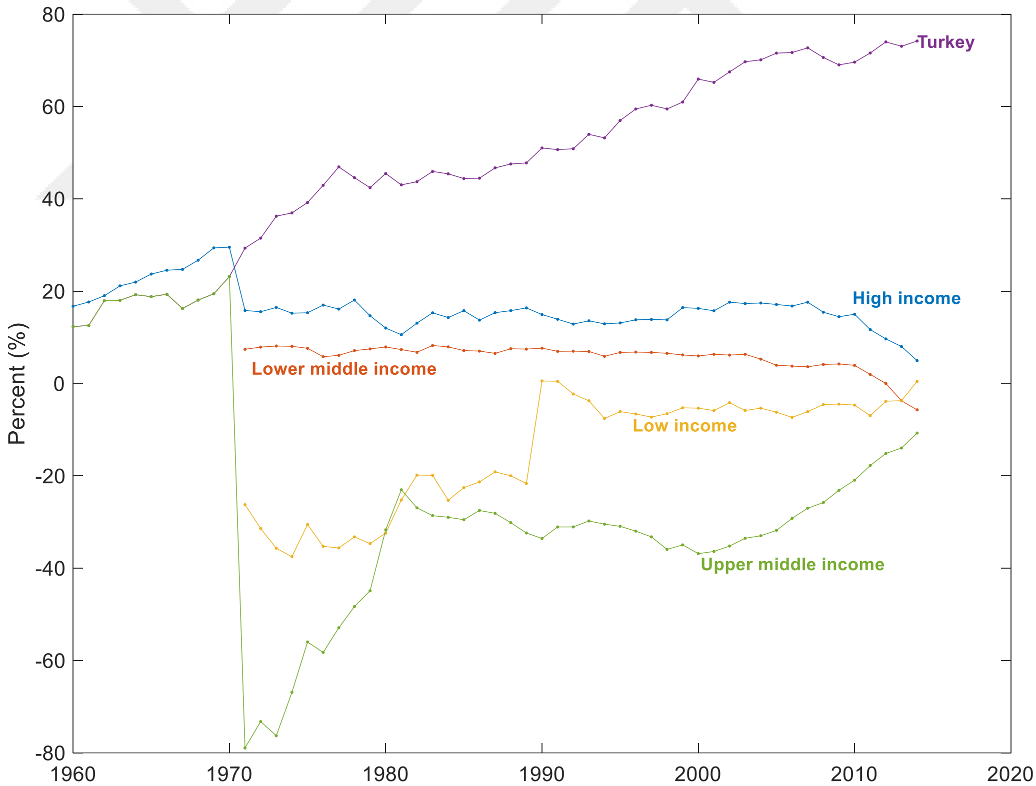


Figure 1.5: Energy net import/export ratios by yearly (Positive ratios are imports and negative ones are exports)

Türkiye also has considerable potential for renewable energy resources. Solar, wind, and geothermal energy resources have rapidly increased their portion in electricity generation in the last years with the help of government incentives. However, existing technologies for renewable energy resources do not solve intermittency problems completely. This intermittency problem risks energy supply security and slows the energy transition from fossil fuels to renewable energy resources. Hence, Türkiye should make a long-term plan carefully to increase renewable energy resources in total consumption.

For instance, Denmark is one country that generates a high portion of electricity from renewable energy resources, but it has encountered high electricity prices in the last years. The highest electricity prices occurred in Germany and Denmark worldwide in the second half of 2021 (*Electricity Prices around the World | GlobalPetrolPrices.Com*, n.d.). These two countries have a high transition speed from fossil fuels to renewable ones, as shown in Figure 1.6. The main reason for these high prices is that solar and wind power have high dominancy in renewable energy resources. Sweden is another country with higher renewable energy resource usage but primarily uses hydropower. Solar and wind power's contribution to electricity generation is significantly less. The price fluctuation and increase are not as high as in Denmark and Germany because of no intermittency problem with hydropower.

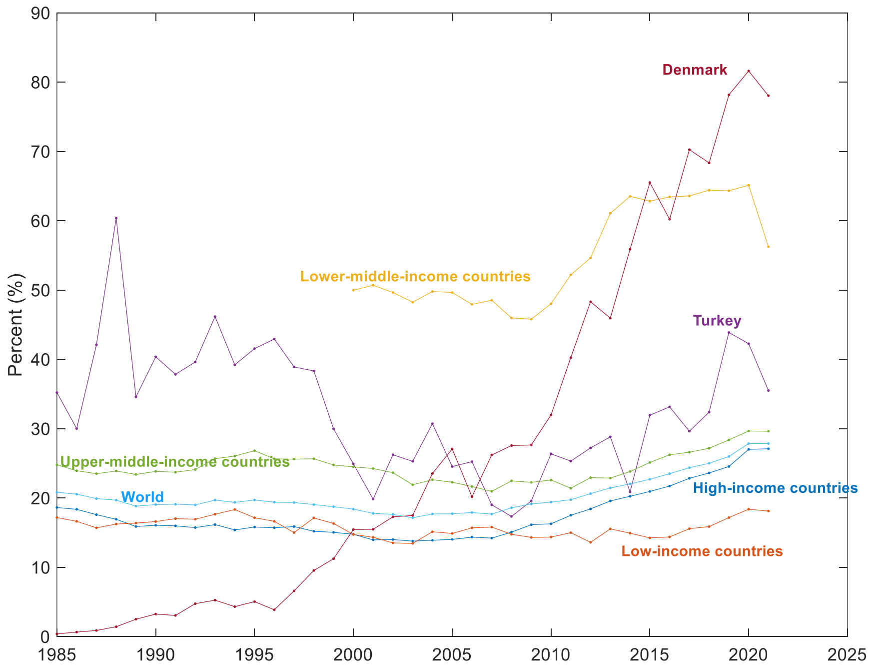


Figure 1.6: Ratio of renewable energy resources in electricity generation

On the other hand, Türkiye is on a very strategic route to transfer Asia and Middle East energy resources to Europe and other countries. This advantage can potentially eliminate the lack of own energy resources. To achieve this, Türkiye developed and implemented various energy transfer projects:

- Trans Anatolian natural gas pipeline: Natural gas transfer between Azerbaijan and the EU
- Baku – Tiflis – Ceyhan, and Iraq – Türkiye oil pipelines: Oil transfer from Azerbaijan, Turkmenistan, Kazakhstan, and Iraq to Ceyhan port to deliver to other countries with sea-borne
- Russia – Türkiye natural gas pipeline (West)
- Blue stream (with Russia)
- Iran – Türkiye natural gas pipeline
- Baku – Tiflis – Erzurum natural gas pipeline
- Türkiye – Greece natural gas pipeline (not completed)
- TurkStream (with Russia)

Türkiye has also continued hydrocarbon and natural gas exploration in the Black and Mediterranean Seas. These efforts will decrease dependency on imported energy resources and ensure energy security.

1.4. The History of Deregulated Electricity Markets

In the early 90s, governments produced, distributed, and transmitted electricity in most countries. Traditionally, the electricity market has been developed and operated as a framework strictly organized and vertically integrated into all or most activities, from production to retail trade. However, with the free-market trends starting in the 90s, most countries privatized the assets, started free trades, and gradually developed competitive markets and various support services for operating energy systems with the effects of liberalization trends. These reforms aimed to lead to more efficiency and improvements of electricity supplies in many aspects, including operation cost and pricing.

These improvements can also be classified as having short-term and long-term implications. Short-term expected advantages of privatized markets are increased competition, decreased retail prices, and lower operating costs. Long-term expected advantages of privatized markets are to develop new technologies with increasing competition, new investments to increase capacity, and reliable transmission systems to provide and secure continuous electricity supply for end users-consumers and decrease environmental effects. According to the energy outlook reports by International Energy Agency, they expect downward pressure on wholesale electricity prices with the increasing capacity of zero-emission generations (*How Will the Electricity Market of the Future Work? – Analysis - IEA*, n.d.).

With the liberalization of electricity markets, governments aim to increase public welfare, reduce prices, secure a sustainable electricity supply, and encourage the private sector for new investments to increase the efficiency and use of renewable energy resources. As highlighted before, electricity operation has three segments, generation, transmission, and distribution. Due to natural limits, increasing competition in transmission and distribution would not be possible (Willems & Ehlers, 2008). Transmission and distribution companies operate based on location. However, billing, contracting, or managing other services for end-user is another area called the supply side (Pollitt, 1997).

Supply firms are responsible for marketing and service quality. Market liberalization policies encourage new firms to enter and invest in new technologies. Market barriers are changed to make small firms' entrance to the market easier because existing market firms can resist applying new technologies to prevent additional short-term costs. However, small firms focus on a specific technology to increase efficiency. If they become a part of markets, they decrease operation costs, and existing firms must follow and adopt new technologies. As a result, some governments make policies to encourage new investments in renewable energy resources and technologies. These supports are also crucial for decreasing greenhouse gas emissions (Kim & Jeong, 2016). They provide feed-in tariffs or give incentives due to environmental concerns.

On the other hand, competition on the supply side provides an alternative for the customer to choose their supplier. Deregulated markets allow all or some consumers to make contracts directly with suppliers and fix electricity prices. Changing supplier ratio is another factor that shows the success of the deregulation process. The United Kingdom and Ireland have more switching ratios than all EU countries (Shin & Managi, 2017).

To summarize energy market developments around the globe, we have selected some countries according to the importance of their market reform progress and income level and examined their market developments individually to see different approaches and results.

1.4.1. Chile

Chile is known as the first country that reforms energy markets. Chile National Energy Commission was the regulator in charge of setting up transmission and distribution tariffs; in 1980, broad economic reforms started, and these movements also affected the energy market structure. The first legislation for vertical and horizontal unbundling was enacted in 1982. The privatization process was completed in 1989 for transmission, distribution, and generation. During the same time horizon, the compliance board was formed with the legislation to audit electricity and fuel markets.

In 1995, An agreement was made with Argentine to import natural gas, and the privatization process continued until 2005. There were two underlying crises in 1998 and 2005. The first was due to the drought in 1998, which suddenly caused hydropower generation to decrease, and the second was the supply problem of natural gas from Argentina. The high dependency on imported natural gas from Argentina directly affects generation costs. In 2005, distribution companies started to buy electricity from the open market with public auctions.

1.4.2. Argentina

Argentina is shown as one of the examples of the energy market reforms. Reforms and their implementations set an example for the reforms in other countries' energy markets and are considered a road map to achieving full-scale power market reforms.

With government legislation in 1991, market unbundling started for transmission, distribution, and generation. The following year, wholesale markets started operating, and energy prices were determined with public auctions. Privatization of transmission, distribution, and generation continued until the 2000s. Argentina faced an economic and political crisis at the start of the 2000s. As a result, the Argentine currency was devalued and directly affected the functioning of free markets. Foreign exchange trading was restricted by law. This economic crisis caused a shortage in gas extraction, and the government started importing gas from

Bolivia and limiting gas exports to Chile. Energy markets answered the crisis well with an increased number of generation resources. There are also improvements in different indicators, such as operating cost, capacity index, and electricity prices (Vagliasindi & Besant-Jones, 2013).

1.4.3. Brazil

Brazil is another pioneer country that showcases the changes from a vertical market structure to a free market. Previously, the energy market was dominated by hydropower and controlled by government companies vertically. In 1995, the first legislation was made for the unbundling market structure, and Brazilin Electricity Regulatory Commission was founded in the following year. While generation and transmission privatizations continued between 1995 and 2000, many different laws and commissions were established for a competitive market in the same period. The wholesale electricity market launched in 1998 was established, and Centrais Elétricas Brasileiras S.A. (Eletrobras) split into six holdings and 14 companies. These companies joined the market for generation and transmission (*OECD Reviews of Regulatory Reform: Brazil 2008*, 2008).

Brazil grew rapidly in the 1990s, and the increase in private electricity generation capacity stayed behind demand growth in this period. As a result, the government launched a thermal priority program to augment the diversification of electricity generation resources and decrease the dominancy of hydropower. The great drought in 2002 also directly affected the electricity market. Other market reforms were implemented in 2003 and 2004 due to the slowdown in new investments. According to these new reforms, the long-term electricity energy market was supported, and distribution companies were required to purchase long-term electricity from this market. In this way, the risks that might arise for new investments were reduced. Finally, in 2004 and 2005, primary and secondary auction market mechanisms were established for distribution companies.

1.4.4. South Korea

South Korea is another country that changed its energy market gradually in the 1990s. The electricity operation was managed by KEPCO (Korean Electric Power Company) for generation, transmission, and distribution in a vertically integrated market structure. The dominance of KEPCO made privatization difficult, so they created a pool market, and KEPCO participated as a retail seller in the generation, transmission, and distribution. These policies did not lead to free markets as planned because the dominancy of KEPCO created substantial entry barriers for new markets. The Korean government announced a market restructuring plan in 2001, and KEPCO split into six-generation companies. However, distribution unbundling was canceled in 2003, and the market restructuring was stopped.

1.4.5. Japan

Electricity market reforms were in 1990-2000, similarly in Japan. The market structure was different from other countries. Before liberalization trends, governments or state-owned companies operated most of the energy markets. However, Japanese energy market utilities have operated privately since 1939. In 1951 one generation and nine distribution companies participated in the market. Hence, only an open market was created, and distribution companies started bidding on electricity with public auctions. Consumers started to get the right to choose their retail companies for electricity suppliers according to their consumption level. The consumption level to be eligible for choosing supplier companies decreased; in 2016, all consumers could choose their suppliers. Yearly changes in consumption limit to choose suppliers are presented below in Figure 1.7 (Shinkawa, 2018). Also, The Japanese Ministry of Economy, Trade, and Industry started applying a feed-in tariff for solar PV in 2012 to increase renewable energy resources usage. The incentives create an opportunity for foreign companies to enter Japan's energy markets. Many companies invested to benefit from the incentives, making Japan's energy market one of the biggest solar PV markets.

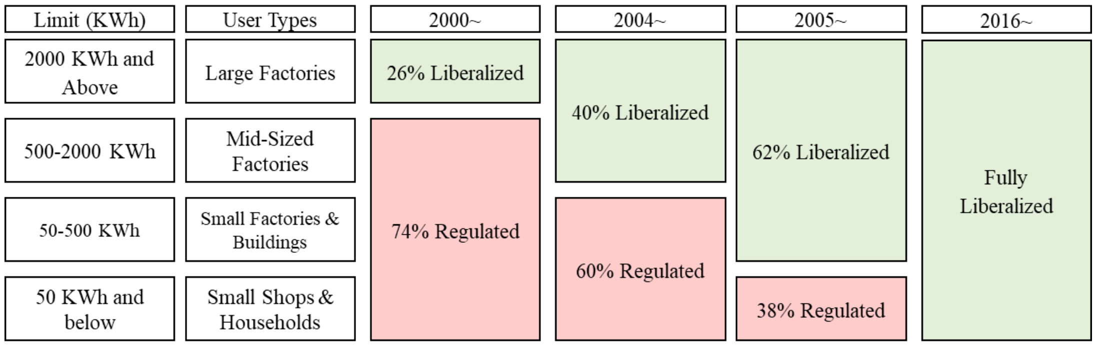


Figure 1.7: Percentage of consumer eligible to choose their supplier

1.4.6. Market Reforms in Other Countries

Energy market liberalization continued across the globe in the 1990s. Sweden and New Zealand could be considered pioneer countries in Europe. They completed the deregulation of electricity markets in 1994 and 1996. Then Norway gave customers a right to choose their suppliers and energy markets combined with Sweden and called "Nord Pool," the first multinational energy market (Flatabo et al., 2003). In the same period, England, Wales, Germany, and other European countries also restructured their energy markets to allow private companies' participation and increase competition between generators.

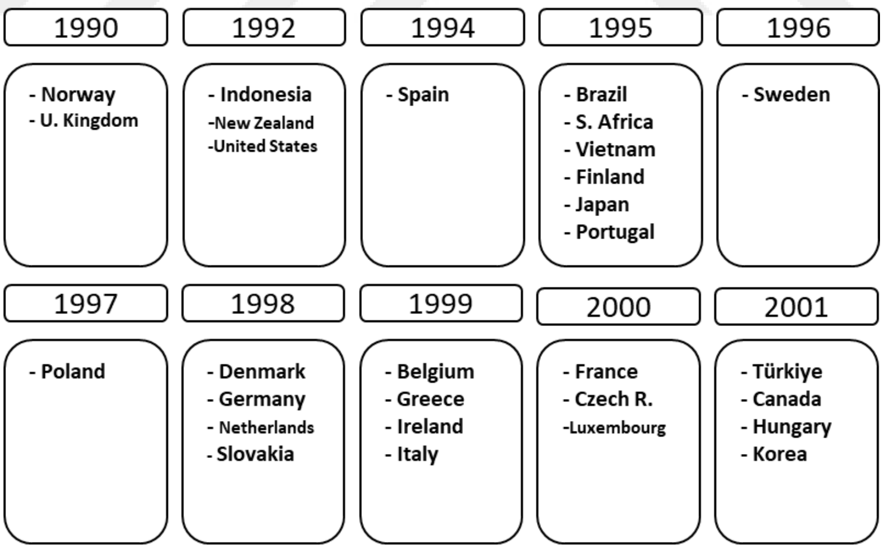


Figure 1.8: The energy market liberalization years of some countries

The reforms differ and change according to countries' existing market structures, but the policy frames and targets are similar. Liberalization trends continued in the U.S., but Asian countries did not follow this trend and continued operating their regulated markets until the 2000s. The electricity market liberalization years are given in Figure 1.8 for some countries(ALSUNAIDY & GREEN, 2006; Vagliasindi & Besant-Jones, 2013).

Despite the success story of the United Kingdom and Ireland's deregulation process, the benefits of the deregulated markets are not the same for every country (Nagayama, 2007). Although the deregulated energy markets expect to decrease retail prices and operation costs for end users with higher service levels (Erdogdu, 2011), these are not the natural outcomes of deregulation. After the deregulation of the electricity markets in Estonia, significant price rises have been observed (Vihalemm & Keller, 2016). In some circumstances, restructuring the market makes stronger existing natural monopolies. Transmission and distribution operations are mostly natural monopolies in most countries. In Chile, transmission and distribution companies increase their monopoly activities by entering the retail market to supply electricity to large customers (Palacios M. & Saavedra P., 2017). Intense monopoly activities, especially in different services, negatively affect market operations.

On the other hand, independent transmission, distribution, and generation functions cause other problems with energy reliability because electricity cannot be stored, and demand and supply should be balanced. Hence, some countries cannot prefer multiple transmission companies to control generation and supply from a single point. They also supported new-generation companies in entering the market by making long-term supply agreements. High competition creates problems in operations as well. Competitors may focus only on cost and selling price and manipulate price with changing production plans. California electricity market deregulation is commonly known as the catastrophic application of this situation. In 2000, a price hike occurred due to insufficient supply in the summertime and continued until September 2001 (Sweeney, n.d.). Many other states in the U.S. started to revise their policies for liberalization and open spot markets to prevent the adverse effects of monopolies and high competition.

1.5. Türkiye Electricity Market

The first step toward the free electricity market in Türkiye was the Turkish Electricity Authority (TEK) split, which managed the generation, transmission, and distribution units, into two generation-transmission and distribution in 1993. TEAŞ was established for electricity generation and transmission, and TEDAŞ was established for distribution. This operational structure continued until 2001. This year, TEAŞ split, and Elektrik Üretim A.Ş. (EÜAŞ), Turkiye Electricity Transmission A.Ş. (TEİAŞ), and Türkiye Electricity Trade and Commitment A.Ş. (TETAŞ) were established. The first license for electricity trade was given to TETAŞ. The Energy Market Regulatory Authority (EMRA) was founded as the regulator, and the legislation, number 4628, was enacted for the new energy market.

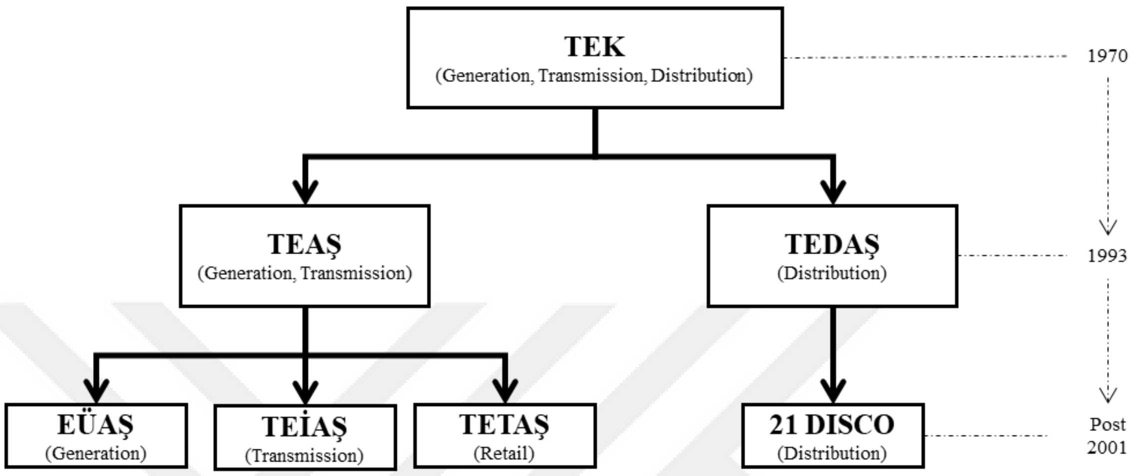


Figure 1.9: Türkiye Electricity organization changes

In 2004 and 2005, the energy sector strategy document and the law "Balancing and Conciliation" were announced. Then, balancing market and cash conciliation started in 2006 and 2007. The "Market Financial Settlement Center" (PMUM) was established to calculate settlements based on the differences between the realized purchases and sales and the contracted amounts. Then, debts or credits are announced for each market participant.

Between 2008 and 2010, ancillary services and day-ahead planning markets opened with the privatization of rivers to increase installed power. Moreover, the Türkiye electricity grid was connected to the European Network of Transmission System Operators (Entso-e) in 2011. 2011 and 2012 were busy years for setting up the free energy market structure; EMRA enacted many legislations and regulations to ensure market sustainability. The government announced improved and detailed legislation to support renewable energy investments based on the 2005 renewable energy legislation (YEKDEM).

Forward & futures electricity and intra-day markets were opened, and EÜAŞ generation capacity was handed to private companies. Finally, Energy Exchange Istanbul (EXIST) was established, and individual markets were collected under this name. In the following years, the government continued to private generations and distributions function and supported private companies to increase the number of participants in the market. The milestones and essential development steps of Türkiye Energy Markets are shown in Figure 1.10 according to the years.

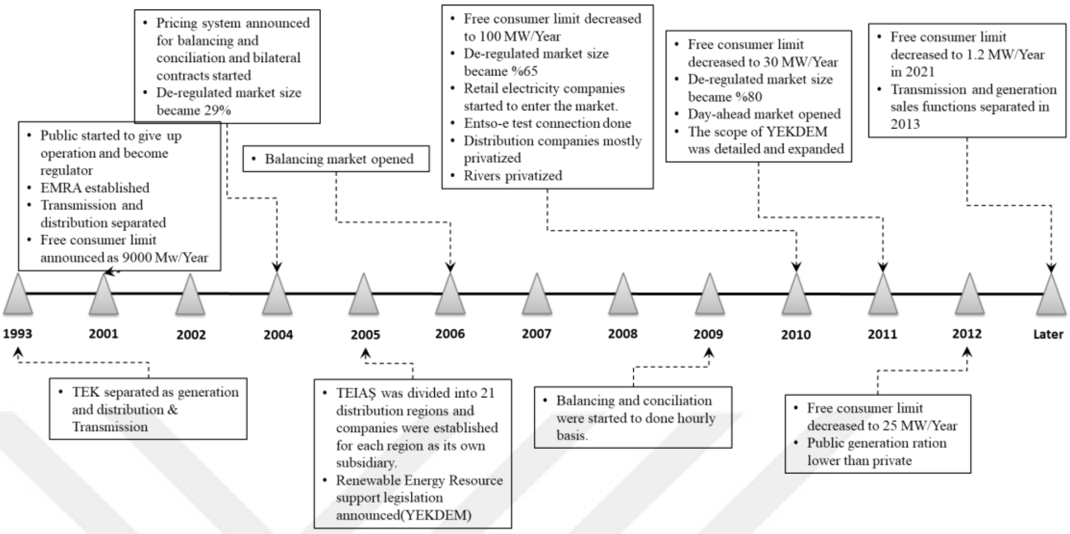


Figure 1.10: Yearly developments of the energy market in Türkiye

The electricity markets can be classified as disorganized, organized, and real-time markets. Organized markets consist of day-ahead and intra-day markets, and the disorganized market is only the bilateral contracts between suppliers and large-scale residential and industrial customers. Real-time markets balance demand and supply and ensure electricity transmission quality and transmission lines efficiency. The market structure is presented in Figure 1.11.

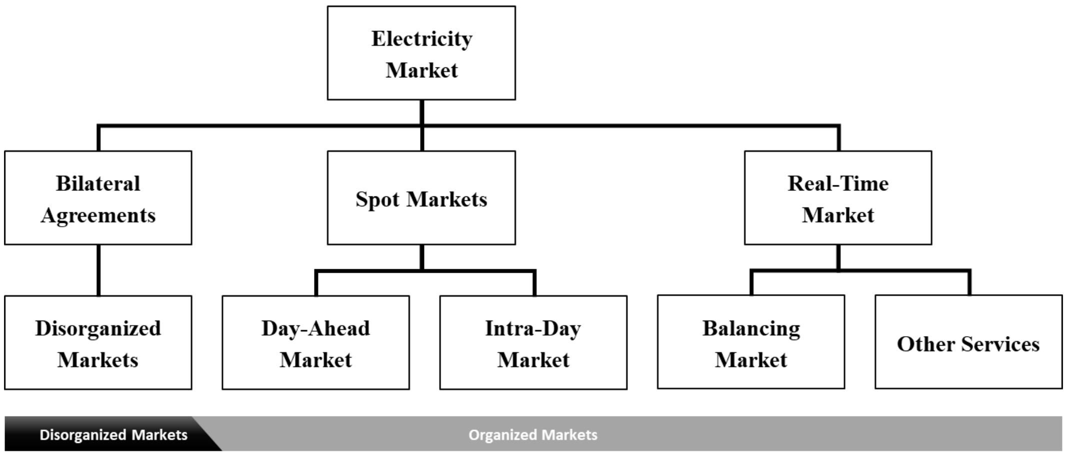


Figure 1.11: Electricity markets based on operation and organization

Each market type has different operating hours and regulations. There is no limitation to making bilateral agreements for a specific day, but spot and real-time markets have a strictly defined operation schedule. In Figure 1.12, the operation schedules of each market are shown. As seen in the figure, suppliers and customers can make an agreement anytime before the supply day. The day-ahead market starts on the previous day for the next day and continues at a particular time. According to the current operation schedules, system operators collects bid until 12:30 and announces accepted bid and equilibrium prices at 13:30. All participants can check and object to the announced quantity and prices until 13:50. The system operator finalize the following day's planning schedule at 14:00.

With the completion of the day-ahead market, the intra-day market starts and continues until two hours ahead for each hour. Participants should check their consumption/supply status and must balance it. For instance, suppliers/customers must buy/sell an additional quantity in the intra-day market if they have a supply/demand shortage or vice-versa. They have to keep their position on supply-demand equilibrium as net zero. If they have not accomplished balancing their position until two hours ahead, balancing markets make adjustments automatically by using pre-collected bids and offers for balancing. The balancing markets collected bids and offers must be activated in 15 minutes.

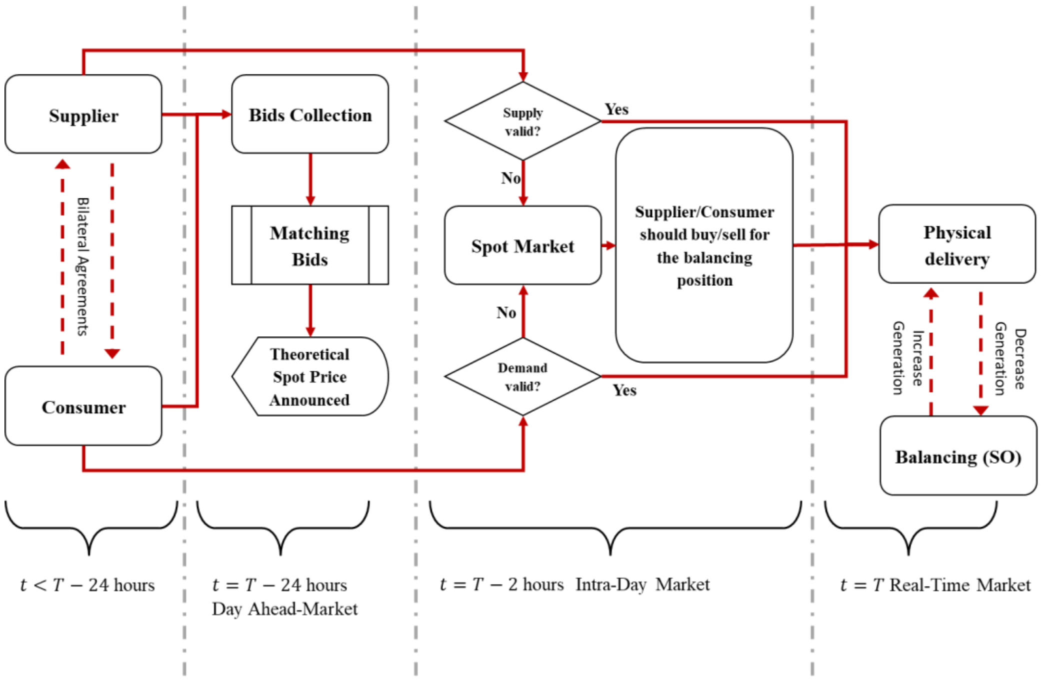


Figure 1.12: Operation structures of the Türkiye electricity market

1.6. Results of Electricity Market Liberalization in Türkiye

As highlighted before, liberalization trends affected not only the energy markets but also different markets to increase quality and efficiency by using competition between companies. In Türkiye, With the liberalization of energy markets, the number of large-scale and small-scale power plants and auto producers increased with the government's financial support. The main target is to decrease the dependency on imported energy resources that cause to higher national account deficit. There are various support mechanisms to increase competition and resource diversity. One of the critical incentives is the feed-in tariff application to increase renewable energy installed capacity and generation within the scope of national energy politics. The effect of legislation and liberalization on installed capacity is seen in Figure 1.13.

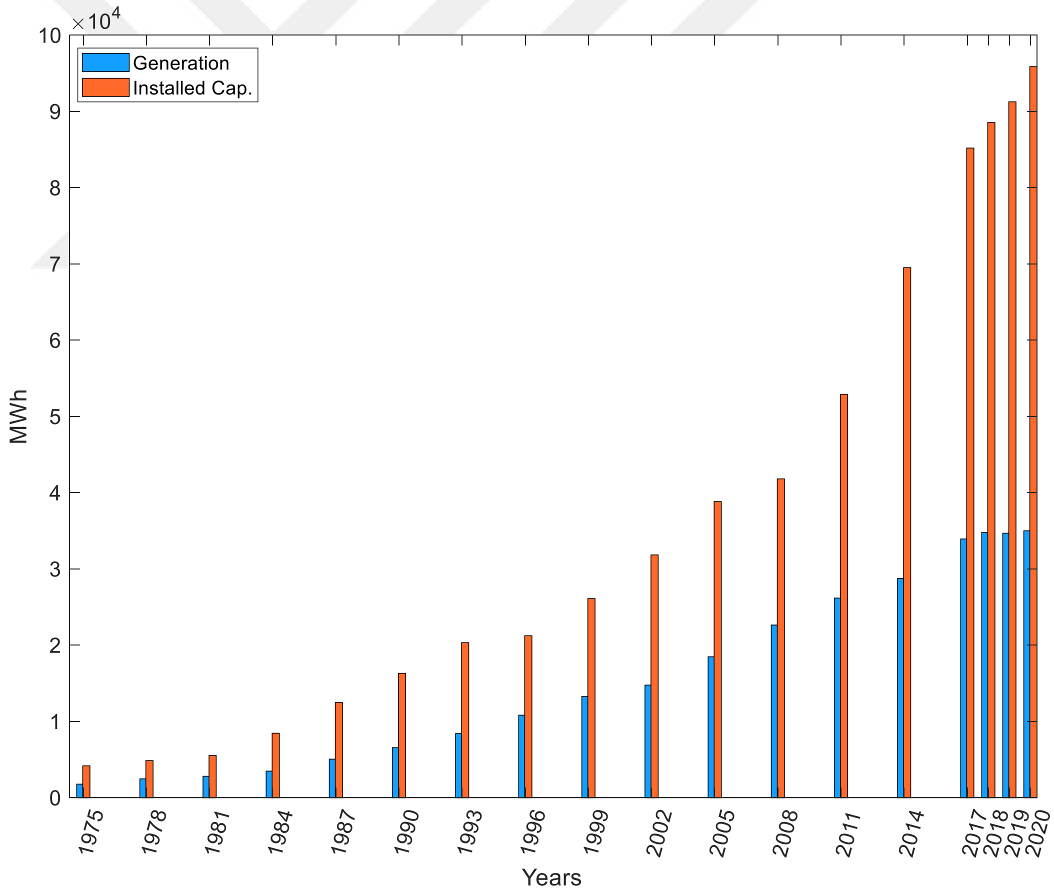


Figure 1.13: Türkiye electricity installed capacity and yearly generation changes by year

As seen in the figure, installed capacity increased higher than the average hourly generation in the last years. However, this additional capacity comes from mostly renewable power plants. The most common renewable energy resources in Türkiye are the river, hydro, wind, and solar. With recent technologies, wind and solar cannot be considered reliable energy resources due to intermittency. Wind and solar have a crucial role in reaching the zero-emission target and decreasing dependency on imported energy resources. However, energy security and reliability are vital for countries with higher industrial and population growth, like Türkiye. There has been no rapid change in installed capacity in the last three years, but wind and solar continue to increase their share in installed capacity (*Kurulu Güç Raporları*, n.d.).

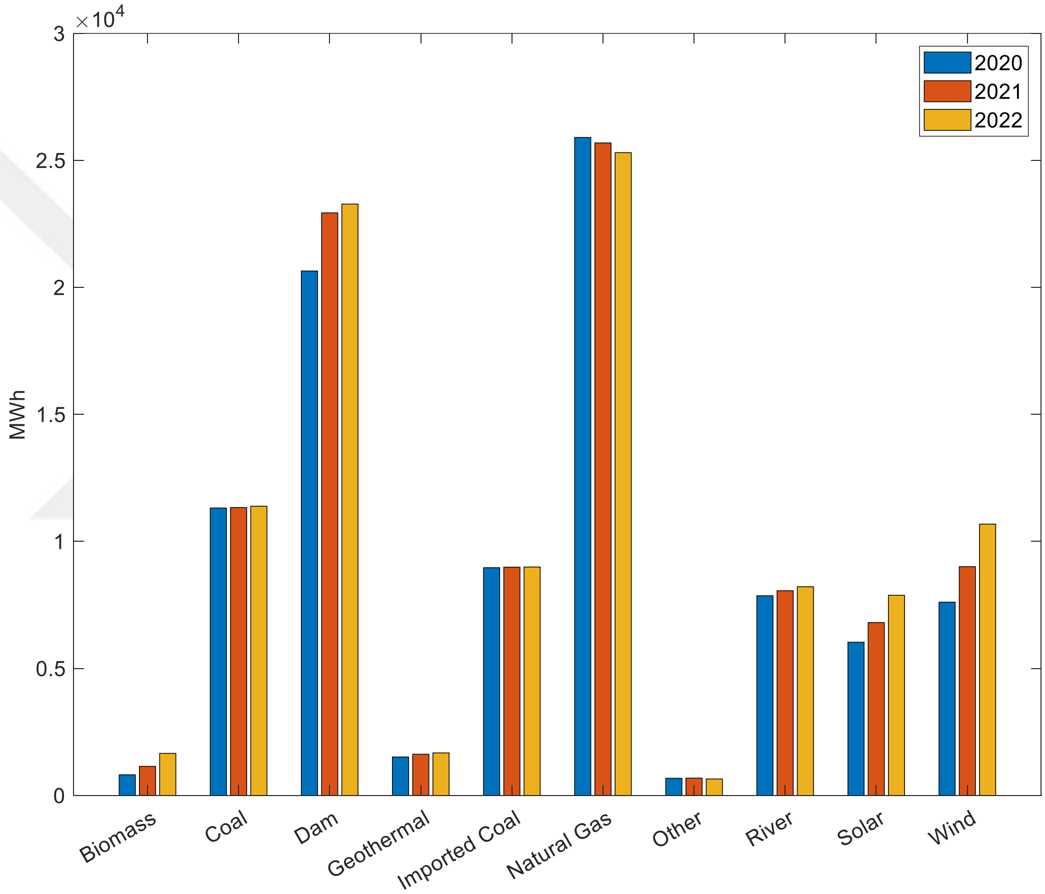


Figure 1.14: Yearly installed capacity changes by source

Figure 1.13 and Figure 1.14 show that liberalization supports investments in renewable energy resources but does not provide energy security with existing technologies. When we compare the hourly average load with installed capacity, installed capacity increases higher than consumption, but the contribution of renewable energy resources to energy security is limited. Moreover, a rapid increase in the installed capacity of renewable energy resources causes a decrease in capacity utilization, leading to an increase in energy prices (*Steady Decline in Capacity Utilisation in the German Electricity Sector*, n.d.).

Türkiye is above world average electricity generation from renewables, mainly including hydropower. However, the general renewable energy generation trend in the world gets slower due to shifting demand during the pandemic period and energy price fluctuations resulting from the political issue between Russia and Europe. As seen in Figure 1.15, the average electricity generation from renewables in the world goes up slightly on a yearly basis, but it is even around 25% in the European Union, which is the pioneer region of the energy transition from fossil fuels to renewable ones.

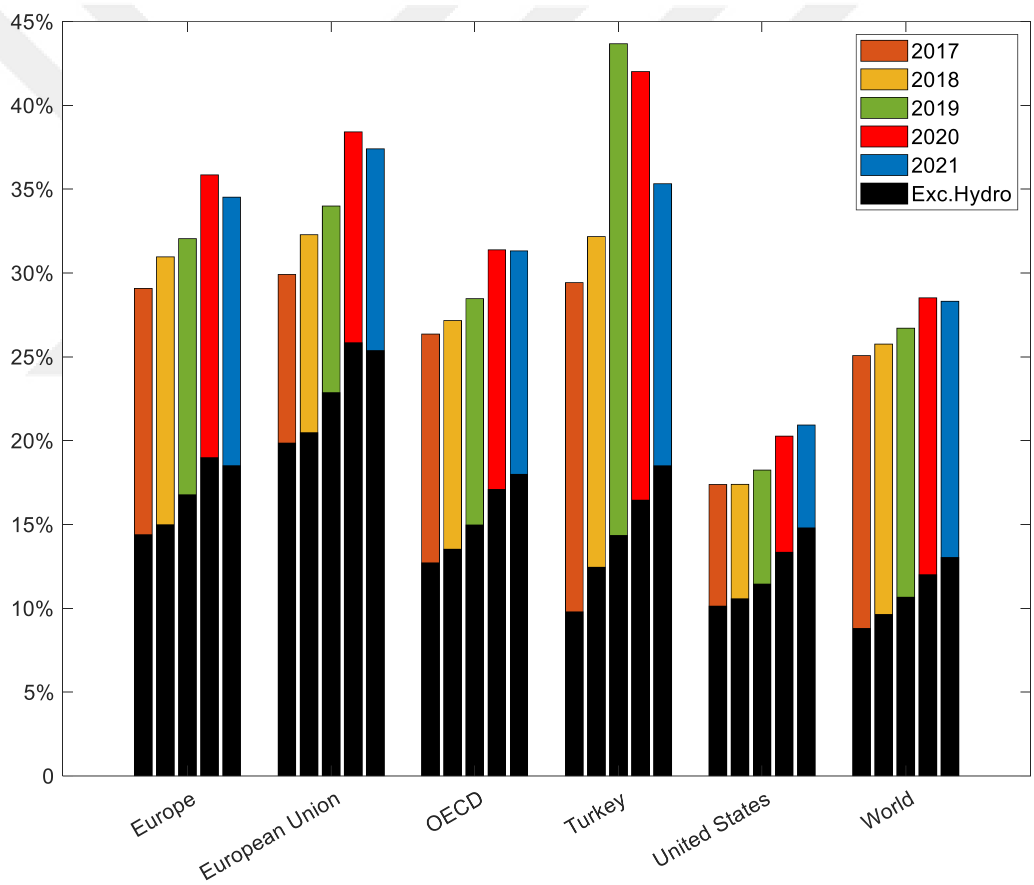


Figure 1.15: Percentage of electricity generation from renewable energy resources

As a result, European Union is the most affected region by the Russian invasion of Ukraine because of its high dependency on Russian fossil fuels. The energy prices rapidly increased with limited energy supply from Russia due to economic embargos that Europe and the United Nations applied. These political crises clearly show that the transition from fossil fuel to renewable energy resources should be planned carefully, and energy security is as essential as energy sustainability. Most European countries started activating their conventional coal-nuclear power plants again to solve the energy crisis (*Germany's Scholz Says Switch Back to Coal and Oil 'temporary' | News | DW | 16.07.2022*, n.d.).

On the other hand, with liberalization, price decrease due to high competition and diversification in energy resources is one of the expected results. Recent energy price trends still tend to get higher because of government incentives, less competition, and the intermittency of renewable energy resources. Investing in alternative energy resources could not decrease the dependency on fossil fuels globally and not provide a considerable advantage on electricity generation costs. As seen in Figure 1.16, when we checked the household electricity prices in the second half of 2021, Denmark and Germany cannot differ in electricity generation costs even though they have high renewable energy capacity (*Electricity Price Statistics - Statistics Explained*, n.d.).

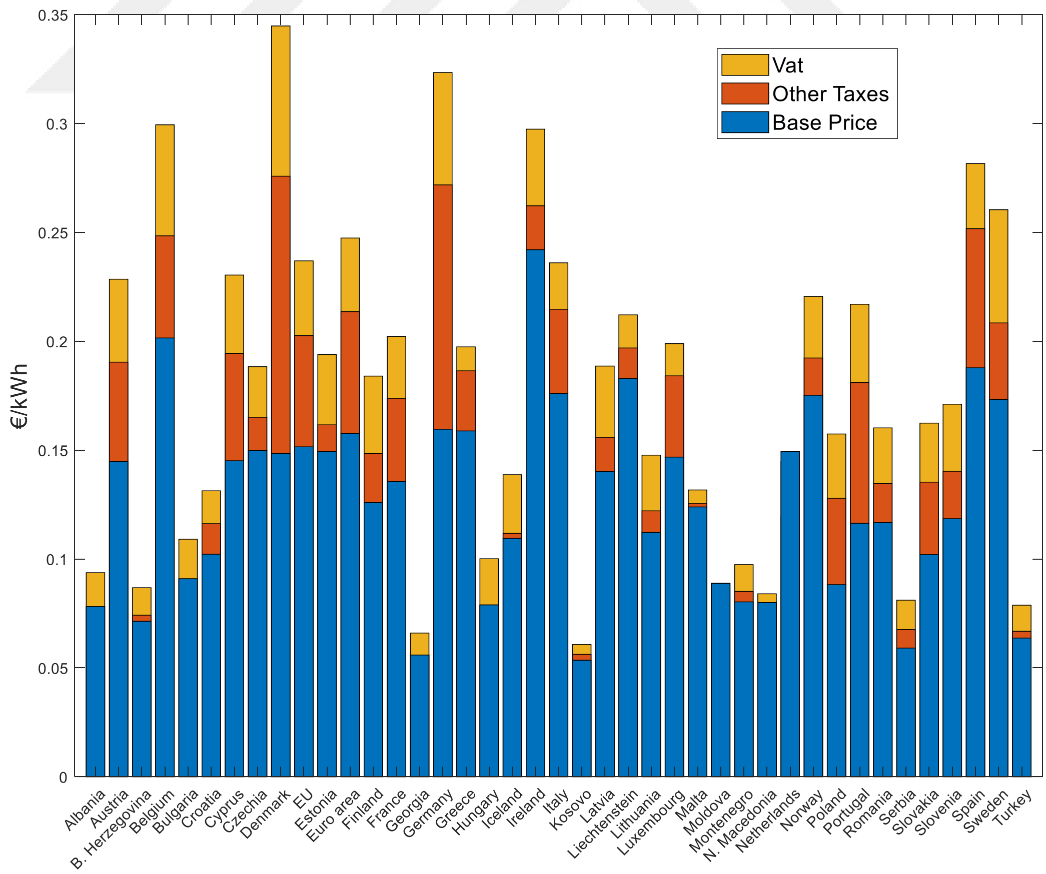


Figure 1.16: Electricity prices per kWh in the second half of 2021 by country

The trend in electricity prices is the same in Türkiye. Even though the installed capacity of renewable energy resources has increased, base electricity prices do not decrease. Figure 1.17 presents US dollar electricity price changes for residential, commercial, and industrial consumers after market liberalization. The prices were higher for residential and commercial consumers at the beginning. However, in the following years, industrial and commercial electricity prices increased more than residential because of economic problems and fluctuations in exchange rates. The government prevented residential electricity tariff increases due to political concerns and decreased taxes for residential consumers in March 2022, while commercial and industrial consumers were directly affected by the global energy crisis in 2022. In addition, electricity prices in dollars seem to decrease between 2013-2019, but this is not a result of liberalization. The main reason is that the economic policy applied by the government caused fluctuations in the exchange rate.

Figure 1.17: The electricity prices change for residential, commercial, and industrial consumers in Türkiye.

Renewable energy capacity cannot provide flexibility or decrease energy crisis effects in this period because of ineffective renewable energy mechanisms and energy security problems of renewable energy resources.

As mentioned before, Türkiye applied a feed-in tariff to support and increase installed capacity and generation share of renewable energy resources. All these buy-back guarantees were given in foreign currency. Türkiye economically is too sensitive to exchange rates due to high energy import dependency. This feed-in tariff in foreign currency causes the transfer of economic fragility to energy markets entirely. Figure 1.18 illustrates the share of energy resources in electricity generation. Türkiye electricity generation comes from natural gas and coal mainly. Both resources share in the generation is approximately 65-70% of the total electricity generation. With the liberalization of existing hydropower plants, all these renewable energy resources can benefit from feed-in tariffs in foreign currency. Incentives and support based on foreign currencies economically make 85% of electricity generation dependent on outside incidents. These cause every global political and economic problem directly affecting Turkish energy market prices. The average incentive was around 25% initially, but in 2018 and 2019, approximately 80% of an overpayment was made compared with realized spot prices due to exchange rate fluctuations (Orhan Aytaç, 2019). Both ineffective liberalization methods, global economic recession, and inflation pressure on the exchange rate will continue to become a risk Türkiye Energy Market.

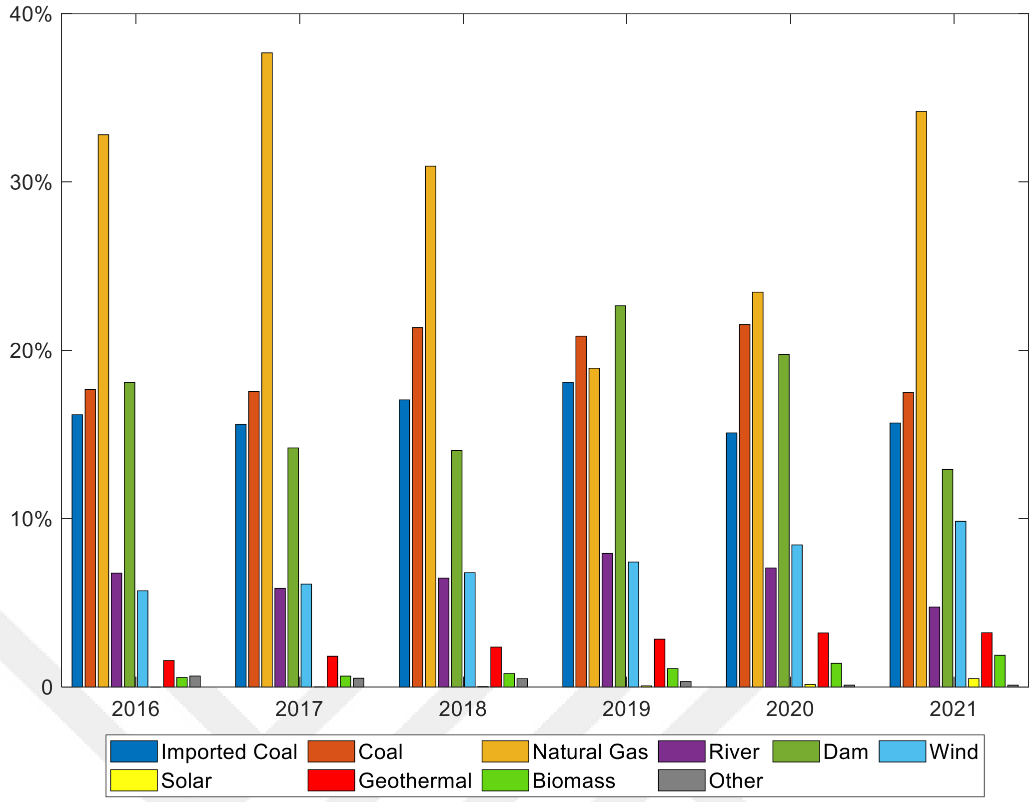


Figure 1.18: Yearly electricity generation share by energy resources

1.7. ENTSO-E Common Grid System

As mentioned in the beginning, with the development of human civilization, energy supply, supply security, environmental effects, and efficiency became challenging. Reaching sustainable growth with limited resources and decreasing adverse environmental effects have become the main target of industrialization to secure continuous growth. The European electricity transmission network is one of the largest networks and developed over a very long period and the integration of many countries. Hence, European countries are pioneers in the energy transition from fossil fuels to renewable ones. Many countries across the world follow their policies and plans.

At the beginning of the 19th century, Europe's growth rate increased energy needs. Power grids constantly expanded from the local to the international scale, and countries started power exchanges to decrease energy costs, optimize allocation, and deliver energy to more areas. France, Sweden, and Italy established one of the first joint grids in 1906 (*Cross-Border Interconnections on Every Continent*, 2010).

They started the cross-border transmission for some areas. With the joining of Belgium and Germany, they planned to expand grids to cover more areas and also efficient resource allocation (Lagendijk, n.d.). Europe's growth continued, and energy consumption increased by approximately 70 percent from 1960 to 1990 (*Global Comparison: How Much Energy Do People Consume? - Our World in Data*, n.d.).

The developments of grid expansion and energy generation continued, especially between France and Germany, due to nuclear energy capacity of France. In 1951, France, Germany, Austria, Belgium, Netherlands, Italy, and Luxemburg started synchronous grid connections under the Union for the Coordination of Production and Transmission of Electricity (UCTE). In the following years, Portugal, Spain, and the former Yugoslav connected to the UCTE, and expansion continued year by year with the joining of other European countries. However, interconnected transmission grids developed a regional basis. Especially at the end of the 90s, the Association of the Transmission System Operators of Ireland (ATSOI), the Baltic Transmission System Operators (BALTSO), and the UK Transmission System Operators Association (UKTSOA) were founded and managed on a regional basis. The Internal Electricity Market (IEM) foundation by European Union created continental transmission management that includes all European countries(*Former Associations*, n.d.). To accomplish this, in 1999, the European Transmission System Operators (ETSO) was founded 1999 to create a shared network and cross-border electricity trade, which also conforms with the competitive market concept. In 2001 the number of member countries reached 15, with 32 independent European Transmission System Operators (TSO).

With the success of the ETSO in operation and compensation, in 2008, 40 TSOs signed a protocol to expand this shared network. At the end of that year, European Transmission System Operation for Electricity (ENTSO-E) was founded. All ETSO activities were transferred to ENTSO-E in 2009. ENTSO-E consists of five Regional Security Coordinators (RSCs), Coreso, TSC, SCC, Nordic, and Baltic, and all members operate under these RSCs. Figure 1.19 shows the territory of the RSCs and adhered countries across Europe(*Power Regions*, n.d.).

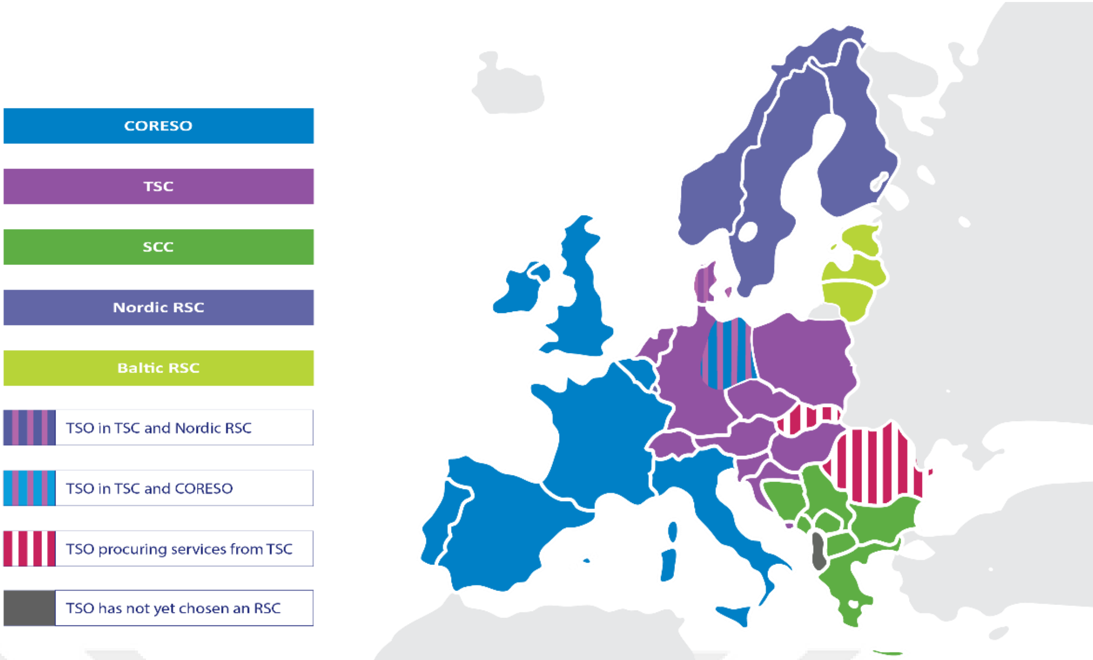


Figure 1.19: ENTSO-E RSCs and countries map (https://www.entsoe.eu)

Five primary tasks of RSCs based on the latest regulations are shown in Table 1.1. (ENTSO-E, 2019a). RSCs are responsible for supporting short- and medium-term adequacy and outage planning coordination. Moreover, capacity calculation and coordination and regional planning according to capacity potential are the tasks to secure and build a typical grid for the participant TSOs.

Table 1.1: Primary responsibilities of the RSCs and their benefits

Task	Benefits for TSOs and market participants
Regional operational security coordination	Identify operational security violations in the active planning phase. Identify the most efficient remedial actions and recommend them to the concerned TSOs.
Regional outage coordination	Detect outage planning incompatibilities and the solutions to solve the incompatibilities.
The coordinated capacity calculation for CACM	Calculate available electricity transfer capacity across borders (using flow-based or net transfer capacity methodologies). Maximize the capacity offered to the market.
Regional adequacy assessment	Provide TSOs with short (day-ahead) to medium (up to week-ahead) adequacy forecasts to foresee possible critical grid situations and deal with these accordingly.
Building a common grid model	Provide the standard grid model for all timeframes and applications to all TSOs that an RSC serves.

1.8. Research Objectives and Motivation

As stated in the study's introduction, electricity was used only for heating and lighting purposes at the beginning. However, after the industrial revolution, it started to be used as a labor force. For this reason, the need for energy in societies has constantly increased, and a linear relationship has emerged between energy and development. The dominance over energy resources has been one of the determining factors of societies' political and economic power. Sustainable energy and energy security are vital for all societies since traditional energy sources are not infinite, and their distribution is unequal.

Efficient use of resources and correct resource planning are essential due to environmental and economic effects, as well as equal access to existing resources. Although the use of alternative energy sources has increased with developing technologies, the issue of energy security is still a critical issue for developed and developing countries. With the existing technologies, the costs of renewable energy sources such as solar, wind, and geothermal have decreased, and their efficiency has increased. However, sufficient technological applications have not yet been developed regarding planning and continuity. Today, all machines and vehicles dependent on fossil fuels are replaced with electrically powered ones. Since many countries use fossil fuels such as natural gas in electricity generation, the effect of this change is small, but breaking the direct link between fossil fuels and the end user is significant for the alternative energy transition in the future. As a result, the harmful environmental effects of fossil resources will also decrease.

The biggest expectation in this regard is the development of large-scale energy storage technologies. The development of energy storage technologies will ensure energy security and reduce the cost of supply-demand imbalances. Today, all electricity markets work on balancing supply and demand, and they are faced with high costs due to any problem in the system or a bottleneck in any energy transmission or transportation. At the same time, situations where the supply is high cause inefficient use of resources. Therefore, demand or consumption forecasting became vital for efficiently managing a competitive market. The electricity must be produced and consumed simultaneously with a limited storage capacity. In order to plan production, the demand levels should be known in advance. The fuel and operation planning of the generation source will be organized accordingly.

On the other hand, the transmission line has a limited capacity, and the power needs to be distributed to required locations using the transmission and distribution infrastructure. The electricity demand has stochastic characteristics with many unknown or uncertain parameters. The climate, work habits, daily routines of people, weekday-weekend consumptions, holiday seasons, special days, and consumption of industrial and commercial facilities must be evaluated carefully. Hence, demand segregation might be needed to characterize the demand better. The electricity demand forecasting methodologies have always been an important issue, but long-term forecasting for different time periods has not been studied extensively.

In addition, the increase in large and small-scale participants in privatized energy markets and the development of renewable energy technologies caused an increase in the volume of the intraday electricity market. However, day-ahead planning is crucial for reliable production, transmission, and distribution mechanism. Energy needs to be planned for efficient and effective use of existing resources and transmission systems. The increase in intra-day market volume directly affects the quality of day-ahead planning. For this reason, storage technologies will be a critical function in preventing imbalances and inefficiencies that may arise from the participants' desire to trade in the intraday market. The deviation between day-ahead and real-time demand leads to imbalances that cause extra costs, eventually imposed on the end consumer. In addition, to use better short- and long-term demand forecasting methodologies, an efficient mechanism can be developed to increase market efficiency and lead to better planning, eventually leading to lower costs for market participants. This dissertation's research objectives can be summarized below in light of these problems;

1. Building forecasting models for both short-term and long-term forecasts based on past data without incorporating any external information,
2. Analyzing consumption trends and proposing methodologies to classify countries into groups based on their consumption objectives and temperatures
3. Developing an innovative algorithm that effectively reduces real-time imbalance costs leads to more efficient markets and increases social benefit.

The methodology part of this study consists of chapters 3,4 and 5. After giving information about the methods to be used in the second part, the data set was introduced in the third part with the applied steps in making the data set ready for the models. In chapter 4, short and long-term demand forecasting models are created for all countries, and the results are compared. The sensitivities of countries' electricity consumption in different geographies to the weather conditions are shown, and classifications are made. Chapter 5 examines the Turkish electricity market, and a new strategy is proposed to reduce costs due to the supply-demand imbalance. Simulations have demonstrated the results of the proposed strategy in different energy storage capacities. We provide discussions and conclusions at the end.

2. TIME SERIES METHODS FOR DEMAND FORECASTING

The increasing number and variety of data with the development of technology have made using traditional statistical methods essential and widespread. However, thanks to advanced computers, computer learning algorithms have become applicable, and various hybrid models have been created.

Time series analysis is also widely used because forecasting problems often involve time as well. In this way, it is used in a wide variety of areas such as sales, stock, weather, earthquake, or financial market forecasts.

Although time series methods are applicable, many internal and external factors can affect the accuracy of the forecasts or model results. Achieving a better and more reliable result remains a challenge. Many different time series methods are used, such as exponential smoothing, autoregressive methods, multiple linear regression, and artificial neural networks.

Sets of data collected with fixed time intervals or timestamps are called time series. Mathematically Y time series is a set of collected data, $Y = (y_1, y_2, y_3 \dots \dots \dots y_i)$, where y is collected data at time i. If the data occurs in fixed time intervals like monthly, daily, or hourly, that is called a "Regular time series," and if data occurs in random time intervals, that is called an "Irregular time series." Typically, time series analyses are performed on a fixed time interval assumption. However, in some variables, such as demand or natural events, data sets may not occur with fixed intervals. For this reason, studies are carried out by converting such data sets into fixed interval data sets using methods such as interpolation.

2.1. Time Series Components

Time series are basically composed of four main components: trend (T), seasonal (S), cyclic (C), and random (R) movements (residuals), as shown in the below equation. Some sources consider seasonal and cyclical movements together and call them as periodic movements. Although seasonality and cyclicality are similar concepts, they are different. Seasonality can be defined as periodic movements caused mainly by external and continuous influences, while cyclicality can be defined as changes caused by internal influences. Changes due to cyclicality are not constant in time, while seasonality is observed in the same periods. For example, for a company, the variability created by campaigns to increase sales is cyclical, but the variation in sales by month is seasonal.

$$Y = f(T, S, C, R) \tag{2.1}$$

Trend and periodic components are deterministic; after removing the deterministic part from the function, residuals remain, and the series becomes stationary.

2.2. Time Series Forecasting

There are two different time series analyses in the literature: time series forecast (TSF) and time series classification (TSC). TSF is mainly used to model how the data set will behave in the short or long term by considering each data point. TSC can be used for different purposes. It is generally used to group similar data points into different classes and to name data points in existing data. Therefore, TSC is seen as classification or grouping, while TSF is seen as regression analysis.

TSF can be used in many different areas. Some of the most common uses are language translation applications using natural language processes (NLP) and load estimation models, which is one of the topics of this thesis. This study will use TSC methods to analyze energy consumption characteristics while creating load forecasting models.

Natural language processes are a subfield of computer science. By using artificial intelligence, the interaction of the computer with the human language has been ensured. In this way, many different information and services have been revealed by analyzing human language data. The most prominent features are speech recognition, speech segmentation, automatic text completion, and translation features. The basis of all these algorithms is to guess the next word or letter in the sequence and reach the whole word or paragraph. While machine learning was widely used for NLP initially, significant progress has been made with Neural network algorithms in recent years. In particular, recurrent neural networks are widely used in character-based language models.

TSF is widely used in different disciplines to get an idea about the future. As in this study, it is used for different periods in energy demand forecasting modeling. These methods can be developed for different periods, from the shortest to the longest.

2.3. Linear Models

Linear models consist of auto-regressive models and simple moving average models applied to stationary time series. Linear equations are shown as the equation below is the most straightforward equation for the independent and identically distributed data set.

$$Y = AX + B \tag{2.2}$$

Where A and B are denoted as slope and intercept, respectively, if we define the observation set, it becomes as below.

$$Y_t = \theta X_{t-1} + \varepsilon_t \tag{2.3}$$

The equation is an example of simple autoregressive series when θ constant for every observation. Each observation is affected by the previous one with the size of $|\theta|<1$ and error term $(\mu=0,\sigma^2)$. In the time series lag operator, L is used to show previous observations. It includes basic mathematical operations like those shown in the below equation.

$$L^2X_t = LLX_t = LX_{t-1} = X_{t-2} \tag{2.4}$$

2.4. Moving Averages

Moving averages are one of the simplest time series analyses for stationary series. The method ignores the errors and calculates the next data point as the average of the past q data points. The size of the q number determines how much of the historical data set the predicted data point will be affected by. A small number of q takes into account more current data, while a large number of q takes into account more past data

$$Y_t = \sum_{i=t-q}^{t-1} \frac{X_i}{q} \quad \{\, q, t \in Z^+ \; and \; t > q\} \tag{2.5}$$

Weighted moving averages, unlike simple moving averages, are formed by assigning different weights(w) to each data point according to the characteristics of the data. These weights can be optimized using the least squares method using historical data points or added to the model as a constant.

$$Y_t = \sum_{i=t-q}^{t-1} w_i \frac{X_i}{q} \quad \{ q, t \in Z^+, t > q \text{ and } \sum w_i = 1\} \quad (2.6)$$

Another method is exponential smoothing. Brown, Holt, and Winters introduced the exponential smoothing method in 1950 (Holt, 2004; R. G. Brown, 1959; Winters, 1960). The following data point is estimated by decreasing the weights given to the historical data from the recent to the old one. In short, the closest data has the highest weight, while the farthest data from the target point has the lowest weight. This approach is straightforward and quickly applicable to various problems from different fields. There are three types of exponential smoothing methods according to their complexity. First, it is the simple exponential smoothing method, which is the simplest according to its complexity. Simple exponential smoothing applies only to series with no apparent trends or seasonality.

$$Y_t = LX_t\alpha + L^2X_t\alpha(1-\alpha) + L^3X_t\alpha(1-\alpha)^2 \ldots \quad (2.7)$$

In the general equation above, α is the correction factor between 0<α<1. it determines how the previous data will affect the result. Large α values increase the impact of current data points on the result, while small α values include a long-range data set in the calculation. It is crucial to choose the α value according to the characteristics of the data set. If the effect of sudden changes close to t is significant, the α value is close to 1, and if the long-term change of the data set is essential, the alpha value is close to zero.

Other exponential smoothing methods have also been developed to model trend and seasonality effects. Standard exponential smoothing is applied in the above equation. However, in data sets such as electricity consumption, trend and especially seasonality constitute an essential part. Seasonality, trend, and standard smoothing parameters are handled separately, and a forecast model is created from their combination (Holt, 2004). The forecast Y_{t+h} is expressed as the equation below if we name the trend, seasonality, and regular changes as b_t, c_t, and s_t.

$$Y_{t+h} = s_t + h * b_t + c_{t+h-n} \quad (2.8)$$

The smoothing coefficients for all parameters are defined separately, and all $0 < \alpha, \beta, \delta < 1$, similar to the simple exponential smoothing coefficient. These components have a recursive relationship to calculate each next Y value. The n

value in the equation refers to the length of the seasonal period. The recursive equations and components of Y_{t+h} are shown in the below equation set.

$$s_t = \alpha(x_t - c_{t-n}) + (1-\alpha)(s_{t-1} + b_{t-1})$$
$$b_t = \beta(s_t - s_{t-l}) + (1-\beta)(b_{t-1}) \tag{2.9}$$
$$c_t = \delta(x_t - s_t) + (1-\delta)(c_{t-n})$$

As can be seen from the equation, the parameter must be defined first. However, initial value assignment can lead to bias in models that give too much weight to past estimates. Therefore, the selection of the first value is essential. Various techniques have been advised for selecting the initial value and reducing the bias effects of the initial value selection on the model.

2.5. Autoregressive Methods

Autoregressive (AR) models are used to predict or model the value at time t based on the previous values of any x variable. As the name suggests, only past values are used as inputs for the model. Since it is assumed that past values contain information about all variables, so it is widely used in price estimation, especially in the financial sector. For example, only the last price changes are used to predict the price change of a financial instrument at time t; past prices include the effect of all market developments and company news. Therefore, no additional data is used.

In autoregressive models, the output values are linearly dependent on the previous values. For this reason, the equations are expressed as stochastic difference equations. The following equation shows the $AR(p)$ model, where c, φ, and $\varepsilon \sim N(0, \sigma^2)$ are the model parameters and white noise components, respectively.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{2.10}$$

This model is based on the principle of data being stationary, which means that stationary time series has a constant mean, variance, and autocovariance and change only with a time lag L. The order, p of the model, is determined beforehand according to the data set characteristic, and it is not changed (Taylor, 2010; Taylor & McSharry, 2007). The typical model for short-term forecasting is a combination of AR and MA called Autoregressive Moving Average. The ARMA method expresses a stochastic process in terms of autoregressive and moving average values

(Wilson, 2016). The $MA(q)$ process expression is shown below the equation where μ, θ, ε_{t}, $\varepsilon_{t-i}$$\sim N(0, \sigma^2)$ are the expected value of X_t, parameters of the model, and white noise terms, respectively.

$$X_t = c + \varepsilon_t + \sum_i^p \theta_i \varepsilon_{t-i} \tag{2.11}$$

Hence $ARMA(p, q)$ models are the sum of AR(p) and MA(q), and μ is usually assumed zero as shown below equation.

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t + \sum_i^p \theta_i \varepsilon_{t-i} \tag{2.12}$$

The generalized $ARIMA$ is a prevalent method in the literature. Its extensions ARIMAX (can be added external features) and SARIMA (considering seasonality or cycles) are commonly used in short-term modeling data. In the real world, load data or financial data are generally non-stationary, so $ARMA$ model can have integration factors and become $ARIM$(p,d,q), called Autoregressive Integrated Moving Average.

2.6. Fourier Series Expansion (FSE)

A Fourier series expansion is a mathematical method used to represent a periodic function as a sum of sine and cosine functions. This can be useful for analyzing the properties of the function, such as its frequency content and amplitude at different frequencies, or for approximating the function with a finite number of terms.

To create a Fourier series expansion for a function, we first take the Fourier transform of the function. This converts the function from the time domain to the frequency domain, resulting in a set of complex numbers that can be written in the form $a + bi$, where a is the real part, and b is the imaginary part. Next, we write each complex number as the sum of a sine and cosine function as $a/cos(x) + b/sin(x)$ and the resulting Fourier series expansion will be a sum of these sine and cosine functions, with coefficients that depend on the original function.

One of the main advantages of using a Fourier series expansion is that it allows us to analyze the frequency content of a periodic function. By examining the coefficients of the sine and cosine functions in the expansion, we can determine the relative amplitudes of the different frequency components in the original function.

This can be useful for understanding the behavior of a system or for designing filters to remove unwanted frequencies from a signal.

Another advantage of Fourier series expansions is that they can be used to approximate a periodic function with a finite number of terms. This can be useful for creating simplified models of a system or for analyzing the behavior of a function using fewer terms than would be required with other methods.

A Fourier series expansion is a mathematical representation of a periodic function as a sum of sine and cosine functions. It is typically written in the following form:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\, cos(\frac{2\pi nt}{T}) + b_n\, sin(\frac{2\pi nt}{T})\right) \tag{2.13}$$

where a_0, a_n, b_n are the Fourier coefficients, which are determined by the function $f(t)$. The variable t represents time, and T is the period of the function, which is the length of one complete cycle.

The Fourier coefficients can be calculated using the following formulas:

$$a_0 = \frac{1}{T}\int_T f(t)dt \tag{2.14}$$

$$a_n = \frac{1}{T}\int_T f(t)\cos(\frac{2\pi nt}{T})dt \tag{2.15}$$

$$b_n = \frac{1}{T}\int_T f(t)\sin(\frac{2\pi nt}{T})dt \tag{2.16}$$

The a_0 coefficient is the average value of the function $f(t)$ over one period, and the a_n and b_n coefficients represent the amplitude and phase of the harmonic components of the function. The Fourier series expansion is a way to represent the periodic function as a sum of these harmonic components.

In general, the more terms included in the Fourier series expansion, the more accurately it will approximate the original function $f(t)$. However, calculating the Fourier coefficients can be computationally intensive, so the number of terms included in the series is typically chosen based on the desired accuracy and computational limitations.

To illustrate the concept, consider the simple function $f(t) = 1$ for $0 \le t \le 1$ and $f(t) = 0$ for $1 < t < 2$. This function has a period of $T = 2$ and can be represented as a Fourier series as follows:

$$f(t) = \frac{1}{2} + \frac{1}{2}\cos(\frac{2\pi t}{2}) \quad (2.17)$$

The a_0 coefficient is calculated as:

$$a_0 = \frac{1}{2}\int_0^1 1\, dt = \frac{1}{2} \quad (2.18)$$

The a_1 coefficient is calculated as:

$$a_1 = \frac{1}{2}\int_0^1 1\cos(\frac{2\pi t}{2})dt = 0 \quad (2.19)$$

Since a_1 is zero, the Fourier series becomes:

$$f(t) = \frac{1}{2} \quad (2.20)$$

This means that the function $f(t)$ is constant and equal to $1/2$ over one period.

In general, the more terms included in the Fourier series, the more accurately it will approximate the original function $f(t)$. However, calculating the Fourier coefficients can be mathematically intensive, so the number of terms included in the series is typically chosen based on the desired accuracy and computational limitations. Fourier series are used in many applications, including signal processing, image and video compression, and analysis of periodic phenomena in physics and engineering

Fourier series expansion is a mathematical method named after the French mathematician and physicist Joseph Fourier. Fourier was interested in studying heat transfer and the way that heat flows through solid objects. In his work, he developed a mathematical model to describe how heat diffuses through an object over time.

Fourier’s model was based on the idea that any periodic function, such as a wave or a vibration, could be represented as the sum of a series of sine and cosine functions. He derived a set of equations, known as the Fourier series, to describe this relationship mathematically.

Fourier’s work was published in his book “The Analytical Theory of Heat” in 1822, and it quickly became a key concept in the field of mathematics. It was later applied to other fields, such as physics, engineering, and signal processing, where it is used to analyze and represent periodic functions.

Today, Fourier series expansions are a widely used tool in many different fields, and the concepts introduced by Joseph Fourier continue to be an important part of modern mathematics and science.

2.7. Neural Networks

Neural networks are a type of machine learning algorithm that is particularly well-suited for time series forecasting. They are a popular method for time series forecasting because they are able to automatically learn and identify complex patterns in the data, which makes them more accurate and efficient than traditional methods such as ARIMA.

To understand how neural networks can be used for time series forecasting, it is helpful first to understand the basics of how neural networks work. A neural network is a type of machine learning algorithm that is designed to recognize patterns in data. It consists of multiple layers of interconnected nodes, where each node represents a unit of computation. The nodes are connected by weights, which determine the strength of the connection between the nodes. To train a neural network, we feed it a large dataset of known inputs and outputs. The neural network uses this data to adjust the weights between the nodes, in order to learn to predict the output based on the input. Once the neural network has been trained, we can use it to make forecasts on new data by feeding it the input and letting it use the learned weights to generate an output. When it comes to time series forecasting, we can use a neural network to predict the future values of a time series based on its past values. This is done by training the neural network on a dataset that contains the past values of the time series, along with the corresponding future values that we want to predict. Once the neural network has been trained, we can feed it new past values, and it will generate a forecast for the future value.

There are many different types of neural networks that can be used for time series forecasting, including feedforward neural networks, recurrent neural networks, and convolutional neural networks. Each type of neural network has its own strengths and weaknesses, and which type to use will depend on the specific characteristics of the time series data. Some of the most common types of neural networks include:

- Feedforward neural networks: This is the most basic type of neural network, where the data flows in one direction from the input layer to the output layer without looping back.
- Recurrent neural networks: This type of neural network includes loops in the connections between the nodes, which allows it to process input data in a sequence or temporal order. This makes it well-suited for tasks such as natural language processing and time series forecasting.
- Convolutional neural networks: This type of neural network is designed to process data with a grid-like structure, such as an image or a video. It uses convolutional layers, which apply a filter to the input data to extract features and reduce the dimensionality of the data.
- Deep neural networks: This type of neural network includes multiple hidden layers, which allows it to learn more complex patterns in the data. Deep learning algorithms are popular for many machine learning tasks, including image and speech recognition.
- Generative adversarial networks: This type of neural network consists of two parts: a generator and a discriminator. The generator produces synthetic data, while the discriminator attempts to distinguish the synthetic data from real data. The two parts compete against each other, allowing the network to learn to generate realistic data.There are many other types of neural networks, and new types are constantly being developed. Which type of neural network to use will depend on the specific characteristics of the data and the task at hand.

One of the key advantages of using neural networks for time series forecasting is that they are able to automatically learn and identify complex patterns in the data. This means that they can handle nonlinear relationships and seasonality in the data, which are challenging to model using traditional methods. Additionally, neural networks are able to adapt to changes in the underlying data distribution, which makes them more robust and accurate over time. Another advantage of neural networks is that they can be easily fine-tuned to improve their performance. This can be done by adjusting the network architecture, such as the number of layers and nodes, or by using regularization techniques to prevent overfitting.

The number of hidden layers and neurons in a neural network can have a significant impact on its performance. In general, a neural network with more hidden layers and neurons can learn more complex patterns in the data. However, it also requires more computational resources and is more susceptible to overfitting. Therefore, choosing the right number of hidden layers and neurons is a trade-off between model complexity and performance.

One common approach to determining the number of hidden layers and neurons is to start with a simple network and gradually increase the number of hidden layers and neurons until the desired performance is achieved. This can be done through a process of trial and error, where the network is trained and evaluated using different settings, and the best-performing configuration is selected.

Alternatively, some researchers have developed mathematical methods for estimating the optimal number of hidden layers and neurons in a neural network. These methods typically involve complex equations and assumptions about the data and the network, and they may not always be applicable or accurate in practice. Ultimately, the right number of hidden layers and neurons will depend on the specific characteristics of the time series data and the desired forecast accuracy. There is no one-size-fits-all solution, and it may be necessary to experiment with different configurations to find the best-performing network.

The mathematical representation of a neural network depends on the specific architecture of the network and the type of data it is being applied to. However, in general, a neural network can be represented as a series of computational units, or nodes, that are connected by weights. Each node in the network represents a computation unit and performs a mathematical operation on the input data. This operation typically involves combining the input data with the weights of the incoming connections to produce an output value. The output value is then passed to the next layer of nodes, where it is processed in the same way. This process continues until the output of the final layer of nodes is produced, which is the neural network’s output.

The weights between the nodes are the key to a neural network’s ability to learn and make forecasts. They determine the strength of the connections between the nodes and are adjusted during the training process to minimize the forecast error. The final set of weights, once the network is trained, represents the learned knowledge of the network, and it is used to make forecasts on new data.

Mathematically, the computation performed by a node can be represented by a vector of input values, a vector of weights, and a scalar bias term. The node's output is then calculated as the dot product of the input vector and the weight vector plus the bias term. This output value is then passed to the next layer of nodes, where it is processed in the same way.Overall, the mathematical representation of a neural network is a complex and multi-dimensional system of equations that describes the operations performed by the nodes and the connections between them. It is a powerful tool for learning and predicting complex patterns in data, and it forms the basis of many machine learning and artificial intelligence systems.

This study uses the Nonlinear Autoregressive with Exogenous Input method to estimate possible energy deficiency using the main forecast results and electricity production plan. The NARX method is another type of neural network that is specifically designed for time series forecasting. It is a variant of the traditional autoregressive (AR) model, which is a statistical method for predicting future values of a time series based on its past values. The NARX model extends the traditional AR model by adding exogenous input variables, which are external variables that may affect the time series. This allows the NARX model to capture the influence of these external variables on the time series and the internal dependencies within the time series itself.

The NARX model is implemented as a feedforward neural network with multiple hidden layers and an output layer. The input layer consists of the past values of the time series and the exogenous input variables. The hidden layers are used to learn the complex patterns in the data, and the output layer produces the forecast for the next time step. To train the NARX model, we need to provide it with a dataset of past values and exogenous input variables, along with the corresponding future values that we want to predict. The model uses this data to learn the relationships between the inputs and the outputs and to adjust the weights between the nodes. Once the model has been trained, we can use it to make forecasts on new data by feeding it the past values and exogenous input variables, and letting it generate a forecast for the next time step.

The NARX model has several advantages for time series forecasting. It can capture nonlinear relationships and seasonality in the data and adapt to changes in the underlying data distribution. Additionally, the use of exogenous input variables allows the NARX model to capture the influence of external factors on the time series, which can improve the accuracy of the forecasts. However, it is essential to carefully select the input variables and the model architecture and adequately tune the model parameters to achieve good performance with the NARX model.

The history of neural networks can be traced back to the 1940s and 1950s when researchers started exploring the idea of using networks of artificial neurons to simulate the behavior of the human brain. The first neural network was developed by Warren McCulloch and Walter Pitts in 1943, and it was called the McCulloch-Pitts neuron. Over the next few decades, researchers continued to develop new types of neural networks and explore their potential applications. In the 1960s, the Perceptron, a type of single-layer neural network, was developed by Frank Rosenblatt. This was followed by the development of multi-layer neural networks in the 1980s, which paved the way for the modern deep learning algorithms that are widely used today.

In the 1990s, the field of neural networks experienced a resurgence of interest due in part to the availability of powerful computers and the development of new learning algorithms. This led to the widespread use of neural networks in various applications, including image and speech recognition, natural language processing, and time series forecasting. Today, neural networks are a key component of many machine learning and artificial intelligence systems and continue to be an active area of research and development.

In summary, neural networks are a powerful and flexible method for time series forecasting. They are able to automatically learn and identify complex patterns in the data and can be easily fine-tuned to improve their performance. While they may require more data and computational resources than traditional methods, they offer significant advantages in terms of accuracy and robustness.

3. ELECTRICITY DATA PROCESSING AND MODELLING FOR ENTSOE

Electricity is a crucial commodity, and data is collected regularly. However, the data might be in different formats and for different periods. Hence special methods need to be used for data processing and modeling. In this chapter, the data for ENTSO-E will be explored, and proposed methodologies will be discussed.

The ENTSO-E electricity load data consists of hourly consumption and covers 2006-2018 (ENTSO-E, 2019b). In order to observe the effects and analyze the relation between weather conditions and electricity load, we bought hourly weather data corresponding to each country's analysis period. To make the collection of weather data easy and consistent in the calculations, we obtained weather data only for the capital cities of each country. All data sets were checked carefully, arranged in the same time domain, and shown in the same time domain length to standardize analysis throughout the study. Missing or incorrect data on an annual base were excluded from the study. The countries list was presented in APPENDIX J.

3.1. Handling Missing Data

Data quality was checked for each country, and the whole year was excluded in case of extensive missing and inaccurate data. Data for Ukraine is excluded because the whole country's consumption is unavailable; their grid data record is divided into the western and eastern parts of Ukraine. Data for Albania, North Macedonia, Northern Ireland, Cyprus, Luxembourg, Montenegro, and Iceland are either incomplete or display irregularities are excluded. Among these, part of the data for Cyprus and Iceland will be used for specific purposes. Cyprus is a typical example of dominant summer consumption for cooling. Iceland is located at high latitudes, and it is an example of winter consumption but not for heating purposes. After the initial data control, we made systematic checks and processes for each country, as listed below;

- Combine each year's data and make single time series data with corresponding time axis for each country and some data excluded annually due to extensive missing points or inaccurate records by plotting each year for each country.

- The synchronization of all country data's time domain with the study's time domain 2006-2018; filled the empty data points with NaN.
- Some zero-value data points, due to wrong records or an hour shift for daylight saving, were changed with the closest data point of the same hour of the same day. Figure 3.1 is a sample and shows data points for Estonia we marked change.

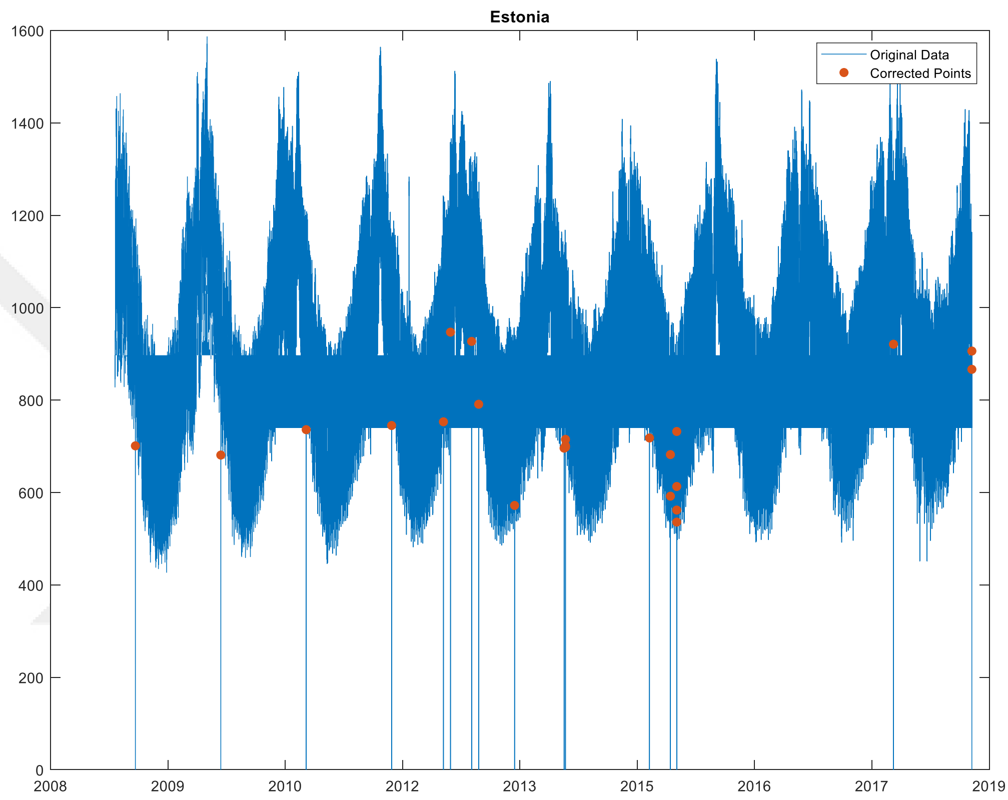


Figure 3.1: Estonia's original data and corrected ones

3.2. Main Consumption Profiles

Electricity consumption is broadly classified as household and industrial/commercial/office. Other consumption areas such as railroad transportation, street illumination, agricultural needs, and electric vehicles are not directly observable from the data; hence they are included in the basic types above. The household component consists of the consumption due to household appliances, illumination, heating, and cooling needs. Appliances work more or less continuously, but illumination needs are determined by the (deterministic) daylight cycle, depending on the latitude. The consumption for heating or cooling needs is much more complicated; it depends on weather conditions and social habits that determine comfortable temperatures and even display memory effects arising from the heating of buildings.

Commercial electricity use is mainly confined to daylight hours, but it may be effective during weekends. On the other hand, offices are usually off during weekends. Industrial electricity usage is a crucial component of non-household consumption, and it is hard to estimate because certain plants may work uninterruptedly. Nevertheless, certain countries may have holidays or vacation periods during which all (non-crucial) plants are shut down. In such cases, it is possible to estimate the share of purely household consumption from data. The relative proportions of weekday and weekend consumption are also an indicator of industrial activities. Different consumption patterns are observed when all countries' electricity consumption data are examined. The data for Sweden, Croatia, Germany, and Poland are displayed in Figure 3.2 as examples.

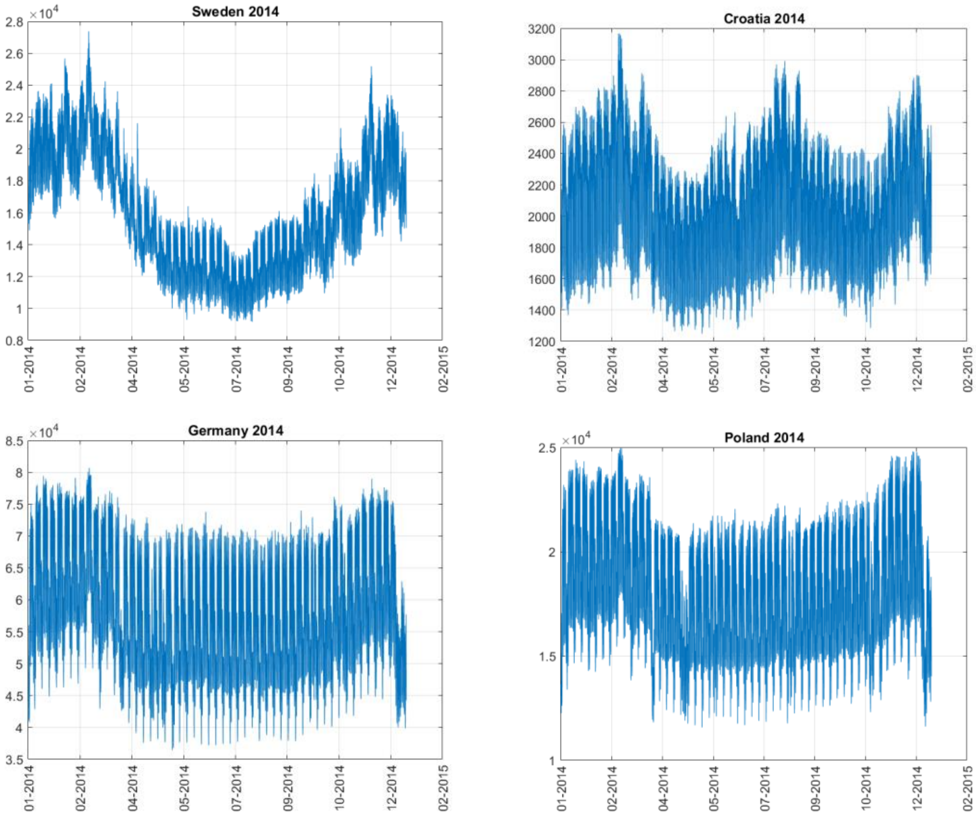


Figure 3.2: Comparison of consumption patterns for Sweden, Croatia, Germany, and Poland

The consumption of Sweden is high and irregular in winter, indicating the use of electricity for heating. For Croatia, increased summer consumption indicates cooling needs, and summer and winter consumption is higher than spring and autumn due to the same heating and cooling capacity. The seasonality of electricity consumption in Germany and Poland is less dominant, indicating a higher weight of industrial consumption. Substantial weekly variations are also indicators of industrialization. The consumption patterns for Germany and Poland are at different scales but follow a similar pattern throughout the year.

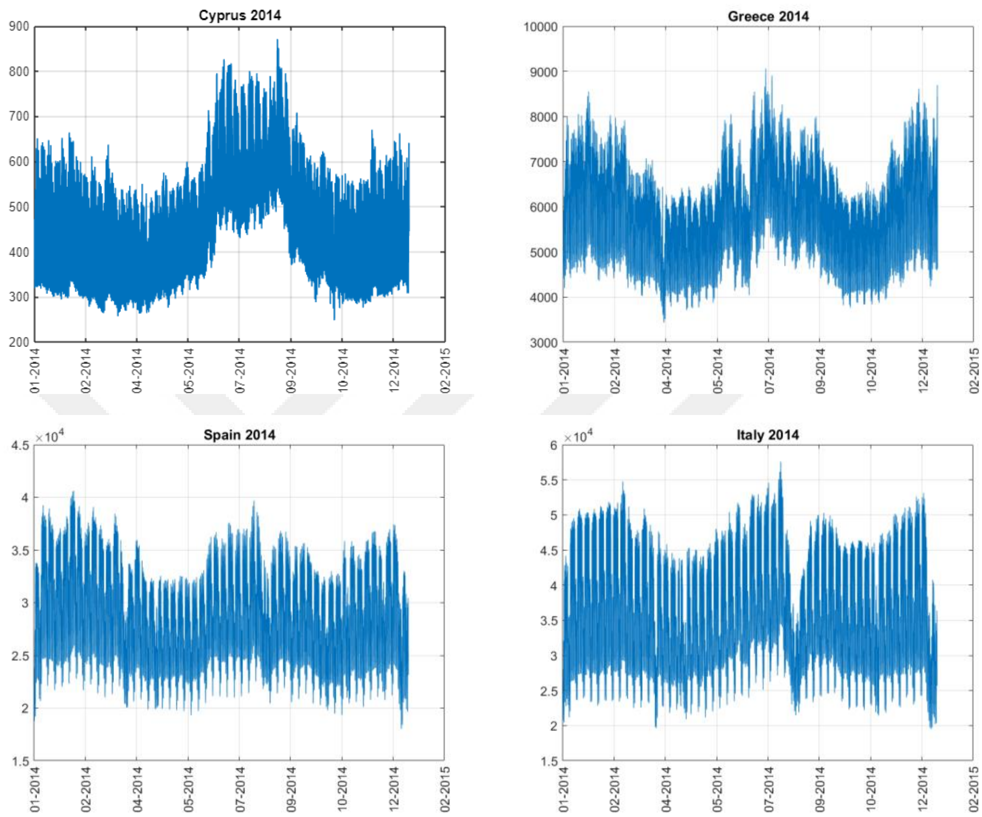


Figure 3.3: Comparison of consumption patterns for Cyprus, Greece, Spain, and Italy

Figure 3.3 presents the hourly electricity consumption of four Mediterranean countries, Cyprus, Greece, Spain, and Italy, for 2014. Cyprus has a typical Mediterranean climate with mild winters and hot summers. Thus, electricity consumption is expected to be higher in summer. The climate in Greece is more moderate than in other selected countries, and summer and winter consumptions are comparable. The climate in Spain and Italy are also mixed, and industrial electricity usage has a higher share for these countries, so the increase in electricity consumption in summer is less dominant. In Italy, the low consumption period in summer corresponds to the shutdown of industrial plants due to summer vacation.

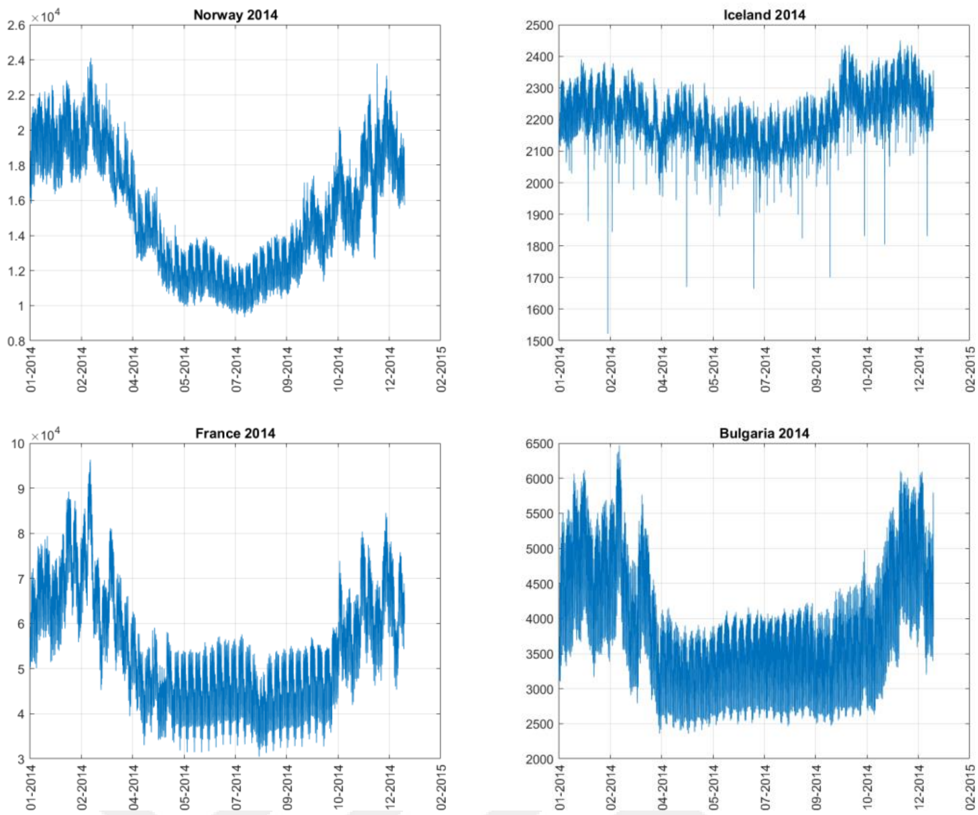


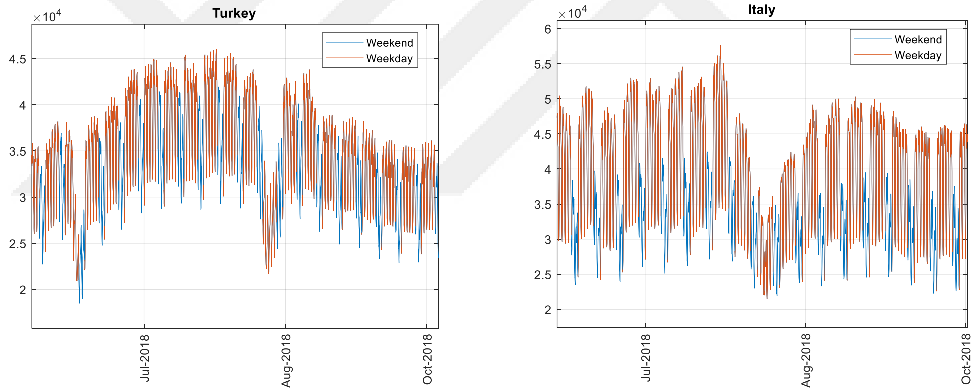
Figure 3.4: Comparison of consumption patterns for Norway, Iceland, France, and Bulgaria

Increased electricity consumption in winter is related to both the diminishing of daylight hours and the use of electricity for heating. Hence, increased electricity consumption in winter may not be the sole indicator of its usage in heating. In Figure 3.4, we present electricity consumption for 2014 in the high-latitude countries Sweden and Iceland and the mid-latitude countries France and Bulgaria. Although Sweden and Iceland are high-latitude countries, the increase in winter consumption for Iceland is moderate compared to Sweden.

3.3. Special Event & Holiday Detection

Special days and events are crucial for determining the daily base household consumption profiles, hence estimating the share of commercial and industrial consumption. National or religious holidays or annual industrial shutdowns that occur more or less on fixed dates can be considered an anomaly in data, but these are typical consumption pattern characteristics. Time series can also be used for the special day and event detection (Zhang et al., 2019). Most existing studies for detecting events focus on the end-user level consumption to explain different load profiles, energy efficiency for household appliances, and energy management (Eibl et al., 2018; Gajowniczek et al., 2017; Goia et al., 2010; Goodwin & Yazidi, 2014). There are different approaches and applied methods to detect special events. However, they mostly use smart meter data gathered from specialized devices for specific conditions or users and aim to detect occupancy (Akbar et al., 2015; Becker & Kleiminger, 2018).

Holidays and special events can be country-specific or regional, like New Year's Eve, Christmas holidays, or time changes for daylight savings throughout Europe. In Muslim countries, the dates of religious holidays shift earlier every year because they are determined according to the lunar calendar. Also, those countries with higher industrial activities have industrial shutdown periods in summer. In the below figure, we display a close-up of the data for Türkiye and Italy as typical examples. In the first one, we display the low consumption period in Türkiye, corresponding to religious holidays.



(a) (b)

Figure 3.5: Holiday consumption changes. (a) Türkiye, low consumption periods coincide with religious holidays; (b) Italy, low consumption period corresponds to the shutdown of industrial facilities in mid-summer

The dates of these religious holidays based on the lunar calendar shift back by ten days each year relative to the Gregorian calendar. In such cases, FSE fails to detect and predict these periods, and special treatment for these events is needed. The decrease in consumption in Figure 3.5 corresponds to an industrial shutdown period in Italy.

It is challenging to map days to events directly as there is incomplete information and uncertainty about the aggregate data for a specific daily consumption.

Therefore, the impact of an event on daily or hourly consumption is difficult to estimate accurately. Hence, an unsupervised model to distinguish special/exceptional events from data is proposed to determine whether a specific day is special/exceptional or not. First, the data structure is changed, and consumption patterns are evaluated for each hour individually so that the daily average does not disguise the irregularities in consumption. Generally, consumption is meager and stable at midnight compared to intraday. That is why, when the data is analyzed based on average consumption in a day, most of the special/exceptional events are not noticed.

On the other hand, when the weekly averages are considered and compared based on weekly average load data, every Sunday's load becomes special/exceptional as it is lower than that of other days in a week. The data is split into weekdays and weekends to prevent misleading results, providing a logical comparison as the same profile is expected every weekday according to consumers' daily routines. However, there are also significant differences between night and daytime consumption, preventing the correctly detecting special/exceptional events.

For each country, the time span of the data is arranged as starting at midnight hour 00:00 of a Monday of the first year until ending at hour 23:00 of the Sunday of the last year under consideration. If the data covers N weeks, this time series is arranged in the form of a matrix S_{total} with 24 rows and $7 \times N$ columns. With this rearrangement, the columns with index $6 + 7k, k = 0,1, \dots, N$, correspond to Saturdays, the columns with index $7k, k = 1,2, \dots, N$ correspond to Sundays, and the remaining columns correspond to weekdays. Then, matrices corresponding to weekdays, Saturdays, and Sundays are denoted by S_w, S_{sat}, S_{sun}, respectively. With this setup, each row of S_{total} is a time series for a given hour and S_w, S_{sat}, S_{sun} are subset of S_{total} .

Generally, consumption is meager and stable at midnight compared to intraday, as we highlighted before. Let $t_{h,d}$ be the consumption at the hour h of the day d in S_w,. In order to decide whether $t_{h,d}$ is an outlier or not, we compare it to the mean of the same hours in the days, covering two days before and two days after the current days d. The consumption ratio, $O_{h,d}$, is defined as:

$$O_{h,d} = \frac{t_{h,d}}{(t_{h,d-2}+t_{h,d-1}+t_{h,d}+t_{h,d+1}+t_{h,d-2})/5} \tag{3.1}$$

We compare an hour on weekdays within the previous and later weekdays and an hour on Saturday and Sundays within the previous or next week's Saturday and Sunday.

We choose a threshold value $T_{h,d} = 0.95$ to decide whether the hour h of day d is an exceptional hour or not. We define the index $E_{h,d} = 1$ if $O_{h,d} < T_{h,d}$, otherwise $E_{h,d} = 0$ to label each hour. Thus, depending on the threshold, each hour is classified as an "exceptional hour" ($E_{h,d} = 1$) or an "ordinary hour" ($E_{h,d} = 0$). Clearly, each day can have between 0-24 exceptional hours.

In order to decide whether a day is an ordinary or an exceptional day, we count the number of exceptional hours on that day as:

$$O_d = \sum_{h=1}^{24} E_{h,d} \tag{3.2}$$

The classification of a day as an "ordinary day" or an "exceptional day" is based on a choice of a threshold T_d. If $O_d > T_d$, then the day d is classified as an "exceptional day"; otherwise, it is classified as an "ordinary day." Night hours tend to be ordinary, even if a day is extraordinary. Based on this observation and preliminary analysis, we have chosen $T_d = 10$ means that if approximately 85% of intra-day hours are exceptional, the day also exceptional too. The threshold value can be adjusted according to the different analysis purposes. A smaller threshold can provide a detailed analysis of changes in daily demand or consumption, like power plant failure or any incident affecting national daily routines.

Finally, after determining the exceptional dates for each year in whole data on a country basis, we count and compare these dates' occurrences. If a particular day of a particular month appears consistently unusual, that date is a nationally important day or a national holiday for the country concerned. With his method, we were able to identify several exceptional days from data, including Christmas, Easter Monday, and All Saints and some specific holidays such as "Epiphany" (the sixth of every January) in Germany or "All Saints' Day" (first of every November) in certain countries. Also, it is found from the data that there is no holiday on May 1 as a "Labour Holiday" in the United Kingdom, but the "May Bank Holiday" is celebrated on the first Monday of May each year. In the below table, we presented the results for Australia. As seen in Table 3.1 fixed holiday frequency of marked extraordinary is nearly equal to the data length. Some missed ones, like National Day or All Saints' Day, are due to weekend effects.

Table 3.1: Special event detection result for Australia

Date	**Count the date when $T_d > 10$**	**Data Length (Year)**	**Day Info** (*Austria Public Holidays 2022 - Qppstudio.Net*, n.d.)
24-Dec	12	12	Christmas Eve bank holiday
25-Dec	12	12	Christmas Day
26-Dec	12	12	St. Stephen`s Day
15-Aug	11	12	Assumption Day
06-Jan	10	12	Epiphany Day
01-May	10	12	Labor Day
26-Oct	10	12	National Day
01-Nov	10	12	All Saints` Day
08-Dec	10	12	Conception Day
01-Jan	9	12	New Year

We also note that the method explained above can automatically detect special/exceptional days when such holidays are on fixed days of the calendar year or variable days, such as holidays based on the lunar calendar.

4. ELECTRICITY DEMAND MODELING AND FORECASTING

Electricity demand has an increasing trend component, climatic effects, and stochastic characteristics. A linear model of a modulated FSE was used to forecast hourly electricity demand over a 1-year horizon. This method is handy in cases where periodic variations are dominant and electricity is used predominantly for illumination, i.e., negligible heating and cooling-related demand. The method can be used to forecast and understand how electricity consumption is affected by external factors or changes (Yukseltan et al., 2017, 2020).

The modulated FSE model can be summarized as follows. A periodic function $f(t)$ with period T can represent an infinite sum of cosine and sine functions with periods T/n. Those sinusoidal functions with periods T/n are called the "*harmonics*" of the main variation. In the time series for the hourly electricity demand, the dominant component is the daily variation within 24 hours. The harmonics of this variation have periods of $12, 8, 6, and\ 24/n$ hours.

In addition to these "*fast*" variations, there are weekly and seasonal variations. The weekly variation reflects the weekend effect, i.e., industrial and office consumption shutdown. Seasonal variations have components arising from illumination, heating, and cooling needs. The change in the demand due to the changes in the daylight hours can be incorporated into the FSE by adding the "modulation" of the high-frequency variations (i.e., sinusoids with periods of $24/n$ hours, $n \in \mathbb{Z}^{+}$) and the low-frequency variations (i.e., the harmonics of the seasonal variation with periods of 365/n days, $n \in \mathbb{Z}^{+}$). This "modulation" considers the variations in the amplitude of the 24-hour variation throughout the year. It is obtained by adding the products of high- and low-frequency harmonics as regressors. The demand arising from heating and cooling needs is modeled by adding the deviations from comfortable temperatures as a regressor to the linear model. In addition to these periodic components, overall trends arise from demographics and economic growth. Finally, it should be noted that electricity demand on special days, such as holidays, has to be treated separately.

Other time-series methods, such as AR or ARMA, focus on short-term electricity consumption forecasts. However, the modulated FSE provides hourly forecasts over a long-term horizon, such as a year, within good modeling and forecast errors, particularly in cases where illumination and industrial usage are dominant.

The number of regressors to be included in the modulated FSE is limited theoretically by the "Sampling Theorem," stating that the shortest period that can be included in the expansion is twice the sampling interval, in this case, 2 hours. Hence, the number of regressors should be large enough to capture the essence of the data but should be moderate to avoid the inversion of matrices with high condition numbers.

In the model, the daily electricity demand is denoted by S. A constant vector (represented by 1) and a linear term (represented by t) are used for the linear trend in the data. Periodic variations consist of X_n (the n^{th} harmonics of sinusoidal functions with a period of one year, i.e., 364/n days), Z_m (the mth harmonics of one week, i.e., 7/m days) and of Y_k (the kth harmonics of sinusoidal functions with 24 hours, i.e., 24/k hours). The regressors that represent the modulation of the high-frequency variations (Z_m and Y_k) by the low-frequency variations (X_n) is included by the component-wise product of the corresponding vectors, denoted as X_nZ_m and X_nY_k. The number of this last group of regressors should be moderate to avoid over-learning. The effect of climatic conditions is represented by $T_\delta = \mathrm{abs}(T_c - T)$, which measures the deviation from a threshold temperature T_c, that people start to use electricity for cooling or heating. The representation of the model is as follows.

$$F = \left[1, t,\ X_1\ X_2 \ldots X_N,\ Z_1\ Z_2 \ldots Z_m,\ Y_1\ Y_2 \ldots Y_K,\ X_1Z_1 .. X_iZ_j .. X_1Y_1 .. X_kY_l\ T_\delta\right] \quad (4.1)$$

Then, the coefficient vector a and model vector y can be calculated as below.

$$a = (F^tF)^{-1}F^tS \quad (4.2)$$

$$y = Fa \quad (4.3)$$

The model is adopted for the forecast as follows. Data is split into "training" and "test" periods. Recall that the regression coefficients are obtained from Equation (18), where S is the data and F is the matrix, whose columns are the model functions, and the best fit to the data in the mean square sense is given by Equation (8). The data splitting into training and test periods corresponds to the splitting of the matrix F and the vector S as, $F = [F_1F_2]$,

$S = [S_1 S_2]$, where F_1 and S_1 cover the training period. The model coefficients are computed in terms of F_1 and S_1 as follows.

$$a_1 = (F_1^t F_1)^{-1} F_1^t S_1 \tag{4.4}$$

The forecast for the test period is obtained from the equation below.

$$y_2 = F_2 a_1 \tag{4.5}$$

The forecast error is the norm of the difference between the forecast for the test period, y_2, and the data for the test period, S_2; $|F_2 a_1 - S_2|$.

4.1. Long Term Modelling and Forecasting

We applied the model we created for all countries. While applying the model, we used 2-year hourly data as the observation period and predicted the next year according to this data.

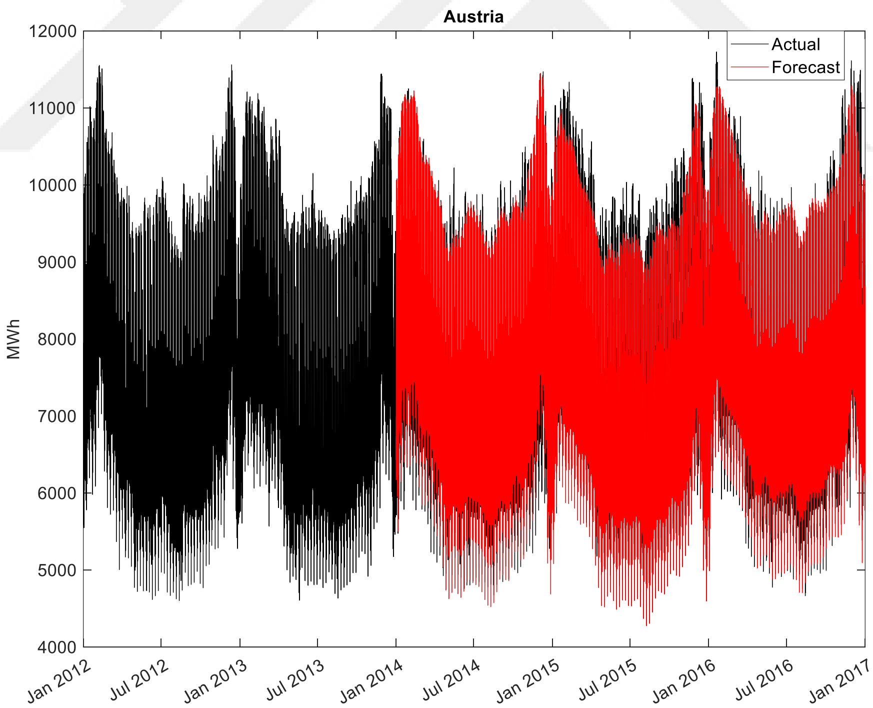


Figure 4.1: Austria's actual consumption and forecast between 2012-2017.

In short, we estimated each year based on the data of the previous two years. We can easily change the training data and estimation period with the modulated Fourier series to forecast different periods. Although one-year estimates have been compared to assess the results, this period may be extended for different purposes. Long-term forecasts play an important role in deciding the medium and long-term projections of electricity installed capacity and consumption and determining the investments to be made. As seen in Figure 4.1, we presented actual consumption and forecast result for Austria in the period between 2012-2017 years. The yearly forecast errors for all countries are shown in Appendix B.

We adopted the same method, and we updated the model parameters using data from the previous 2-year period to predict the next day and repeat this procedure for each day. The method can be interpreted as viewing data over a sliding observation window of two years (2x52x7 days) and applying the forecast with a 1-day roll-over period.

Every day, the model is updated using the latest available data. Figure 4.2 presents the day ahead forecast and the histogram of the error distribution for the day ahead forecast model.

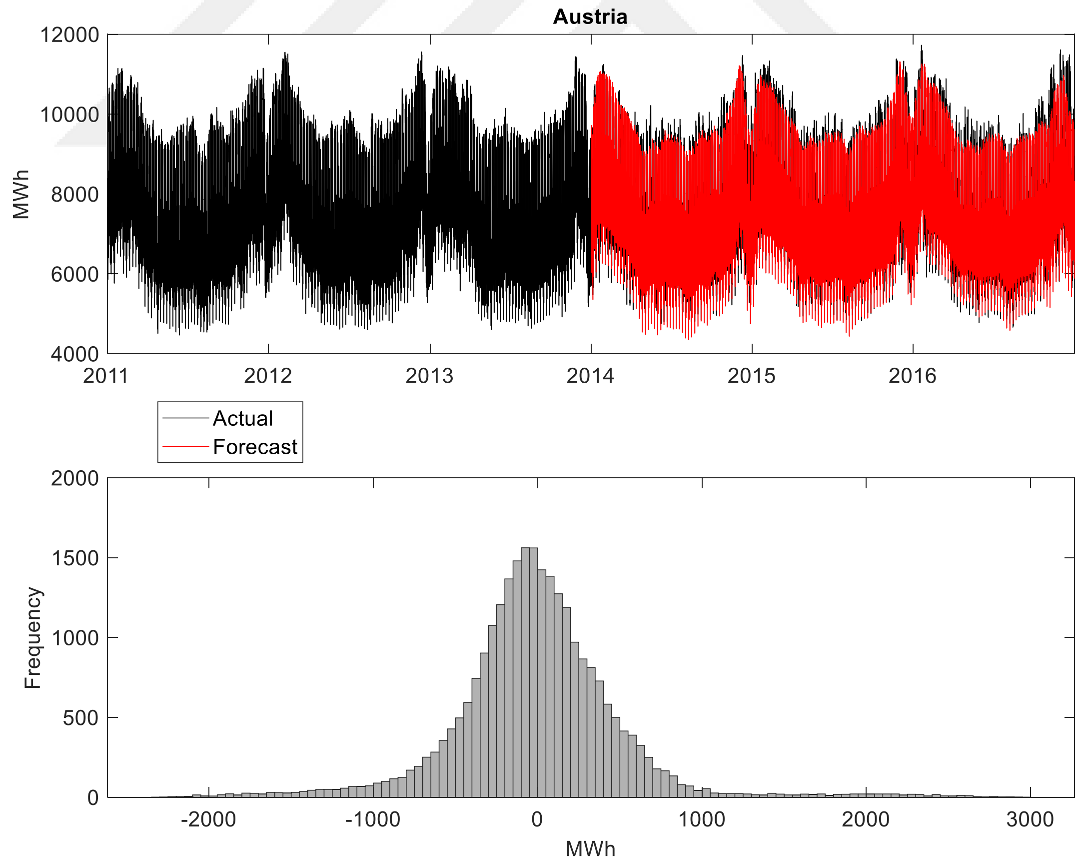


Figure 4.2: Data and day-ahead forecast for 2012-2017 (upper panel) and histogram (lower panel)

The same methodology can be used for day-ahead forecasts. MAPE values for the day-ahead forecast errors are given in Appendix C. The daily rolling forecast has a lower error rate compared with the yearly forecast results. However, these results show how consistent the results obtained for the long-term projection are and have great flexibility in making forecasts for different periods. This model can be quickly revised and applied to forecast periods such as weekly, monthly, quarterly, or semi-year.

4.2. Short Term Modelling and Forecasting

In the previous section, we presented the results for predicting hourly consumption over 1-year and 1-day horizons in modulated Fourier expansion. The actual consumption and daily forecast discrepancies are mainly due to weather conditions or unplanned extended holidays. Although it would be possible to incorporate climactic information into the model, we prefer to adopt a data-independent approach. In this section, we obtain a 1-hour ahead forecast for consumption as a correction term to the 1-day horizon forecasts. This simple approach is the "feedback" method that consists of adding a multiple of the forecast error for the previous hour as a correction term.

In order to determine the best feedback parameter, the model is run with feedback parameters in a particular range, and the forecast errors are calculated. The value leading to the lowest MAPE and RMSPE values is chosen as the feedback parameter. The feedback parameter based on the whole period is k=0.96. In order to study the time variation of the feedback coefficient, this procedure is repeated every quarter, and the resulting feedback coefficients are plotted below figure, together with their bounds. Although the coefficient change is essential, the determined 0.96 coefficient gives a reasonable result. Continuous coefficient updating or using a fixed coefficient may vary depending on the area where the results to be obtained will be used. For general planning, a quick result can be obtained using a fixed coefficient, or for more detailed planning, the coefficient can be updated in different short periods.

The correction coefficient can be applied quickly once the predicted period's actual value is formed. If we call the new forecast value $z(n)$, the feedback is expressed as follows: S and y are the observed and predicted values, respectively, and k is the feedback coefficient.

$$z(n) = y(n) + k\ (S(n-1) - y(n-1)) \qquad n > 1 \qquad (4.6)$$

Instead of the correction coefficient, a more sophisticated autoregressive model can be applied. However, determining the coefficients with the autoregressive method and also determining which periods have systematic errors requires more complex calculations. For comparison, estimation errors were also analyzed with the autoregressive method. Figure 4.5 shows partial autocorrelation graphs before and after the autoregressive model is applied.

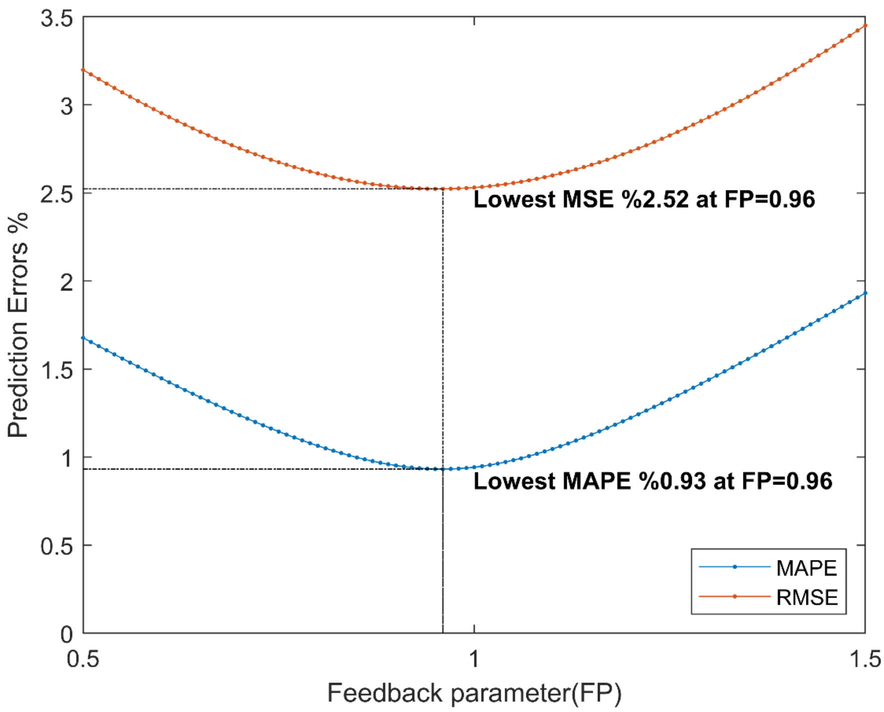


Figure 4.3: Best feedback coefficient based on six years of data

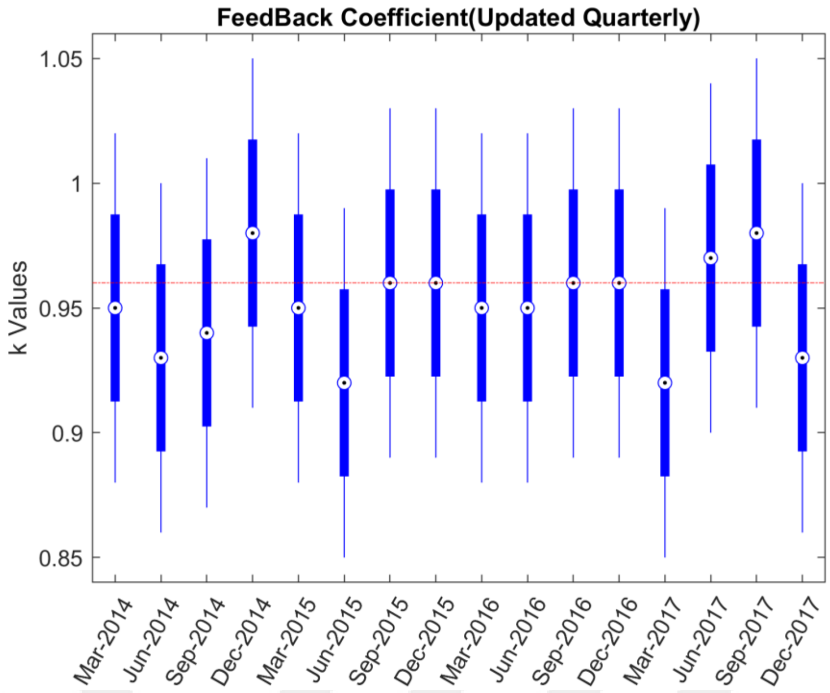


Figure 4.4: Best feedback coefficients based on 3-month periods.

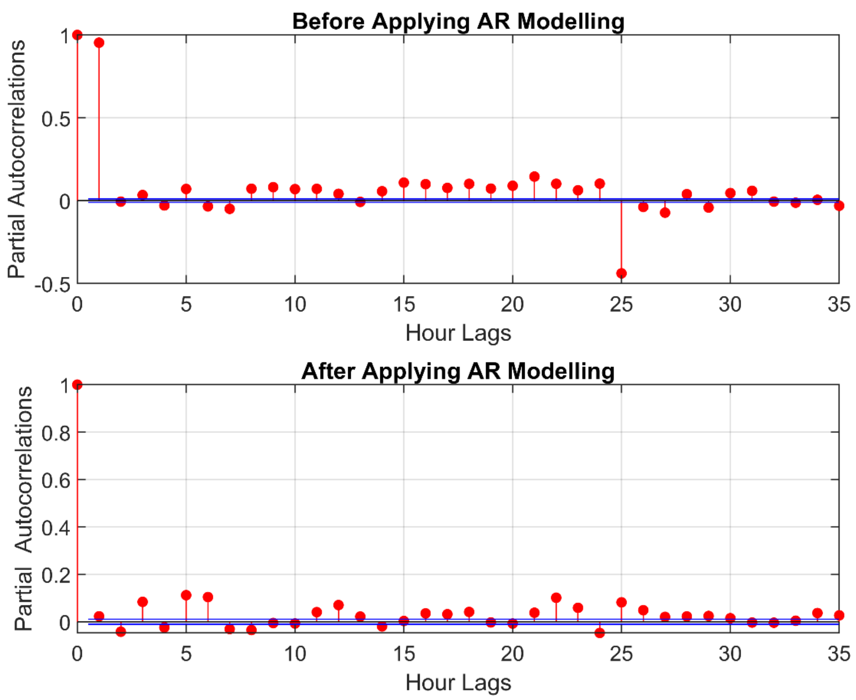


Figure 4.5: The partial autocorrelation function for the difference between data and the day-ahead forecast before and after applying AR modeling.

As can be seen from the original partial autocorrelation graph, lags 1 and 25 have a dominant effect on the latest error. Likewise, it can be said that delay number 24 has a limited effect on the final error. Therefore, we took the autoregressive modeling ($AR(p)$), p value as 25 for analysis. However, autoregressive models generally calculate the coefficient of all lags whose p value is less than itself. For $AR(25)$, values for all lags between 1 and 25 are calculated. In the analysis, only the coefficients of lags 1,24, and 25 were calculated using the MATLAB library. In the autoregressive model, the coefficient of the first lag regressor is c_1= 0.900, which is close to the value found by the best feedback coefficient, k=0.96. Coefficients for lag 24- and 25-hour are c_{24}= 0.476 and c_{25}= -0.440, respectively. The model for the forecast $z(n)$ is given by

$$z(n) = y(n) + c_1 h(n-1) + c_{24} h(n-24) + c_{25}\, h(n-25) \quad n > 25 \qquad (4.7)$$

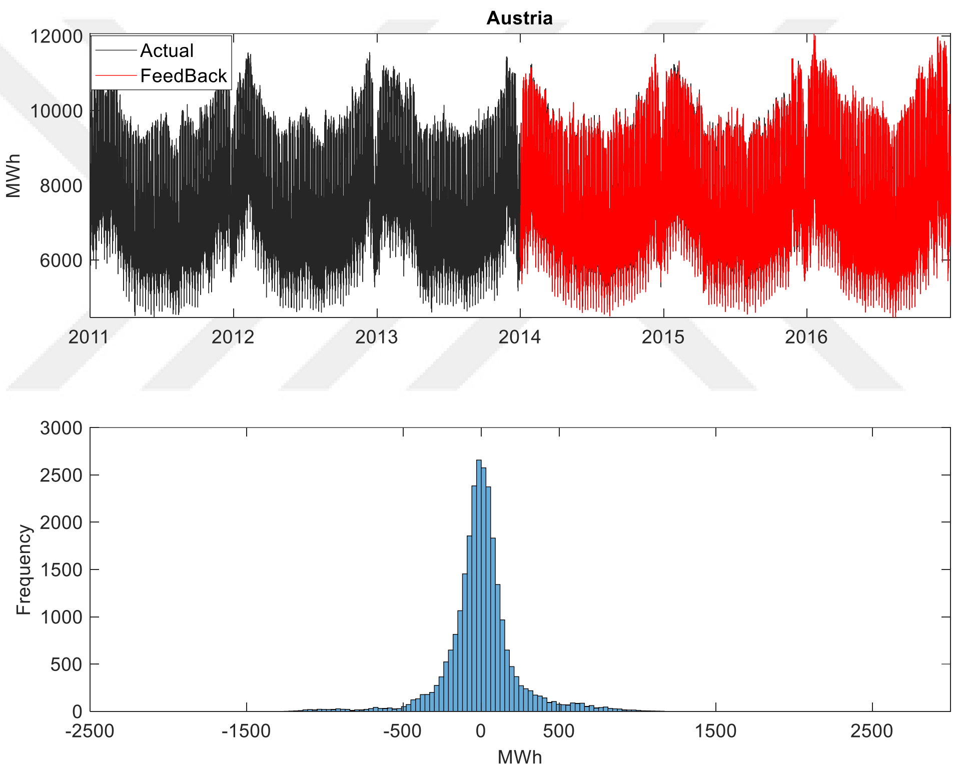


Figure 4.6: Data and day-ahead forecast with feedback for 2012-2017 (upper panel) and histogram (lower panel).

$y(n)$ is the day-ahead linear regression model, $h(n)$ differences between observation and forecast ($S(n) - y(n)$). The autoregressive model applied to the difference between the data, and the day-ahead forecast improves the simple feedback by adding correlations at multiple lags. The forecast made with the feedback and the error distribution graph is given in Figure 4.6 for Austria as an example. Although the error rate is higher when compared with the autoregressive method, as shared in Appendix D and E for each country, it can be said that the increase in the error rate is at a reasonable level according to its easy and fast applicability.

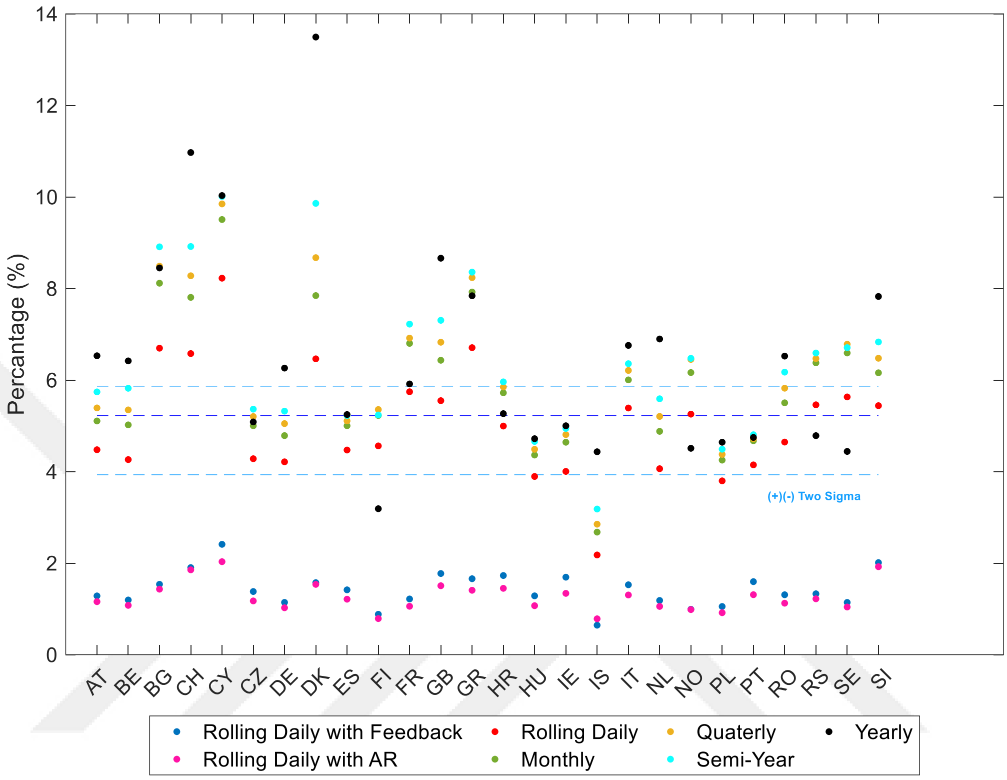


Figure 4.7: Error change for different forecast horizons by countries

In order to compare the long-term and short-term model results for different periods, the error changes of each country are plotted below. Figure 4.7 shows that the model forecast results get better for most countries when the forecast horizon gets shorter. For some countries yearly forecast performs better than the semi-year because of the dominancy of a yearly cycle. Autoregressive and feedback method results are very close to each other. A simple feedback parameter efficiently clears the systematic errors and helps obtain residuals in white noise form. Another significant property of the model, there is no high variance in the different forecast horizons. The model effectively answers forecast horizon changes within the acceptable error range. If the autoregressive and feedback correction models are excluded, we can say that daily, monthly, quarterly, semi-year, and yearly forecasts are distributed in two-sigma bounds.

4.3. Analysis of Weather Conditions on Consumption

Temperature is one of the primary external information that affects electricity consumption. The sensitivity of the consumption depends on countries' resource allocation for heating and cooling. If the electricity is used for heating, consumption during winter is high and irregular. On the other hand, if electricity is used for cooling, summer consumption is higher and irregular. Furthermore, residential use of electricity for heating or cooling may lead to a difference in the weekend and weekday consumption. In general, irregularity indicates electricity usage for heating and cooling purposes. In such cases, the data must be supplemented by meteorological information to make a reliable model and forecast. We calculated each country's average cooling and heating days to compare countries' climatic conditions. We separately calculated the hourly temperature deviation from the comfortable temperature for winter and summer. Comfortable temperatures for winter (T_w) and summer (T_s) are determined as 18.5º and 23º Celsius, respectively. There are other comments about the ambient temperature level in winter and summer, but these values are acceptable to mark the start of the heater or cooler use. We calculated the cooling (CDD) and heating (HDD) degree requirements for each hour i with the equations below.

$$HDD_i = Max(T_w - T_i, 0) \tag{4.8}$$

$$CDD_i = Max(T_i - T_s, 0) \tag{4.9}$$

Figure 4.8 presents the heating and cooling needs of the European countries in the grid. As seen in the figure, while Northern European countries such as Iceland, Norway, and Sweden have a higher need for heating needs, central European countries such as Romania and Hungary have a more balanced climate, and it can be said that their cooling and heating needs are equal. In contrast to Northern Europe, countries in the Mediterranean region, such as Greece and Cyprus, need more cooling needs. The difference in heating and cooling energy needs directly affects the electricity consumption characteristics of countries. Especially the direct and individual use of electrical energy in heating and cooling causes too much variability in consumption.

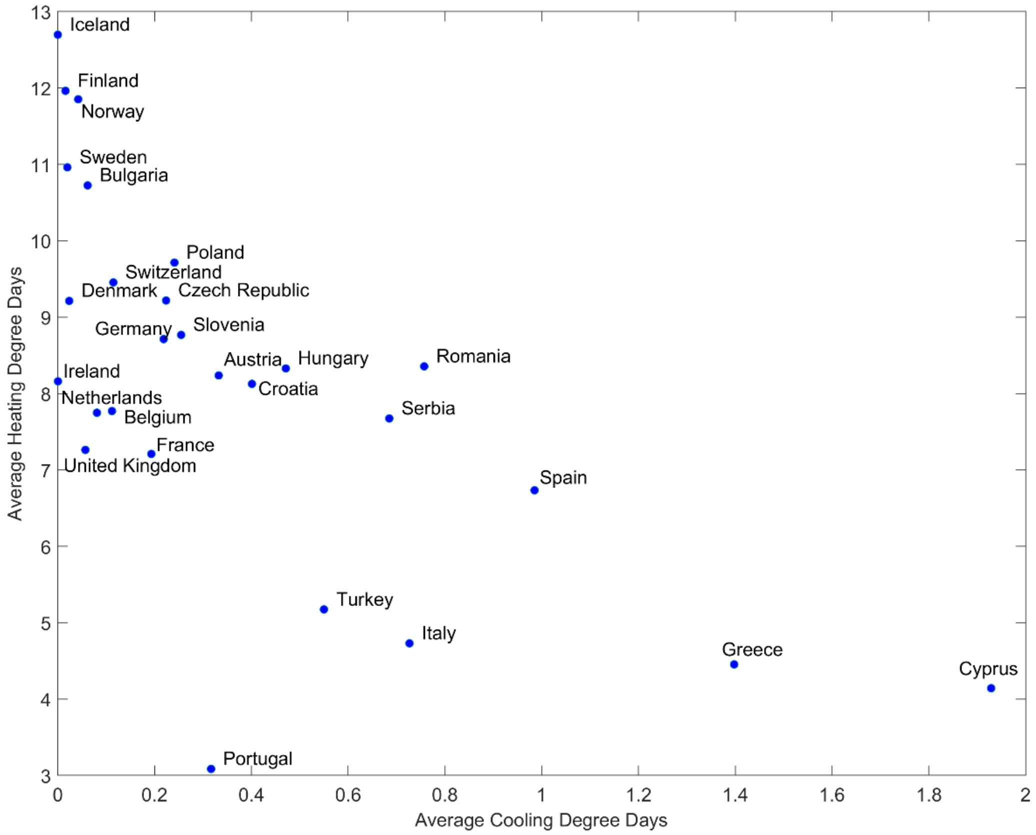


Figure 4.8: Average heating and cooling degree days for European countries in the ENTSO-E grid

We recall that electricity consumption for illumination and industrial purposes is more or less deterministic, and the modulated Fourier series expansion provides entirely satisfactory models. On the other hand, electricity consumption for heating and cooling is temperature dependent and cannot be determined solely from the periodicities in the data. Firstly the model is developed using solely modulated Fourier series expansion to detect the share of electricity consumption for heating or cooling. Then we supplemented the model with deviations from comfortable temperatures as an additional regressor. The average forecast errors range from 3% to 11% for the selected countries with different consumption patterns. The model can be used to predict hourly consumption over a year, with an accuracy of around 3% in cases where the usage of electricity for heating purposes is not dominant.

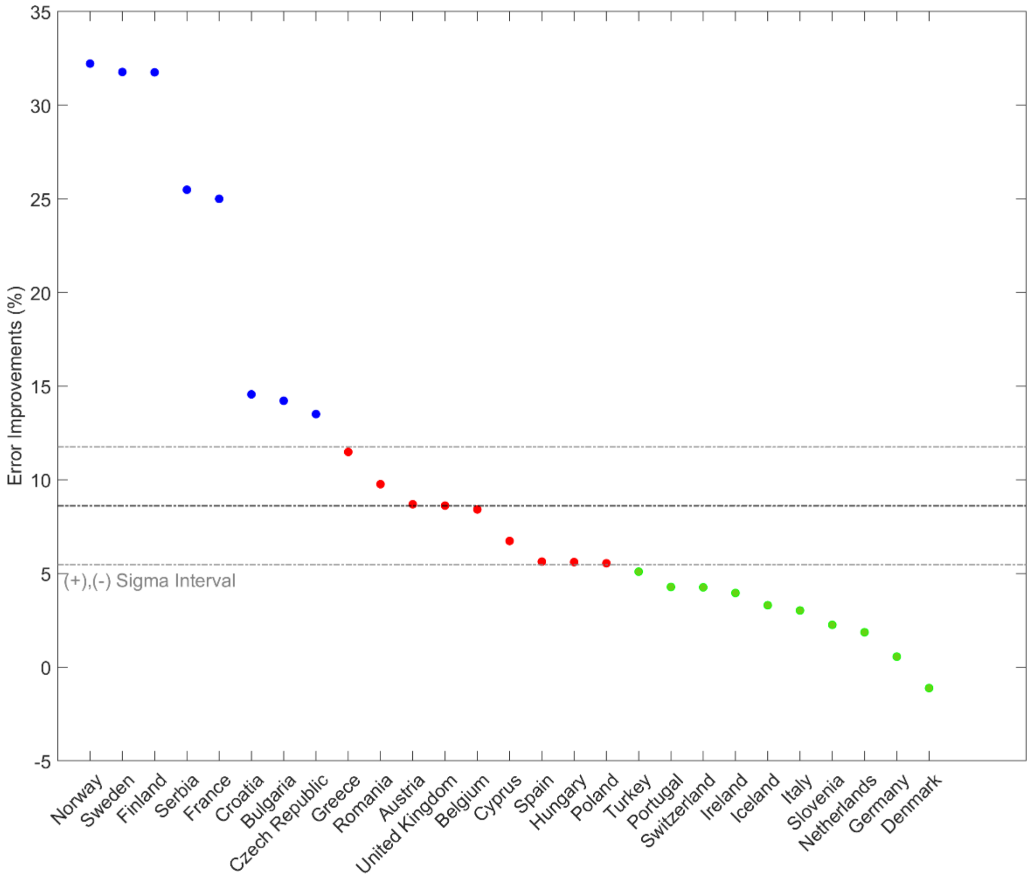
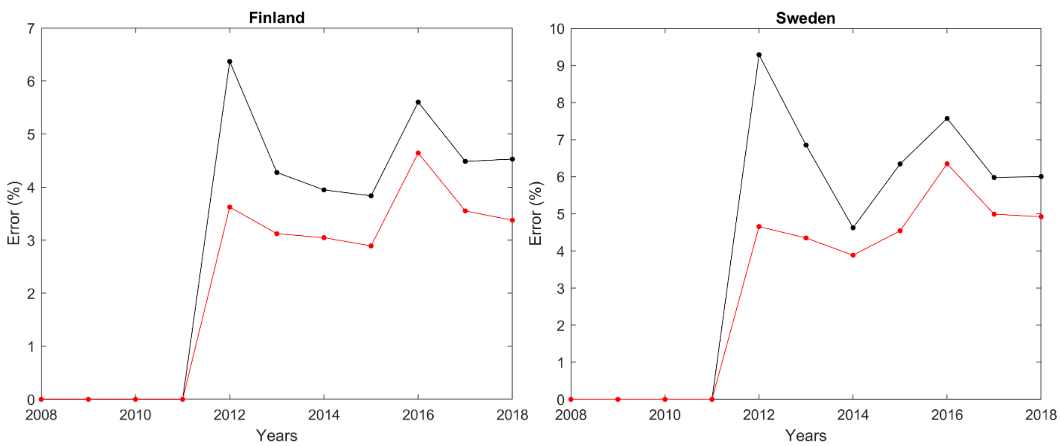


Figure 4.9: Improvement in the modeling error after adding deviations from comfortable temperatures as a regressor

The improvement in the modeling error when using temperature deviation from comfortable temperature as a regressor is shown in Figure 4.9. From Figure 4.9, one can see that Norway, Sweden, Finland, Serbia, and France form a cluster of countries for which electricity consumption is weather dependent. For the remaining countries, the improvement in the error decreases gradually. It is noteworthy that this improvement is the lowest for Germany and Denmark.



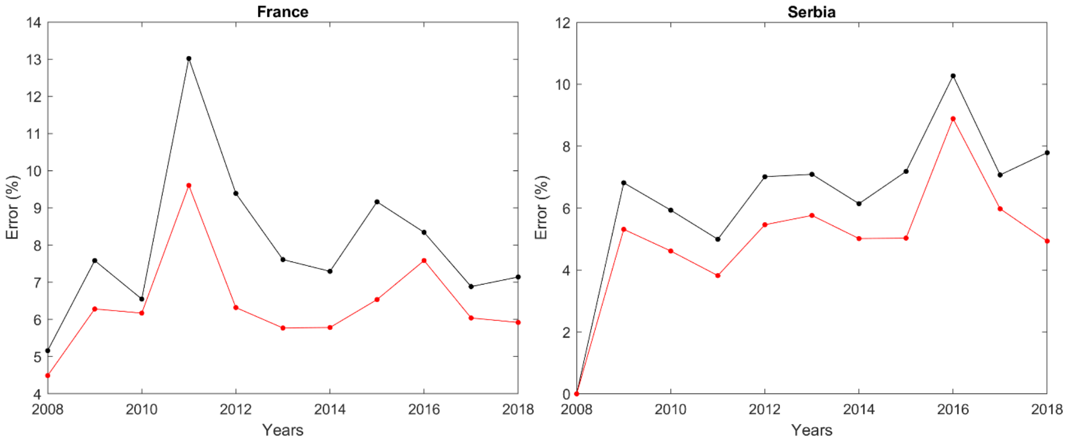
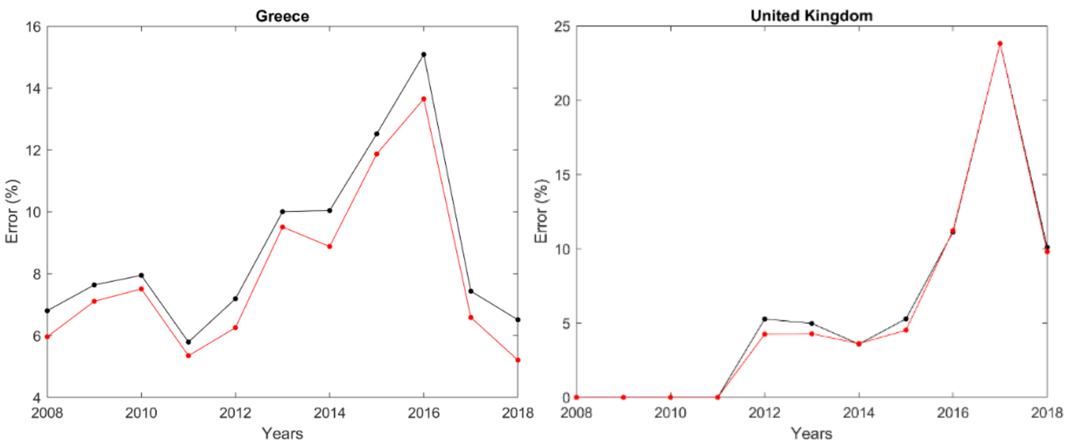


Figure 4.10: Improvement in the modeling error for Finland, Sweden, France, and Serbia (Black series: Standart Model, Red series: Using temperature deviation as a regressor)

In Figure 4.10, we present the improvement in the errors for Finland, Sweden, France, and Serbia as examples of the most significant improvement in the error. We note that a significant improvement obtained by introducing temperature information is not enough to conclude that electricity is used for heating. For example, in Sweden, electricity is not used directly for heating; there is a central heating system. The system pumps hot water, and electricity is used to operate these pumps for large regions (Werner, 2017).

Figure 4.11 shows error improvements using temperature for other selected countries, Greece, the United Kingdom, Spain, and Germany. Although Spain and Greece are in the warm climate zone, there has been a slight improvement in the model error rate. However, it cannot be said that the model in which heat is used makes a difference in Germany and the United Kingdom, which have high industrial consumption. It can be said that the effect of electricity used for heating and cooling on total consumption in these countries is limited.



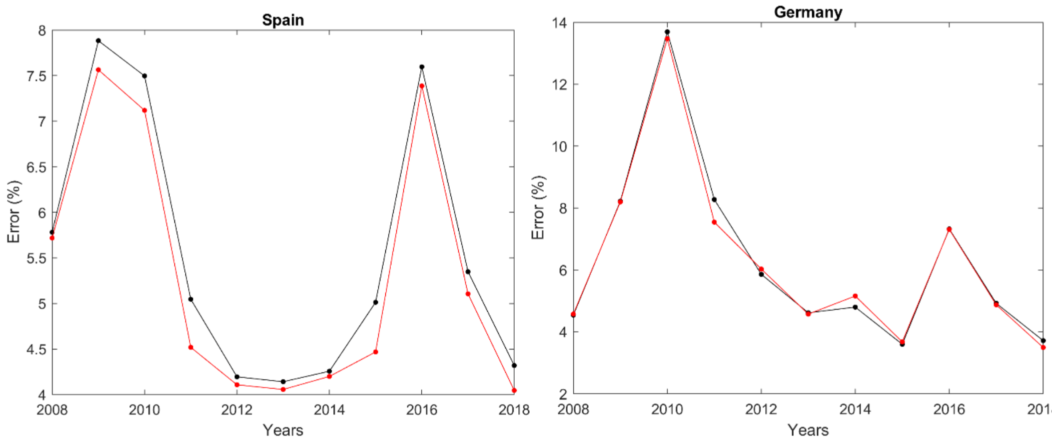


Figure 4.11: Improvement in the modeling error for Greece, the United Kingdom, Spain, and Germany (Black series: Standart Model, Red series: Using temperature deviation as a regressor)

In the previous figures, the relationship of the model created with temperature for some selected countries was given in detail by years. These figures show that Finland and Sweden are well-defined clusters of countries whose electricity consumption strongly depends on cold temperatures. The cluster consisting of Greece and Spain has low heating needs and does not benefit from additional climatic information.

In order to create a clearer picture, the model output can be used to see the relationship of all countries with temperature, and the average CDD and HDD values of all countries are calculated. The average model error improvement versus CDD and HDD are plotted as follows.

In Figure 4.12 and Figure 4.14, we see the other countries similar to selected countries, like Italy and Portugal, or see clearly how Norway, Sweden, and France's electricity consumption response to temperature changes. Iceland also shows different characteristics, even having higher heating requirements. They use district heating systems, and electricity is not the primary resource for heating in Iceland. Hence, electricity consumption cannot affect the model result much. We got this information directly from the consumption data and corroborated it from Iceland's official sources. According to the ministry of the environment and climate, 65% of total energy consumption came from geothermal resources in 2016.

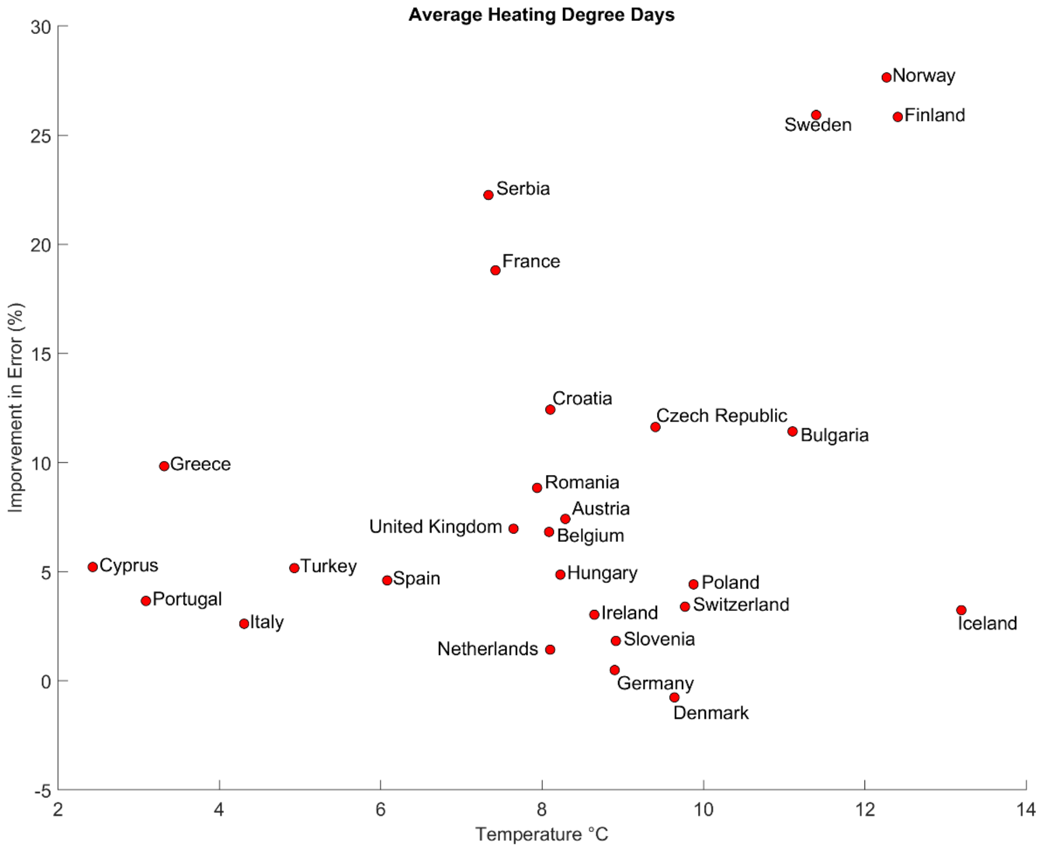


Figure 4.12: Percentage improvement in the modeling error as a function of average heating degree days

Similarly, the percentage improvement in the error as a function of average cooling degree days is shown in the figure below. This graph shows data for Cyprus and Türkiye as examples of Mediterranean countries with higher cooling needs. In this figure, Cyprus and Greece appear as a cluster of countries with high cooling needs and benefit from introducing weather information. The cluster in the upper left, consisting of Norway, Sweden, and Finland, significantly improves the modeling error as previously. However, as they have low cooling needs, this improvement cannot be tied to information about high temperatures.

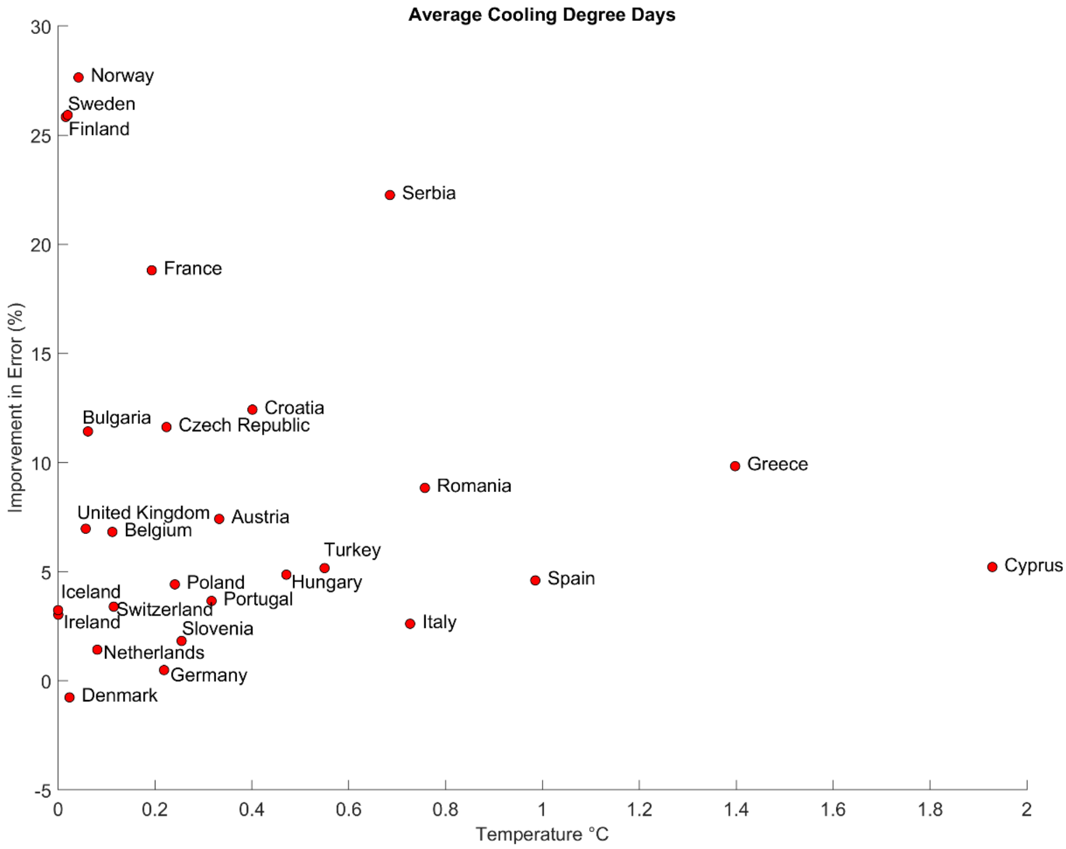


Figure 4.13: Percentage improvement in the modeling error as a function of average cooling degree days

As a further confirmation of the use of electricity for heating or cooling and deciding whether the improvement obtained by introducing temperature information is related to high or low temperatures, the scatter plots of the electricity consumption together with a piecewise linear fit to the data are given.

The following figures (Figure 4.14-40) present the scatter plots of the mean hourly consumption for Cyprus, Spain, Germany, and Finland, respectively, for weekdays and weekends, for 2016-2018. We also note that these scatter plots indicate that each country's threshold for “comfortable temperatures” differs. In fact, these graphs show at which temperature level the electricity consumption characteristics of the countries change. This ensures that the comfort temperature, which we previously took as the same for all countries, is calculated differently for each country.

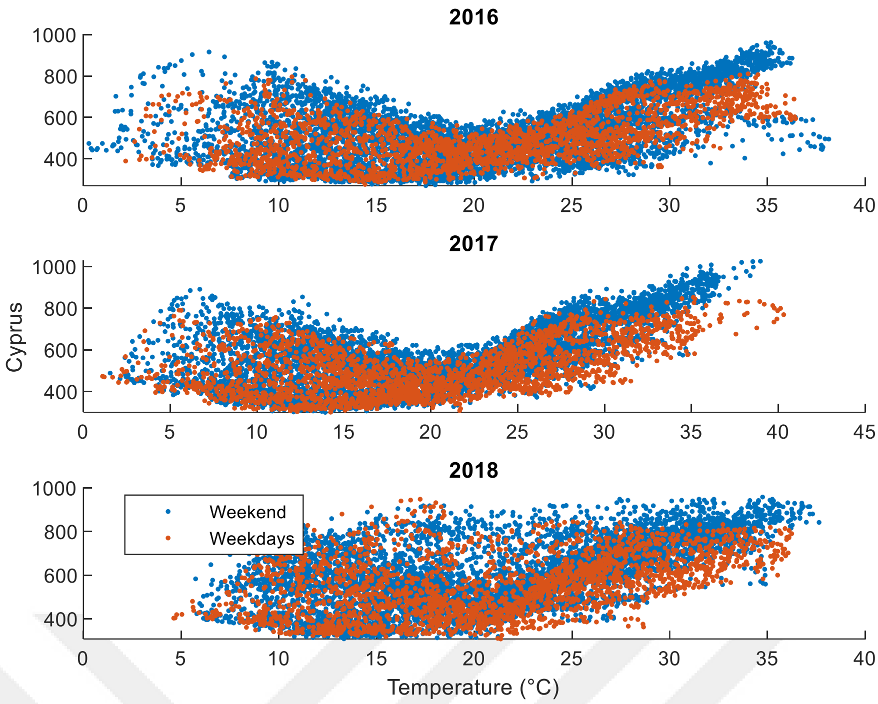


Figure 4.14: Mean hourly consumption (MW) with respect to temperature for the years 2016-2018, Cyprus

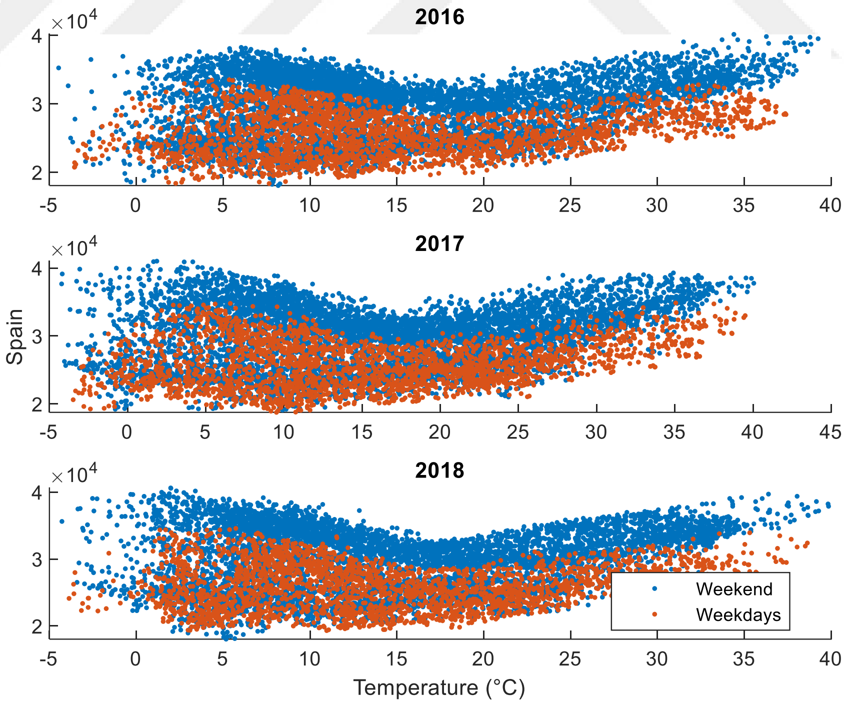


Figure 4.15: Mean hourly consumption (MW) with respect to temperature for the years 2016-2018, Spain

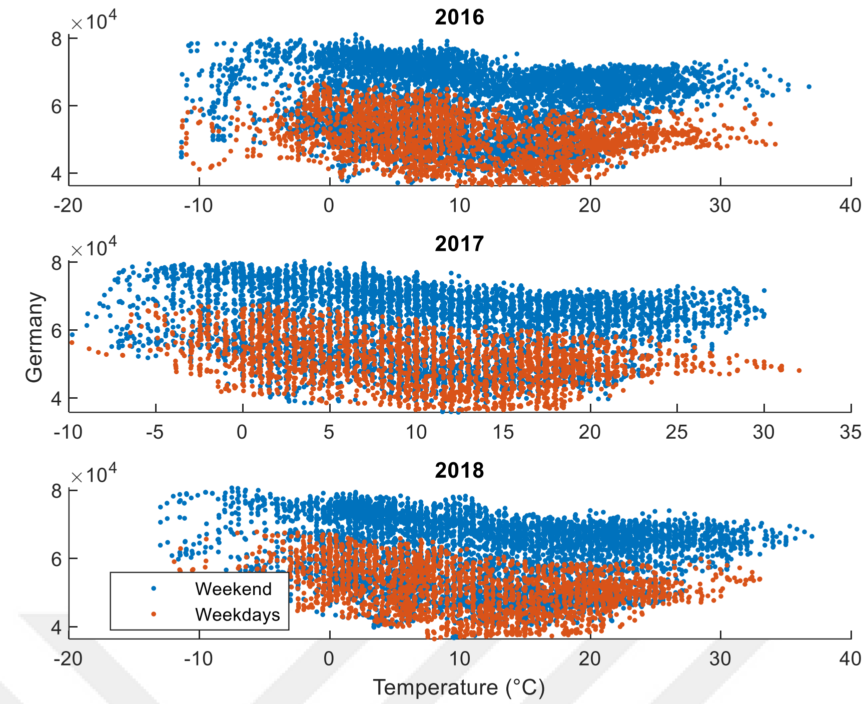


Figure 4.16: Mean hourly consumption (MW) with respect to temperature for the years 2016-2018, Germany

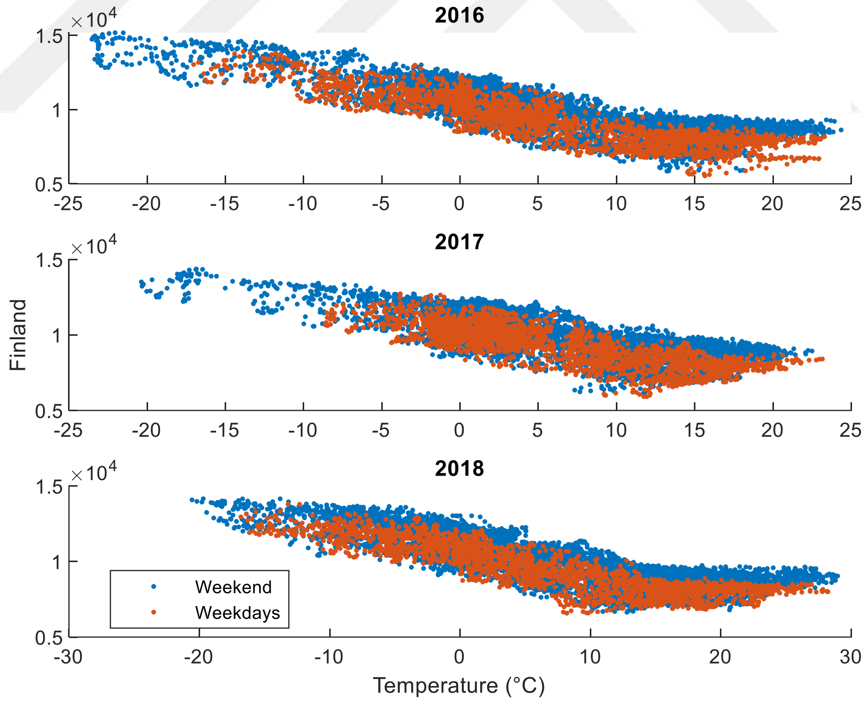


Figure 4.17: Mean hourly consumption (MW) with respect to temperature for the years 2016-2018, Finland

These figures show that in Cyprus, it is clearly seen that electricity is used for cooling after a threshold slightly higher than 20º Celcius. The relatively low difference in consumption between weekends and weekdays indicates a low level of industrialization. For Spain, the mean hourly consumption as a function of temperature indicates a moderate use of electricity for cooling. The difference between weekend and weekday consumption is a good example of a higher level of industrialization. In Germany, another excellent example of higher industrial consumption, electricity is used moderately for heating and cooling. However, industrial consumption dominates because the maximal consumption is almost independent of temperature, particularly during weekdays. The data for Finland shows that electricity consumption increases linearly with the absolute value of the difference from temperatures around 10º Celcius.

In order to quantify the electricity consumption dependence on heating and cooling, we used a piecewise linear fit to the scatter plot of the electricity consumption as a function of temperature. A one-dimensional optimization estimates the junction of the lines with positive and negative slopes by minimizing the least-squares error of the regression lines. Typical results for the year 2014 are shown in Figure 4.18 for selected countries.

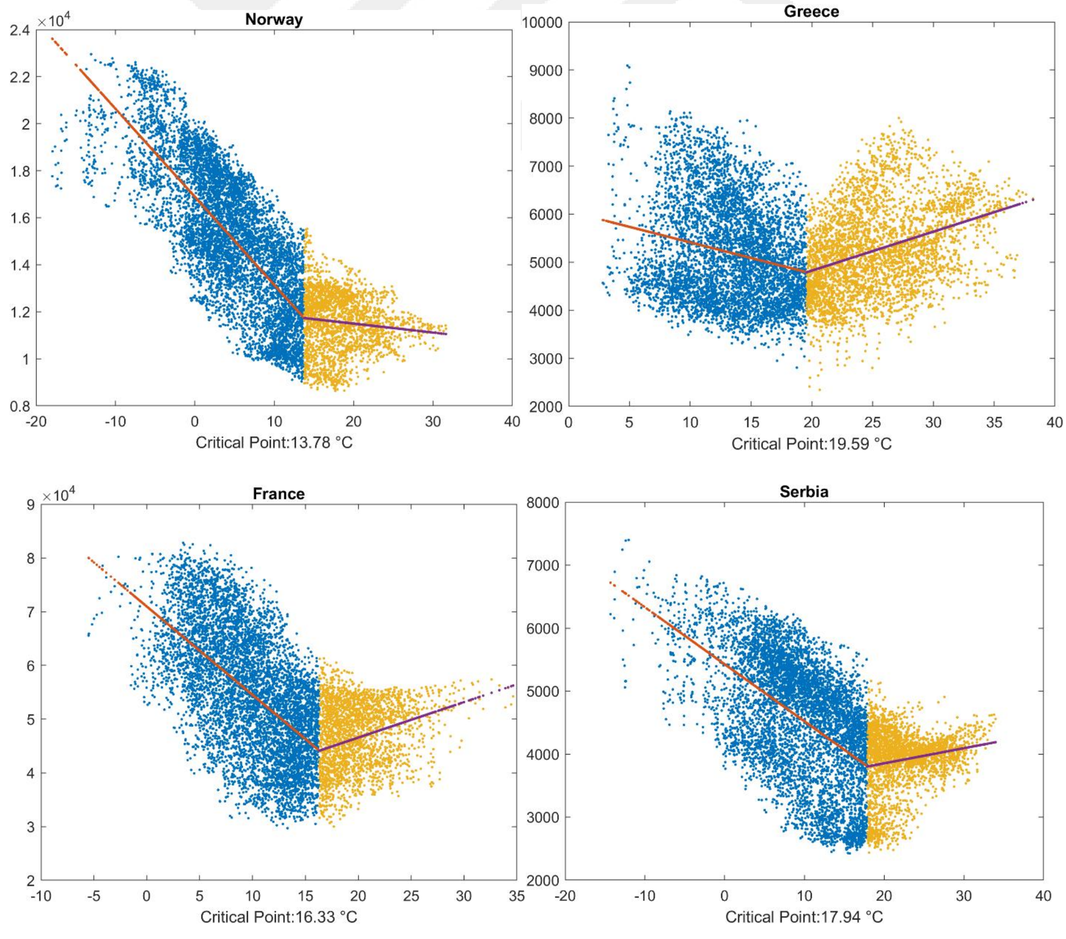


Figure 4.18: Piecewise linear fits to the electricity consumption data for the year 2014 for Norway, Greece, France, and Serbia

Recall that Norway and Greece are expected to use electricity for heating and cooling needs, respectively. This is consistent with steeper slopes at low and high temperatures, respectively. For France and Serbia, the slopes at colder temperatures are steeper; thus, we can conclude that the use of electricity for heating purposes is more common in these two countries.

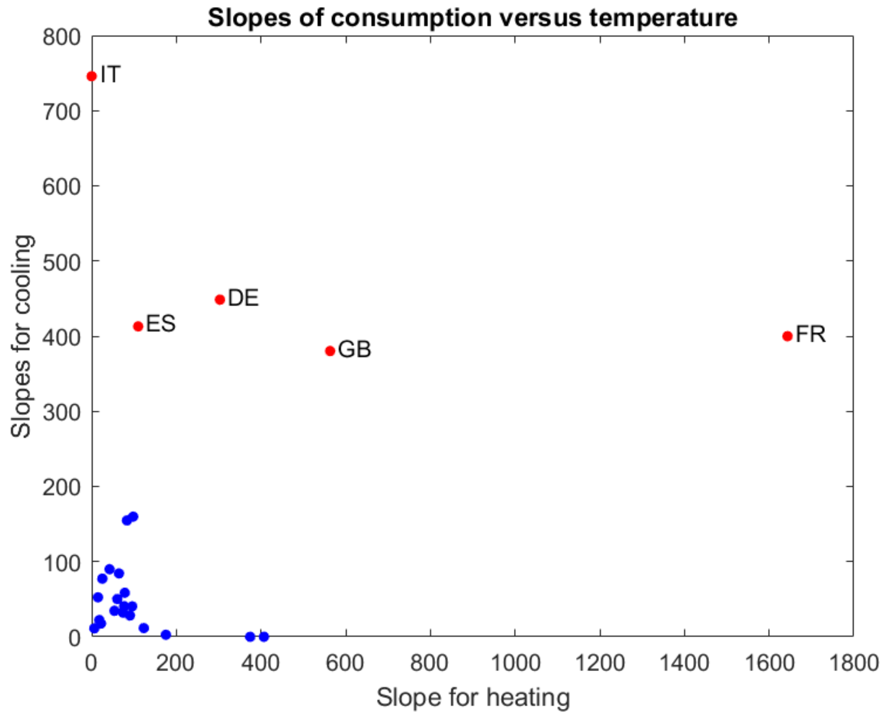


Figure 4.19: Slopes of piecewise linear fits for heating and cooling needs

To obtain a clustering of the ENTSO-E countries according to their use of electricity for heating and cooling purposes, we plotted the slopes of the regression lines at high and low temperatures as a scatter plot. Note that for Norway, the slope of the regression line is slightly negative; such cases have occurred in a few countries, and these exceptional values were replaced by zero. This information is displayed in Figure 4.19 and Figure 4.20.

Those countries with a large slope, Italy, Spain, Germany, the United Kingdom, and France, form a cluster. Italy and France are characterized by steeper slopes for cooling and heating needs, respectively. Close-ups excluding these countries are presented at below.

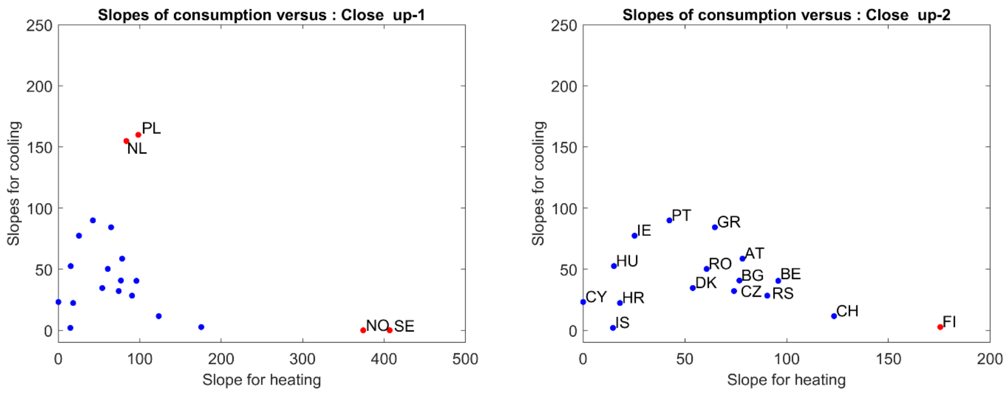


Figure 4.20: Close-ups on slopes of piecewise linear fits for heating and cooling needs

As seen from Figure 4.20, at a first close-up, Norway and Sweden appear as countries whose electricity consumptions are sensitive to cold temperatures. In contrast, Netherlands and Poland appear to be sensitive to warm temperatures. The sensitivity of Finland to low temperatures is seen only in the second close-up.

A high level of consumption in winter may not be tied directly to heating because electricity consumption can be much higher in winter compared to summer, due to shorter daylight hours, especially at high latitudes. Nevertheless, a well-defined, linear increase in consumption as a function of the deviations from comfortable temperatures clearly indicates the use of electricity for heating or cooling, as in the case of Norway, Sweden, and Finland.

4.4. Demand Segregation

Electricity consumption differs on holidays from typical consumption patterns as the closure of industrial facilities is common for such nationwide events. Public and industrial routines change during these periods, and these changes can be used to estimate the portion of industrial and household consumption on total demand.

Consumption patterns during holidays occurring in different seasons may also provide information about heating and cooling demand. During holidays in spring or autumn, heating and cooling needs are minimal, and electricity consumption consists of primary household usage. For holidays occurring in winter or summer, the consumption would be the sum of primary household usage and household heating and cooling. Thus holidays occurring at different seasons of the year for the same country can provide information on the portion of household electricity consumption for heating and cooling.

To further clarify the methodology, the consumption components are classified below.

Household usage (H_0): This component consists of the electricity used in residential units on a 24-hour basis, such as refrigerators, computers, and other home appliances, and the consumption for illumination purposes. This component of electricity consumption is relatively deterministic and consists of a constant term superimposed by variations of 24 hours. Due to shorter daylight hours, these variations have higher amplitudes in winter, i.e., the diurnal variations are modulated by seasonal variations. Also, the shape of the daily variation curves gives clues about the work habits in society.

Household heating-cooling (H_1, H_2): This component is the consumption due to household heating, and it appears as high and irregular in winter and summer. It has a strong dependence on deviations from comfortable temperatures. The heating-cooling requirement in social areas is the same throughout the week, i.e., for both weekdays and weekends, and they are considered as part of household heating-cooling.

Industrial (Non-stop production) and Base consumption (I_0): This component consists of the consumption for the industrial sectors that never stop their production, even on holidays. The sum of H_0 and I_0 represents the lowest consumption periods in the consumption profile of a country.

Industrial & Commercial consumption (I_1): This component consists of electricity used by industrial plants and commercial facilities that work during weekdays. Office consumption is considered part of industrial consumption. This component can be estimated from the difference between weekday and weekend consumption.

Industrial & Commercial heating-cooling (I_2, I_3): This component consists of heating and cooling consumption in the industry; seasonal differences on holidays and weekends can be used to estimate a portion of this consumption.

The relation between these components based on the season is shown in Table 4.1, where we divide electricity consumption into its components according to season and type of day. We do not use Saturday consumption data even though it is similar to Sunday. Saturday can have mixed industrial and household components because there is no specific regulation that makes Saturday a working day or not. However, it can also be added as a weekend, which generalizes the algorithm for every country. There will be no such significant differences as annual averages are used.

Table 4.1: Electricity components based on the season and day type

	Type of Day			
Season	**Holiday**	**Weekday**	**Sunday**	**Holiday Nighttime (Hours 00:00-05:00)**
Winter	$H_0+I_0+H_1$	$H_0+I_0+I_1+I_2+H_1$	$H_0+I_0+H_1$	I_0+H_1
Spring	H_0+I_0	$H_0+I_0+I_1$	H_0+I_0	I_0
Summer	$H_0+I_0+H_2$	$H_0+I_0+I_1+I_3+H_2$	$H_0+I_0+H_2$	I_0+H_2
Autumn	H_0+I_0	$H_0+I_0+I_1$	H_0+I_0	I_0

As seen in the Table, we take average consumption at night time in spring and autumn as a reference base consumption because we assumed that there is no or negligible consumption for heating-cooling in this period. Also, household consumption is negligible at night due to consumers' daily routines. The exact period of winter and summer differs because of geographical places.

If we denote the hourly consumption on holidays as P_{ni}, where i represents the hour of day n, and k represents the total number of holidays in a given period, base consumption I_0 will be calculated as follows;

$$I_0 = \frac{\sum_{n=1}^{k}\sum_{i=1}^{5} P_{ni}}{5*k} \tag{4.10}$$

The weights of other variables can be calculated after determining the average base consumption and setting it as 100% in order. Each component's weights are shown in Table 4.2 for Türkiye and Germany for 2016-2018 and 2006-2018, respectively. Cooling consumption in summer is more than heating in winter because Türkiye mostly depends on natural gas for heating in winter. Moreover, the cooling requirement is higher for social areas and households than for commercial and industrial areas. On the other hand, Germany's household heating and cooling consumption are minimal because the primary resource for heating is natural gas, and the cooling requirement is relatively less. The results can give an idea about how any consumption purposes contribute to total electricity consumption in different seasons.

Table 4.2: Weights of electricity consumption components for Türkiye and Germany (2006-2018)

Country	H_0	H_1	H_2	I_0	I_1	I_2	I_3
Türkiye	12.21%	6.84%	13.71%	100%	23.32%	1.69%	9.02%
Germany	20.62%	%~0	1.05%	100%	31.35%	7.08%	%~0

From a production planning perspective, demand segregation can be vital because it efficiently schedules electricity generation resources. Suppose the electricity demand is adequately segmented. In that case, production planning can be carried out to ensure that the appropriate generation resources are used to meet the needs of each class of consumption, or forecast results can be adjusted according to the same days in different seasons. This can help to reduce costs, improve efficiency, and ensure that power is delivered to customers reliably and consistently.

The developed approach that automatically estimates demand type could be a valuable tool for utilities and other organizations involved in the electricity market. It could help them better to understand the demand patterns of different classes of customers and make more informed decisions about how to allocate resources to meet those demands. Moreover, it can be directly used as input for any forecast model to detect and treats special days differently without spending much efforts. However, it is important to note that demand segregation can be a complex issue that depends on various factors. Any algorithm used to estimate it must be carefully designed and validated to ensure that it produces accurate and reliable results according to the aim of usage.

5. TÜRKİYE ELECTRICITY MARKET STRUCTURE AND IMBALANCE REDUCTION METHODOLOGY

5.1. Data and Bidding Types

Turkish Electricity Markets consist of three different markets, and these are Day-Ahead Market (DAM), Intra-Day Market (IDM), and the Balancing Market (BM). These markets are in the organized market category and regulated by the system operator. There is a schedule to participate in every market with a specific deadline.

Participants in the Turkish electricity market should give their offer for the next day until 13:00. At 13:30, the system operator informs every participant of accepted offers and quantities. The participants have the right to object between 13:30 and 13:50, and these objections are evaluated until 14:00 by the system operator. Finally, the system operator announces a conclusive price and quantity for the next day.

After completing the DAM process, the participants provide the requested information to qualify for the IDM. At 18:00, IDM becomes active. The participants can make transactions for each hour of the next day; orders can be given or updated until an hour before the physical delivery. Matching logic is the same with stock exchange markets.

If the participants cannot balance their position during intraday, the system operator puts imbalanced markets into automatic processes.

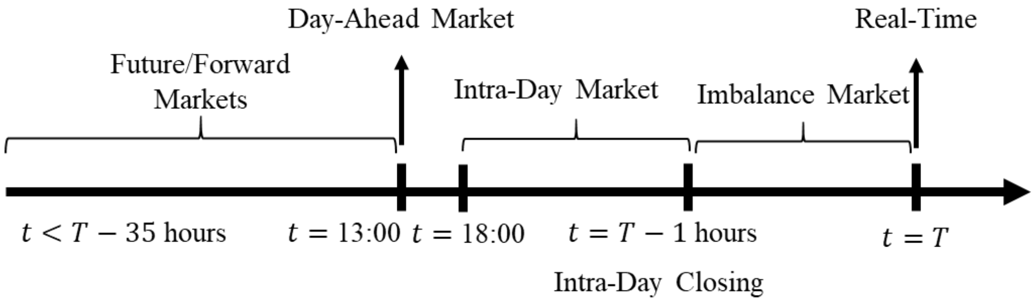


Figure 5.1: Turkish Electricity Markets operation schedule

There are three bid types in the current market: hourly, block, and flexible. Each offer consists of at least one price-quantity pair. Price-quantity pairs can be given in two directions: buying or selling. If the given amount is negative, this pair is in the sell direction; if it is positive, it is in the buy direction. The participant who gives a binary where f represents the price and m the quantity indicates that he/she has given the maximum/minimum f unit price (₺ / MWh) to buy/sell m quantity. Amounts are given in lot units, and one lot equals 0.1 MWh.

Hourly Bids: In this type of bid, participants bid price-quantity pairs (breakpoint) for the hours of the next day to buy or sell electricity. The amounts of price breakdowns on the sell side are given as negative. Due to the nature of the economy, market players want to buy goods at low prices and sell them at high prices. Therefore, as price breakdowns increase in hourly offers, the corresponding quantities should decrease (i.e., the buying quantity should decrease, and the absolute value of the selling quantity should increase). Although the hourly bid is a set of price-quantity pairs, it allows for linear interpolation to fill in between the breakpoints. The result is a piecewise linear function. The match amount for an hourly bid is also calculated on this piecewise linear function. It is determined as the amount corresponding to PTF.

Block offers: Block offers are the second most common type in the market after hourly offers. Block bids can be considered consecutive hourly bids that cannot be broken down on the timeline. However, in this offer type, there is only one price-quantity pair for the validity period. In this case, the price offered is given on the condition of buying/selling the whole amount. In addition to the price and quantity information, the number of consecutive hours to buy/sell electricity for the next day is specified. A block bid is either entirely accepted or rejected when it is active.

Linked block bids, which are a type of block bids, are also a type of bid used in the day ahead market. Linked block bids can consist of child and parent bids like decision trees. Accordingly, the child block offer is also not accepted if the parent bid is not accepted. Up to 3 block bids can be linked together in this block bid type. Also, block proposals cannot be linked in a loop; this means if block bid A is tied to block bid B, then block bid B cannot be tied to block bid A. However, interlinked block bids must be in the same direction (sell or buy).

Flexible bids: A flexible bid is given hourly and can be accepted at any hour. It can only be given in the sales direction in the current market. Flexible offers are either wholly accepted or rejected. The time the offer is accepted does not have to be the hour with the highest PTF, but if the bid price is below the highest PTF, the offer is accepted for a reasonable hour.

Many data sets are available in API systems, and we downloaded the below data to analyze market status.

- Hourly DAM (PTF) and Imbalance (SMF) Price (2016-2021)
- Hourly real-time consumption (2016-2021)
- Hourly scheduled production by resource type (SCHP) (2016-2021)
- Hourly real-time production (2016-2021)
- Hourly forecast of System Operator (2016-2021)
- Hourly load up and down (2016-2021)
- Hourly positive and negative imbalance quantity (2016-2021)
- Hourly positive and negative imbalance cost (2016-2021)
- Hourly supply and demand curve (2019-2021)

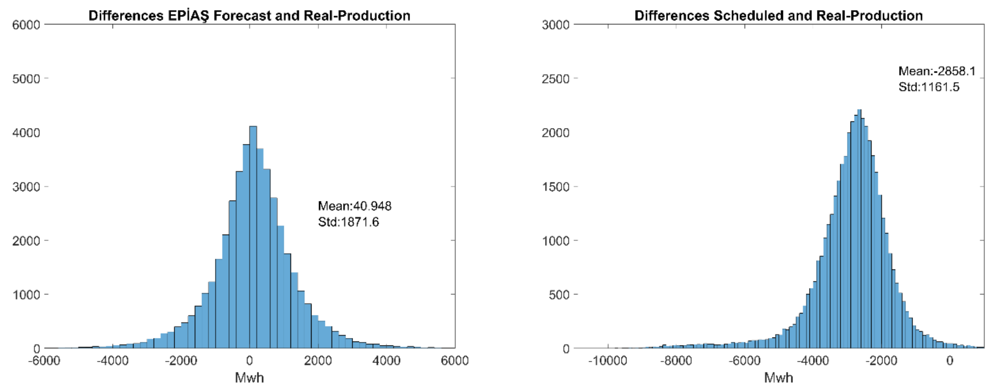
5.2. Analyzing Hourly Imbalance Quantity

An imbalance in the electricity market refers to the difference between the actual electricity supply and demand at any given time. This imbalance can occur for a variety of reasons, including unplanned outages or failures of power plants or transmission lines, unexpected changes in consumer demand, or errors in forecasting or scheduling.

There are several ways in which imbalances can affect the electricity market. One of the most significant impacts is on the price of electricity. When there is an imbalance between supply and demand, the price of electricity can fluctuate significantly, which can lead to price volatility and market instability. This can be particularly problematic for consumers, as they may face unexpected changes in their energy bills.

Imbalances can also have an impact on the reliability of the electricity grid. If there is not enough electricity being generated to meet demand, it can lead to blackouts or brownouts, which can be disruptive and costly for businesses and households. On the other hand, if an excess of electricity is generated, it can be difficult for utilities to find customers to sell the excess power to, which can lead to stranded assets and financial losses for power generators.

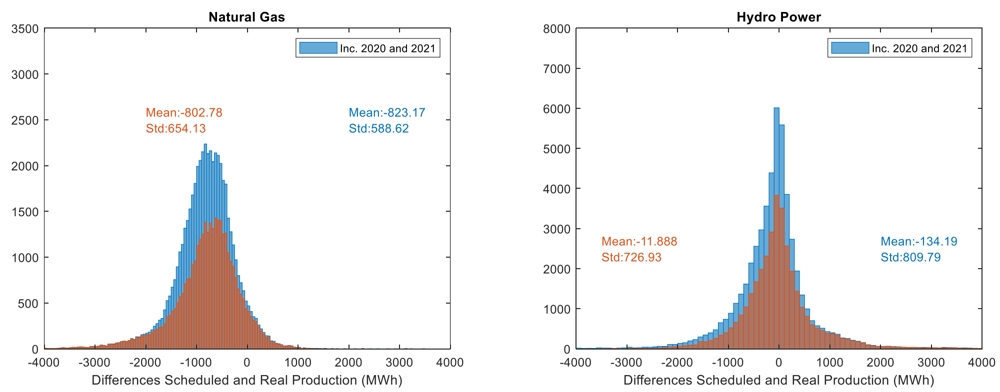
At first, we thought that low forecast accuracy was the reason for the high imbalance on an hourly basis. However, when we checked the distribution of an hourly error of SO forecast in the below figures, we saw that forecast is quite good. We do not observe the same thing between scheduled production and actual production. Figure 5.2 shows that even though forecast error distribution is close to normal, scheduled production is consistently lower than actual production. There can be other reasons that create this shortage, for example, renewable energy volatility and intra-day market volume volatility. Intra-day market volume is one of the main reasons for the differences. However, negative differences between scheduled and actual production show that market players consciously leave some of their trading capacities to the intraday market.



(a) (b)

Figure 5.2: Differences between a) SO forecast- Actual production, b) Scheduled Production and Actual Production

The differences between scheduled-actual production plans are presented in Figure 5.3 for each energy resource to see the effects of renewable energy resources' intermittency on planning.



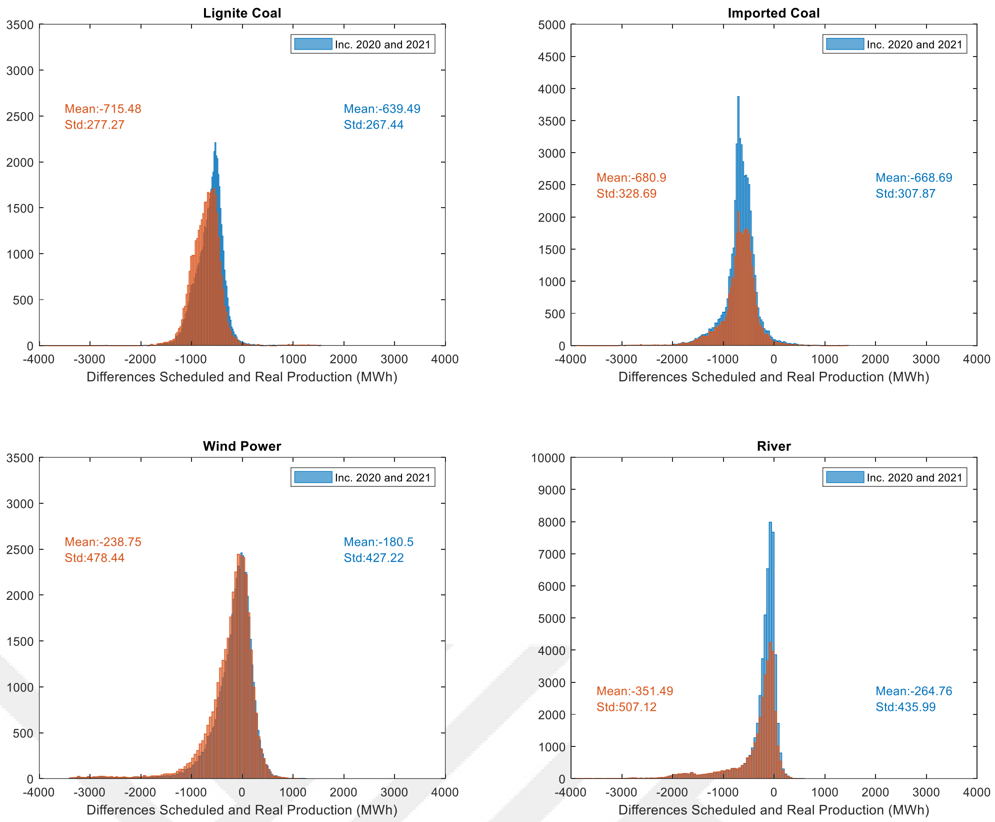


Figure 5.3: Differences between scheduled plans and actual production by resources type

The distributions seem random when we examine the renewable energy source, wind, and river graphs. Although the average is close to zero, it can be said that the planned and the actual production are not aligned. This consistent situation shows that there may be a problem in resource-based production planning; participants are likely to make a defensive estimation for these resources due to the variability in renewable energy resources. The reason can be that the purchase guarantee incentive is applied for renewable energy sources independent of the planned production.

The more striking point here is that the production of coal and natural gas power plants is constantly higher than the planned production. Natural gas power plants can be used in case of energy deficit due to variations in other sources or power plant failures due to the short start-up time. However, this is not possible for coal. Again, when the distribution of the differences is analyzed, the fact that coal and natural gas power plants produce more than planned continuously shows that this is a strategy. When we compare the planned and actual production of fossil fuels, approximately 5% of the total daily production is continuously released to the intraday market.

Figure 5.4 presents hourly data distribution in the same x-axis domain to analyze relations between imbalance quantity and other data sets.

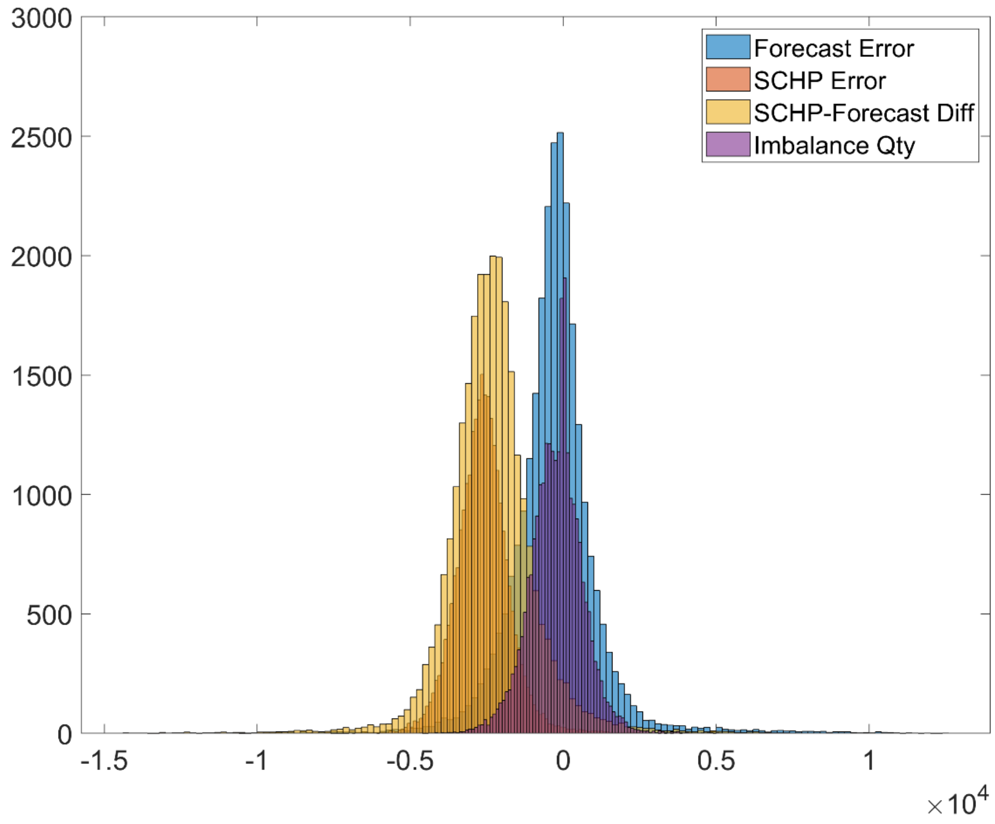


Figure 5.4: Histogram of data sets (axis x:MWh and y:Frequency)

Although the imbalance data seem as a normal distribution pattern, the cost-based effect is not this way. Especially in cases of energy deficit, extraordinary price spikes may occur in the imbalance market. Again, as we can see from the EPİAŞ data, even if the net load up and down difference is negative for the same hour, and there is excess energy, the imbalance cost can be high and in different directions. This means that the load-up and load-down costs per MWh are different.

Before analyzing the imbalance cost, the data sets provided by EPİAŞ need to be introduced. Two different imbalance data are published on the EPİAŞ data portal. Although both are hourly, one is announced daily according to the hourly values. The other is announced monthly according to the basis for settlement values. For the imbalance data announced monthly, the cost of each positive and negative imbalance is published hourly in the same format. Although the two data sets are similar, as seen in the figure below, we preferred to use the monthly data since we were going to make cost analyses in our study.

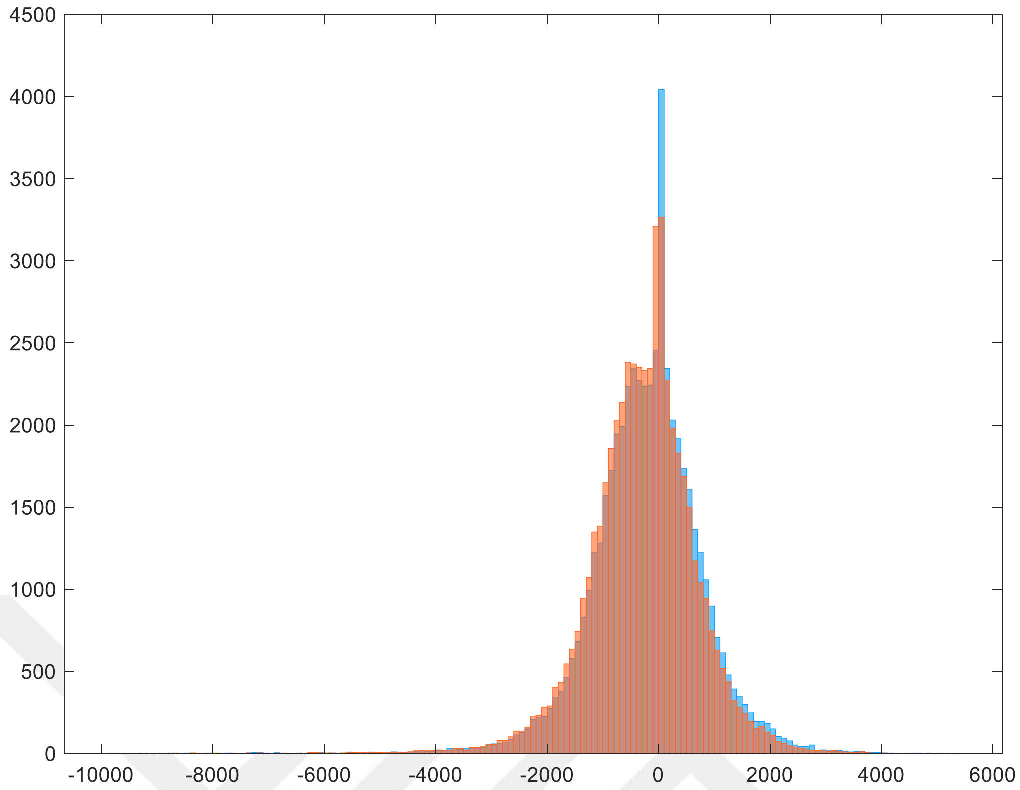


Figure 5.5: EPİAŞ imbalance data sets comparison (axis x: MWh and y: Frequency)

Also, as we mentioned earlier, different imbalance costs may occur for the same quantity of imbalances in different hours. This leads to an independent cost in terms of imbalance direction. In other words, when an excess of energy has occurred for a specific hour, the cost of additional load purchases due to imbalance is much more than the load-down cost. In this case, it can be said that there is no perfect competition in the current market and that the producers make more profit than usual in case of an energy deficit.

When we checked the data sets and distributions, SCHP was always less than the actual production. We can explain this with market participants' intentions. There is no storage option in the electricity markets, and sellers know that buyers have to buy specific amounts and buyers know that sellers cannot produce less than a certain amount due to ramp-up costs. Hence, buyers intend to secure the lower bound of their forecast because it is the critical limit and has to be supplied to the end user. If they need more, they can trade in the intra-day market. The reason of that if they buy more electricity and do not use all of them, their expected cost will be DAM price for an MWh. On the other hand, if they purchase electricity from the intra-day market, the additional cost that they must stand for would be $BM - DAM$ prices. In this circumstance, supplying additional quantities from IDM is less risky as long as $BM \leq 2 * DAM$.

The perception of sellers is also similar to buyers. They intend to sell the critical quantity in DAM and leave the remaining capacity for IDM and BM to sell at higher prices. As seen below, even imbalance quantity distribution is similar to forecast error distribution

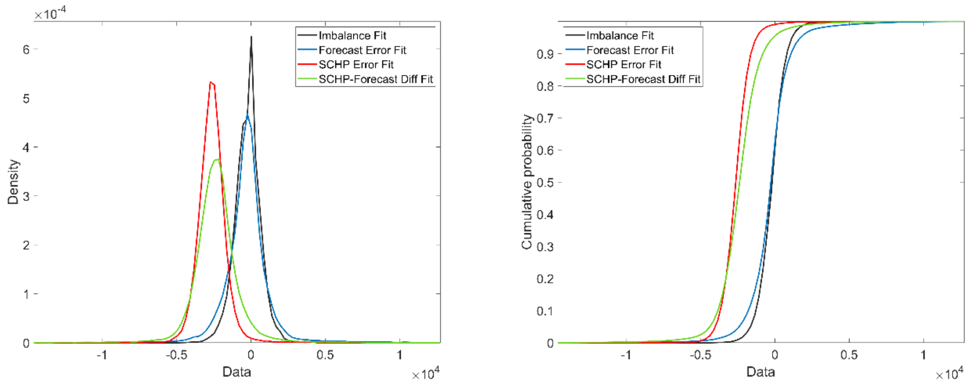


Figure 5.6: Probability and cumulative density functions of data sets

Both buyer and seller intend to increase negative imbalance, and their strategies seem logical individually. However, this strategy causes a high price on a shortage at the aggregate level. As seen in the graphs above, the distribution of the imbalance quantities and the distribution of the estimation errors are similar. Both data sets have equal samples on the negative and positive sides, and it is impossible to talk about any systematic pattern. Again, the distributions are similar when compared with the planned, actual, and forecast results. This gives us a clear picture of how consistent and realistic our estimates are.

The strangest thing about the imbalance arises when we examine the actual costs. Table 5.1 shows previous years' positive, negative and net imbalance costs. As can be seen from the table, the cost of negative imbalances has been higher in all years. This shows that the effect of the energy deficit on the total cost is much more than the energy surplus. Based on this information, our method for decreasing negative hourly imbalances will be presented in the next section.

Table 5.1: Yearly negative, positive and net imbalance amount

Years	Positive Amt(TL)	Negative Amt(TL)	Net Imb Amt(TL)
2016	1,246,169,755	-2,569,888,340	-1,323,718,584
2017	1,273,704,106	-2,296,230,162	-1,022,526,056
2018	1,675,090,087	-2,847,074,980	-1,171,984,893
2019	1,680,333,209	- 2,774,274,616	-1,093,941,407
2020	1,413,836,387	-2,837,607,944	- 1,423,771,557
2021	2,473,154,094	-6,402,714,654	-3,929,560,559

5.3. New Approach to Decrease Imbalance in Day-Ahead Market

If we give information about how our forecast model has been revised to use this algorithm before getting to know the imbalance decrease algorithm, as we mentioned earlier, each market has clearly defined start and end times in the actual electricity market.

Since the forecast made with the planned consumption in this algorithm will be used directly, the forecast periods should be compatible with the day-ahead market. To recall, for the day-ahead market, the participants should submit their hourly offers for the next day until 12:30. They should check the accepted plan by the system operator between 13:30 and 13:50 and make objections if necessary. As of 14:00, the final plan will be announced. Accordingly, the model has been revised to predict the next day at noon every day. For this reason, a 36-hour forecast period was created, and the 12th and 35th values of this forecast were recorded as the forecast for the next day.

The FSE model forecast period was used as 35 hours to obtain the forecast for the following days. Although the forecast results with the FSE model, like the forecast section impressions, are very reliable, feedback or error correction with the autoregressive model gives better results in short-term forecasts. For this reason, an autoregressive model was created at different lags for each hour in the forecast period. Lag information and values are presented in Appendix F in detail.

Returning to the market operation, we have both the predicted consumption and the official day-ahead production plan hourly for the next day as of 13:30, after completing our forecast. With these data, an imbalance estimation can be made for each point. In order to do this, the NARX model was created using the MATLAB library. The default settings of the MATLAB library are used for the model. The schema of the NARX model created below is shared.

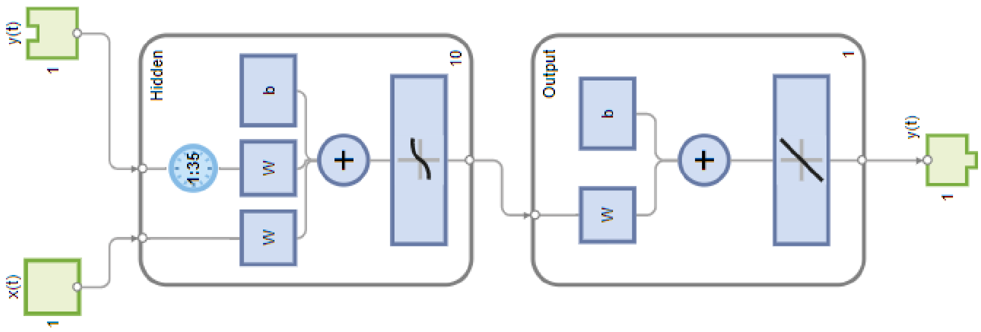


Figure 5.7: Nonlinear Autoregressive with exogenous model schema

In the nonlinear neural network model created, similar to the autoregressive model, the difference between planned production(SCHP) and forecast and the historical data of the imbalance data form the model's inputs as external variables. The model layer is created as size ten and time delay two. For the model's training, as in the general practice, 2018 (8760 Hours), data was divided by 70%, 15%, and 15% and used as training, validation, and test data, respectively. The model was trained only by 2018 data. The resulting model was used for step-ahead imbalance estimation for the years 2018, 2019, and 2020.

The imbalance estimated by NARX will be directly used in the imbalance reduction algorithm. This algorithm aims to shift the market equilibrium point that occurs in the day ahead market to the right, to get additional capacity for negative imbalances that may occur during the day. In electricity markets like ours, where producers are dominant, producers deliberately trade in certain volumes in the day-ahead market. Because these companies, whose primary purpose is to make a profit, can use the imbalance market to their advantage. As we saw in the previous section, although the distributions of negative imbalances and positive imbalances are equal, there is a massive difference in costs. The basic concept and steps of the algorithm are shown in Figure 5.8.

$m_i = market\ equlibrium\ price\ at\ i^{th}\ hour(TL)$

$t_i = market\ equlibrium\ price\ index\ at\ i^{th}\ hour\ in\ orders$

$n_i = new\ price\ at\ i^{th}\ hour\ after\ reserve\ capacity\ bought\ (TL)$

$deficiency_i = calculated\ deficieny\ quantity\ for\ i^{th}\ hour\ (MWh)$

$r_i = additional\ accepted\ supplys\ for\ i^{th}\ hour\ (MWh)$

$p_i = imbalance\ price\ for\ i^{th}\ hour\ (TL)$

$rmb_i = real\ imbalance\ \ quantity\ for\ i^{th}$

$rmp_i = real\ imbalance\ price\ for\ i^{th}\ hour\ (TL)$

L= $price\ ratio(\frac{n_i}{m_i})\ limits\ to\ prevent\ buying\ at\ the\ higher\ prices$

Step 1. Collect the bids for supply and demand for the hour

Step 2. Calculate/Forecast the net imbalance quantity and price per MWh for the hour

Step 3. If there is insufficient supply(deficiency) for the hour, go to step 4; otherwise, do nothing

Step 4. $While\ r_i\ <\ -deficiency_i\ \ and\ \ n_i/m_i < \mathrm{L}$

$r_i = -orders.supply(t_i + c) + orders.supply(t_i)$

$c = c + 1$

Step 5. Calculate the profit or loss using realized imbalance cost

$$profit = \min(r_i, \max(-rmb_i, 0)) * (rmp_i - n_i) - \max(0, r_i + \min(0, rmb_i)) * n_i$$

Figure 5.8: Simple Psuedo code of the algorithm

The reason for the price ratio limit (L) in the algorithm is to prevent considering imbalance quantity as the only constraint while shifting the market equilibrium point to the right. In this way, if the price for the additional capacity is high compared to the equilibrium price, the model will secure some part of the additional capacity or not at all.

To optimize L and see the algorithm's success, firstly, actual imbalance data is used to test. The algorithm run with different L values and profit changes of each year is shown in the below figure. As seen in the figure, profit changes in the years 2019 and 2020 are more smooth for different L values. However, profit change in 2021 more dramatically decreasing with higher L. The reason for that is a change in buyers' and sellers' hourly bid distribution. In other words, the angle between supply and demand is smaller in 2019 and 2020 than in 2021. A higher angle causes a higher price for additional capacity when the equilibrium point shifts right. The optimum L value is chosen as 1.17, and yearly profits with different L values are presented in Appendix G.

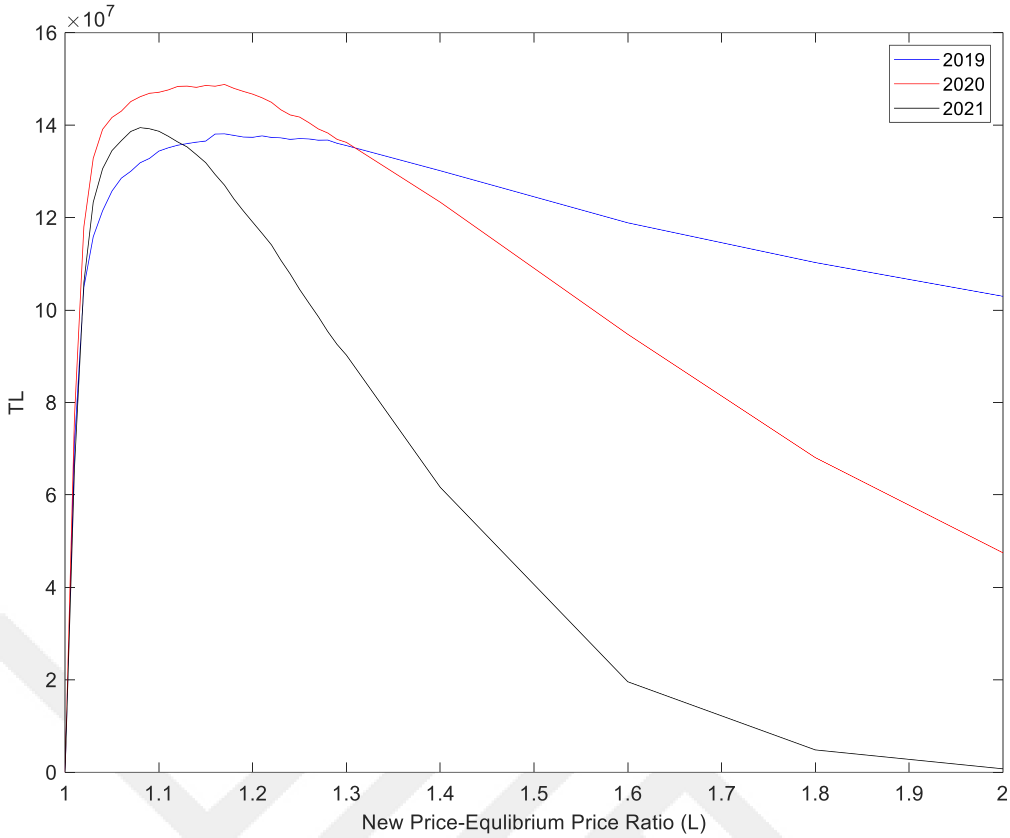


Figure 5.9: The algorithm performance with different ***L*** values

We used the same coefficient and ran the model with a calculated imbalance with the NARX model. The results are shown below in the table. As seen in Table 5.2, forecasted data success on decreasing imbalance cost is similar in 2019-2020 but dramatically decreases in 2021. The reason is that day-ahead offers are not regular as in 2019-2020, and buying extra capacity is more expensive than in previous years.

Table 5.2: Algorithm profit result with the actual and forecasted imbalance and total net imbalance cost in TL

Year	Profit with Actual Data	Profit with Forecast	Total
2019	138,134,915	59,311,041	1,288,887,787
2020	148,817,491	71,343,673	1,585,867,839
2021	139,476,236	30,365,007	4,171,561,251

We conducted a weekly performance analysis to evaluate the overall performance of our strategy. We randomly selected 156 weeks from a three-year simulation.

While choosing the weeks, we took the starting days randomly. The 156-week performance is shown in the table below. The 38 weeks of a total of 156 weeks have negative profit, and the 118 weeks algorithm has a positive profit.

Table 5.3: Algorithm weekly performance between 2019-2021.

Definition	Value	The Lowest	The Highest
Average Profit per week (TL)	5610.47	-27478.20	51544.38
Std of Profit (TL)	10671.65	-	-
Average Profit$^{(+)}$ per week (TL)	9315.72	106.67	51544.38
Std of Profit$^{(+)}$(TL)	9553.83		
Average Profit$^{(-)}$ per week (TL)	-5195.43	-27478.20	-8.07
Std of Profit$^{(-)}$(TL)	6642.10		
Average # of Hours Algorithm applied in a week	69.12	10	125

5.4. Imbalance Optimization with large-scale storage

Large-scale energy storage refers to the use of various technologies to store large amounts of electricity for later use. This can be useful for addressing the intermittent nature of renewable energy sources such as wind and solar power, as well as for providing backup power during outages or other emergencies.

There are several different technologies that are used for large-scale energy storage, including pumped hydroelectric storage, compressed air energy storage, and battery storage. Each of these technologies has its own advantages and disadvantages. The most appropriate technology for a given situation will depend on factors such as the location, the amount of energy to be stored, and the length of time it needs to be stored.

Pumped hydroelectric storage is one of the most common technologies used for large-scale energy storage. It involves using excess electricity to pump water from a lower reservoir to an upper reservoir, where it is stored until it is needed. When the stored energy is needed, the water is released back down through a turbine, generating electricity. This technology has been in use for many decades, and there are currently more than 100 pumped hydroelectric storage facilities around the world.

Compressed air energy storage (CAES) is another technology that is used for large-scale energy storage. It involves using excess electricity to compress air, which is then stored in underground caverns or other types of storage containers. When the stored energy is needed, the compressed air is released and heated, expanding and driving a turbine to generate electricity. CAES technology has been in use since the 1970s, and there are currently two CAES facilities in operation around the world.

Battery storage is a relatively newer technology that is also used for large-scale energy storage. It involves using batteries to store excess electricity, which can be released back into the grid when needed. Battery storage technology has advanced significantly in recent years, and there are now many large-scale battery storage facilities in operation around the world.

Overall, large-scale energy storage is a critical technology for enabling the widespread adoption of renewable energy sources and for improving the reliability of the electric grid. While there are still many challenges to be overcome, such as the high cost of some energy storage technologies and the need for more efficient and longer-lasting batteries, the future of large-scale energy storage looks promising

- Improved reliability and stability of the power grid: Large-scale energy batteries can provide backup power during times of high demand or when there are disruptions in the power supply. This can help prevent blackouts and ensure a stable and reliable power supply.
- Increased integration of renewable energy sources: Energy batteries can store excess energy generated by renewable sources, such as solar and wind, and release it when needed. This can help increase the use of renewable energy and reduce reliance on fossil fuels.
- Enhanced grid flexibility: Energy batteries can help balance the supply and demand of electricity on the grid, allowing for more flexible and responsive power generation and distribution.
- Reduced greenhouse gas emissions: Energy batteries can help reduce carbon emissions by reducing the need for fossil fuel-based power generation and enabling the use of renewable energy sources.
- Cost savings: Large-scale energy batteries can provide cost savings by reducing the need for expensive peaker plants and other backup power generation systems. They can also help utilities avoid expensive grid upgrades and improve their overall operational efficiency

There are several recent and planned projects involving large-scale energy batteries;

- In California, the Pacific Gas and Electric Company recently announced plans to deploy the world's largest lithium-ion battery storage system, which will have a capacity of 1,200 megawatt-hours.
- In Australia, the Tesla Powerpack project has installed a 100-megawatt/129-megawatt-hour battery system to provide backup power for the South Australia grid.
- In the UK, the National Grid has plans to build a 40-megawatt battery storage system to help manage the country's electricity demand.
- In Germany, the grid operator TenneT is planning to build a 300-megawatt battery storage system to help integrate renewable energy sources into the power grid.

The development of large-scale battery possibilities will not only increase the share of renewable energy sources in energy production. However, it will also be vital to manage the inefficiency caused by imbalance, independent of demand-supply forecasting. For this reason, we made revisions and applied the developed algorithm for different storage capacities and scenarios.

Our algorithm was used in the same way, but the following changes were applied for the specified scenarios.

Scenario 1: We aim to run our algorithm for different storage capacities according to the net imbalance quantity in this scenario. After deciding to purchase additional capacity according to the day-ahead forecast, we added the storage capacity to the net gain function when this capacity was realized. The new profit function (Step 5) is expressed as follows. Also, we need to consider excess supply profit as well because of using it to fill storage. To clarify, if we calculated energy deficiency for a specific hour and our forecast was wrong, storage will provide extra cushion. If we do not forecast energy deficiency and excess supply occurs, we can use the excess quantity to fill storage, and the profit will be the total load-down amount.

Recall Step 5. Calculate the profit or loss using realized imbalance cost

$S = Storage\ Capacity(MW)$

$s_i = Available\ power\ in\ storage\ at\ starting\ of\ i^{th}\ hour(MW)$

$$profit = \min(r_i, \max(-rmb_i, 0)) * (n_i - rmp_i) - \max(0, r_i + \min(0, rmb_i) - (S - s_i)) * n_i + \min(-\min(-rmb_i, 0), (S - s_i)) * n_i$$

Figure 5.10 shows the profit amounts for different storage capacities by year. The profit generated at the zero storage point is the same as the result of the first algorithm. The impact of storage for 2019 and 2020 is limited compared to 2021. This is another point that shows us that there has been a differentiation in managing imbalances over the years, and the costs are increasing faster than usual. The increase in storage capacity affects profit less after a certain point. Although this point changes according to years, it can be said 3000 MW, as seen roughly from the graph. The storage capacity is used according to the imbalance estimate made in this scenario. The storage is charged with additional unused capacity during the hours when the positive imbalance is predicted and the hours when the negative imbalance is incorrectly predicted.

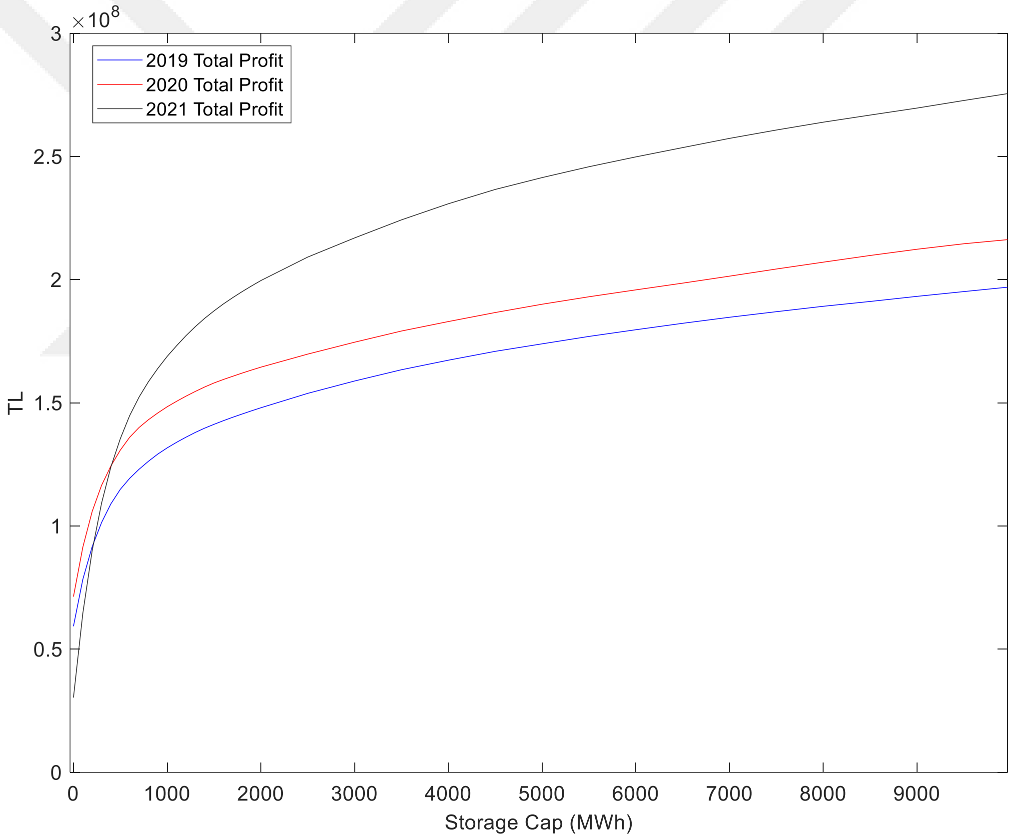


Figure 5.10: Total profit change with different storage levels by years

Scenario 2: This scenario will be applied for comparison. The algorithm was run for different storage capacities with actual data, regardless of the estimates. If the net imbalance amount in any hour is positive, electricity is stored; if it is negative, it is used. The imbalance costs of the stored electricity and the electricity used from the storage are considered as the total profit.

Similar to the previous chart, while the profits of 2019 and 2020 are similar, the year 2021 is very different. The increase in storage capacity leads to greater profits for the year 2021. This means that, as we mentioned in the previous chart, there is an imbalance structure outside of the norms in 2021. In this graph, the profit for zero storage is also zero for all years because only the effect of storage was calculated using the real imbalance. In this scenario, as long as there is no storage, it is impossible to reduce the costs of positive and negative imbalances.

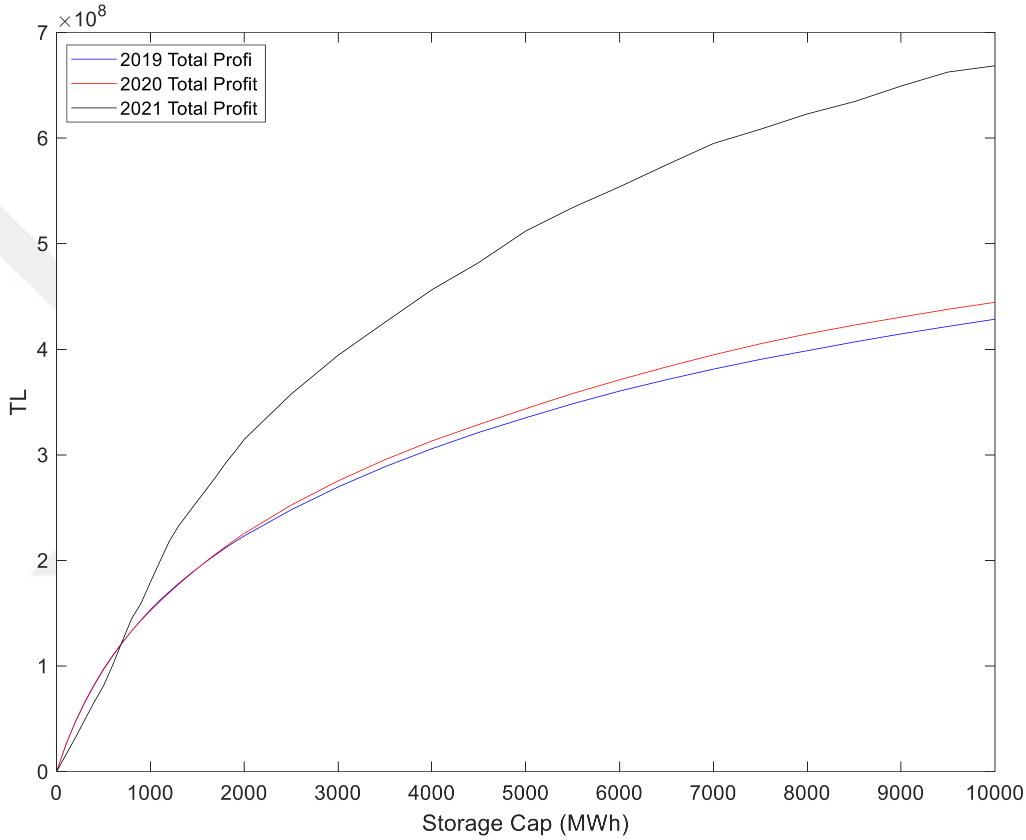


Figure 5.11: Total profit change with different storage levels by years using an actual net imbalance

The result of the first and second scenarios was shared in APPENDIX H and I.

6. DISCUSSION

One of the challenges of the electricity markets is the potential misuse of monopolies or strategic behavior among electricity market participants. Suppliers and buyers in a free electricity market may engage in monopolistic or strategic behavior to maximize their profits. For example, a supplier may withhold electricity from the market in order to drive up prices, or a buyer may attempt to drive down prices by threatening to switch to a different supplier. This behavior can lead to market inefficiencies and make it difficult for utilities to reliably plan for future electricity needs. To mitigate these challenges, market participants can adopt a variety of strategies. For example, suppliers may enter into long-term bilateral contracts with buyers to reach more predictable and stable electricity prices. On the other hand, buyers may choose to diversify their electricity sources to reduce their reliance on a single supplier and increase their bargaining power. Additionally, regulatory bodies may implement rules or market mechanisms to promote more efficient behavior and reduce the potential for strategic behavior.

The integration of large-scale battery storage can also potentially change the strategies of market participants in the day-ahead market. For example, with large-scale storage availability, suppliers may offer more flexible and responsive services to buyers, potentially leading to more efficient outcomes in the market. On the other hand, buyers may be able to use large-scale storage to smooth out fluctuations in electricity demand, reducing their need to rely on traditional generation sources and potentially leading to cost savings. Overall, the integration of large-scale battery storage into electricity systems has the potential to bring significant benefits. However, it is essential to address the challenges of cost and strategic behavior in order to realize these benefits.

Accurate forecasting is another critical function in the electricity market, as it allows market participants to make informed decisions about generation, transmission, and consumption. With the increasing adoption of renewable energy sources and the integration of large-scale battery storage, the importance of forecasting is likely to increase. In the current electricity market, forecasting is used to predict future electricity demand, which allows utilities to plan for sufficient generation capacity to meet that demand. With the adoption of smart grids, which enable more flexible and responsive electricity systems, the need for accurate forecasting is likely to increase. Smart grids rely on a combination of distributed energy resources, such as rooftop solar panels, small-scale battery storage, and centralized generation sources. Accurate forecasting is necessary to ensure that these resources are used efficiently and effectively to meet changing electricity demand.

The integration of large-scale battery storage into smart grids can also benefit from accurate forecasting. For example, accurate electricity demand forecasts can help utilities determine when it is most cost-effective to charge and discharge large-scale batteries. This can help utilities optimize the use of these batteries and reduce the need for expensive fossil fuel generation. Additionally, accurate forecasts of renewable energy generation, such as solar and wind, can help utilities determine when it is most cost-effective to use these resources and when it is necessary to rely on other sources of generation.

Overall, accurate forecasting is essential for electricity markets' efficient and effective operation. It is likely to become even more important as the adoption of renewable energy sources and the integration of large-scale battery storage continue to increase.

Although real market conditions and constraints are tried to be fully considered for the algorithms and approaches developed in this study, there are some limits. In the forecast model developed for the imbalance reduction algorithm, the data of the last actual consumption hours used to revise the hourly forecasts are not always available at the specified time. Moreover, the algorithm's first purpose was to increase the efficiency of the day ahead market and decrease balancing market volume rather than directly reducing the cost. The algorithm may not purchase any additional capacity even if there is an energy deficit within the defined constraints, even when an energy deficit is forecasted. The reason is that the changes in offer characteristics of sellers or irregular price steps correspond to different production levels. For this reason, the implementing mechanism should apply this algorithm within the framework of explicit rules.

7. CONCLUSION

Although there are many studies in the literature about electricity consumption modeling, it is a constantly evolving field. Since electricity is a crucial input in many fields and no function is sustainable without electricity in today's life, optimizing electricity production and consumption is very important in terms of efficient use of resources. Therefore, analyzing and forecasting electricity consumption is very important for long and short-term projections.

With the FSE model developed in this study, consumption profiles of countries were created using only the consumption data, and country-specific characteristics were determined. Additionally, long and medium-term forecasts were produced successfully without changing model complexity. These data-driven analyses can provide valuable information about countries for the ENTSO-E grid expansion. The countries in the ENTSO-E region were classified by comparing the temperature data of the model outputs on a country basis.

With the flexibility of the FSE model, short-term forecasts were also made from the same model. In order to make these estimations more reliable, the AR model and Feedback model are suggested as a hybrid model. While the AR model needs complex calculations for the coefficients, the feedback model provides convenience in terms of applicability. Obtaining stable results for countries of different scales shows that the model can be applied to different levels of planning.

Seller and buyer activities in free electricity markets directly affect market efficiency. Especially in underdeveloped markets, less or limited competition allows the seller or buyer to make more profit. In such markets, sellers' and buyers' desire to maximize profits leads to more imbalances and costs. For this reason, the hybrid forecast method created with the FSE and AR methods was applied to the Turkish Electricity Market under realistic conditions and used as an input to the algorithm created to reduce the imbalance cost.

In order to use the model in the day-ahead market, 35 different AR models were created to adjust the 35-hour forecast of FSE model. Thus, day-ahead forecasts were made with a very low error rate. The differences between the generated forecasts and the planned production are used for the hourly imbalance forecast. For this forecast, the NARX method was applied using the MATLAB library. The imbalance estimation was made using the differences calculated between scheduled production and forecast with the historical imbalance data as input. The algorithm aims to reduce the high costs in the imbalance market by taking additional capacity for the hours for which the imbalance is estimated in the day ahead market. According to the results, when the algorithm is run with actual values, it reduces the total cost by approximately 5.5%, 5.01%, and 0.77% for 2019-2021. When we run the algorithm considering that there is a storage option, it gives much more effective results.

With the support of a very accurate forecast, buying additional capacity to prevent high prices intra-day and balancing the market by adjusting equilibrium prices in the day-ahead market can help increase market efficiency. By having an accurate demand forecast, the system operator can effectively plan and schedule the generation resources and ensure that adequate power is available to meet customers' needs. This can help to reduce the costs of power generation and improve the reliability of the power supply. Additionally, by having an accurate forecast of demand, the system operator can more effectively manage the balancing of supply and demand, which can help to reduce the need for expensive balancing services, such as buying additional capacity on the spot market. Hence, high market efficiency provides to make better decisions and optimize the use of resources, which can lead to a more efficient and reliable electricity market and ultimately benefit the consumers.

The imbalance reduction algorithm provides flexibility to manage the market condition and supports the increase of day-ahead market volume and efficiency without any storage capacity. Moreover, if it is combined with storage availability with advancing storage technologies, it becomes a more useful tool and provides more gain even in low storage scenarios.

With the hybrid methods created in this study, electricity consumption scenarios can be created from hourly to annual. With the algorithms developed in our study, many characteristic features of the consumption profiles can be obtained by using only electrical energy data. In this way, it becomes a fast and reliable indicator of the decision-making for any planned or unplanned changes in production, transmission, or distribution strategies.

In summary, energy is an issue whose importance is constantly increasing from past to present and directly affects the welfare level of societies. In this area, it is vital to use resources efficiently as well as to have them. This is one of the issues that will not fall off the agenda regarding the countries' interests and environmental effects. It will be more and more critical to develop good forecasts and analyze the current structure of consumer profiles as the primary input of planning. While understanding the consumption habits of countries with the forecast models developed in this study, we also showed that these forecasts could be used as the primary input of imbalance reducing algorithms for our own electricity market. Therefore, our study has provides many and valuable insights that can be applicable by standard systems on different consumers levels. Each algorithm can be used stand-alone or together to get information about electricity markets and consumers or support decision making systems to reach more efficient and effective markets.

REFERENCES

Akbar, A., Nati, M., Carrez, F., & Moessner, K. (2015). Contextual occupancy detection for smart office by pattern recognition of electricity consumption data. *IEEE International Conference on Communications*, *2015-September*, 561–566. https://doi.org/10.1109/ICC.2015.7248381

Allerhand, A. (2017). A Contrarian History of Early Electric Power Distribution [Scanning Our Past]. *Proceedings of the IEEE*, *105*(4), 768–778. https://doi.org/10.1109/JPROC.2017.2677558

ALSUNAIDY, A., & GREEN, R. (2006). Electricity deregulation in OECD (Organization for Economic Cooperation and Development) countries. *Energy*, *31*(6–7), 769–787. https://doi.org/10.1016/j.energy.2005.02.017

Austria Public Holidays 2022 - qppstudio.net. (n.d.). Retrieved December 3, 2022, from https://www.qppstudio.net/publicholidays2022/austria.htm

Becker, V., & Kleiminger, W. (2018). Exploring zero-training algorithms for occupancy detection based on smart meter measurements. *Computer Science - Research and Development*, *33*(1–2), 25–36. https://doi.org/10.1007/s00450-017-0344-9

Brown, M. H., & Sedano, R. P. (2004). *Electricity Transmission A Primer National Council on Electricity Policy*. www.ncouncil.org

Cross-Border Interconnections on Every Continent. (2010). www.geni.orgpeter@geni.org

Eibl, G., Burkhart, S., & Engel, D. (2018). Unsupervised Holiday Detection from Low-resolution Smart Metering Data. *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, 477–486. https://doi.org/10.5220/0006719704770486

Electricity price statistics - Statistics Explained. (n.d.). Retrieved September 24, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics

Electricity prices around the world | GlobalPetrolPrices.com. (n.d.). Retrieved August 11, 2022, from https://www.globalpetrolprices.com/electricity_prices/

ENTSO-E. (2019a). *ENTSO-E Annual Work Programme 2019*. https://www.entsoe.eu/Documents/Publications/ENTSO-E general publications/AWP_2019.pdf

ENTSO-E. (2019b). *Power Statistics*. https://www.entsoe.eu/data/power-stats/

Erdogdu, E. (2011). What happened to efficiency in electricity industries after reforms? *Energy Policy*, *39*(10), 6551–6560. https://doi.org/10.1016/J.ENPOL.2011.07.059

Flatabo, N., Doorman, G., Grande, O. S., Randen, H., & Wangensteen, I. (2003). Experience with the Nord Pool design and implementation. *IEEE Transactions on Power Systems*, *18*(2), 541–547. https://doi.org/10.1109/TPWRS.2003.810694

Former Associations. (n.d.). Retrieved October 13, 2022, from https://www.entsoe.eu/news-events/former-associations/#union-for-the-coordination-of-the-transmission-of-electricity-ucte

Gajowniczek, K., Nafkha, R., & Ząbkowski, T. (2017). *Electricity peak demand classification with artificial neural networks*. 307–315. https://doi.org/10.15439/2017F168

Germany's Scholz says switch back to coal and oil 'temporary' | News | DW | 16.07.2022. (n.d.). Retrieved September 24, 2022, from https://www.dw.com/en/germanys-scholz-says-switch-back-to-coal-and-oil-temporary/a-62498429

Gilbert, W. (1958). *De Magnete*. Dover Publications.

Global comparison: how much energy do people consume? - Our World in Data. (n.d.). Retrieved October 9, 2022, from https://ourworldindata.org/per-capita-energy

Goia, A., May, C., & Fusai, G. (2010). Functional clustering and linear regression for peak load forecasting. *International Journal of Forecasting*, *26*(4), 700–711. https://doi.org/10.1016/j.ijforecast.2009.05.015

Goodwin, M., & Yazidi, A. (2014). *A Pattern Recognition Approach for Peak Prediction of Electrical Consumption* (pp. 265–275). https://doi.org/10.1007/978-3-662-44654-6_26

Guarnieri, M. (2012). Birth of Amplification Before Vacuum Tubes [Historical]. *IEEE Industrial Electronics Magazine*, *6*(4), 57–60. https://doi.org/10.1109/MIE.2012.2221357

Hellström, S. (1998). The discovery of static electricity and its manifestation. In *ESD — The Scourge of Electronics* (pp. 1–20). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-80302-4_1

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, *20*(1), 5–10. https://doi.org/10.1016/J.IJFORECAST.2003.09.015

How will the electricity market of the future work? – Analysis - IEA. (n.d.). Retrieved August 14, 2022, from https://www.iea.org/commentaries/how-will-the-electricity-market-of-the-future-work

International Energy Agency, I. (n.d.). *The Power of Transformation - Wind, Sun and the Economics of Flexible Power Systems*. Retrieved July 24, 2022, from http://www.iea.org/termsandconditionsuseandcopyright/

Kim, D., & Jeong, J. (2016). Electricity restructuring, greenhouse gas emissions efficiency and employment reallocation. *Energy Policy*, *92*, 468–476. https://doi.org/10.1016/J.ENPOL.2016.02.009

Kurulu Güç Raporları. (n.d.). Retrieved September 18, 2022, from https://www.teias.gov.tr/kurulu-guc-raporlari

Lagendijk, V. (n.d.). *Europe's Rhine power: connections, borders, and flows*. https://doi.org/10.1007/s12685-015-0140-z

Majstrović, G. (2020). *Transmission and Distribution Systems*.

Nagayama, H. (2007). Effects of regulatory reforms in the electricity supply industry on electricity prices in developing countries. *Energy Policy*, *35*(6), 3440–3462. https://doi.org/10.1016/J.ENPOL.2006.12.018

Newton, I., William, A., & Fierz, M. (1951). *Optik*. Springer.

OECD Reviews of Regulatory Reform: Brazil 2008. (2008). https://doi.org/10.1787/9789264042940-EN

Orhan Aytaç. (2019). *ÖZEL ELEKTRİK ÜRETİM ŞİRKETLERİNE SAĞLANAN "ÖZEL VE GÜZEL" DESTEKLER*.

Palacios M., S., & Saavedra P., E. (2017). Alternative policies for the liberalization of retail electricity markets in Chile. *Utilities Policy*, *49*, 72–92. https://doi.org/10.1016/J.JUP.2017.06.009

Pollitt, M. G. (1997). The impact of liberalization on the performance of the electricity supply industry: an international survey. *Journal of Energy Literature*, *3*, 3–31.

Power Regions. (n.d.). Retrieved October 13, 2022, from https://www.entsoe.eu/regions/

Pumfrey, S. (2004). *Hauksbee, Francis (bap. 1660, d. 1713), natural philosopher and scientific instrument maker* (Vol. 1). Oxford University Press. https://doi.org/10.1093/ref:odnb/12618

R. G. Brown. (1959). *Statistical forecasting for inventory control* (First).

Shin, K. J., & Managi, S. (2017). Liberalization of a retail electricity market: Consumer satisfaction and household switching behavior in Japan. *Energy Policy*, *110*, 675–685. https://doi.org/10.1016/J.ENPOL.2017.07.048

Shinkawa, T. (2018). *Electricity System and Market in Japan*.

Steady decline in capacity utilisation in the German electricity sector. (n.d.). Retrieved September 22, 2022, from https://www.dbresearch.com/servlet/reweb2.ReWEB?rwsite=RPS_EN-PROD&rwobj=ReDisplay.Start.class&document=PROD0000000000495243

Sweeney, J. L. (n.d.). *The California Electricity Crisis: Lessons for the Future*.

Taylor, J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, *204*(1), 139–152. https://doi.org/10.1016/j.ejor.2009.10.003

Taylor, J. W., & McSharry, P. E. (2007). Short-Term Load Forecasting Methods: An Evaluation Based on European Data. *IEEE Transactions on Power Systems*, *22*(4), 2213–2219. https://doi.org/10.1109/TPWRS.2007.907583

V. Experimental researches in electricity. (1832). *Philosophical Transactions of the Royal Society of London*, *122*, 125–162. https://doi.org/10.1098/rstl.1832.0006

Vagliasindi, M., & Besant-Jones, J. (2013). *Power Market Structure*. The World Bank. https://doi.org/10.1596/978-0-8213-9556-1

Vihalemm, T., & Keller, M. (2016). Consumers, citizens or citizen-consumers? Domestic users in the process of Estonian electricity market liberalization. *Energy Research & Social Science*, *13*, 38–48. https://doi.org/10.1016/J.ERSS.2015.12.004

Werner, S. (2017). District heating and cooling in Sweden. *Energy*, *126*, 419–429. https://doi.org/10.1016/j.energy.2017.03.052

Willems, B., & Ehlers, E. (2008). Cross-Subsidies in the Electricity Sector: *Http://Dx.Doi.Org/10.1177/178359170800900301*, *9*(3), 201–227. https://doi.org/10.1177/178359170800900301

Wilson, G. T. (2016). Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. *Journal of Time Series Analysis*, *37*(5), 709–711. https://doi.org/10.1111/jtsa.12194

Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Undefined*, *6*(3), 324–342. https://doi.org/10.1287/MNSC.6.3.324

World Bank Country and Lending Groups – World Bank Data Help Desk. (n.d.). Retrieved August 7, 2022, from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Yukseltan, E., Yucekaya, A., & Bilge, A. H. (2017). Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation. *Applied Energy*, *193*, 287–296. https://doi.org/10.1016/j.apenergy.2017.02.054

Yukseltan, E., Yucekaya, A., & Bilge, A. H. (2020). Hourly electricity demand forecasting using Fourier analysis with feedback. *Energy Strategy Reviews*, *31*, 100524. https://doi.org/10.1016/j.esr.2020.100524

Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., Zong, B., Chen, H., & Chawla, N. v. (2019). A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data. *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, 1409–1416. https://doi.org/10.1609/aaai.v33i01.33011409

Zhixin, Z., & Xin, R. (2011). Causal Relationships between Energy Consumption and Economic Growth. *Energy Procedia*, *5*, 2065–2071. https://doi.org/10.1016/j.egypro.2011.03.356

Zhou, J., Ma, Z., Wei, T., & Li, C. (2021). Threshold Effect of Economic Growth on Energy Intensity—Evidence from 21 Developed Countries. *Energies*, *14*(14), 4199. https://doi.org/10.3390/en14144199

APPENDIX

APPENDIX A. Country Classification (Source: Our World Bank Data)

According to the World Bank, countries are classified in three ways: regional wise, yearly income per person, and landing groups. We used income classification as shown below table in this study. (HI- High Income, UPM- Upper-Middle Income, LMI-Lower Middle Income, and LI-Lower Income)

Country	Class	Country	Class	Country	Class	Country	Class
Afghanistan	LI	Benin	LMI	Chile	HI	Sri Lanka	LMI
Albania	UPM	Bermuda	HI	China	UPM	St. Kitts and Nevis	HI
Algeria	LMI	Bhutan	LMI	Colombia	UPM	St. Lucia	UPM
American Samoa	UPM	Bolivia	LMI	Comoros	LMI	St. Martin (French part)	HI
Andorra	HI	Bosnia and Herzegovina	UPM	Congo, Dem. Rep	LI	St. Vincent and the Grenadines	UPM
Angola	LMI	Botswana	UPM	Congo, Rep.	LMI	Sudan	LI
Antigua and Barbuda	HI	Brazil	UPM	Costa Rica	UPM	Suriname	UPM
Argentina	UPM	British Virgin Islands	HI	Côte d'Ivoire	LMI	Sweden	HI
Armenia	UPM	Brunei Darussalam	HI	Croatia	HI	Switzerland	HI
Aruba	HI	Bulgaria	UPM	Cuba	UPM	Syrian Arab Republic	LI
Australia	HI	Burkina Faso	LI	Curaçao	HI	Taiwan, China	HI
Austria	HI	Burundi	LI	Cyprus	HI	Tajikistan	LMI
Azerbaijan	UPM	Cabo Verde	LMI	Czech Republic	HI	Tanzania	LMI
Bahamas, The	HI	Cambodia	LMI	Denmark	HI	Thailand	UPM
Bahrain	HI	Cameroon	LMI	Djibouti	LMI	Timor-Leste	LMI
Bangladesh	LMI	Canada	HI	Dominica	UPM	Togo	LI
Barbados	HI	Cayman Islands	HI	Dominican Republic	UPM	Tonga	UPM
Belarus	UPM	Central African Republic	LI	Ecuador	UPM	Trinidad and Tobago	HI
Belgium	HI	Chad	LI	Egypt, Arab Rep.	LMI	Tunisia	LMI
Belize	UPM	Channel Islands	HI	El Salvador	LMI	Türkiye	UPM
Equatorial Guinea	UPM	Guatemala	UPM	Kazakhstan	UPM	Turkmenistan	UPM
Eritrea	LI	Guinea	LI	Kenya	LMI	Turks and Caicos Islands	HI
Estonia	HI	Guinea-Bissau	LI	Kiribati	LMI	Tuvalu	UPM
Eswatini	LMI	Guyana	UPM	Korea, Dem. People's Rep	LI	Uganda	LI
Ethiopia	LI	Haiti	LMI	Korea, Rep.	HI	Ukraine	LMI
Faroe Islands	HI	Honduras	LMI	Kosovo	UPM	United Arab Emirates	HI
Fiji	UPM	Hong Kong SAR, China	HI	Kuwait	HI	United Kingdom	HI
Finland	HI	Hungary	HI	Kyrgyz Republic	LMI	United States	HI
France	HI	Iceland	HI	Lao PDR	LMI	Uruguay	HI

French Polynesia	HI	India	LMI	Latvia	HI	Uzbekistan	LMI
Gabon	UPM	Indonesia	LMI	**Lebanon**	LMI	Vanuatu	LMI
Gambia, The	LI	Iran, Islamic Rep	LMI	Lesotho	LMI	Vietnam	LMI
Georgia	UPM	Iraq	UPM	Liberia	LI	Virgin Islands (U.S.)	HI
Germany	HI	Ireland	HI	Libya	UPM	West Bank and Gaza	LMI
Ghana	LMI	Isle of Man	HI	Liechtenstein	HI	Yemen, Rep.	LI
Gibraltar	HI	Israel	HI	Lithuania	HI	**Zambia**	LI
Greece	HI	Italy	HI	Luxembourg	HI	Zimbabwe	LMI
Greenland	HI	Jamaica	UPM	Macao SAR, China	HI		
Grenada	UPM	Japan	HI	Madagascar	LI		
Guam	HI	Jordan	UPM	Malawi	LI		
Malaysia	UPM	New Caledonia	HI	**Romania**	HI		
Maldives	UPM	New Zealand	HI	Russian Federation	UPM		
Mali	LI	Nicaragua	LMI	Rwanda	LI		
Malta	HI	Niger	LI	Samoa	LMI		
Marshall Islands	UPM	Nigeria	LMI	San Marino	HI		
Mauritania	LMI	North Macedonia	UPM	São Tomé and Principe	LMI		
Mauritius	UPM	Northern Mariana Islands	HI	Saudi Arabia	HI		
Mexico	UPM	Norway	HI	Senegal	LMI		
Micronesia, Fed. Sts.	LMI	Oman	HI	Serbia	UPM		
Moldova	UPM	Pakistan	LMI	Seychelles	HI		
Monaco	HI	**Palau**	UPM	Sierra Leone	LI		
Mongolia	LMI	**Panama**	HI	Singapore	HI		
Montenegro	UPM	Papua New Guinea	LMI	Sint Maarten (Dutch part)	HI		
Morocco	LMI	Paraguay	UPM	Slovak Republic	HI		
Mozambique	LI	Peru	UPM	Slovenia	HI		
Myanmar	LMI	Philippines	LMI	Solomon Islands	LMI		
Namibia	UPM	Poland	HI	Somalia	LI		
Nauru	HI	Portugal	HI	South Africa	UPM		
Nepal	LMI	Puerto Rico	HI	South Sudan	LI		
Netherlands	HI	Qatar	HI	Spain	HI		

APPENDIX B. Yearly forecast errors of each country (ND: No Data, T: Training)

Year	AT	BE	BG	CH	CY	CZ	DE	DK	ES	FI	FR	GB	GR
2006	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2007	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2008	3.41	3.61	4.87	10.68	ND	4.34	4.59	3.97	5.70	ND	4.10	ND	5.86
2009	5.87	7.79	8.66	9.72	ND	7.32	8.20	6.03	7.51	ND	5.83	ND	7.06
2010	6.43	13.95	6.70	7.30	ND	10.03	13.48	44.86	7.08	T	5.93	T	7.37
2011	13.62	11.24	6.40	7.34	ND	4.23	7.47	45.31	4.39	T	8.81	T	5.38
2012	15.80	3.92	9.19	8.92	ND	4.01	6.04	4.47	4.11	3.32	5.71	4.00	6.07
2013	3.30	5.87	6.32	8.14	ND	4.41	4.56	9.01	4.00	2.94	5.23	4.05	9.38
2014	3.47	4.57	4.92	5.38	T	3.50	5.25	11.61	4.19	2.83	5.06	3.71	8.72
2015	3.57	3.84	6.35	25.75	T	3.52	3.69	6.15	4.39	2.75	5.90	4.41	11.58
2016	7.63	6.12	12.49	25.32	11.34	5.94	7.30	7.95	7.34	4.36	7.36	11.17	13.44
2017	4.80	4.14	8.22	7.30	6.94	4.70	4.87	5.05	5.04	3.23	5.72	23.63	6.43
2018	3.93	5.55	18.81	4.82	11.81	3.93	3.42	4.01	3.95	2.89	5.41	9.66	4.96

Year	HR	HU	IE	IS	IT	NL	NO	PL	PT	RO	RS	SE	SI
2006	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2007	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2008	4.14	4.60	T	ND	5.78	6.35	ND	4.39	4.02	5.31	6.47	ND	13.89
2009	5.17	5.77	T	ND	6.87	12.37	ND	5.40	4.49	10.99	5.13	ND	5.63
2010	4.20	3.78	6.81	T	11.95	11.34	T	8.57	4.35	12.56	4.39	T	19.54
2011	4.25	7.79	4.14	T	6.20	4.23	T	3.43	5.33	3.36	3.59	T	5.60
2012	5.27	3.84	5.05	1.37	4.66	4.88	6.22	3.88	3.60	5.86	5.18	4.16	6.67
2013	4.68	3.92	4.45	1.94	6.20	8.99	4.51	3.30	5.39	5.10	5.54	3.97	6.20
2014	4.28	3.48	3.90	5.83	4.54	7.82	2.83	3.42	3.36	6.33	4.85	3.46	6.00
2015	5.40	3.55	4.45	8.30	5.87	4.67	3.73	3.30	3.69	5.29	4.69	4.29	4.40
2016	8.66	6.49	8.61	7.13	10.20	6.44	4.66	6.99	9.13	7.73	8.74	6.13	8.02
2017	5.93	4.31	4.54	4.55	6.90	4.74	4.85	4.81	4.94	5.32	5.87	4.65	5.36
2018	5.94	4.37	3.06	1.91	5.15	4.03	4.76	3.56	3.90	3.90	4.65	4.42	4.77

APPENDIX C. Daily Rolling forecast errors of each country (ND: No Data, T: Training)

Year	AT	BE	BG	CH	CY	CZ	DE	DK	ES	FI	FR	GB	GR
2006	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2007	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2008	3.69	3.36	5.94	6.01	ND	4.75	4.01	3.83	4.96	ND	5.37	ND	5.65
2009	4.39	4.91	5.49	6.40	ND	4.43	4.77	3.79	5.47	ND	5.14	ND	5.33
2010	4.08	5.22	6.42	6.35	ND	4.88	5.30	13.74	4.87	T	5.37	T	5.84
2011	5.86	4.83	5.38	5.75	ND	3.86	4.72	14.19	3.84	T	5.27	T	5.42
2012	5.73	4.21	6.97	5.83	ND	3.87	4.38	3.71	3.99	5.56	5.89	3.75	6.14
2013	3.50	3.92	7.00	6.23	ND	4.31	3.87	7.62	3.88	4.38	6.15	4.35	8.06
2014	3.70	3.73	4.97	5.27	T	3.83	3.73	6.10	3.50	3.83	6.11	3.99	6.83
2015	3.45	3.22	6.57	11.06	T	3.38	2.86	3.89	4.20	3.26	4.95	4.11	9.23
2016	6.37	4.83	10.37	9.39	9.84	5.17	5.69	6.48	6.19	4.56	6.90	7.91	9.21
2017	4.26	3.45	7.69	5.28	7.29	4.42	3.70	3.84	4.29	3.65	6.00	8.89	6.64
2018	4.22	5.19	6.86	4.79	10.60	4.19	3.31	3.89	3.99	3.99	6.03	6.67	5.42

Year	HR	HU	IE	IS	IT	NL	NO	PL	PT	RO	RS	SE	SI
2006	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2007	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2008	5.00	3.87	T	ND	5.05	5.12	ND	4.00	3.83	5.13	6.09	ND	6.00
2009	4.02	3.85	T	ND	5.56	6.59	ND	3.60	4.42	4.75	4.97	ND	5.41
2010	4.29	3.59	4.13	T	5.68	4.40	T	3.91	4.11	4.97	5.41	T	8.32
2011	4.44	4.08	3.95	T	4.64	2.75	T	3.40	3.54	3.65	4.56	T	4.96
2012	5.06	3.65	5.23	1.19	4.68	2.57	5.87	3.37	3.63	3.91	5.20	6.16	5.58
2013	4.71	3.65	3.24	1.38	4.79	3.99	6.46	3.56	3.56	4.02	5.58	6.11	5.72
2014	4.71	3.34	3.27	3.63	4.53	3.70	4.88	3.55	3.42	4.54	5.09	5.24	4.98
2015	4.49	3.73	3.41	4.57	5.36	3.19	3.86	3.09	3.66	4.97	4.20	3.87	3.77
2016	7.45	5.51	6.53	2.06	8.42	4.99	5.43	5.82	7.42	6.95	7.63	6.45	6.38
2017	5.28	3.76	3.48	2.06	5.83	3.78	4.88	4.09	4.05	4.66	6.06	4.72	4.30
2018	5.49	3.79	2.79	1.73	4.74	3.60	4.44	3.39	3.98	3.53	5.26	5.74	4.41

APPENDIX D. Hourly Rolling forecast with feedback correction errors of each country (ND: No Data, T: Training)

Year	AT	BE	BG	CH	CY	CZ	DE	DK	ES	FI	FR	GB	GR
2006	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2007	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2008	1.01	0.93	1.32	2.32	ND	1.47	1.00	1.32	2.11	ND	0.93	ND	1.31
2009	1.09	1.01	1.36	1.81	ND	1.47	1.08	1.43	1.76	ND	0.93	ND	1.32
2010	1.01	0.99	1.44	1.87	ND	1.43	1.07	1.79	1.27	T	0.88	T	1.30
2011	1.34	1.11	1.24	1.66	ND	1.17	1.10	1.43	1.17	T	0.92	T	1.34
2012	1.15	1.09	1.30	1.69	ND	1.33	1.08	1.18	1.18	0.75	0.93	1.40	1.48
2013	1.08	1.19	1.29	1.70	ND	1.57	1.02	1.83	1.14	0.72	0.94	1.19	2.07
2014	1.08	0.98	1.18	1.75	T	1.11	0.91	1.21	1.08	0.70	0.94	1.21	1.65
2015	1.04	0.89	1.16	1.68	T	1.01	0.84	1.22	1.03	0.67	0.90	1.34	2.05
2016	2.73	2.33	3.85	2.38	3.34	2.24	2.41	3.13	2.57	1.71	3.26	3.46	2.99
2017	1.45	1.30	1.68	2.13	1.80	1.28	1.18	1.60	1.29	0.89	1.55	2.03	1.56
2018	1.15	1.32	1.10	1.93	2.09	1.10	0.87	1.18	0.99	0.73	1.21	1.78	1.18

Year	HR	HU	IE	IS	IT	NL	NO	PL	PT	RO	RS	SE	SI
2006	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2007	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2008	1.26	1.83	T	ND	1.36	1.23	ND	0.90	1.51	1.33	0.88	ND	1.67
2009	1.28	1.69	T	ND	1.36	1.34	ND	0.91	1.53	1.02	1.18	ND	1.95
2010	1.42	1.56	1.46	T	1.36	1.12	T	0.87	1.42	0.98	1.23	T	2.36
2011	1.56	1.03	1.57	T	1.21	0.95	T	0.87	1.41	1.00	1.15	T	2.04
2012	1.61	0.92	1.55	0.61	1.23	0.93	0.87	0.86	1.38	0.91	1.15	1.02	2.20
2013	1.55	0.90	1.33	0.56	1.20	0.92	0.98	0.85	1.35	0.98	1.14	0.92	2.48
2014	1.55	0.90	1.62	0.67	1.20	0.90	0.91	0.87	1.33	1.14	1.12	0.85	1.81
2015	1.58	0.86	1.84	0.65	1.28	0.92	0.85	0.85	1.33	1.73	1.12	0.81	1.52
2016	3.79	2.31	3.09	0.65	3.56	2.30	1.51	2.47	3.71	3.06	3.16	2.02	3.06
2017	1.94	1.19	1.63	0.65	1.76	1.38	1.00	1.24	1.48	1.41	1.42	1.30	1.62
2018	1.49	0.95	1.16	0.72	1.29	1.04	0.83	0.89	1.10	0.84	1.09	1.07	1.42

APPENDIX E. Hourly Rolling forecast with AR correction errors of each country (ND: No Data, T: Training)

Year	AT	BE	BG	CH	CY	CZ	DE	DK	ES	FI	FR	GB	GR
2006	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2007	T	T	T	T	ND	T	T	T	T	ND	T	ND	T
2008	0.93	0.87	1.32	2.22	ND	1.26	0.93	1.15	1.71	ND	0.86	ND	1.21
2009	0.99	0.96	1.33	1.78	ND	1.27	1.01	1.25	1.72	ND	0.87	ND	1.18
2010	0.96	1.02	1.48	1.82	ND	1.18	1.10	2.17	1.15	T	0.83	T	1.19
2011	1.25	1.10	1.24	1.61	ND	1.03	1.03	2.27	1.01	T	0.92	T	1.15
2012	1.22	1.00	1.31	1.63	ND	1.13	1.02	1.03	1.04	0.76	0.91	1.20	1.27
2013	1.01	1.09	1.29	1.64	ND	1.30	0.95	1.83	1.01	0.69	0.88	1.04	1.72
2014	1.02	0.91	1.10	1.68	T	1.00	0.82	1.14	0.92	0.66	0.90	1.05	1.52
2015	0.99	0.84	1.10	1.82	T	0.92	0.74	1.14	0.89	0.64	0.85	1.18	1.65
2016	2.01	1.70	2.73	2.20	2.54	1.69	1.78	2.26	1.84	1.28	2.27	2.49	2.14
2017	1.32	1.18	1.61	2.10	1.70	1.16	1.11	1.52	1.18	0.83	1.36	2.00	1.42
2018	1.03	1.16	1.22	1.90	1.85	0.97	0.80	1.14	0.86	0.67	0.96	1.59	1.02

Year	HR	HU	IE	IS	IT	NL	NO	PL	PT	RO	RS	SE	SI
2006	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2007	T	T	ND	ND	T	T	ND	T	T	T	T	ND	T
2008	1.12	1.46	T	ND	1.17	1.15	ND	0.82	1.26	1.20	1.02	ND	1.67
2009	1.09	1.42	T	ND	1.21	1.35	ND	0.82	1.23	1.03	1.14	ND	1.92
2010	1.22	1.18	1.23	T	1.20	1.00	T	0.83	1.18	1.01	1.18	T	2.53
2011	1.32	0.99	1.27	T	1.07	0.83	T	0.80	1.17	0.92	1.10	T	2.00
2012	1.38	0.81	1.32	0.69	1.05	0.82	0.98	0.78	1.17	0.85	1.09	0.95	2.16
2013	1.33	0.81	1.07	0.65	1.08	0.87	1.01	0.76	1.14	0.91	1.07	0.92	2.48
2014	1.31	0.77	1.23	0.85	1.06	0.84	0.91	0.78	1.13	1.03	1.04	0.83	1.62
2015	1.35	0.76	1.37	0.90	1.14	0.81	0.87	0.76	1.14	1.31	1.05	0.79	1.49
2016	2.74	1.66	2.22	0.79	2.57	1.73	1.25	1.76	2.68	2.06	2.28	1.55	2.34
2017	1.78	1.07	1.36	0.77	1.64	1.24	1.01	1.15	1.35	1.30	1.40	1.19	1.58
2018	1.30	0.84	0.99	0.84	1.18	0.95	0.87	0.83	0.98	0.77	1.05	1.07	1.36

APPENDIX F. AR Model Coefficient for each hour

Predicted Hour	Lag 1	Lag 1 Hour	Lag 1 Value	Lag 2	Lag 2 Hour	Lag 2 Value	Lag 3	Lag 3 Hour	Lag 3 Value
13:00	t-1	12:00	0.96	t-24	13:00	0.49	t-25	12:00	-0.46
14:00	t-2	12:00	0.82	t-24	14:00	0.17	t-25	13:00	-0.02
15:00	t-3	12:00	0.71	t-24	15:00	0.28	t-25	14:00	-0.03
16:00	t-4	12:00	0.60	t-24	16:00	0.37	t-25	15:00	-0.02
17:00	t-5	12:00	0.54	t-24	17:00	0.50	t-25	16:00	-0.08
18:00	t-6	12:00	0.46	t-24	18:00	0.51	t-25	17:00	-0.03
19:00	t-7	12:00	0.40	t-24	19:00	0.60	t-25	18:00	-0.07
20:00	t-8	12:00	0.36	t-24	20:00	0.61	t-25	19:00	-0.04
21:00	t-9	12:00	0.33	t-24	21:00	0.65	t-25	20:00	-0.05
22:00	t-10	12:00	0.30	t-24	22:00	0.66	t-25	21:00	-0.04
23:00	t-11	12:00	0.28	t-24	23:00	0.67	t-25	22:00	-0.05
00:00	t-12	12:00	0.26	t-24	00:00	0.67	t-25	23:00	-0.03
01:00	t-13	12:00	0.24	t-24	01:00	0.69	t-25	00:00	-0.03
02:00	t-14	12:00	0.22	t-24	02:00	0.69	t-25	01:00	-0.03
03:00	t-15	12:00	0.20	t-24	03:00	0.67	t-25	02:00	0.00
04:00	t-16	12:00	0.19	t-24	04:00	0.65	t-25	03:00	0.03
05:00	t-17	12:00	0.19	t-24	05:00	0.62	t-25	04:00	0.06
06:00	t-18	12:00	0.19	t-24	06:00	0.60	t-25	05:00	0.08
07:00	t-19	12:00	0.19	t-24	07:00	0.59	t-25	06:00	0.08
08:00	t-20	12:00	0.20	t-24	08:00	0.57	t-25	07:00	0.08
09:00	t-21	12:00	0.23	t-24	09:00	0.50	t-25	08:00	0.12
10:00	t-22	12:00	0.29	t-24	10:00	0.45	t-25	09:00	0.11
11:00	t-23	12:00	0.36	t-24	11:00	0.38	t-25	10:00	0.10
12:00	t-24	12:00	0.90	t-24	12:00	0.07	t-25	11:00	-0.17
13:00	t-25	12:00	0.95	t-48	13:00	0.06	t-49	12:00	-0.23
14:00	t-26	12:00	0.91	t-48	14:00	-0.03	t-49	13:00	-0.12
15:00	t-27	12:00	0.87	t-48	15:00	-0.01	t-49	14:00	-0.13
16:00	t-28	12:00	0.85	t-48	16:00	0.00	t-49	15:00	-0.15
17:00	t-29	12:00	0.83	t-48	17:00	0.00	t-49	16:00	-0.17
18:00	t-30	12:00	0.81	t-48	18:00	0.00	t-49	17:00	-0.18
19:00	t-31	12:00	0.78	t-48	19:00	0.02	t-49	18:00	-0.20
20:00	t-32	12:00	0.76	t-48	20:00	-0.02	t-49	19:00	-0.16
21:00	t-33	12:00	0.75	t-48	21:00	-0.13	t-49	20:00	-0.05
22:00	t-34	12:00	0.74	t-48	22:00	-0.18	t-49	21:00	0.00
23:00	t-35	12:00	0.74	t-48	23:00	-0.20	t-49	22:00	0.02

APPENDIX G. Yearly profit (TL) changes with different new and original equilibrium price ratios.

Ratio	2019	2020	2021
1.00	-	-	-
1.01	70,953,222	76,892,837	65,740,164
1.02	104,913,293	117,997,441	105,892,001
1.03	115,858,077	132,828,963	123,327,648
1.04	121,487,981	139,108,849	130,586,024
1.05	125,749,071	141,680,420	134,505,110
1.06	128,546,400	143,045,543	136,657,082
1.07	130,021,800	145,110,475	138,641,797
1.08	131,868,610	146,158,326	**139,476,236**
1.09	132,817,235	146,905,664	139,230,936
1.10	134,397,957	147,122,564	138,692,520
1.11	135,105,728	147,633,401	137,609,542
1.12	135,666,208	148,383,182	136,383,027
1.13	136,015,725	148,455,697	135,314,559
1.14	136,327,213	148,189,593	133,633,656
1.15	136,568,463	148,608,826	131,880,332
1.16	138,090,599	148,435,814	129,335,392
1.17	**138,134,915**	**148,817,491**	127,018,196
1.18	137,789,863	147,976,900	124,050,738
1.19	137,447,516	147,317,527	121,473,626
1.20	137,377,664	146,721,305	119,030,464
1.21	137,695,011	145,935,761	116,609,267
1.22	137,348,303	144,957,984	114,117,593
1.23	137,267,158	143,346,109	110,810,036
1.24	136,930,183	142,223,302	107,820,326
1.25	137,108,170	141,764,049	104,495,750
1.26	137,027,021	140,547,462	101,514,705
1.27	136,751,708	139,194,154	98,584,478
1.28	136,788,003	138,324,597	95,385,549
1.29	136,088,057	136,943,263	92,573,323
1.30	135,589,026	136,280,934	90,248,642
1.40	130,154,483	123,352,210	61,662,208
1.60	118,880,139	94,764,924	19,581,014
1.80	110,285,180	68,075,017	4,832,302
2.00	102,982,427	47,487,583	743,238

APPENDIX H. The result of the first scenario that uses forecasted imbalance quantity with different storage levels

Storage (MW)	2019			2020			2021		
	Positive	Negative	Total	Positive	Negative	Total	Positive	Negative	Total
0	-	59.3 M	59.3 M	-	71.3 M	71.3 M	-	30.4 M	30.4 M
100	1.9 M	76.4 M	78.3 M	2.3 M	89.2 M	91.4 M	3.8 M	60.7 M	64.6 M
200	3.9 M	87.8 M	91.7 M	4.6 M	101.3 M	105.9 M	8.0 M	82.0 M	90.1 M
300	6.1 M	95.3 M	101.4 M	7.1 M	109.4 M	116.5 M	12.5 M	96.8 M	109.3 M
400	8.5 M	100.6 M	109.0 M	9.7 M	114.7 M	124.4 M	17.1 M	107.0 M	124.1 M
500	10.9 M	104.0 M	114.9 M	12.2 M	118.6 M	130.8 M	21.8 M	113.6 M	135.5 M
600	13.3 M	106.1 M	119.4 M	14.6 M	121.5 M	136.0 M	26.3 M	118.5 M	144.8 M
700	15.5 M	107.5 M	123.0 M	16.8 M	123.2 M	140.0 M	30.6 M	121.7 M	152.4 M
800	17.8 M	108.5 M	126.3 M	19.0 M	124.2 M	143.2 M	34.6 M	123.9 M	158.6 M
900	19.9 M	109.3 M	129.2 M	20.9 M	125.0 M	145.9 M	38.5 M	125.5 M	164.0 M
1000	22.0 M	109.7 M	131.7 M	22.9 M	125.6 M	148.4 M	42.2 M	126.6 M	168.9 M
1100	24.0 M	110.0 M	134.0 M	24.7 M	126.0 M	150.6 M	45.7 M	127.5 M	173.2 M
1200	25.8 M	110.2 M	136.1 M	26.4 M	126.3 M	152.7 M	49.1 M	128.2 M	177.2 M
1300	27.6 M	110.4 M	138.0 M	28.0 M	126.7 M	154.6 M	52.2 M	128.8 M	180.9 M
1400	29.2 M	110.5 M	139.7 M	29.5 M	126.9 M	156.4 M	55.1 M	129.2 M	184.3 M
1500	30.7 M	110.6 M	141.3 M	31.0 M	127.1 M	158.1 M	57.9 M	129.4 M	187.3 M
1600	32.1 M	110.7 M	142.8 M	32.3 M	127.2 M	159.5 M	60.6 M	129.6 M	190.2 M
1700	33.5 M	110.7 M	144.2 M	33.6 M	127.2 M	160.8 M	63.1 M	129.7 M	192.7 M
1800	34.8 M	110.7 M	145.5 M	34.8 M	127.3 M	162.1 M	65.4 M	129.8 M	195.2 M
1900	36.0 M	110.7 M	146.8 M	36.0 M	127.3 M	163.3 M	67.6 M	129.9 M	197.5 M
2000	37.3 M	110.7 M	148.0 M	37.1 M	127.4 M	164.5 M	69.7 M	129.9 M	199.7 M
2500	43.1 M	110.8 M	153.9 M	42.3 M	127.5 M	169.8 M	79.0 M	130.2 M	209.2 M
3000	48.0 M	110.9 M	158.9 M	47.0 M	127.7 M	174.6 M	86.6 M	130.3 M	217.0 M
3500	52.6 M	110.9 M	163.4 M	51.3 M	127.9 M	179.2 M	93.5 M	130.8 M	224.3 M
4000	56.4 M	111.0 M	167.3 M	55.1 M	127.9 M	183.0 M	99.9 M	131.0 M	230.8 M
4500	59.7 M	111.2 M	170.9 M	58.7 M	128.0 M	186.7 M	105.5 M	131.2 M	236.6 M
5000	62.7 M	111.3 M	174.0 M	62.0 M	128.0 M	190.0 M	110.1 M	131.3 M	241.4 M
5500	65.7 M	111.3 M	177.0 M	64.9 M	128.1 M	193.1 M	114.1 M	131.8 M	245.9 M
6000	68.4 M	111.3 M	179.7 M	67.7 M	128.1 M	195.9 M	118.0 M	131.8 M	249.8 M
6500	71.0 M	111.3 M	182.3 M	70.5 M	128.1 M	198.6 M	121.8 M	131.8 M	253.6 M
7000	73.4 M	111.3 M	184.8 M	73.3 M	128.1 M	201.4 M	125.4 M	132.0 M	257.4 M
7500	75.6 M	111.4 M	187.0 M	76.2 M	128.2 M	204.3 M	128.8 M	132.0 M	260.8 M
8000	77.6 M	111.6 M	189.2 M	79.0 M	128.2 M	207.1 M	131.9 M	132.0 M	263.9 M
8500	79.6 M	111.6 M	191.2 M	81.6 M	128.2 M	209.8 M	134.8 M	132.0 M	266.8 M
9000	81.7 M	111.6 M	193.2 M	84.1 M	128.3 M	212.4 M	137.7 M	132.0 M	269.7 M
9500	83.6 M	111.6 M	195.2 M	86.1 M	128.5 M	214.6 M	140.6 M	132.2 M	272.7 M
10000	85.5 M	111.6 M	197.1 M	87.9 M	128.5 M	216.4 M	143.2 M	132.5 M	275.7 M

APPENDIX I. The result of the second scenario that uses actual imbalance quantity with different storage levels

Storage (MW)	2019			2020			2021		
	Positive	Negative	Total	Positive	Negative	Total	Positive	Negative	Total
0	-	-	-	-	-	-	-	-	-
100	6.3 M	19.4 M	25.7 M	5.8 M	20.2 M	26.0 M	10.5 M	5.6 M	16.1 M
200	12.1 M	34.8 M	46.9 M	11.1 M	36.3 M	47.4 M	19.9 M	12.1 M	32.0 M
300	17.3 M	48.0 M	65.3 M	15.8 M	50.1 M	65.9 M	28.3 M	20.8 M	49.1 M
400	22.0 M	59.7 M	81.8 M	20.2 M	62.3 M	82.5 M	35.8 M	29.9 M	65.7 M
500	26.3 M	70.2 M	96.6 M	24.1 M	73.1 M	97.2 M	42.6 M	38.7 M	81.2 M
600	30.2 M	79.6 M	109.8 M	27.7 M	82.7 M	110.4 M	48.9 M	52.5 M	101.3 M
700	33.7 M	88.2 M	121.9 M	31.0 M	91.4 M	122.4 M	54.8 M	68.7 M	123.5 M
800	37.0 M	96.2 M	133.2 M	34.0 M	99.2 M	133.3 M	60.2 M	84.7 M	144.9 M
900	40.0 M	103.6 M	143.6 M	36.8 M	106.4 M	143.2 M	65.3 M	94.1 M	159.4 M
1000	42.9 M	110.4 M	153.3 M	39.3 M	113.1 M	152.4 M	70.1 M	109.4 M	179.5 M
1100	45.5 M	116.6 M	162.2 M	41.7 M	119.4 M	161.1 M	74.6 M	124.2 M	198.8 M
1200	48.0 M	122.3 M	170.3 M	44.0 M	125.4 M	169.4 M	79.0 M	139.0 M	218.0 M
1300	50.3 M	127.8 M	178.1 M	46.2 M	131.1 M	177.3 M	83.0 M	149.6 M	232.6 M
1400	52.6 M	132.9 M	185.5 M	48.3 M	136.6 M	184.9 M	86.8 M	157.6 M	244.4 M
1500	54.7 M	137.8 M	192.5 M	50.3 M	142.0 M	192.3 M	90.3 M	165.9 M	256.2 M
1600	56.8 M	142.4 M	199.2 M	52.3 M	147.3 M	199.6 M	93.7 M	174.1 M	267.7 M
1700	58.8 M	146.8 M	205.6 M	54.2 M	152.2 M	206.4 M	96.9 M	182.6 M	279.5 M
1800	60.8 M	151.1 M	211.9 M	56.0 M	157.0 M	213.0 M	100.0 M	192.1 M	292.1 M
1900	62.6 M	155.0 M	217.6 M	57.8 M	161.6 M	219.4 M	102.6 M	200.4 M	303.0 M
2000	64.4 M	158.8 M	223.2 M	59.5 M	166.0 M	225.5 M	105.2 M	209.6 M	314.7 M
2500	72.5 M	175.4 M	247.9 M	67.0 M	185.4 M	252.4 M	116.0 M	241.5 M	357.6 M
3000	79.5 M	190.1 M	269.6 M	73.4 M	201.9 M	275.3 M	125.4 M	268.9 M	394.3 M
3500	85.8 M	202.9 M	288.8 M	79.0 M	216.4 M	295.4 M	133.7 M	291.9 M	425.6 M
4000	91.3 M	214.5 M	305.8 M	84.1 M	229.0 M	313.2 M	141.1 M	315.3 M	456.4 M
4500	96.1 M	225.2 M	321.3 M	88.8 M	240.0 M	328.8 M	148.2 M	334.0 M	482.2 M
5000	100.3 M	234.8 M	335.1 M	93.1 M	250.6 M	343.7 M	154.9 M	357.1 M	512.0 M
5500	104.6 M	243.8 M	348.3 M	97.1 M	261.0 M	358.1 M	160.9 M	373.3 M	534.2 M
6000	108.3 M	252.1 M	360.4 M	100.6 M	270.4 M	371.0 M	166.4 M	387.5 M	553.9 M
6500	111.7 M	259.5 M	371.1 M	104.1 M	279.2 M	383.3 M	171.4 M	403.2 M	574.7 M
7000	114.8 M	266.4 M	381.2 M	107.3 M	287.4 M	394.7 M	175.9 M	418.9 M	594.9 M
7500	117.7 M	272.6 M	390.4 M	110.4 M	294.8 M	405.2 M	180.0 M	428.3 M	608.4 M
8000	120.4 M	278.2 M	398.6 M	113.2 M	301.3 M	414.5 M	183.7 M	439.2 M	622.9 M
8500	123.0 M	283.8 M	406.9 M	115.7 M	307.1 M	422.8 M	187.3 M	447.2 M	634.5 M
9000	125.5 M	289.0 M	414.5 M	118.1 M	312.3 M	430.4 M	190.5 M	458.8 M	649.3 M
9500	127.8 M	293.8 M	421.7 M	120.3 M	317.6 M	437.9 M	193.5 M	469.0 M	662.6 M
10000	130.1 M	298.3 M	428.4 M	122.4 M	322.1 M	444.5 M	196.1 M	472.4 M	668.5 M

APPENDIX J. Countries in the ENTSO-E system and their basic information

Country	Code	MHC (All Available Years)	MHC (Last Available Year)	Population
Türkiye (1)	TR	32576	33352	84340
Germany	DE	59840	59085	83191
United Kingdom	GB	38391	38831	67886
France	FR	55097	54303	67287
Italy	IT	36634	36777	59258
Spain	ES	29076	28946	47431
Ukraine (2)	UA	630	634	41902
Poland	PL	16785	18468	38268
Romania	RO	6065	6402	19266
Netherlands	NL	12760	13302	17425
Belgium	BE	9886	9734	11493
Czech Republic	CZ	7285	7620	10702
Greece	GR	5873	5875	10689
Sweden	SE	15943	16095	10379
Portugal	PT	5737	5810	10305
Hungary	HU	4702	4928	9770
Austria	AT	7440	8132	8901
Switzerland	CH	5973	6874	8637
Bulgaria	BG	4254	3876	6917
Serbia	RS	4499	4476	6908
Denmark	DK	3426	3899	5823
Finland	FI	9692	9977	5536
Slovakia	SK	3244	3369	5460
Norway	NO	14753	15460	5368
Ireland	IE	3048	3299	4995
Croatia	HR	1994	2092	4047
Bosnia and Herzegovina (4)	BA	1382	1424	3281
Albania (3)	AL	814	818	2878
Lithuania (4)	LT	1250	1382	2795
Slovenia (4)	SI	1504	1640	2100
North Macedonia (3)	MK	913	779	2083
Latvia (4)	LV	815	838	1902
Northern Ireland (3)	NI	1025	1003	1890
Estonia (4)	EE	920	959	1331
Cyprus (3)	CY	526	586	1207
Luxembourg (3)	LU	734	729	632
Montenegro (3)	ME	408	388	622
Iceland (3)	IS	2026	2207	366

(1) Data excluded because Türkiye is not part of the ENTSOE grid
(2) Data excluded because only west Ukraine data is available
(3) Data excluded due to insufficiency
(4) Data excluded because of low consumption

CURRICULUM VITAE

Personal Information

Name and surname : Ergün YÜKSELTAN

Academic Background

Bachelor's Degree Education : Enerji Sistemleri Mühendisliği (2006-2011)

Bachelor's Degree Education : Endüstri Mühendisliği (2007-2012)

Post Graduate Education : Finans Mühendisliği (2013-2015)

Foreign Languages : İngilizce

Work Experience

Institutions Served and Their Dates:

2012-2016 Samsung Electronics Türkiye

2022-… Arçelik A.Ş.