

APPLICATIONS OF TIME-SERIES METHODS IN ELECTRICITY PRICE FORECASTING

A Thesis

by

Toygar Ülgen

Submitted to the
Graduate School of Sciences and Engineering
In Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in the
Department of Electrical and Electronics Engineering

Özyeğin University
January 2021

Copyright © 2021 by Toygar Ülgen

APPLICATIONS OF TIME-SERIES METHODS IN ELECTRICITY PRICE FORECASTING

Approved by:

Asst. Professor Göktürk Poyrazoğlu
Advisor
Department of Electrical and Electronics
Engineering
Özyeğin University

Assoc. Professor Cenk Demiroğlu
Department of Electrical and Electronics
Engineering
Özyeğin University

Asst. Professor Emre Çelebi
Department of Industrial Engineering
Yeditepe University

Date Approved: 18 January 2021



To my family...

ABSTRACT

The prediction of day-ahead electricity prices with higher accuracy is always helpful for the market players of the power exchange. Models were designed in the first place to find out the best time-series estimation method for the selected 12 European countries before and during COVID-19 using exponential smoothing family and Naïve processing. The classification approach is followed by 33 different features of each country to answer the question of the best method for other countries. Also, 24 hours are estimated without any independent variables using seasonal autoregressive integrated moving average. Apart from the statistical forecasting model, long-short term memory is used for each European country. The multiple linear regression is used on electricity price forecasting. Different predictors are analyzed to reduce the mean absolute percentage error. The training data includes the dates from the day-ahead electricity market in Turkey. It is proved that the lagged electricity prices such as previous one day, one week, and lagged moving average prices play a key role in electricity price estimation. Aside from other valuable coefficients, natural gas, oil, and coal prices are tested in the forecasting model. The error rates of the fuel prices are noticeably decreased by using multiple linear regression that generates more accurate results and crucial variables influencing hourly electricity price have been determined. Different training data length is an essential part of decreasing the error proportions in the electricity price estimation. Also, it is analyzed that there is no big difference regarding the error rates if it is compared to the Regular and Dynamic Regression models in the forecast of electricity prices. Regular and Dynamic Feedforward Neural Network Methods were utilized for the model to give better error performance and more accurate results were obtained.

ÖZETÇE

Gün öncesi elektrik fiyatlarının daha yüksek doğrulukta tahmin edilmesi, enerji borsasının piyasa oyuncuları için her zaman yararlıdır. Modeller, üstel tahminleme metotları ve naive işleme kullanılarak COVID-19 öncesi ve boyunca seçilen 12 Avrupa ülkesi için en iyi zaman serisi tahmin yöntemini bulmak için ilk olarak tasarlandı. Daha sonra, diğer ülkeler için en iyi yöntemin hangisi olduğu sorusuna cevap vermek için her ülkenin 33 farklı özelliği bir sınıflandırma yaklaşımını takip eder. Ayrıca, mevsimsel otoregresif entegre hareketli ortalama kullanılarak herhangi bir bağımsız değişken olmaksızın 24 saat tahmin edilir. İstatistiksel tahmin modelinin yanı sıra, her Avrupa ülkesi için uzun-kısa süreli bellek kullanılmaktadır. Elektrik fiyatı tahmininde çoklu doğrusal regresyon kullanıldı. Ortalama mutlak yüzde hatasını azaltmak için farklı tahmin ediciler analiz edilir. Eğitim verileri, Türkiye'deki gün öncesi elektrik piyasasından alınan tarihleri içermektedir. Elektrik fiyat tahmininde önceki bir gün, bir hafta gibi gecikmeli elektrik fiyatları ve gecikmeli hareketli ortalama fiyatların önemli rol oynadığı kanıtlanmıştır. Tahmin modelinde diğer değerli katsayıların yanı sıra doğalgaz, petrol ve kömür fiyatları test edilmektedir. Daha doğru sonuçlar üreten çoklu doğrusal regresyon kullanılarak akaryakıt fiyatlarının hata oranları belirgin şekilde düşürülmüş ve saatlik elektrik fiyatını etkileyen önemli değişkenler belirlenmiştir. Farklı eğitim veri uzunluğu, elektrik fiyatı tahmininde hata oranlarını düşürmenin önemli bir parçasıdır. Ayrıca elektrik fiyatları tahmininde Düzenli ve Dinamik Regresyon modeli ile karşılaştırıldığında hata oranlarında büyük bir fark olmadığı analiz edilmektedir. Modelde daha iyi hata performansı vermesi için Düzenli ve Dinamik İleri Beslemeli Sinir Ağı Yöntemleri kullanılmış ve daha doğru sonuçlar elde edilmiştir.

ACKNOWLEDGEMENTS

This work would have been impossible without the contributions of my family and friends. I also would like to thank my thesis committee for their valuable time. And finally, thanks to my family for their massive support.



TABLE OF CONTENTS

DEDICATION	iii
ABSTRACT	iv
ÖZETÇE	v
ACKNOWLEDGEMENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	xi
I INTRODUCTION	1
II TIME-SERIES PREDICTION METHODS	8
2.1 Data Pre-Processing For Naive and Smoothing Methods	8
2.2 Naive Method	10
2.3 Simple Exponential Smoothing Method	12
2.4 Holt's Method	15
2.5 Holt-Winter's Method	16
2.6 Results of Smoothing Methods and Naive Method	18
2.7 Impact of COVID-19 on Electricity Prices in European Countries . .	23
2.7.1 Impacts of COVID-19 on Prediction Methods	27
2.8 Seasonal Auto-Regressive Integrated Moving-Average	29
2.8.1 Data Pre-Processing For SARIMA	30
2.9 Long Short-Term Memory (LSTM)	34
2.9.1 Data Pre-Processing For LSTM	36
III CLASSIFICATION METHODS	42
3.1 K-Nearest Neighbors (kNN)	45
3.2 Decision Tree	45
3.3 Support Vector Machine (SVM)	45
3.4 Gradient Boosting Tree Classifier	45

3.5	AdaBoost Classifier	45
3.6	Stochastic Gradient Descent	46
3.7	Random Forest	46
3.8	Bagging Classifier	46
3.9	Ridge Classifier	46
3.10	Results of Classification Methods	46
IV	MULTIPLE LINEAR REGRESSION	49
4.1	Data Pre-Processing For Multiple Linear Regression	49
4.2	Impact of Historical Electricity Prices	51
4.3	Impact of Fuel Prices	52
4.4	Impact of Training Data Length	54
4.5	Dynamic Forecast Method	55
4.6	Feed-forward Neural Network	59
V	CONCLUSION	64
	REFERENCES	66

LIST OF TABLES

1	Statistics of Naive Day Prediction	11
2	Statistics of Naive Week Prediction	11
3	Statistics of Simple Exponential Smoothing Day Prediction	12
4	Statistics of Simple Exponential Smoothing Week Prediction	12
5	Statistics of Holt's Method Day Prediction	15
6	Statistics of Holt's Method Week Prediction	16
7	Statistics of Holt-Winter's Day Prediction	18
8	Statistics of Holt-Winter's Week Prediction	18
9	Real Prices of Portugal and Spain (EUR/MWh)	20
10	Shares of Best Time-series Prediction Methods	21
11	RMSE of Day Predictions For Last 14 Days of 2019	21
12	RMSE of Week Predictions For Last 14 Days of 2019	22
13	MSD of Day Predictions For Last 14 Days of 2019	22
14	MSD of Week Predictions For Last 14 Days of 2019	23
15	RMSE of Day Predictions For Last 14 Days of March 2020 During COVID-19	24
16	RMSE of Week Predictions For Last 14 Days of March 2020 During COVID-19	25
17	MSD of Day Predictions For Last 14 Days of March 2020 During COVID-19	25
18	MSD of Week Predictions For Last 14 Days of March 2020 During COVID-19	26
19	Share of Best Time-Series Prediction Methods During COVID-19 . .	27
20	Comparison of Best Forecasting Techniques During COVID-19	28
21	Best SARIMA Model For Each European Country	33
22	Error Performance of SARIMA Prediction	34
23	Statistics of SARIMA Prediction	34
24	Hyperparameters of Long-Short Term Memory	37

25	RMSE Values of LSTM Each European Country	38
26	Statistics of LSTM (Look-back=24)	39
27	Statistics of LSTM (Look-back=48)	39
28	Classification Features	44
29	Best Parameters of Classification Methods	47
30	Accuracy of Different Classification Methods	48
31	Error Performance of Rolling Forecasting Period	58
32	Statistics of Regular and Dynamic Forecast Method	59
33	Statistics of Regular and Dynamic Forecast FFNN	62

LIST OF FIGURES

1	Exponential Smoothing Family	9
2	Model Time Period	9
3	Naive Method Processing	11
4	Real and Naive Day/Week Forecasts (Last 7 Days of 2019)	13
5	Real and Simple Exponential Smoothing Day/Week Forecasts (Last 7 Days of 2019)	14
6	Real and Holt's Method Day/Week Forecasts (Last 7 Days of 2019) .	17
7	Real and Holt-Winter's Day/Week Forecasts (Last 7 Days of 2019) .	19
8	SARIMA Model Steps	29
9	SARIMA Model Time Period	31
10	Autocorrelation Figures	32
11	Partial Autocorrelation Figures	32
12	SARIMA Model Prices vs Actual Prices	35
13	LSTM Model Period	36
14	Actual and LSTM Predicted Prices (Look-back = 24)	40
15	Actual and LSTM Predicted Prices (Look-back = 48)	41
16	Regular Forecast Method Rolling Forecast Structure	50
17	Error Proportions of Non-Lagged Prices vs Lagged Prices	52
18	Error Rates of Fuel Price Predictors	54
19	Error Rates of Various Training Data Lengths	55
20	Dynamic Forecast Method Rolling Forecast Model Time Period . . .	56
21	Estimated Prices vs Real Price	57
22	A Schematic Diagram of A Three-Layered Feedforward Network Model	60
23	The Error Rate Histograms of Four Different Forecasting Methods . .	61
24	Regular Rolling Forecasting FFNN Prices vs. Actual Prices	62
25	Dynamic Rolling Forecasting FFNN Prices vs. Actual Prices	63

CHAPTER I

INTRODUCTION

Through several decades, the worldwide energy market sector was controlled by either a governmental authority or monopolistic actions. Since the 1990s, the competitive markets have started to replace the traditional authorities, shaping a new future for electricity markets. Nowadays, the worldwide energy market rules are operating electricity trades by spot and derivative contracts. Unlike other commodities, electricity is still non-storable due to the lack of maturity of large-scaled energy storage equipment in terms of cost and technology. However, a reliability criterion forces the electricity market to maintain the generation balance throughout a volatilized demand [1, 2]. Many factors affect the dynamic pattern of electrical demand. The main two factors are weather (temperature, wind speed, humidity, precipitation, etc.) and everyday industrial and residential activities (off-peak hours, on-peak hours, holidays, sudden events i.e., COVID-19, etc.). The vast variations in the electrical demand due to the above characteristics drive the electricity price dynamics and the energy market as a whole [3]. Electricity price forecasting has gained attention in the academic field as well as at the corporate level where the forecasting results became a major input to the decision-making system of energy companies [4, 5]. There are a lot of attempts in the literature to develop statistical electricity price models that can forecast from several hours till few months ahead with varying degrees of success. A “similar day” method has been applied to electricity price forecasting at the beginning of this decade [1, 6]. It is simply to mimic the characteristics of a historical day such as weekday, holiday type, electrical consumption, weather feature, etc. to the forecasted day. The historical values will be assumed constant for the forecasted values of the similar day.

Exponential smoothing is another method, used by [7, 8] to forecast electricity prices. It is also one of the simple pragmatic tools in this field. It is to use the weighted average of past observations to predict future prices. Holt-Winter is also an effective method that is used a lot in the literature for electricity price forecasting purposes. A comparison approach between Holt-Winters and other methods such as exponential smoothing, ARIMA, and neural networks, was studied in the literature as well [9, 10, 11]. They used the time-series historical data to forecast future electricity prices while a complex algorithm framework is also discussed in another study [12]. Firstly, it modeled regression including wind power and hourly demand as predictors to forecast electricity spot prices. Then, Holt-winter's and autoregressive were modeled to identify the autocorrelation and seasonal behavior of the residuals. Autoregressive model (AR) and its derivatives such as; ARIMA, SARIMA, ARIMAX, etc. are considered as the backbone of electricity price forecasting, especially for short-term horizons. It is important to hybridize ARIMA models with exogenous variables such as; temperature, load, etc. [3]. ARIMA models have been applied [13, 14] for short-term forecasting of the German Electricity price market while SARIMA was applied by [15]. It is noticed by [16] that, different AR models with different lags such as 24, 28, and 168 hours while modeling each hour separately has released better performance than a single SARIMA with all hours. Computational intelligent approaches (artificial neural networks, fuzzy logic, support vector machines, etc.) have been used widely in electricity price forecasting. They are considered flexible and handle complex and nonlinear problems [3]. A comparison between ARIMA and artificial neural network (ANN) to forecast for the Turkish day-ahead electricity prices has been conducted in [17]. The network performance has been studied based on different topologies, input sets, and training algorithms. An algorithm that includes different forecasting models such as neural networks, fuzzy regressions, and a standard regression was developed by [18]. The algorithm is used to switch between the models forecasting throughout

a long-timescale.

A comparison Machine learning algorithms have shown vast usage in power systems economics especially when combined with time-series method for the sake of electricity price and load forecasting. An algorithm selection approach was introduced [19] to effectively forecast residential electricity demand. The aim is to specify the best forecasting method for a specific home with special features. The different time series prediction methods; ARIMA, Holt-Winters, LSTM [20], and TESLA were used in the first level to identify which method fits into which house and it became a class for the load profile. In the second level, the classes found before and the houses' features where the input of different classifiers; regression, stochastic gradient descent (SGD), k-nearest neighbor (kNN), random forest (RF), and support vector machine (SVM) to identify the accurate prediction method for a specified house of features. Ref [21] has used the core vector machine (CVM) to classify electricity prices like high prices and low prices. They used a feature subset selector and applied the method to Austria market data. Simulations have shown reliable results as well as a quantitative level of confidence for the classification. A recursive learning algorithm was applied in [22] to classify electricity prices 24 hours ahead. The classes used the classification methods included where Decision Trees (DT), Multilayer Perceptron (MLP), Bayes, and K-nearest neighbor (KNN). The same approach used in a demand-side management case study as well and a performance-based comparative investigation was conducted. Instead of the single support vector machine (SVM), a multiple SVM is applied by [23] the midterm electricity market-clearing price. The method was suggested as a two-stage algorithm where the corresponding price zones were separated in the first stage using a single SVM. In the second stage, four parallel SVMs designed to forecast the electricity price in four different price zones. The proposed algorithm was tested on PJM interconnection data and more accurate results have revealed compared with the traditional method. An integrated machine learning algorithm based on Bayesian

Clustering by Dynamics (BCD) and Support Vector Machine (SVM) is proposed [24] to forecast electricity prices in short time horizons. The combined algorithm starts by clustering the input data set, using the BCD classifier, into many subsets in an unsupervised way. Then, for the electricity prices of the next day, 24 grouped SVMs where apply each training subset in a supervised manner. The proposed forecasting system has been tested on New England electricity prices historical data and showed a high degree of accuracy compared to methods. Three different decision tree algorithms; (Classification and Regression Tree) CART, (Reduced Error Pruning Tree) REPTree and Decision Stump have been designed by [25] to solve the electrical load forecasting problem. The results have shown that REPTree Decision Tree has better performance with lower error metrics and higher correctly classification instances when forecasting for electric power load, compared to the other Decision trees. With competitive advances in electricity sectors in many countries, market structures have dramatically adjusted daily. Electricity markets are mostly operated on an hourly basis in real-time balancing and day-ahead markets. These markets also operate based on the marginal pricing principle applied in supply-demand balancing. Prices in the day-ahead market are in great significance for consideration in investment decisions. From this point of view, it can particularly be said that electricity prices are predictable in the best possible way. There is a fluctuation of Euro/Turkish Lira parity in electricity markets in Turkey nowadays. It is a well-known fact that many fundamental factors are correlated with electricity prices. Electricity is co-integrated with commodities such as natural gas (NG), coal, and crude oil both in developed and developing countries because they are import commodities in Turkey. After the NG market has established in Turkey, the NG prices are more predictable in the industry as a commodity. It is said that electricity and NG prices have a robust correlation; however, their price dynamics may not be similar. Thus, NG prices cannot solely be

sufficient to predict electricity prices. Moreover, coal is still popular as fuel and indispensable for electricity generation. It plays a crucial role in short and medium-term development and energy economy. Aside from coal price, the crude oil price is determined by supply and demand globally. High electricity demand can automatically raise the demand for crude oil, and higher costs to generate electricity. It is said that energy commodities co-integrating oil and natural gas are substitutes in consumption [26]. Consequently, it is examined that natural gas, coal, and crude oil prices affect the electricity price. Similar expectations can be assumed in lots of countries. For example, natural gas, coal, and crude oil prices are the major factors of Spanish electricity prices [27]. Apart from the effects of fuel prices, it is important for several types that utilize to predict electricity prices. They can be categorized as short, medium or long-term forecasting in the literature [28]. Short-term price forecasting plays a prominent role in the day-ahead market. Also, the medium-term is different from other terms, because it does not generally include a lot of independent variables. It is mostly affected by days such as day-of-week, holiday and hours. Also, long-term price forecasting generally includes the extended range from months to years. In the literature, there are numerous methods to estimate the electricity price. It is used some of the methods, which include multiple linear regression (MLR), autoregression integrated moving average (ARIMA), fuzzy logic and generalized autoregressive conditional heteroscedasticity (GARCH), in the energy economy and power systems. These methods can readily allow an advantage to predict short-term estimation [29]. Another method is on wavelet transform interconnected with ARIMA and GARCH methods [30]. Outside of these unique methods, due to various factors affecting electricity prices, especially fluctuations in the economy, it is very hard to make sense of a precise forecasting score, thus it is used applications of deep learning. It is said that the pioneer study in energy economics is initially written by Wang and Ramsey in 1998 on price forecasting using the artificial neural network which is an example

of deep learning methods [31]. This pioneering study boosts the evolution of deep learning in the energy economy. Thus, deep learning applications are quite becoming widespread in big data sets to perform the estimation. To illustrate this point, it is also seen that electricity price prediction can be done by an artificial neural network (ANN) [32]. Another method is that Jiang et al. [33] implemented price forecasting for the electricity market by using long short-term memory (LSTM). Additionally, it is a different method for multilayer perceptron in spot electricity prices [34]. At the same time, another method is Support Vector Machines (SVM) in price estimation. It is a new technique for classification and regression [35].

This thesis focuses on three different main models which are statistical models, machine learning methods, and deep learning models for short-term electricity price forecasting. In chapter 2, a new algorithm that adjusts on existing time-series methods is used to solve the electricity price forecasting problem. Four different time-series forecasting models such as Holts, Holts-Winter [36], Simple Exponential Smoothing [37], and Naïve [38], are applied to data for 12 different European countries. The MSD and RMSE are used to evaluate the performance of each forecasting model to identify which model dominates for each country. To do so, section 2.7 provides error performance to change the forecasting period before and during COVID-19. It covers the forecasting period of the last 14 days of 2019 and the period between March 18th, 2020 and March 31st, 2020 in which most of the EU countries start the total lockdown that creates a big impact on the total electricity consumption. Section 2.8 covers a comparison of 12 European countries using the SARIMA forecasting method. The LSTM is used in section 2.9. Chapter 3 includes the definitions of classification model types, moreover, this chapter also covers the details of the 33 features selected for the European countries. The classification models used in this thesis are K-Nearest Neighbors (KNN) [39], Decision Tree (DT) [40], Support Vector Machine (SVM) [41], Gradient Boosting Algorithm [42], AdaBoost Classifier [43], Stochastic Gradient

Descent [44], Random Forest [45], Bagging Classifier [46], and Ridge Classifier [47]. Chapter 4 enables a brief overview of Multiple Linear Regression and its parameters in data preparation. After that, it is described that the impact of the training data length is included one year, 9 months, 6 months and 3, months, respectively in section 4.4. Then, it is examined how NG, oil, and coal prices affect the electricity price one by one and how lagged prices are important in this estimation. At the same time, the error performances of Regular and Dynamic Forecast Methods are measured. Lastly, the comparisons observed on Regular and Dynamic FFNN are discussed in Section 4.6.

CHAPTER II

TIME-SERIES PREDICTION METHODS

Time-series forecasting is the way of the model to estimate future values based on previously observed values. Six different time-series prediction methods are performed in this chapter. The exponential smoothing family which is named Exponential Smoothing, Holt's Method, and Holt-Winter's Method have been implemented. Also, we can see the advances of these methods among themselves in Fig. 1. Then, Naive Method, SARIMA, and Long-Short Term Memory have implemented.

2.1 Data Pre-Processing For Naive and Smoothing Methods

Turkey, Estonia, Finland, France, Hungary, Netherlands, Portugal, Spain, Greece, Latvia, Slovenia, and Belgium are chosen. The selected four different time-series prediction methods are implemented in MATLAB and Python [48]. Also, we can review the entire code analysis in the GitHub Repositories [49]. All data is extracted by means of Application Programming Interface (API) from ENTSO-E [50], but Turkish day-ahead electricity price data set is taken using API from the Energy Exchange Istanbul Transparency Platform (EXIST) [51]. The hourly electricity price data includes the period between Jan. 1st, 2015 and Dec. 31th, 2019 for all countries. We can demonstrate splitting train and test set as the model time period for 12 European countries in Fig. 2.

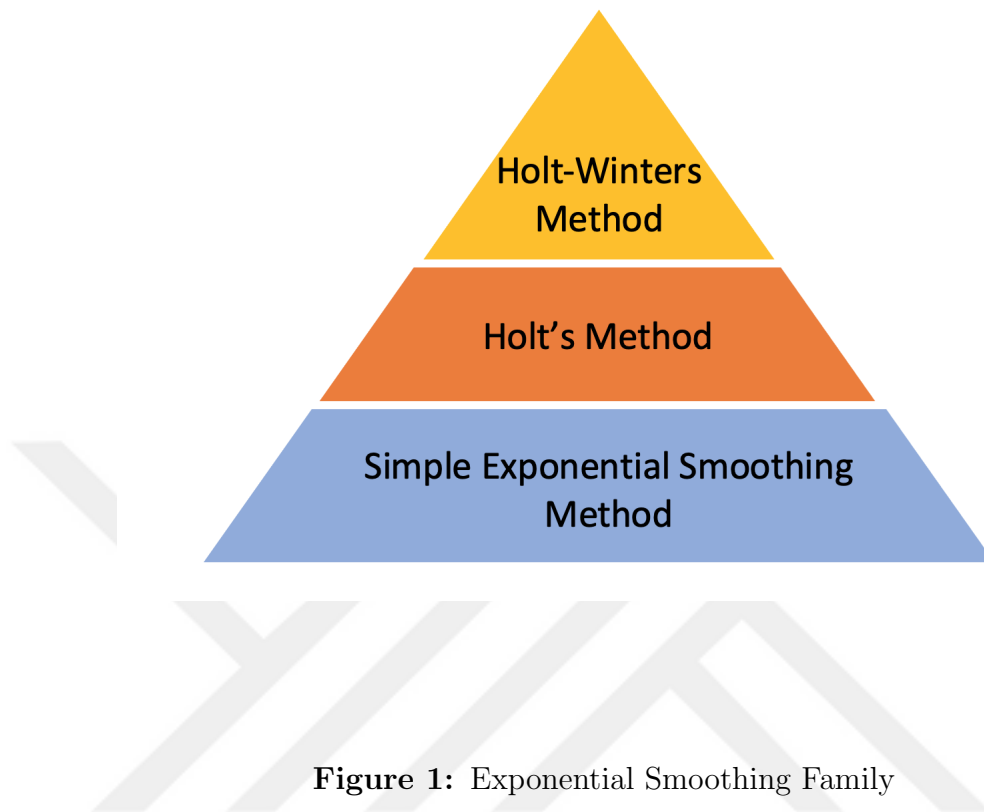


Figure 2: Model Time Period

The selected European countries in this thesis are Estonia, Finland, France, Hungary, Netherlands, Portugal, Spain, Greece, Latvia, Slovenia, Belgium, and Turkey. Those countries are selected due to the availability of the data sets. The performance of the prediction is measured by Mean Absolute Percentage Error, Root Mean Squared Error (RMSE), Mean Signed Deviation (MSD), and Mean Absolute Error (MAE) as given in (1), (2), (3), and (4), respectively.

$$\text{Mean absolute percentage error (MAPE)} = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{(\hat{Y}_t - Y_t)}{y_t} \right| \quad (1)$$

$$\text{Mean Signed Deviation (MSD)} = \frac{1}{n} \sum_{t=1}^n (\hat{Y}_t - Y_t) \quad (2)$$

$$\text{Root mean squared error (RMSE)} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{Y}_t - Y_t)^2} \quad (3)$$

$$\text{Mean absolute error (MAE)} = \sqrt{\frac{1}{n} \sum_{t=1}^n |\hat{Y}_t - Y_t|} \quad (4)$$

2.2 Naive Method

It implies recent data provides the best predictors of the upcoming results. Naïve forecast method is simply valued from the last 24 and 168 hour periods. For instance, the day-ahead electricity price of this Thursday is equal to the electricity price of the previous Wednesday or the last week's Thursday. Ref [52] used a similar Naïve method to forecast electricity prices in the day-ahead market. Naive method processing is shown in Fig. 3. In this section, the Naïve forecast method in day and week algorithms used for predicting electricity price estimation as given in (5):

$$\hat{Y}_t = Y_{t-\tau} \quad (5)$$

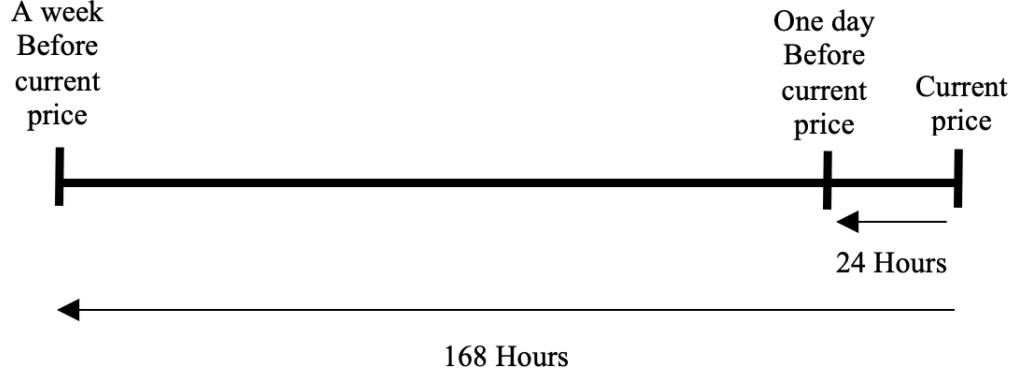


Figure 3: Naive Method Processing

where t is the time of the day, $\hat{Y}_t$ is the vector of forecast values in the future, Y_t is the vector of the historical data set, and τ may take the value of 24 for the previous day method and it takes the value of 168 for the previous week method. We can see Statistics of Naive Day and Week Predictions in Tables 1 and 2, and also the Actual, Naive Day, and Week Method Prices on the last 7 days of 2019 in Fig. 4.

Table 1: Statistics of Naive Day Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	54.56	47.82	47.82	47.82	54.56	51.80	47.83	54.56	58.97	82.88	53.27	50.72
Min.	1.40	0.03	0.03	-0.47	1.40	0.00	-4.14	1.40	9.56	18.00	2.48	6.81
Mean	34.30	27.56	27.56	27.26	34.05	29.95	30.35	34.30	33.97	57.13	30.94	42.05
Std.	7.54	11.63	11.63	11.39	7.46	11.13	10.31	7.54	7.73	15.34	9.96	10.16

Table 2: Statistics of Naive Week Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	58.54	45.35	49.10	54.19	58.54	59.07	55.60	58.54	54.19	78.20	54.19	51.55
Min.	27.40	0.01	0.03	-0.47	27.40	0.00	-5.98	27.40	2.59	0.00	2.37	6.85
Mean	39.97	17.88	18.23	25.28	39.97	31.21	30.52	39.97	34.33	55.61	27.67	43.78
Std.	7.93	15.01	15.21	13.62	7.93	13.27	13.38	7.93	8.17	15.95	11.84	9.31

2.3 Simple Exponential Smoothing Method

The Simple Exponential Smoothing Method (ESM) uses exponentially changing weights for previous observations to estimate the future. [53, 54, 55] The ESM for day and week algorithms is given in (6).

$$\hat{Y}_t = \alpha Y_{t-\tau} + (1 - \alpha) \hat{Y}_{t-\tau} \quad (6)$$

where t is the time in days, $\hat{Y}$ is the vector of predicted values in the future, Y is the vector of the historical data set, α is the smoothing constant that should satisfy $0 < \alpha < 1$. We can see Statistics of Simple Exponential Smoothing Day and Week Predictions in Table 3 and 4, and also the Actual, Simple Exponential Smoothing Day, and Week Method Prices on the last 7 days of 2019 in Fig. 5.

Table 3: Statistics of Simple Exponential Smoothing Day Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	54.54	47.64	47.64	47.64	54.54	51.48	47.66	54.54	56.98	82.82	52.79	50.63
Min.	1.93	0.15	0.15	-0.30	1.93	0.04	-3.93	1.93	9.85	18.44	2.80	6.97
Mean	34.30	27.55	27.55	27.25	34.06	29.94	30.34	34.30	33.97	57.13	30.92	42.06
Std.	7.48	11.56	11.56	11.31	7.41	11.08	10.24	7.48	7.61	15.08	9.85	9.99

Table 4: Statistics of Simple Exponential Smoothing Week Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	58.35	45.21	48.75	53.40	58.35	58.91	55.04	58.35	53.77	78.19	53.55	51.53
Min.	27.95	0.01	0.10	-0.30	27.95	0.04	-5.53	27.95	4.88	2.74	2.58	6.97
Mean	39.97	17.89	18.24	25.29	39.97	31.22	30.53	39.97	34.33	55.61	27.67	43.79
Std.	7.81	14.96	15.15	13.47	7.81	13.19	13.28	7.81	7.98	15.61	11.72	9.20

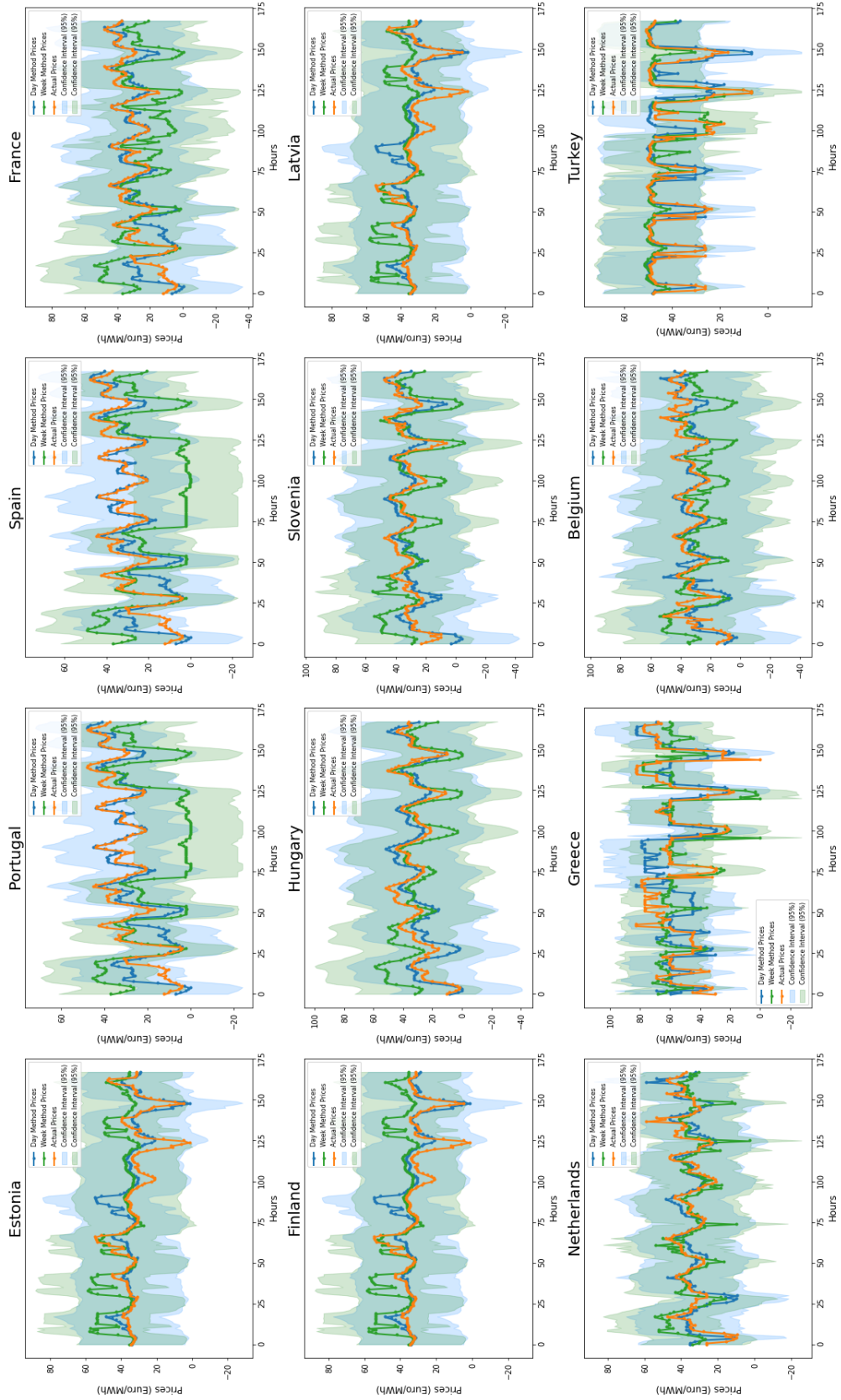


Figure 4: Real and Naive Day/Week Forecasts (Last 7 Days of 2019)

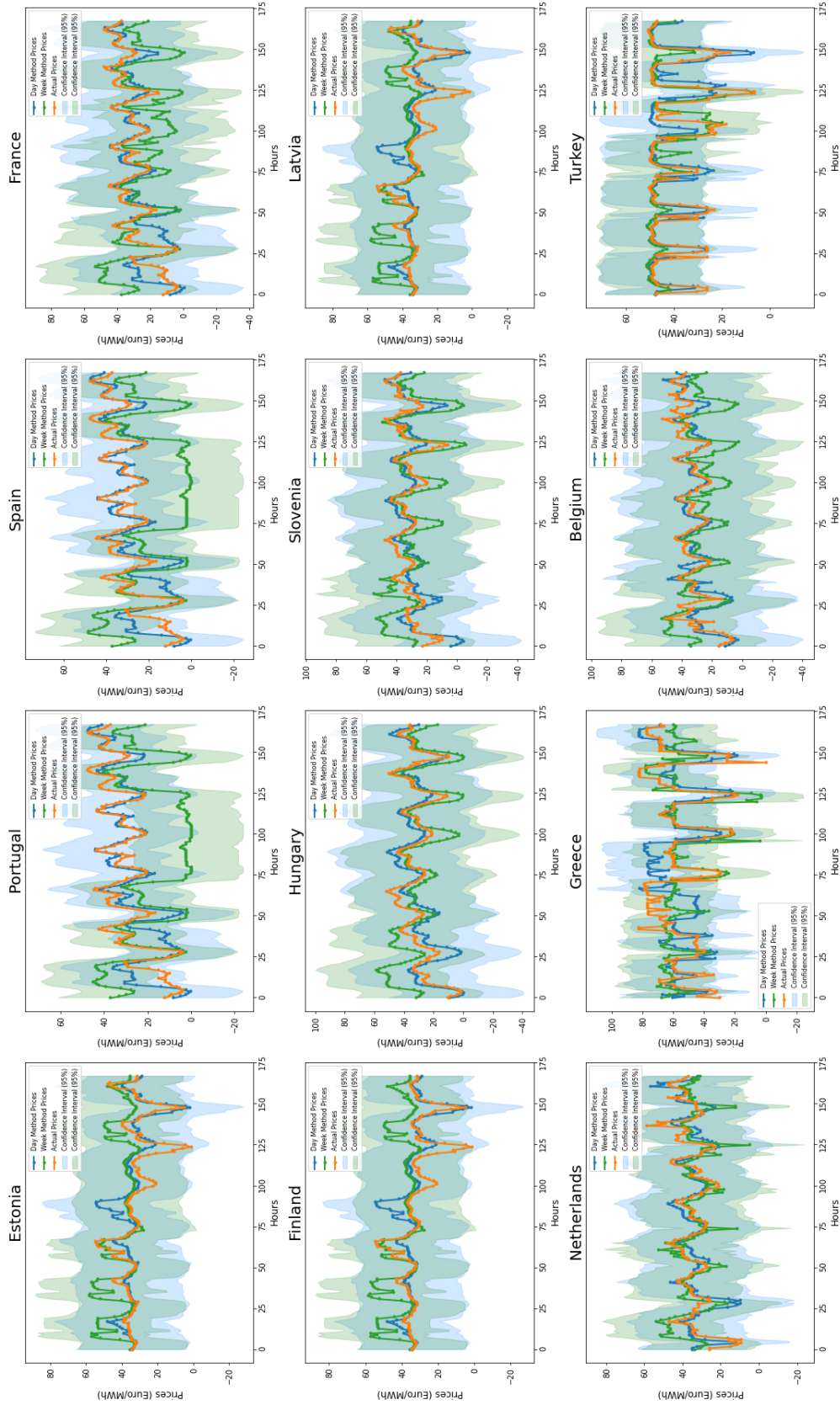


Figure 5: Real and Simple Exponential Smoothing Day/Week Forecasts (Last 7 Days of 2019)

2.4 Holt's Method

Holt's method is designed to deal with time-series data set with a trend. It has two parameters, α , and β . They are previous data observations and trend smoothing constants respectively [56, 57]. The double exponential smoothing method (Holt's Method) in day and week algorithms utilized for evaluating day-ahead electricity price forecasting is given in (7).

$$\begin{aligned} L_t &= \alpha Y_{t-\tau} + (1 - \alpha) \hat{Y}_{t-\tau} \\ T_t &= \beta(L_{t-\tau} + L_{t-\tau-1}) + (1 - \beta)T_{t-\tau-1} \\ \hat{Y}_t &= L_t + T_t \end{aligned} \tag{7}$$

where t is the time in days, $\hat{Y}$ is the vector of forecasted values in the future, Y is the vector of the historical data set, L is new smooth value, α is smoothing constant for the level, T is trend estimate, β is smoothing constant for trend estimate, and τ may take the value of 24 for the previous day method and it takes the value of 168 for the previous week method. The weights α and β should satisfy $0 < \alpha, \beta < 1$. We can see Statistics of Holt's Method Day and Week Predictions in Table 5 and 6, and also the Actual, Holt's Method Day, and Week Method Prices on the last 7 days of 2019 in Fig. 6.

Table 5: Statistics of Holt's Method Day Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	54.71	47.86	47.86	47.85	54.71	51.70	47.84	54.71	57.21	83.11	53.01	50.71
Min.	1.57	-0.07	-0.07	-0.60	1.57	-0.33	-4.33	1.57	9.59	18.03	2.51	6.59
Mean	34.26	27.60	27.60	27.26	34.02	29.93	30.33	34.26	33.97	57.14	30.94	42.05
Std.	7.56	11.67	11.67	11.43	7.48	11.20	10.36	7.56	7.70	15.25	9.95	10.09

Table 6: Statistics of Holt's Method Week Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	58.53	45.30	48.88	53.54	58.53	59.05	55.17	58.53	53.94	78.44	53.72	51.59
Min.	27.79	-0.29	-0.16	-0.60	27.79	-0.33	-5.97	27.79	4.57	2.16	2.29	6.57
Mean	39.96	17.78	18.13	25.20	39.96	31.12	30.45	39.96	34.31	55.57	27.61	43.77
Std.	7.90	15.11	15.29	13.60	7.90	13.31	13.41	7.90	8.07	15.78	11.84	9.30

2.5 *Holt-Winter's Method*

It is generated by handling time-series data with a trend and seasonal variation. It contains three main components that mean α , trend β , and seasonality γ [58]. Ref [59] implemented Holt's Winter method for estimating day-ahead energy prices. In our model, we use the multiplicative version of Holt-Winters. Holt-Winter's method can be valued from the previous 24 and 168 hour periods instead of using the previous one hour. In this part, the Holt-Winter's method in day and week algorithms used for predicting electricity price as given in (8):

$$\begin{aligned}
L_t &= \alpha \frac{Y_{t-\tau}}{S_{t-s-\tau}} + (1 - \alpha)(L_{t-\tau-1} - T_{t-\tau-1}) \\
T_t &= \beta(L_{t-\tau} + L_{t-\tau-1}) + (1 - \beta)T_{t-\tau-1} \\
S_t &= \gamma \frac{Y_{t-\tau}}{L_{t-\tau}} + (1 - \gamma)S_{t-s-\tau} \\
\hat{Y}_t &= (L_t + T_t)S_{t-s}
\end{aligned} \tag{8}$$

where Y is forecast for p periods into the future, p is periods to be forecast into the future, Y_t is a new observation of series in period t , L_t is new smoothed value, α is smoothing constant for the level, T_t is trend estimate, β is smoothing constant for trend estimate, s is the length of seasonality, S_t is a seasonal estimate and γ is smoothing constant for seasonality estimate. We can see Statistics of Holt-Winter's Day and Week Predictions in Table 7 and 8, and also the Actual, Holt-Winter's Day,

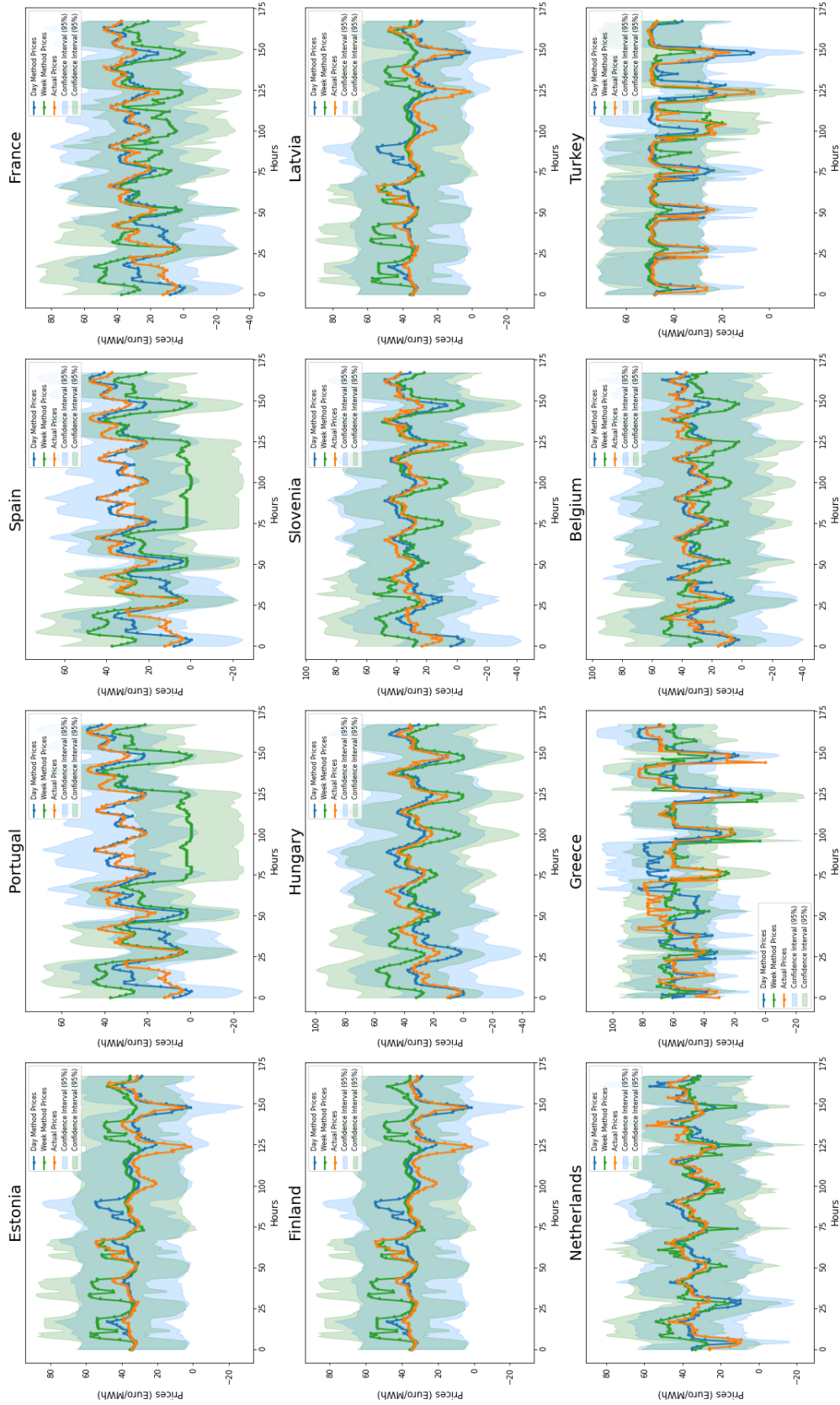


Figure 6: Real and Holt's Method Day/Week Forecasts (Last 7 Days of 2019)

and Week Method Prices on the last 7 days of 2019 in Fig. 7.

Table 7: Statistics of Holt-Winter’s Day Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	54.72	48.04	48.04	48.03	54.72	52.00	47.98	54.72	59.02	83.14	53.48	50.79
Min.	1.19	-0.02	-0.02	-0.83	1.19	-0.01	-4.50	1.19	9.36	17.57	2.27	6.37
Mean	34.26	27.62	27.62	27.28	34.02	29.94	30.35	34.26	33.97	57.14	30.96	42.05
Std.	7.61	11.74	11.74	11.50	7.53	11.23	10.39	7.61	7.81	15.49	10.06	10.26

Table 8: Statistics of Holt-Winter’s Week Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	58.53	45.30	48.88	53.54	58.53	59.05	55.17	58.53	53.94	78.44	53.72	51.59
Min.	27.79	-0.10	-0.25	-0.70	27.79	0.00	-6.01	27.79	4.70	2.34	2.31	6.58
Mean	39.96	17.91	18.13	25.21	39.96	31.13	30.45	39.96	34.31	55.58	27.61	43.77
Std.	7.90	15.00	15.29	13.60	7.90	13.30	13.40	7.90	8.06	15.75	11.84	9.30

2.6 Results of Smoothing Methods and Naive Method

Table 10 represents the share of how many times a time series method is the best based on RMSE and MSD. The RMSE of day and week predictions for the last 14 days of 2019 are given in Table 11 and 12 for the 12 selected countries, respectively. Also, the MSD of day and week predictions for the last 14 days of 2019 are given in Table 13 and 14 for the 12 selected countries, respectively. If we only compare the similar day methods within, then the best time-series method for Turkey, Finland, France, Netherlands, Spain, Greece, Latvia, Slovenia, and Belgium is the Holt-Winters’ method on the RMSE performance. For Estonia, Hungary, and Portugal; it is the Naïve method on the RMSE performance. If we compare the similar week methods, the best time-series method for Turkey, France, Hungary, Spain, Slovenia, and Greece is the Holt-Winters’ method on the RMSE performance. It is the Simple Exponential Smoothing Week Method (ESM) for Estonia, Finland, Netherlands,

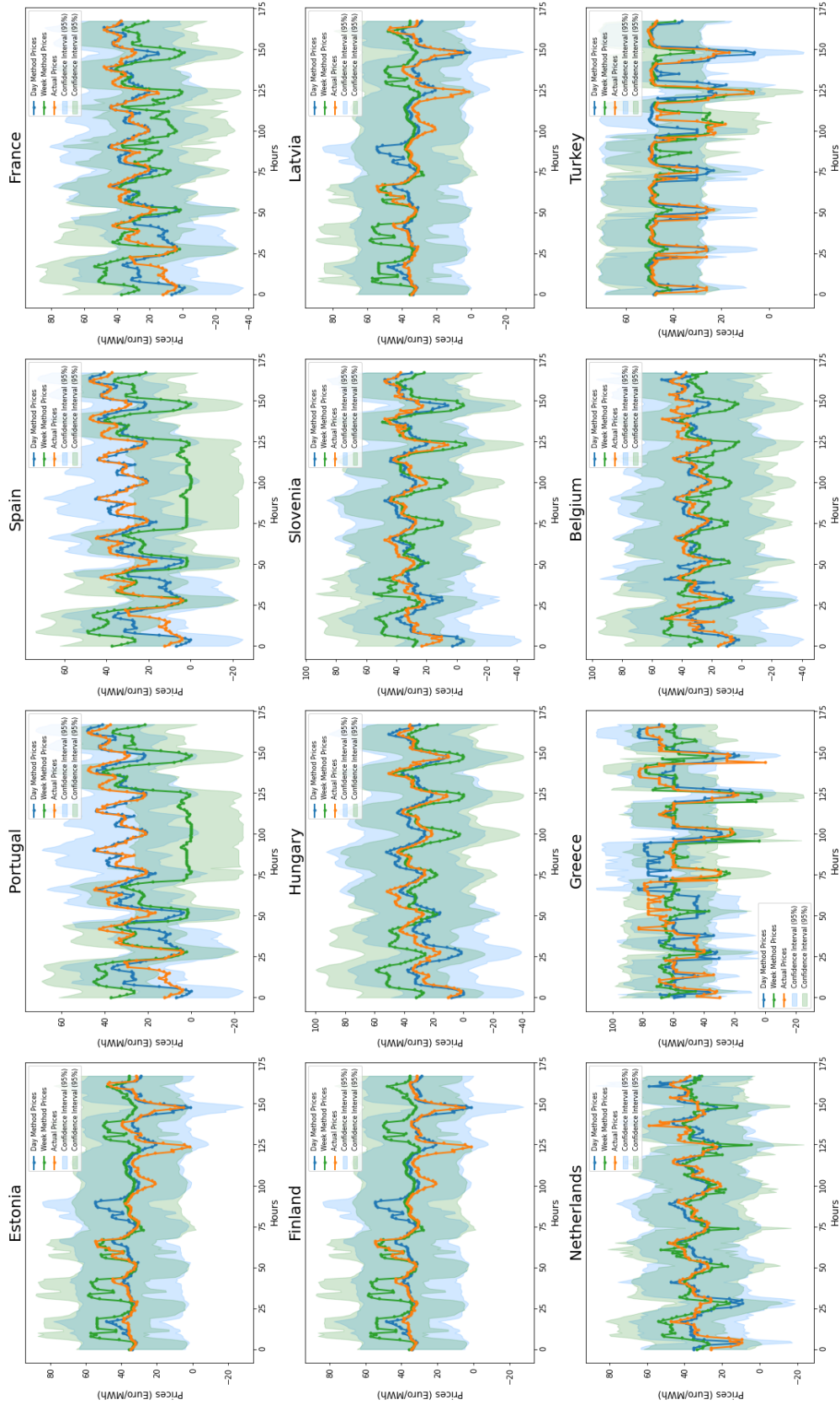


Figure 7: Real and Holt-Winter's Day/Week Forecasts (Last 7 Days of 2019)

Latvia, and Belgium on the RMSE performance. Yet the lowest RMSE within the similar week methods for Portugal is the Naïve method. In all studies, Holt’s method performed the worst for all European countries. Also, there are unexpected estimations in week methods price prediction of Portugal and Spain graphs because of very low real electricity prices. In other words, unexpected predictions are generated as electricity prices were predicted 1 week ago and were close to zero. We can see the real prices of European countries one week ago in Table 9.

Table 9: Real Prices of Portugal and Spain (EUR/MWh)

Date-Time	Portugal	Spain
21.12.2019 22:00	2.3	2.3
21.12.2019 23:00	2.3	2.3
22.12.2019 00:00	0.51	0.51
22.12.2019 01:00	0.5	0.51
22.12.2019 02:00	0.01	0.5
22.12.2019 03:00	0.01	0.1
22.12.2019 04:00	0.01	0.1
22.12.2019 05:00	0.01	0.1
22.12.2019 06:00	0.01	0.2
22.12.2019 07:00	0.01	0.5
22.12.2019 08:00	0.01	0.51
22.12.2019 09:00	0.01	1.02
22.12.2019 10:00	0.2	2.3
22.12.2019 11:00	1	2.3
22.12.2019 12:00	2.05	2.05
22.12.2019 13:00	2.3	2.3
22.12.2019 14:00	2.3	2.3
22.12.2019 15:00	1.1	1.1
22.12.2019 16:00	1.02	1.02
22.12.2019 17:00	2.3	2.3
22.12.2019 18:00	2.3	2.3
22.12.2019 19:00	3.5	3.5
22.12.2019 20:00	5.5	5.5
22.12.2019 21:00	6.6	6.6

Table 10: Shares of Best Time-series Prediction Methods

Similarity	Methods	Share of methods based on RMSE (%)	Share of methods based on MSD (%)
Day	Naïve Method	16.67%	0%
Day	Simple ES.	0%	8.33%
Day	Holt's Method	0%	33.33%
Day	Holt-Winters	33.33%	8.33%
Week	Naïve Method	4.16%	12.5%
Week	Simple ES.	20.83%	0%
Week	Holt's Method	0%	29.16%
Week	Holt-Winters	25.00%	8.33%

Table 11: RMSE of Day Predictions For Last 14 Days of 2019

RMSE	Day ESM	Day HM	Day HWM	Day NM
Turkey	9.54	9.63	9.13	9.14
Estonia	7.01	7.10	6.74	6.71
Finland	6.41	6.46	6.03	6.04
France	10.28	10.35	9.77	9.78
Hungary	8.06	8.11	6.88	6.87
Netherlands	6.68	6.73	6.48	6.54
Portugal	11.97	12.05	11.57	11.51
Spain	11.94	12.02	11.28	11.48
Greece	15.34	15.48	15.26	15.40
Latvia	6.53	6.59	6.11	6.12
Slovenia	7.75	7.80	6.86	6.92
Belgium	8.16	8.20	7.75	7.81

Table 12: RMSE of Week Predictions For Last 14 Days of 2019

<i>RMSE</i>	Week ESM	Week HM	Week HWM	Week NM
Turkey	8.04	8.06	7.50	7.53
Estonia	9.54	9.64	9.68	9.69
Finland	9.29	9.35	9.46	9.48
France	16.05	16.14	15.94	15.97
Hungary	14.87	14.91	14.68	14.72
Netherlands	7.82	7.89	8.13	8.24
Portugal	20.86	20.99	20.75	20.71
Spain	20.81	20.94	20.58	20.69
Greece	13.01	13.07	12.73	12.82
Latvia	8.74	8.80	8.86	8.89
Slovenia	14.14	14.21	14.10	14.15
Belgium	13.19	13.29	13.31	13.36

Table 13: MSD of Day Predictions For Last 14 Days of 2019

<i>MSD</i>	Day ESM	Day HM	Day HWM	Day NM
Turkey	0.51	0.47	0.49	0.48
Estonia	1.19	1.16	1.18	1.19
Finland	1.50	1.46	1.39	1.39
France	0.33	0.24	0.30	0.29
Hungary	1.11	0.98	1.03	1.04
Netherlands	-0.03	-0.06	-0.04	-0.04
Portugal	0.26	0.18	0.28	0.26
Spain	0.28	0.21	0.39	0.28
Greece	0.10	0.07	0.25	0.24
Latvia	1.36	1.32	1.26	1.26
Slovenia	0.59	0.48	0.52	0.52
Belgium	-0.21	-0.26	-0.22	-0.24

Table 14: MSD of Week Predictions For Last 14 Days of 2019

MSD	Week ESM	Week HM	Week HWM	Week NM
Turkey	2.42	2.37	2.39	2.40
Estonia	4.27	4.32	4.27	4.24
Finland	3.68	3.69	3.63	3.62
France	4.39	4.22	4.24	4.26
Hungary	7.44	7.26	7.23	7.25
Netherlands	1.90	1.87	1.87	1.87
Portugal	2.73	2.49	2.55	2.59
Spain	2.82	2.58	2.73	2.68
Greece	1.69	1.62	1.64	1.65
Latvia	3.87	3.87	3.82	3.81
Slovenia	6.09	5.95	5.94	5.95
Belgium	2.05	1.94	1.98	1.99

2.7 *Impact of COVID-19 on Electricity Prices in European Countries*

Electricity prices of European countries dramatically change because of the COVID-19 outbreak. Thus, to understand whether there is any change in the best prediction method for electricity prices in the selected countries right during COVID-19, firstly, we estimate electricity prices between Mar. 18th, 2020, and Mar. 31th, 2020, which is the beginning of several full lockdowns in the selected countries. For example, Italy ordered a lockdown on Mar. 11th, Spain on 14th, France on 17th, Belgium on 18th, Germany on 22nd, and UK on 24th. Four time-series prediction methods are used to estimate the last 14 days of March 2020. Based on the RMSE and MSD values of predictions as given in Table 15, 16, 17, and 18. During the similar day methods for the COVID-19 part, the best time-series method for Finland, Hungary, Netherlands, Portugal, Belgium, and Spain becomes a Naïve Method. The best method for Estonia, France, Latvia, and Slovenia becomes Holt-Winter's method. The best method

for Turkey and Greece becomes the Simple Exponential Smoothing Method. In addition, the best time-series method for Turkey, Estonia, France, Netherlands, Latvia, and Slovenia is the Holt-Winters' method on the RMSE performance, is Simple Exponential Smoothing Method for Greece and Belgium on the RMSE performance, and also is the Naïve Method for Finland, Hungary, Portugal, and Spain in the similar week methods during coronavirus outbreak.

Table 15: RMSE of Day Predictions For Last 14 Days of March 2020 During COVID-

19

<i>RMSE</i>	Day ESM	Day HM	Day HWM	Day NM
Turkey	12.61	12.73	12.82	12.87
Estonia	12.60	12.74	12.36	12.41
Finland	11.35	11.46	11.18	11.16
France	8.37	8.45	7.90	8.02
Hungary	8.12	8.19	7.47	7.43
Netherlands	9.81	9.91	9.82	9.33
Portugal	4.94	4.98	4.82	4.81
Spain	5.17	5.21	5.05	5.04
Greece	16.86	16.98	17.02	17.18
Latvia	12.60	12.74	12.36	12.41
Slovenia	8.40	8.47	8.15	8.19
Belgium	12.10	12.23	12.28	11.98

Table 16: RMSE of Week Predictions For Last 14 Days of March 2020 During COVID-19

<i>RMSE</i>	Week ESM	Week HM	Week HWM	Week NM
Turkey	10.55	10.69	10.51	10.58
Estonia	9.01	9.11	8.32	8.38
Finland	11.73	11.86	11.56	11.54
France	9.04	9.09	9.01	9.03
Hungary	9.73	9.82	9.07	9.02
Netherlands	9.65	9.68	9.51	9.53
Portugal	6.82	6.88	6.79	6.78
Spain	6.95	7.01	6.92	6.91
Greece	17.47	17.58	17.59	17.77
Latvia	9.01	9.11	8.32	8.38
Slovenia	10.69	10.81	10.56	10.59
Belgium	11.01	11.04	11.12	11.05

Table 17: MSD of Day Predictions For Last 14 Days of March 2020 During COVID-19

<i>MSD</i>	Day ESM	Day HM	Day HWM	Day NM
Turkey	0.19	0.20	0.27	0.24
Estonia	0.06	0.04	0.11	0.06
Finland	1.96	-0.67	-0.58	-0.62
France	0.52	0.44	0.48	0.45
Hungary	0.06	0.02	0.05	0.03
Netherlands	0.27	0.17	0.10	0.20
Portugal	0.60	0.54	0.53	0.54
Spain	0.60	0.54	0.53	0.54
Greece	0.46	0.45	0.53	0.47
Latvia	0.06	0.04	0.11	0.06
Slovenia	0.24	0.21	0.21	0.19
Belgium	0.40	0.30	0.27	0.33

Table 18: MSD of Week Predictions For Last 14 Days of March 2020 During COVID-19

<i>MSD</i>	Week ESM	Week HM	Week HWM	Week NM
Turkey	0.82	0.82	0.84	0.80
Estonia	0.75	0.69	0.75	0.72
Finland	1.96	1.92	1.92	1.88
France	3.77	3.67	3.75	3.72
Hungary	1.84	1.75	1.78	1.78
Netherlands	3.86	3.76	3.84	3.84
Portugal	4.46	4.46	4.40	4.40
Spain	4.52	4.53	4.47	4.47
Greece	1.62	1.62	1.64	1.59
Latvia	0.75	0.69	0.75	0.72
Slovenia	1.35	1.24	1.28	1.26
Belgium	4.24	4.13	4.21	4.20

The percentages of time-series methods are compared based on RMSE before and during the COVID-19 pandemic in Table 19. It can be seen an increase in the Naïve Method in similar day and week methods. The Simple Exponential Smoothing increases the share of similar day method, while it decreases the share of similar week method. Holt-Winters' Method decreases percentage during COVID-19 on the share of similar day method while remaining stable in the similar week method. We can imply that Holt-Winter's Method and Naïve Method are a good choice compared to other methods concerning time-series prediction for electricity prices.

Table 19: Share of Best Time-Series Prediction Methods During COVID-19

Similarity	Methods	Share of methods based on RMSE (%)	Share of methods based on MSD (%)
Day	Naïve Method	25%	4.16%
Day	Simple ES.	8.33%	4.16%
Day	Holt's Method	0%	20.83%
Day	Holt-Winters	16.67%	20.83%
Week	Naïve Method	16.67%	20.83%
Week	Simple ES.	8.33%	0%
Week	Holt's Method	0%	29.16%
Week	Holt-Winters	25.00%	0%

2.7.1 Impacts of COVID-19 on Prediction Methods

The world is currently fighting against COVID-19 that is seriously threatening the global economy. The energy market sector has been affected negatively especially when the oil prices dropped to the minimum in decades. Moreover, industrial, commercial, and transportation activities have been stopped in large percentages. Thus, electricity prices curve with its previous characteristics; trends, seasonality, and cycles, has changed and different characteristics have formed on a daily and weekly basis. Each country has its unique electricity price curves including their unique characteristics and the best forecasting technique that fits these characteristics.

Table 20: Comparison of Best Forecasting Techniques During COVID-19

	Before COVID 19 (18 - 31 December 2019)		During COVID 19 (18 - 31 March 2020)	
	Similar Day Method	Similar Week Method	Similar Day Method	Similar Week Method
Turkey	Holt-Winters	Holt-Winters	Simple Exp. Smoothing	Holt-Winters
Estonia	Naive	Simple Exp. Smoothing	Holt-Winters	Holt-Winters
Finland	Holt-Winters	Simple Exp. Smoothing	Naive	Naive
France	Holt-Winters	Holt-Winters	Holt-Winters	Holt-Winters
Hungary	Naive	Holt-Winters	Naive	Naive
Netherlands	Holt-Winters	Simple Exp. Smoothing	Naive	Holt-Winters
Portugal	Naive	Naive	Naive	Naive
Spain	Holt-Winters	Holt-Winters	Naive	Naive
Greece	Holt-Winters	Holt-Winters	Simple Exp. Smoothing	Simple Exp. Smoothing
Latvia	Holt-Winters	Simple Exp. Smoothing	Holt-Winters	Holt-Winters
Slovenia	Holt-Winters	Holt-Winters	Holt-Winters	Holt-Winters
Belgium	Holt-Winters	Simple Exp. Smoothing	Naive	Simple Exp. Smoothing

Before COVID-19 pandemic, electricity prices curves showed trends and seasons where the power demands and so prices were vastly changing within a day and within a week as well. Table 20 represents that it is examined which country's methods change or not. During COVID-19 pandemic many countries; Netherlands, Greece, Spain, Estonia, and Finland show changes in both similar day and week methods in best forecasting models for electricity prices before and during COVID-19. Also, Greece has changed from Holt-Winters Method to the Simple Exponential Smoothing method. The reason behind that is, daily electrical consumption became similar where the residential power demand occupied most of the electricity consumption throughout the week. In other words, the power demand and so price patterns showed smoother on a daily and weekly basis. On the other hand, some countries showed the same methods before and during COVID-19 pandemic like; Portugal, France, and Slovenia did not change their methods as the best models to forecast electricity prices. while other countries like; Turkey, Latvia, Hungary, and Belgium have shown us a change in one method as the best to forecast electricity prices.

2.8 Seasonal Auto-Regressive Integrated Moving-Average

Seasonal Auto-regressive Integrated Moving Average (SARIMA) is another option in time-series modeling. It includes seven crucial components of time series analysis: order of the autoregressive part (p), degree of first differencing involved (d), the order of the moving average part (q), the order of the seasonal autoregressive part (P), degree of seasonal differencing (D), the order of the seasonal moving average part (Q), and time span of repeating seasonal pattern (m). We can see the basic SARIMA in formula 9.

$$(p, d, q)(P, D, Q)_m \quad (9)$$

First of all, the data is visualized to understand which model used. It is important when deciding which type of model to use. If there is not a seasonal and trend in the data, it can be used a regular ARIMA model. At the same time, SARIMA steps are implemented one by one in Fig. 8.

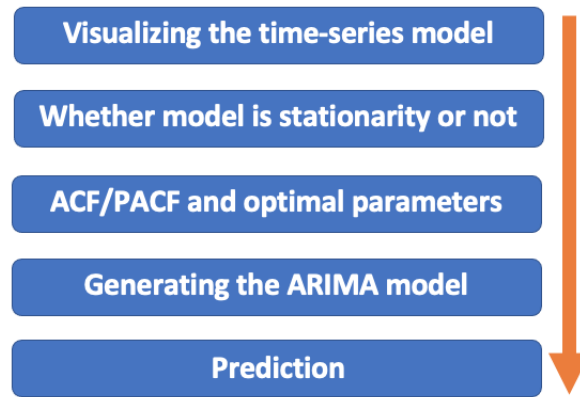


Figure 8: SARIMA Model Steps

2.8.1 Data Pre-Processing For SARIMA

Turkey, Estonia, Finland, France, Hungary, Netherlands, Portugal, Spain, Greece, Latvia, Slovenia, and Belgium are chosen. SARIMA model has been implemented by RStudio and Python. All data is extracted by means of Application Programming Interface (API) from ENTSO-E and EXIST. SARIMA model training set is between Jan. 1st, 2017 and May. 26th, 2019, and the test set is May. 27st, 2019. In other words, May. 27st, 2019 will only be predicted for just the SARIMA model. We can see the SARIMA model time period in Fig. 9.

We start analyzing the data. If the time-series data is stationary or not in the data, it is usually observed at the data and analyzed by lots of accurate assessments, particularly the Augmented Dickey-Fuller (ADF) Test. There are two crucial points at the stationarity; Trend and Seasonality. If we want to stationary our data, our data has not trend and seasonality. Thus, it is implemented the Augmented Dickey-Fuller Test. When it runs the test, the T statistics value and p-value are generated. Our Test Statistic is less than the Critical Value and our P-value is less than 0.05 in comparison to the time series is stationary, according to ADF Test. P-values are obtained less than the significance level (0.05) for each country. Also, T statistics values are negative numbers.

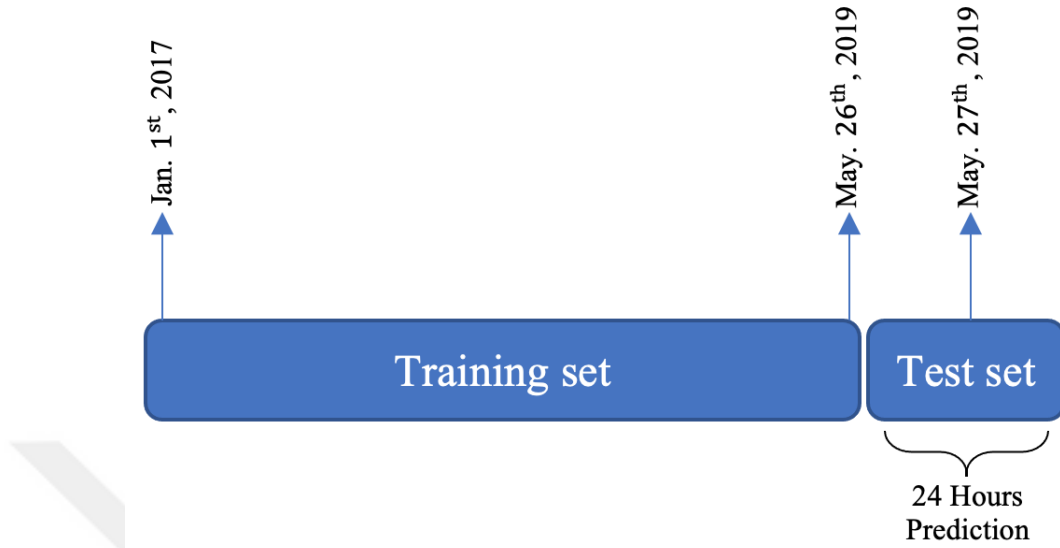


Figure 9: SARIMA Model Time Period

We can analyze whether it is seasonal or not from the autocorrelation and partial autocorrelation graph table. The ACF graph of stationary data drops to zero relatively quickly. The ACF graph of non-stationary data decreases slowly. Also, the model has MA terms at lags 1, 24, and 25. This leads to identifying ACF for the model will have non-zero autocorrelations only at lags 1, 24, and 25. At the same time, the PACF tables show a single significant spike at lag 24. So, m value can choose 24 for each country. The ACF and PACF graphs are shown in Fig. 10 and Fig. 11.

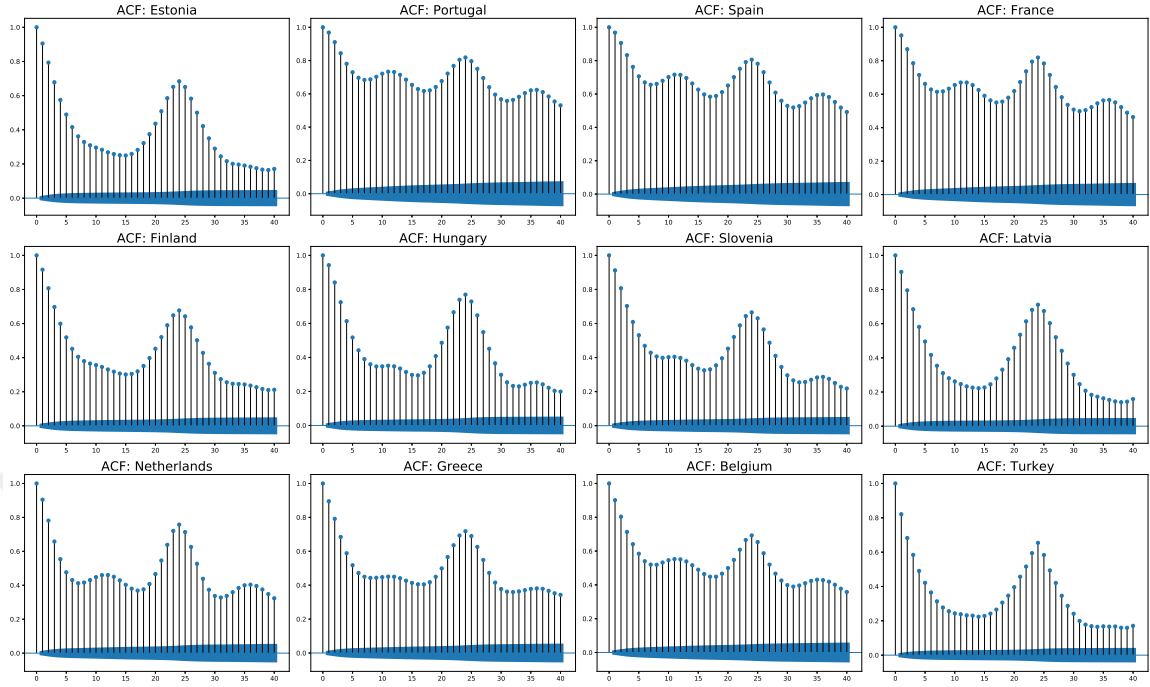


Figure 10: Autocorrelation Figures

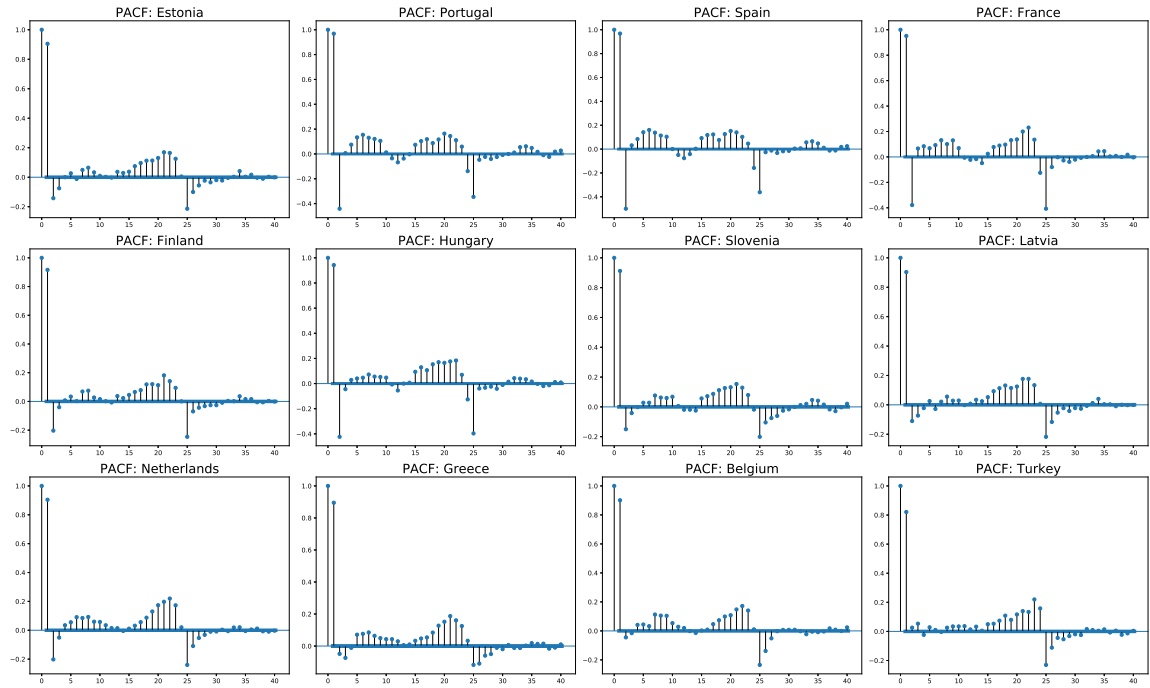


Figure 11: Partial Autocorrelation Figures

The best SARIMA models are easy to identify visualizing the data in ACF and PACF. So, each data has no trend, but each data can visually be seasonality. Lots of the data have visually a seasonality. Therefore, it is implemented Seasonal ARIMA instead of ARIMA. As a result, this should help to eliminate the overall trend and seasonality of the data.

Table 21: Best SARIMA Model For Each European Country

Countries	The Best SARIMA Model
Estonia	$(1, 0, 1)(2, 0, 0)_{24}$
Portugal	$(3, 0, 2)(2, 1, 0)_{24}$
Spain	$(3, 0, 2)(2, 1, 0)_{24}$
France	$(3, 0, 2)(2, 1, 0)_{24}$
Finland	$(1, 0, 1)(2, 0, 0)_{24}$
Hungary	$(5, 0, 1)(2, 1, 0)_{24}$
Slovenia	$(2, 0, 2)(2, 1, 0)_{24}$
Latvia	$(1, 0, 0)(2, 1, 0)_{24}$
Netherlands	$(2, 0, 1)(2, 1, 0)_{24}$
Greece	$(4, 0, 3)(2, 0, 0)_{24}$
Belgium	$(4, 0, 3)(2, 0, 0)_{24}$
Turkey	$(5, 0, 1)(2, 1, 0)_{24}$

SARIMA model parameters are decided by analyzing the autocorrelation and partial autocorrelation graphs. At the same time, `auto.arima()` function helps us to find SARIMA model parameters for each country. In order to find the best model, it is generated the parameters for the model for each country. We can see that each country has different parameters in Table 21. Also, the performance of the estimation is measured by RMSE, MAPE, and MAE as shown in Table 22.

Table 22: Error Performance of SARIMA Prediction

Countries	RMSE	MAPE	MAE
Estonia	20.61	47.68	17.13
Portugal	5.84	9.85	4.99
Spain	6.32	10.56	5.42
France	15.78	36.72	12.64
Finland	20.77	47.55	17.20
Hungary	13.18	26.18	11.00
Slovenia	19.63	37.72	16.35
Latvia	14.05	36.02	11.99
Netherlands	9.10	20.25	7.71
Greece	16.84	15.15	19.65
Belgium	11.63	28.39	9.65
Turkey	18.25	39.12	15.95

The plot of SARIMA method predictions and confidence intervals for the selected 12 European countries is given in Fig. 12, and also we can see Statistics of SARIMA Prediction in Table 23.

Table 23: Statistics of SARIMA Prediction

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	33.02	55.61	56.22	36.79	32.95	45.34	39.78	57.19	40.08	68.60	26.08	36.27
Min.	23.73	40.80	41.19	5.16	23.60	19.60	15.46	24.92	19.03	34.26	9.54	-3.63
Mean	27.32	47.34	47.44	21.37	27.20	28.14	25.65	39.32	30.29	57.35	19.10	17.27
Std.	2.41	4.20	4.02	9.38	2.42	7.46	7.71	9.00	6.11	9.19	4.64	10.87

2.9 Long Short-Term Memory (LSTM)

Long Short-Term Memory network is a variety of RNN architectures. It enables to prevent of vanishing gradient problems in a neural network. Therefore, the LSTM

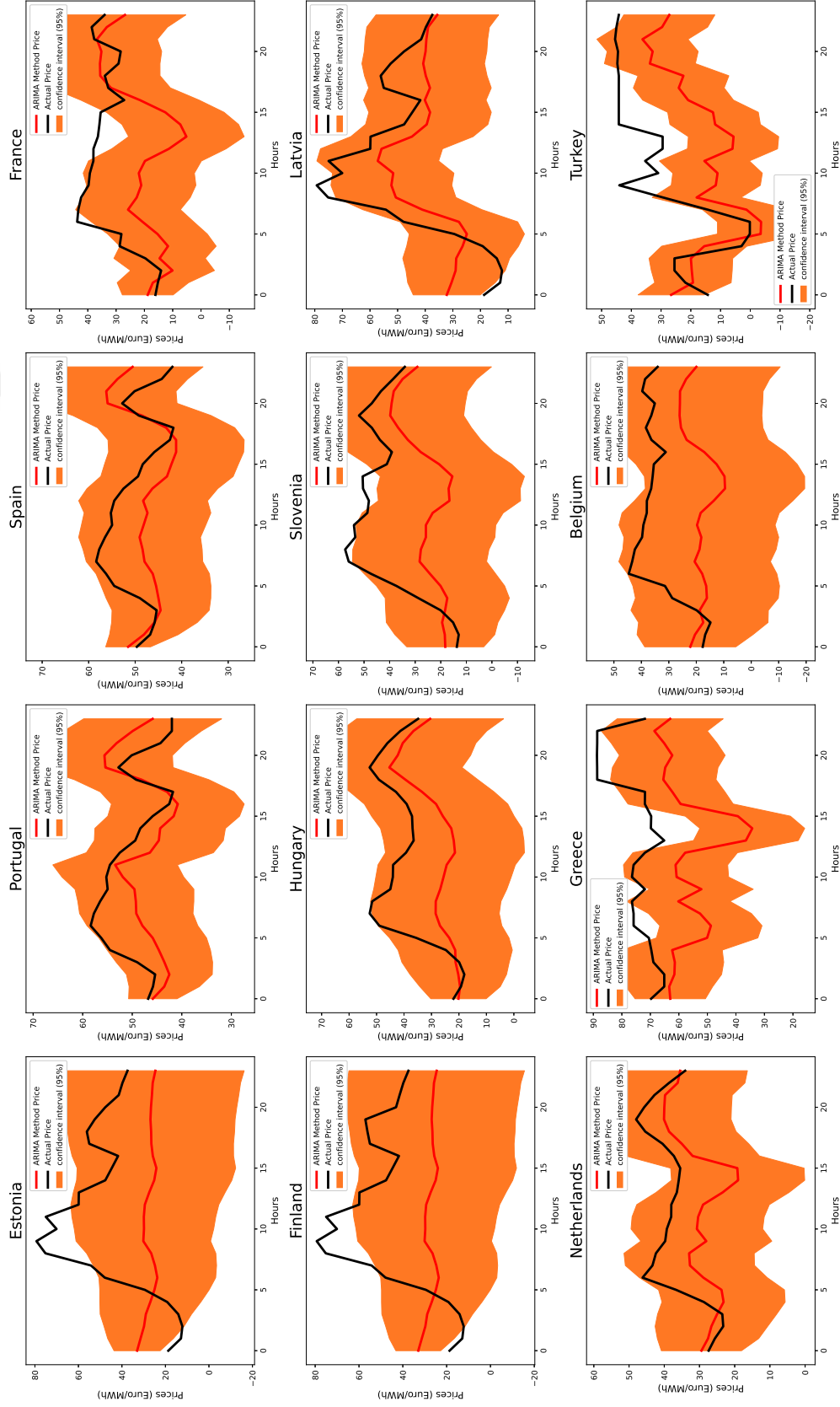


Figure 12: SARIMA Model Prices vs Actual Prices

model attains the expected results when it predicts the time-series algorithm. There are lots of LSTM structures in neural networks. The most important of these architectures include the memory cell which contains three different gates. Also, These special gates are output, forget, and input gates. They include the sigmoid activation function. Also, the sigmoid function takes values between 0 and 1 in these gates, tanh activation takes values between -1 and 1 in these gates.

2.9.1 Data Pre-Processing For LSTM

Turkey, Estonia, Finland, France, Hungary, Netherlands, Portugal, Spain, Greece, Latvia, Slovenia, and Belgium are chosen. LSTM is implemented in Python. All data is extracted by means of Application Programming Interface (API) from ENTSO-E and EXIST. The hourly electricity price data includes the period between Jan. 1st, 2017 and Dec. 24th, 2019 for all countries. We can illustrate splitting train and test set as LSTM model time period for 12 European countries in Fig. 13.

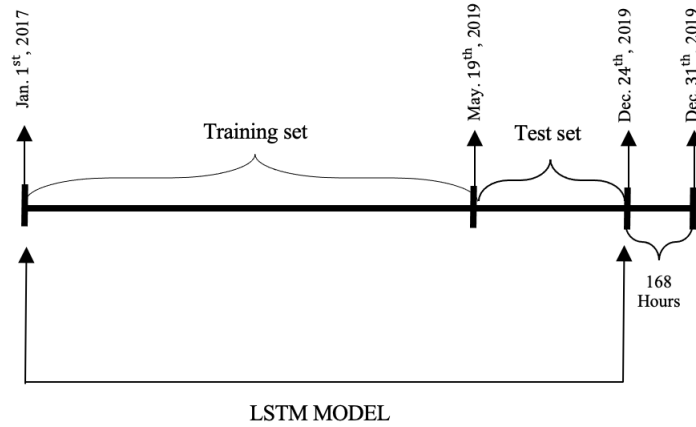


Figure 13: LSTM Model Period

Aside from sigmoid and tanh activation, forget gate decides which information can be thrown away from the previous hidden state. Input modulation gate adjusts the

cell state, and also the hidden state is determined what can be done by the output gate. We can illustrate the formulation of LSTM cells (10).

$$\begin{aligned}
i_t &= \sigma(w_i[h_{t-1}, x_t] + b_i) \\
f_t &= \sigma(w_f[h_{t-1}, x_t] + b_f) \\
o_t &= \sigma(w_o[h_{t-1}, x_t] + b_o) \\
\hat{c}_t &= \tanh(w_c[h_{t-1}, x_t] + b_c) \\
c_t &= f_t * c_{t-1} + i_t * \hat{c}_t \\
h_t &= o_t * \tanh(c_t)
\end{aligned} \tag{10}$$

i_t represents input gate. f_t is forget gate. o_t is output gate. c_t is cell state at time. σ represents sigmoid function. w_i is weight. h_{t-1} represents outputs of the previous LSTM block. x_t is input at current time. b_t is biases for the respective gates. $\hat{c}_t$ represents candidate for cell state at time. The used network hyperparameters for the LSTM model is shown in Table 24.

Table 24: Hyperparameters of Long-Short Term Memory

Name	Values
Batch Size	50
Number of Epochs	15
Optimizer	Adam
Dropout	0.1
LSTM Units	5
LSTM Layer	1

It is implied that the current electricity price is generally similar to the electricity prices of one-week ago. Therefore, one-week electricity prices, which are not included in the model, were compared with the last one-week electricity prices predicted in the test set for each European country. We can see the value of the look-back period is varied from 2 to 168 in Table 25. Five look-back experiments are performed for each

value of look-back periods utilized with the LSTM model and the error performance is measured by the RMSE. There is a noticeable decrease in the error performance results from 24 to 168 look-back sizes, except for Greece. We assume that predicting the electricity prices 1 week ago allows us to get more accurate results in electricity price forecasting. The last one week of the test model that includes Dec. 17th and 24th, 2019 and Dec. 25th and 31st, 2019 that was not included in the LSTM model were compared. We can see the Statistics of LSTM predictions in Table 26 and 27, and also the actual and predicted by LSTM model electricity prices which have some look-back sizes 24 and 48 in Fig. 14 and 15.

Table 25: RMSE Values of LSTM Each European Country

<i>Look-back Size</i>	<i>2</i>	<i>12</i>	<i>24</i>	<i>48</i>	<i>168</i>
<i>Countries</i>	<i>RMSE Values</i>				
Estonia	13.66	13.82	13.43	13.30	12.05
Portugal	21.32	22.87	22.74	24.89	21.69
Spain	20.12	24.05	23.10	21.58	21.41
France	17.02	17.66	17.64	17.35	14.33
Finland	13.48	13.84	13.57	13.92	12.04
Hungary	14.13	14.23	14.72	14.57	12.31
Slovenia	14.33	14.40	14.48	14.63	11.55
Latvia	13.56	13.60	13.37	13.51	11.99
Netherlands	12.55	12.56	12.33	12.27	11.78
Greece	27.14	26.66	27.00	27.17	30.09
Belgium	15.35	13.43	13.70	12.66	13.15
Turkey	14.74	13.51	14.88	13.60	11.74

Table 26: Statistics of LSTM (Look-back=24)

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	58.27	46.91	49.47	54.59	58.02	57.15	52.88	57.27	49.59	77.72	45.32	51.18
Min.	29.86	2.44	-0.64	1.08	30.82	1.88	-0.59	29.47	12.77	17.52	17.00	11.32
Mean	40.68	19.26	18.80	25.48	41.34	30.99	30.56	40.30	34.39	57.22	30.86	40.28
Std.	7.28	13.18	14.17	12.63	7.12	12.24	11.75	7.36	6.84	12.01	6.56	7.46

Table 27: Statistics of LSTM (Look-back=48)

EUR/MWh	Estonia	Portugal	Spain	France	Finland	Hungary	Slovenia	Latvia	Netherlands	Greece	Belgium	Turkey
Max.	58.50	46.34	50.48	53.79	58.11	59.06	51.98	57.60	50.14	78.17	47.15	49.69
Min.	27.87	1.98	-0.10	1.97	26.60	1.39	-3.43	28.46	8.67	11.79	12.57	13.87
Mean	40.41	18.85	20.00	25.42	39.99	31.78	30.03	40.06	34.15	57.47	30.60	42.58
Std.	7.52	13.75	14.54	12.52	7.70	12.54	12.48	7.56	7.06	12.88	7.96	7.35

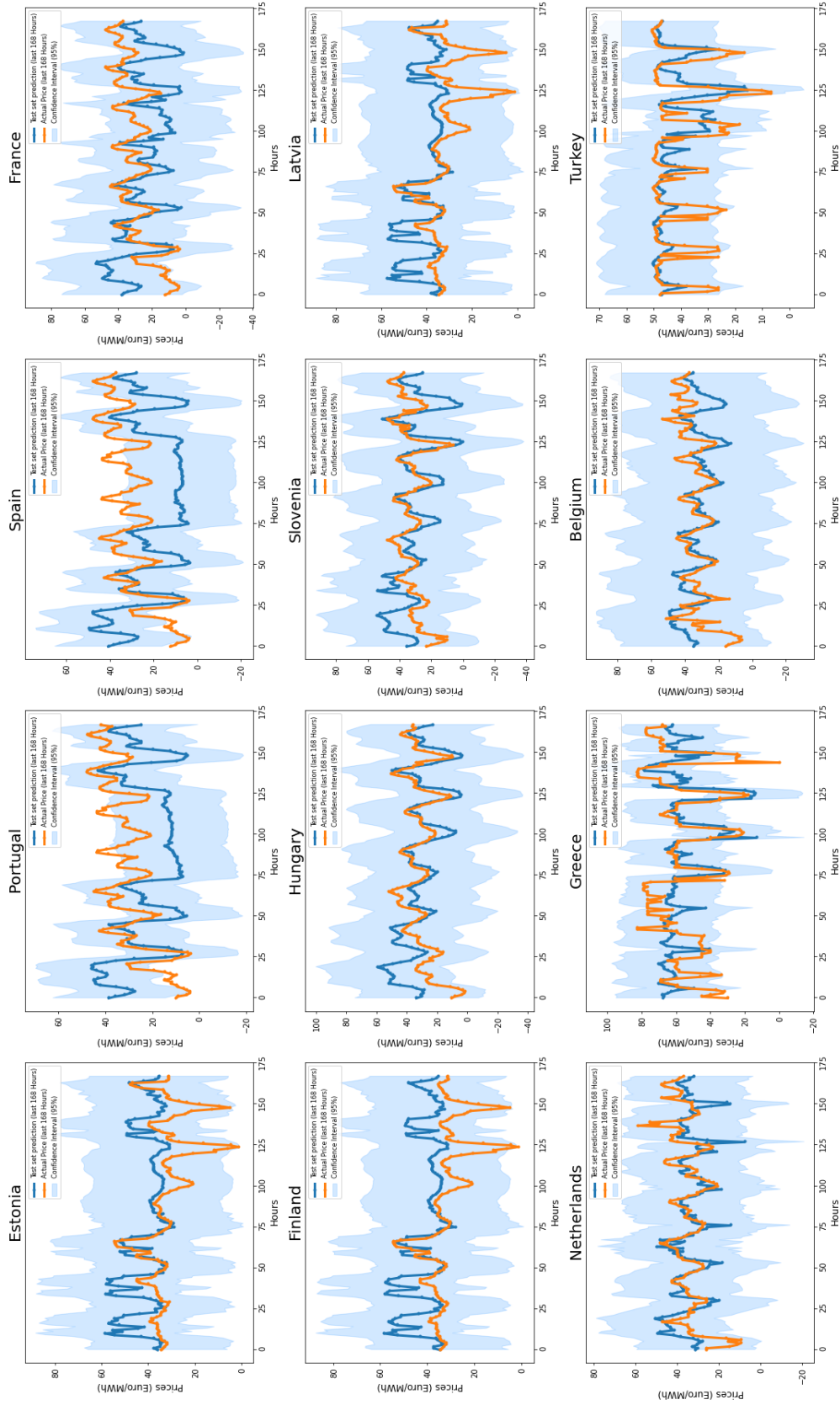


Figure 14: Actual and LSTM Predicted Prices (Look-back = 24)

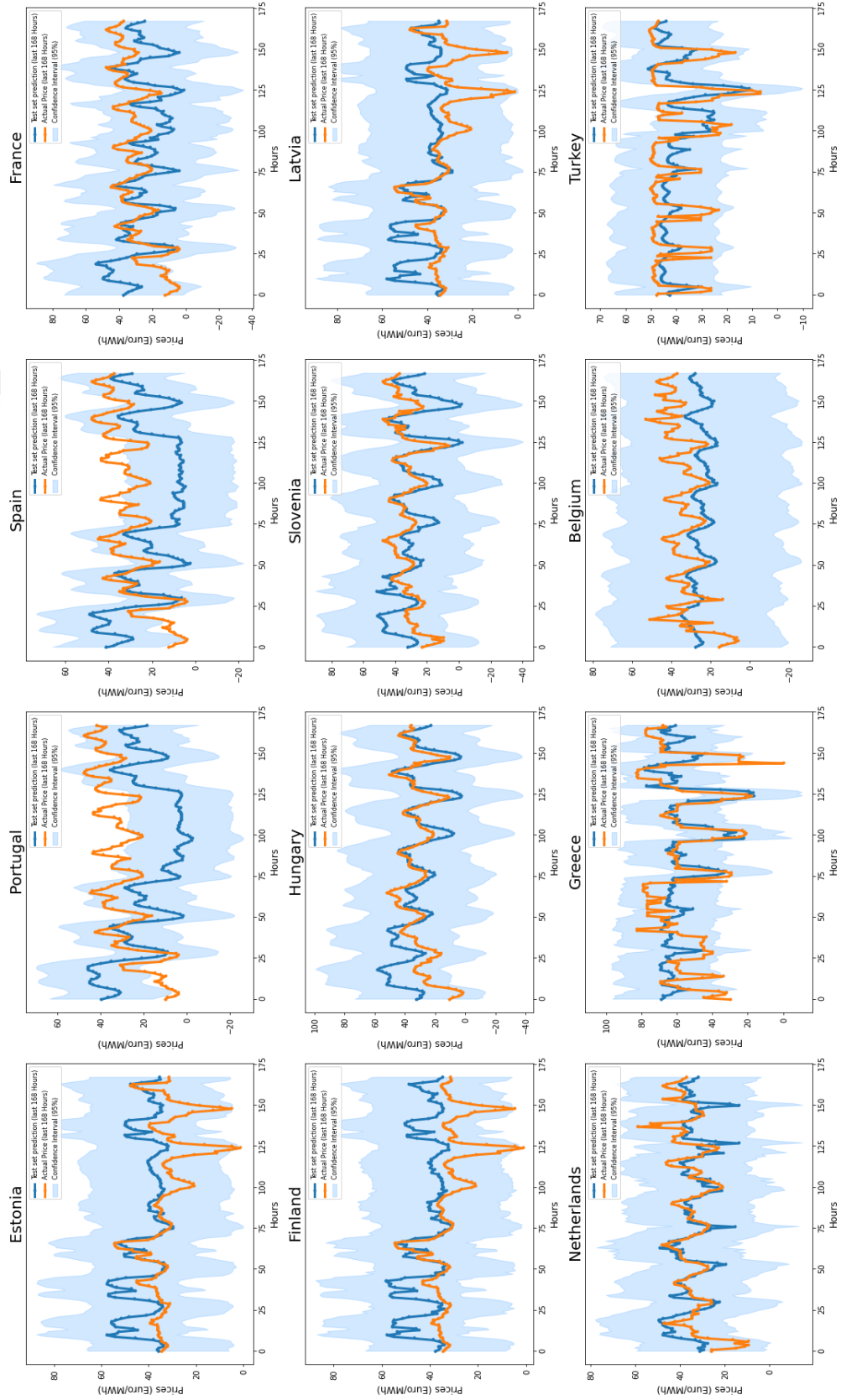


Figure 15: Actual and LSTM Predicted Prices (Look-back = 48)

CHAPTER III

CLASSIFICATION METHODS

The data used for the classification part is extracted from [60]. The raw file includes hundreds of features and their corresponding data that represent the energy and electricity status for all European Union countries until 2019. Only (33) features were considered the most comprehensive and important in terms of their influence on electricity prices. We started with the features that play the main role in maintaining supply-demand balance in any grid and thus control the electricity prices: installed energy generation, final energy consumption, gross electricity generation. Also, [60] has classified final electricity per capita as the main energy indicator in European countries. Due to the strong dependency of electricity generation on fossil fuels, they still play a major role in controlling electricity prices worldwide. Fossil fuels such as solid fuels, oil products, natural gas, and nuclear are different in terms of prices, efficiency, and transportation. The related features used in this thesis are gross electricity generation by fuels: solid fossil fuels, oil and petroleum products, natural gas, nuclear, and others. Moreover, the availability of energy resources in a country is considered another important player in the price portfolio. Some countries have a lack of energy and they used to import it from outside and others have an excess of it where it can be exported and sold outside. Thus, available gross energy, imports, exports, and energy transformation inputs and outputs are considered important indicators and used as features. Renewable energy generation has a big impact on electricity prices as well. Due to its nature of free input, in some countries, higher renewables lower the electricity prices; however, for some countries that may not be the case. This is why we choose six different features to represent a renewable portfolio

of the country: the power of solar and wind capacity, percentage of solar and wind capacity, gross electricity generation by renewables and biofuels, wind gross electricity production, solar gross electricity production. Besides the electricity demand and due to the weather difference in European countries, the heat demand is another key indicator that strongly affects electricity demand and thus electricity prices. Cogeneration or combined heat and power (CHP) represents the simultaneous generation of heat and electricity to enhance covering both electricity and heat demands and especially natural gas prices are considered as a major influencer in electricity prices [61]. The related features used for this purpose are CHP electricity generation, CHP electricity capacity, and CHP in total electricity generation.

There are worldwide trends to estimate and reduce greenhouse gases (GHG) emissions to avoid their fatal consequences on the environment. Some strategies to mitigate these emissions are investing in low carbon technologies, CO₂ trade, emission taxing, etc. Some countries are capable to conduct such strategies while other capabilities are limited. These efforts unevenly affect the electricity market and thus the electricity prices especially in the European countries. The selected features in this thesis to represent the environmental impact are (i) gas emissions from public electricity and heat production, (ii) GHG national total emissions, and (iii) total GHG per capita. The last two features selected to classify electricity prices are the population and Gross Domestic Product (GDP) which play a major role when analyzing any economic aspects in different countries. The previous part represented the time series forecasting methods that are used in the first stage to predict electricity prices in different countries. There was the best forecasting model for each country, and we choose this model to become a class by itself. In this section, we will use those classes to classify countries based on their features that selected and using different classification methods. Thus, classification models will be trained by using the data set above to identify the best time series prediction algorithm for each country. In

other words, the features of countries and their corresponding data will be taken as the input of the classifier and the best prediction model will be released as the output. The overall classification features (data may be found in [60]) are given in Table 28.

Table 28: Classification Features

Energy production (Mtoe)	CHP electricity capacity (GW)
Energy production by renewables and biofuels (Mtoe)	CHP in total electricity generation (%)
Imports of electricity (Mtoe)	Renewable energy generation (%)
Exports of electricity (Mtoe)	Final electricity per capita (KWh per Capita)
Gross available energy (Mtoe)	Gross electricity generation percentages
Transformation input (Mtoe)	- Solid fossil fuels (%)
Transformation output (Mtoe)	- Oil and petroleum products (%)
Energy sector, as electricity (Mtoe)	- Naturel gases (%)
Final energy consumption by electricity (Mtoe)	- Nuclear (%)
Installed Electricity Capacity (MW)	- Renewables and biofuels (%)
Solar and wind power capacity (MW)	- Others (%)
Percentage of solar and wind capacity (%)	Gases Emission
Gross electricity generation (TWh)	- Public electricity and heat production (ton CO2)
Gross electricity generation by Renewables and biofuels (TWh)	- GHG national total emissions (%)
Wind gross electricity production (TWh)	- Total GHG per capita (ton CO2)
Solar gross electricity production (TWh)	Total population (thousands of people)
CHP electricity generation (TWh)	Gross domestic product (GDP) (EUR)

3.1 K-Nearest Neighbors (kNN)

kNN algorithm expresses the sample classified by majority voting from its neighbor, with the sample is assigned to the class widespread. Ref [62] used a kNN method for short-term electricity load and price forecasting.

3.2 Decision Tree

It is generated handling time-series data with a trend and seasonal variation. This model creates that when the predicted result is in the class to which the data belongs [63]. Ref [64] utilized similar research about energy price classification.

3.3 Support Vector Machine (SVM)

SVM is a supervised learning algorithm and it is generally utilized for classification. It creates a hyperplane in dimensional space, so the distance is maximized to the closest data point [65]. Ref [66] utilized support vector machines based on short-term electricity load-price forecasting.

3.4 Gradient Boosting Tree Classifier

Gradient Boosting Algorithm is a method of developing weak learners. New trees are generated based on prediction errors. Iteration continues until the number of decided trees or no further improvement can be made from the model. Ref [67] implemented gradient boosting algorithm based on crude oil price forecasting.

3.5 AdaBoost Classifier

Adaboost is the algorithm that observes more of the training data that a predictor learns missing. First, the classifier is implemented on the training set and predicted. Later, the weight of misclassified training data is boosted. These boosted weights are trained and predicted. The weights are updated again and so on. Ref [68] used A comparison of AdaBoost algorithms in time-series forecasting.

3.6 Stochastic Gradient Descent

Stochastic Gradient Descent observes a random point at each step to change parameters. The points change constantly and aim to reach the optimum point. Bo-Yang et al. [69] proposed a new learning rate scheduler on Stochastic Gradient Descent algorithm.

3.7 Random Forest

Random forest creates multiple decision trees and combines them to get a more accurate forecast. It is a collection of decision trees, often trained in the bagging method. Ref [70] predicted electricity consumption of the USA based on Random Forest.

3.8 Bagging Classifier

The Bagging Classifier utilizes the same training algorithm in each prediction and trains it by separating in different random subsets of the training set. YBagging algorithm allows training samples to be sampled multiple times. Ref [71] implemented the best time-series methods with 33 features for each country using Bagged Trees.

3.9 Ridge Classifier

Ridge Classifier is generally used to analyze multivariate data. It creates a model with all variables, it does not extract unrelated variables on the data, it reduces closer to zero the weights of unrelated variables. Anagh et al. [72] used ridge classifier on comparison of linear discriminant analysis and ridge classifier.

3.10 Results of Classification Methods

The main aim is to classify using labels on the best methods of price estimations for each country. First of all, 33 features are normalized using zero-mean unit variance. The best time-series methods of European countries, classification features are used

as labels. In other words, numbers from 1 to 8 were given for each time-series prediction method. The label of Simple Exponential Smoothing Day Method is 1, Holt's Day Method is 2, Holt-Winter Day Method is 3, Naïve Day Method is 4, Simple Exponential Smoothing Week Method is 5, Holt's Week Method is 6, Holt-Winter Week Method is 7, and Naïve Week Method is 8. After splitting as 80% train set and 20% test set, the best parameter values of 9 different classification methods are found by GridSearchCV function on the data set. Also, the leave-one-out cross-validation method is used as a special case of cross-validation. The best parameter values for each classification method are shown in Table 29. Also, classification algorithms have been implemented by Python.

Table 29: Best Parameters of Classification Methods

Classification Algorithms	The Best Parameters
K-Nearest Neighbors	<i>metric: euclidean, n_neighbors: 2, weights: uniform</i>
Gradient Boosting Classifier	<i>learning_rate: 0.4, n_estimators: 16</i>
AdaBoost Classifier	<i>learning_rate: 0.3, n_estimators: 8</i>
Stochastic Gradient Descent	<i>alpha: 0.05, loss: log, penalty: l1</i>
Support Vector Machine	<i>C: 100, gamma: 0.05, kernel: rbf</i>
Decision Tree	<i>criterion: gini, max_depth: 6, min_samples_leaf: 2, min_samples_split: 8</i>
Random Forest	<i>criterion: gini, max_depth: 6, max_features: log2, n_estimators: 32</i>
Bagging Classifier	<i>base_estimator_max_depth: 20, max_features: 0.5, max_samples: 0.9</i>
Ridge Classifier	<i>alpha: 0.0001, fit_intercept: False, solver: sparse_cg</i>

After finding the best parameters, new labels are predicted based on the best parameters. The accuracy of different classification methods is indicated in Table 30.

Table 30: Accuracy of Different Classification Methods

Methods	Accuracy (%)
AdaBoost Classifier	41.7%
K-Nearest Neighbors	41.7%
Support Vector Machine	33.3%
Decision Tree	25.0%
Random Forest	25.0%
Bagging Classifier	25.0%
Gradient Boosting Classifier	16.7%
Stochastic Gradient Descent	16.7%
Ridge Classifier	16.7%

AdaBoost and K-Nearest Neighbors Classifiers are a good choice in comparison to other methods. On the other hand, Gradient Boosting Classifier, Stochastic Gradient Descent, and Ridge Classifier perform worse compared to others. As a result, Table 30 indicates that using the AdaBoost algorithm gives the minimum prediction error in the hourly day-ahead electricity price, whereas the K-Nearest Neighbors Classifiers approximately outperform a similar percentage as the AdaBoost algorithm.

CHAPTER IV

MULTIPLE LINEAR REGRESSION

Multiple linear regression is one of the common forecasting methods in price forecasting. The MLR model always calculates as a weighted sum of input features. One of the most prominent advantages is linearity. It enables to predict procedure simple. Moreover, it is easy to understand and interpret.

4.1 Data Pre-Processing For Multiple Linear Regression

Turkish day-ahead electricity price is used for estimation in this part. The data set for this thesis has been taken from the Energy Exchange Istanbul Transparency Platform (EXIST) [51] and oil, coal, and natural gas also have been taken from Bloomberg Finance L. P. [73] by means of Application Programming Interface (API). The zero values found in the data set are changed to 0.01, because the MAPE could make the denominator undefined when calculating MAPE because of the zero values.

All observations are hourly data with 21 independent predictors. Some of the features that correlate with electricity price ought to contain like hour-of-day, day-of-week, holidays, forecasted load, lagged electricity price, and lagged NG price, Furthermore, lagged crude oil prices and lagged coal prices are added to culminate in forecasting performance. Also, moving average forecasted load, lagged electricity price, and lagged natural gas price were selected for this model. Tao et al. [74] used a similar moving average model for load estimation. Besides, Regular Forecast Method Rolling Forecast Structure is illustrated in Fig. 16.

Regular Forecast Method - Rolling Forecast maintains 12 Periods												
Sep. 3 rd , 2018 - Sep. 3 rd , 2019	Sep. 3 rd , 2019	Sep. 4 th , 2019	Sep. 5 th , 2019	Sep. 6 th , 2019	Sep. 7 th , 2019	Sep. 8 th , 2019	Sep. 9 th , 2019	Sep. 10 th , 2019	Sep. 11 th , 2019	Sep. 12 th , 2019	Sep. 13 th , 2019	Sep. 14 th , 2019
Actual	Forecast											
Actual	Actual	Forecast										
Actual	Actual	Actual	Forecast									
Actual	Actual	Actual	Actual	Forecast								
Actual	Actual	Actual	Actual	Actual	Forecast							
Actual	Actual	Actual	Actual	Actual	Actual	Forecast						
Actual	Actual	Actual	Actual	Actual	Actual	Actual	Forecast					
Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Forecast				
Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Forecast			
Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Forecast		
Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Forecast	
Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Actual	Forecast

Figure 16: Regular Forecast Method Rolling Forecast Structure

Additionally, the entire model contains lots of functions in the short-term forecasting model as in Model 1 (11). In this thesis, Mean Absolute Percentage Error (MAPE) utilized to measure error rates for evaluating price estimation as given in (1), Root Mean Square Error (RMSE) used error rates for evaluating price estimation as given in (2) and Mean Signed Deviation (MSD) used as given in (3).

$$\text{Model 1 : } f_1(H, W, I, L, t) = \beta_0 + \beta_1 H_t + \beta_2 W_t + \beta_3 I_t + \beta_4 H_t W_t + \beta_5 L_t + \quad (11)$$

$$\beta_6 L_{t-24} + \beta_7 L_{t-168} + \beta_8 \frac{1}{24} \sum_{\tau=1}^{24} L_{t-\tau}$$

$$\textbf{Model 2} : f_2(\hat{P}, t) = f_1(H, W, I, L, t) + \beta_9 \hat{P}_{t-24} + \beta_{10} \hat{P}_{t-168} + \beta_{11} \frac{1}{24} \sum_{\tau=1}^{24} P_{t-\tau} \quad (12)$$

$$\textbf{Model 3} : f_3(\hat{N}G, t) = f_2(\hat{P}, t) + \beta_{12} \hat{N}G_{t-24} + \beta_{13} \hat{N}G_{t-168} + \beta_{14} \frac{1}{24} \sum_{\tau=1}^{24} \hat{N}G_{t-\tau} \quad (13)$$

$$\textbf{Model 4} : f_4(\hat{O}, t) = f_3(\hat{N}G, t) + \beta_{15} \hat{O}_{t-24} + \beta_{16} \hat{O}_{t-168} + \beta_{17} \frac{1}{24} \sum_{\tau=1}^{24} O_{t-\tau} \quad (14)$$

$$\textbf{Model 5} : f_5(\hat{C}, t) = f_4(\hat{O}, t) + \beta_{18} \hat{C}_{t-24} + \beta_{19} \hat{C}_{t-168} + \beta_{20} \frac{1}{24} \sum_{\tau=1}^{24} C_{t-\tau} \quad (15)$$

where p_t is the price estimation at hour t ; β_i are the coefficients by using the regression method; H_t , W_t , I_t are hour-of-day, day-of-week, and whether it is working day or not (a binary variable (0-1)), respectively and $\hat{L}_t$ is forecasted load data at a time; $\hat{P}_t$ is lagged prices at time, $\hat{N}G_t$ is the natural gas price at the time, $\hat{O}_t$ is crude oil price and $\hat{C}_t$ is coal price at the time.

4.2 Impact of Historical Electricity Prices

In the literature of linear regression analysis, the approach of $y=f(x,\beta)$ is the common approach; which indicates the forecast of the variable y with respect to a function of the predictor vector x and the coefficients vector β .

In this section of the thesis, two different models are analyzed to compare the accuracy of the forecast with and without historical electricity prices. Model 1 (non-lagged) consists of predictors that only include calendar data, historical and current load data. To predict hourly prices, $p_t = f_1(x, t)$ equation is used in Model 1. However, it might be a good idea to use historical prices. Therefore, Model 2 (lagged) consists of all the predictors in Model 1 and historical prices. To predict hourly

prices, $p_t = f_1(x, t) + f_2(\hat{P}, t)$ equation is used in Model 2. There is no doubt that lagged prices play a key role in the short-term forecasting model. The peak error of non-lagged electricity price that Sep. 8th, 2019 is almost 14.40%. On the other hand, it is said that there is approximately no difference in Sep. 7th, 2019 between two predictors. It is shown the comparison of non-lagged and lagged electricity prices in Fig. 17.

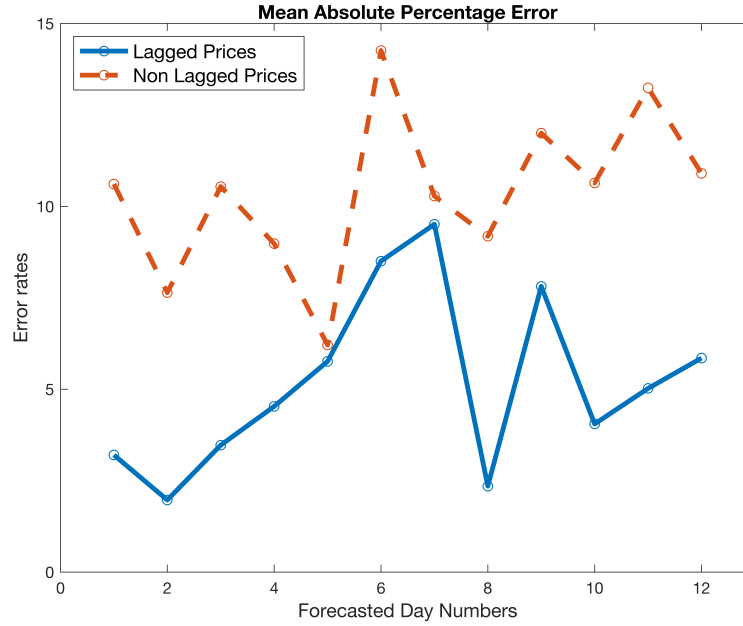


Figure 17: Error Proportions of Non-Lagged Prices vs Lagged Prices

4.3 Impact of Fuel Prices

This section of the thesis focuses mainly on understanding the correlation between fuel prices and day-ahead electricity prices in Turkey. The fuel mix of Turkish power systems has a high dominance by coal and natural gas; however, almost 45 shares is coming from renewable energy sources including hydro, wind, and solar. Therefore, it is important to understand if the electricity prices are correlated with such fuel prices; and whether this impact is useful on future price forecasting. To evaluate the impact of fuel prices, this part of the thesis first calculates the price forecast for 12

days of the testing period by calendar data and the historical day-ahead prices to set a reference forecast. Then the information of historical natural gas (NG), oil, and coal prices are included as additional predictors one by one to the training data set. It has three diverse models that are evaluated to compare the accuracy of the estimate.

First of all, Model 3 consists of predictors which include historical and current natural gas data, calendar data, historical and current load data. To estimate hourly prices, $p_t = f_1(x, t) + f_2(\hat{P}, t) + f_3(\hat{NG}, t)$ equation is used in Model 3, and then Model 4 has predictors which include historical and current oil data, historical and current natural gas data, calendar data, historical and current load data. $p_t = f_1(x, t) + f_2(\hat{P}, t) + f_3(\hat{NG}, t) + f_4(\hat{O}, t)$ equation in Model 4. Lastly, Model 5 contains predictors that have historical and current coal data, historical and current oil data, historical and current natural gas data, calendar data, historical and current load data. $p_t = f_1(x, t) + f_2(\hat{P}, t) + f_3(\hat{NG}, t) + f_4(\hat{O}, t) + f_5(\hat{C}, t)$ equation is used forecasting hourly prices in Model 5.

The average MAPE of Model 1 for the 12 days of the testing period with no consideration of fuel prices is found to be 8.75% without lagged electricity prices. After lagged electricity prices (Model 2) are added, the addition of three predictors from historical Natural gas prices (Model 3) decreased the average MAPE to 4.97% with a maximum daily MAPE of 9.55%, minimum daily MAPE of 2.39%. The addition of historical oil prices (Model 4) later reduced the average MAPE of 12 days to 4.87% with a maximum of 9.83%, a minimum of 2.05%. However, the addition of coal prices (Model 5) later increased the average MAPE to 5.14% with a maximum of 9.25%, a minimum of 2.72%. The daily MAPE of four different models is illustrated in Fig. 18.

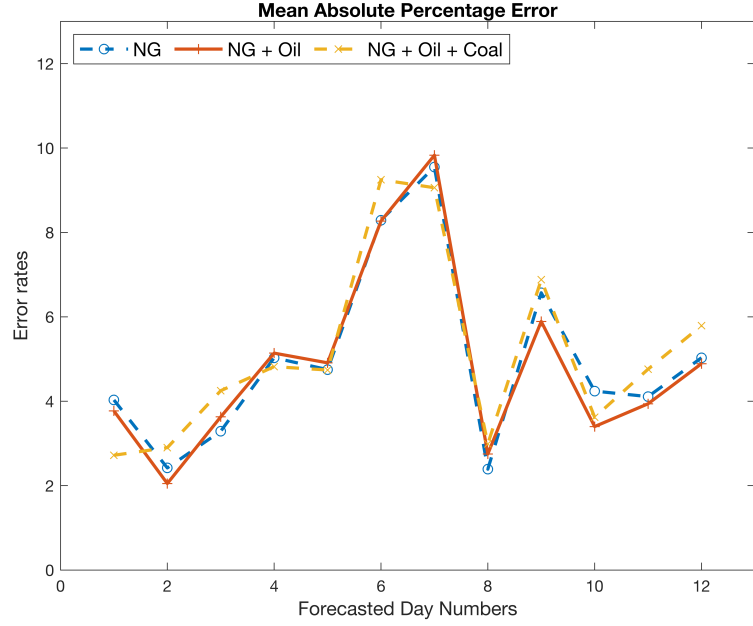


Figure 18: Error Rates of Fuel Price Predictors

4.4 *Impact of Training Data Length*

This part aims to evaluate the impact of historical data length on forecast success such as reducing the error rate. It is used as training data for one year, and then it is declined for 9 months and 6 months. The MAPE ratios are 5.15%, 6.55%, and 7.52% for one year, 9 months, and 6 months respectively. The best result is to use a one-year training set in comparison to other training sets for short-term electricity price forecasting. As a result, using a longer training set can reduce the error rate. For a testing period of 12 days, the MAPE of each day for a different training set is shown in Fig. 19.

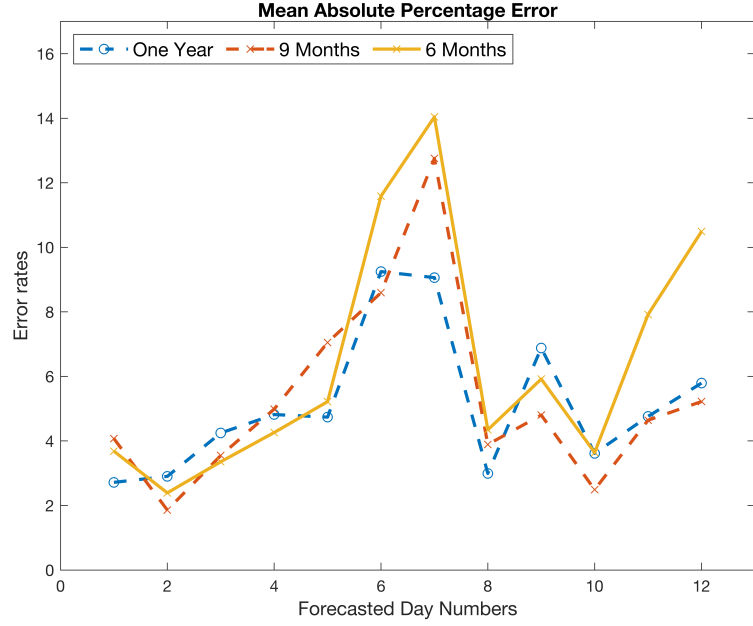


Figure 19: Error Rates of Various Training Data Lengths

4.5 *Dynamic Forecast Method*

In practice, if a multiple linear regression method has lagged values of price as a predictor in the model; this regression method with lagged predictors is called a dynamic regression model. In this section, even the proposed model is only for the next day forecasting of electricity prices, the next 12 days would be forecasted by dynamic forecasting. We execute an iterative process to forecast those lag prices first. We can see the Dynamic Forecast Method Rolling Forecast model time period in Fig. 20.

Dynamic Forecast Method - Rolling Forecast maintains 12 Periods												
Sep. 3 rd , 2018 - Sep. 3 rd , 2019	Sep. 3 rd , 2019	Sep. 4 th , 2019	Sep. 5 th , 2019	Sep. 6 th , 2019	Sep. 7 th , 2019	Sep. 8 th , 2019	Sep. 9 th , 2019	Sep. 10 th , 2019	Sep. 11 th , 2019	Sep. 12 th , 2019	Sep. 13 th , 2019	Sep. 14 th , 2019
Actual	Forecast											
Actual	Forecast	Forecast										
Actual	Forecast	Forecast	Forecast									
Actual	Forecast	Forecast	Forecast	Forecast								
Actual	Forecast	Forecast	Forecast	Forecast	Forecast							
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast						
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast					
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast				
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast			
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast		
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	
Actual	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast	Forecast

Figure 20: Dynamic Forecast Method Rolling Forecast Model Time Period

There are hourly 8760 training data sets of testing until Sep. 2nd, 2019. So, Sep. 3rd, 2019 is predicted hourly. For example, if it is estimated on Sep. 4th, 2019, there is no testing lagged electricity price previous one day for Sep. 4th, 2019. In other words, it is not known in Sep. 3rd, 2019 of electricity price, because it must be taken the previous day, lagged electricity price of Sep. 3rd, 2019 for Sep. 4th, 2019. Therefore, it is taken forecasted lagged electricity prices. In this model, it is initially performed to forecast on Sep. 3rd, 2019 with an existing training set that is an hourly 8760 data set. So, predicted Sep. 3rd, 2019 is created and used lagged electricity price to estimate Sep. 4th, 2019. This period is developed until Sep. 14th, 2019 in the same way. There is no big difference between regular and dynamic forecast electricity

prices in Fig. 21.

The MAPE, RMSE, and MSD results are given for regular forecast electricity prices and dynamic forecast electricity prices in Table 31 day by day. The MAPE results illustrate that regular forecast electricity prices are better in comparison to dynamic forecast electricity prices. Also, we can see the Statistics of Regular and Dynamic Forecast prediction in Table 32.

Dynamic Forecasting Method (% 5.95) has the highest average compared to the Regular Forecast Method (% 5.15) in three error performances. While the maximum error of Dynamic Forecast is in Sep. 14th, with % 9.21 MAPE value, Regular Forecast is in Sep. 8th, with % 9.25 MAPE value. At the same time, the minimum error of Dynamic Forecast is in Sep. 4th, with % 2.55 MAPE value. The Regular Forecast is in Sep. 3th, with % 2.72 MAPE value. The MSD and RMSE values show that Regular Forecast dramatically outperforms Dynamic Forecast.

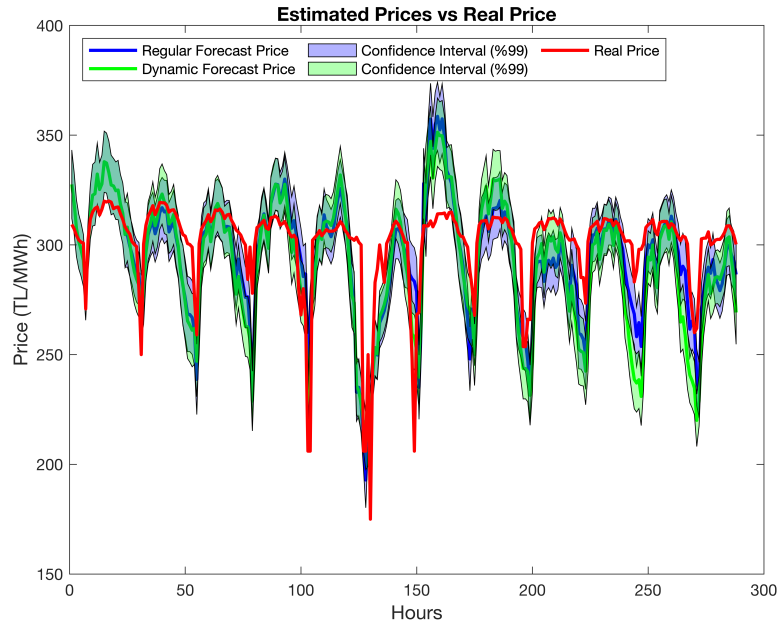


Figure 21: Estimated Prices vs Real Price

Table 31: Error Performance of Rolling Forecasting Period

Days	Mean Absolute Percentage Error		Root Mean Square Error		Mean Signed Deviation	
	<i>Regular Forecast Error</i>	<i>Dynamic Forecast Error</i>	<i>Regular Forecast Error</i>	<i>Dynamic Forecast Error</i>	<i>Regular Forecast Error</i>	<i>Dynamic Forecast Error</i>
Sep. 3, 2019	2.71%	2.72%	14.9	14.9	13.34	13.35
Sep. 4, 2019	2.90%	2.55%	8.87	8.58	-4.16	-1.11
Sep. 5, 2019	4.25%	4.86%	12.89	14.37	-8.39	-8.59
Sep. 6, 2019	4.82%	5.87%	18.63	20.65	-2.03	-4.15
Sep. 7, 2019	4.84%	3.83%	19.44	16.6	10.64	9.09
Sep. 8, 2019	9.25%	8.94%	29.86	28.71	-14.68	-10
Sep. 9, 2020	9.06%	7.46%	31.73	25.24	19.24	12.98
Sep. 10, 2020	2.99%	4.31%	13.45	14.78	-6.50	-1.41
Sep. 11, 2020	6.88%	6.75%	22.48	23.83	-20.20	-20.03
Sep. 12, 2020	3.62%	6.13%	14.64	23.07	-10.63	-18.29
Sep. 13, 2020	4.76%	8.72%	18.45	33.31	-13.96	-26.08
Sep. 14, 2020	5.79%	9.21%	19.41	30.18	-16.94	-26.96
Average Error	5.15%	5.95%	18.88	21.56	-6.40	-9.53

Table 32: Statistics of Regular and Dynamic Forecast Method

EUR/MWh	Regular Forecast Method	Dynamic Forecast Method
Max.	358.58	351.35
Min.	192.87	210.21
Mean	294.47	291.34
Std.	26.07	29.28

4.6 Feed-forward Neural Network

One of the most popular models in deep learning is the feedforward multilayer network. Feedforward neural network includes lots of simple neurons that are generated in layers. All units are connected with units that are generated in the previous layer. Besides, Feedforward Neural Network (FFNN) is an interconnection of perceptrons in the data. These interconnections have various weights that encode in the network information.

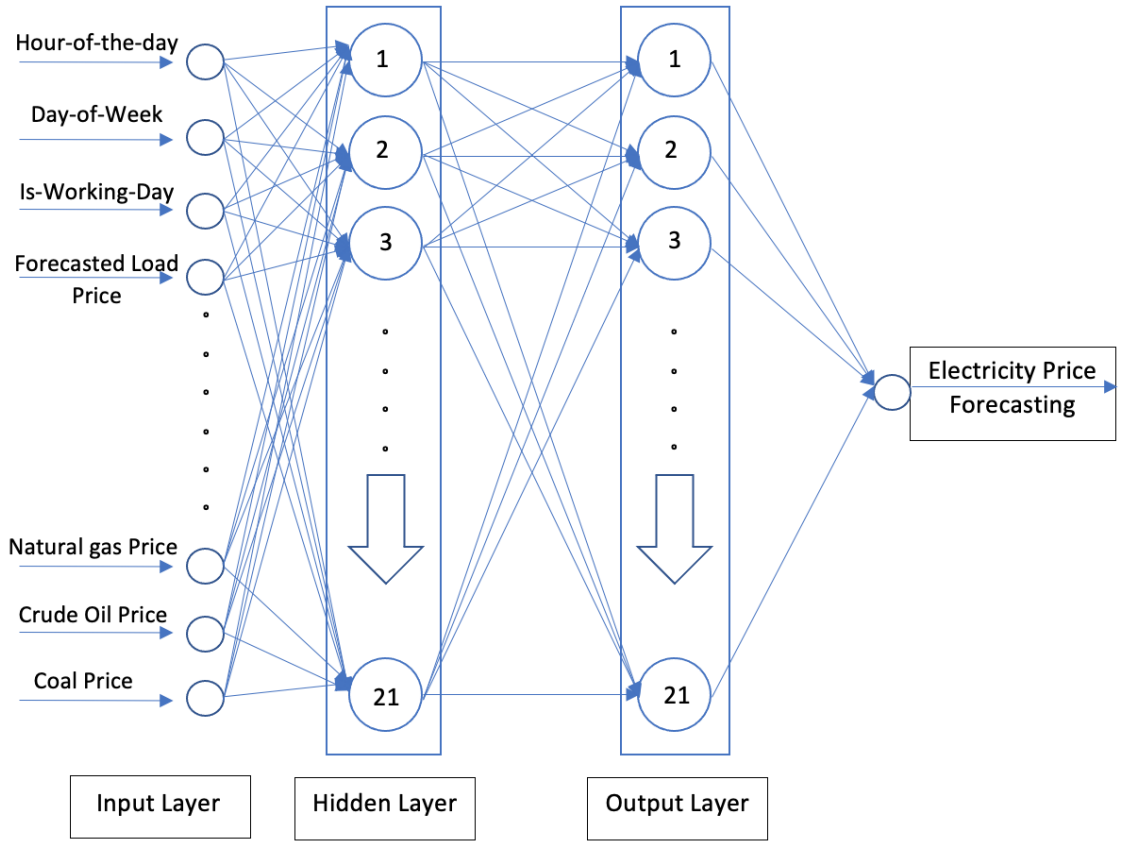


Figure 22: A Schematic Diagram of A Three-Layered Feedforward Network Model

A feedforward network structure has one input layer, one hidden layer, and one neuron in the output layer. Consequently, this model includes a three-layer network (Multi-layer). A three-layered feedforward neural network trained by the scaled conjugate gradient backpropagation for rolling forecasting next-day electricity prices day. The data first enters the input layer. Afterward, the data passes the hidden layer, until it goes into the output layer. The network has 21 input units. These inputs are features that are located in Model 5 at (15). While actual prices are target units in regular forecast FFNN, estimated prices are target units in dynamic forecast FFNN. We can see an example of an input layer with one unit, an output layer with one unit, a hidden layer with one unit in the following Fig. 22. The error performance

is measured day by day. We can see the rolling forecasting error performance by the MAPE values of four different methods in Fig. 23.

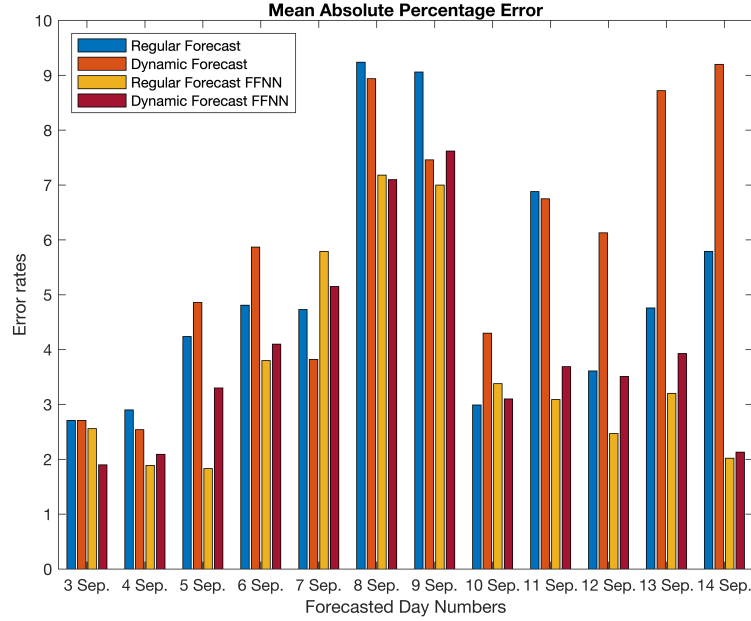


Figure 23: The Error Rate Histograms of Four Different Forecasting Methods

The Error performance of Regular Forecast FFNN and Dynamic Forecast FFNN are close to each other. Regular Forecast FFNN on Sep. 5th, 2019 reaches the lowest 1.83% MAPE and Dynamic Forecast FFNN on Sep. 3rd, 2019 reaches the lowest 1.90% MAPE. Even, the error performance of Regular Forecast Method values shows dramatically poor than Regular Forecast FFNN. The MAPE results demonstrate that Regular Forecast FFNN prices are better than Dynamic Forecast FFNN prices. We can imply that Regular Forecast FFNN prices almost outperform in comparison to other forecast methods.

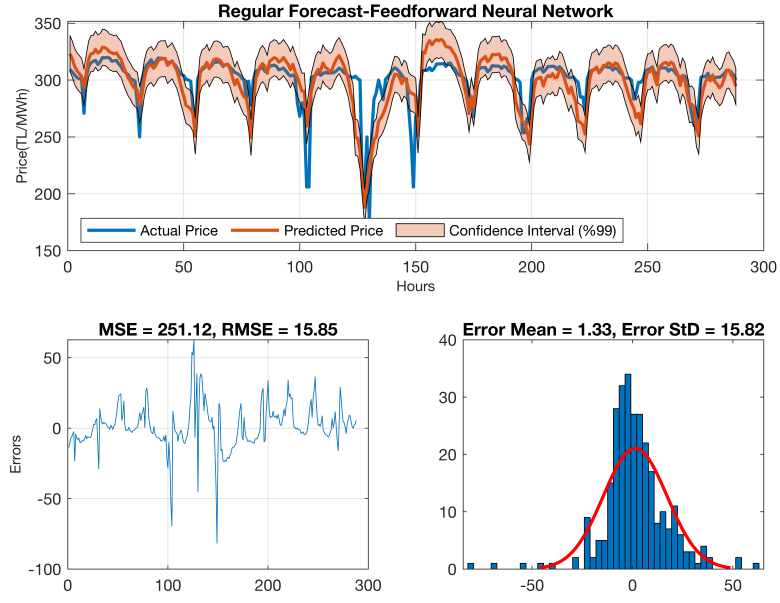


Figure 24: Regular Rolling Forecasting FFNN Prices vs. Actual Prices

Table 33: Statistics of Regular and Dynamic Forecast FFNN

EUR/MWh	Regular Forecast FFNN	Dynamic Forecast FFNN
Max.	335.82	328.72
Min.	188.11	206.95
Mean	299.54	300.16
Std.	23.00	21.80

The average MAPE of Regular Forecast FFNN for the 12 days with consideration of fuel prices (Model 5) is found to be 3.68%, maximum daily MAPE of 7.18%, minimum daily MAPE of 1.83%. Also, the average MAPE of Dynamic Forecast FFNN for the 12 days with consideration of fuel prices (Model 5) is found to be 3.97%, maximum daily MAPE of 7.62%, minimum daily MAPE of 1.90%. Regular and Dynamic Forecast FFNN is illustrated with other error performances and confidence intervals in Fig. 24 and 25. We can also see the Statistics of Regular and Dynamic Forecast FFNN in Table 33.

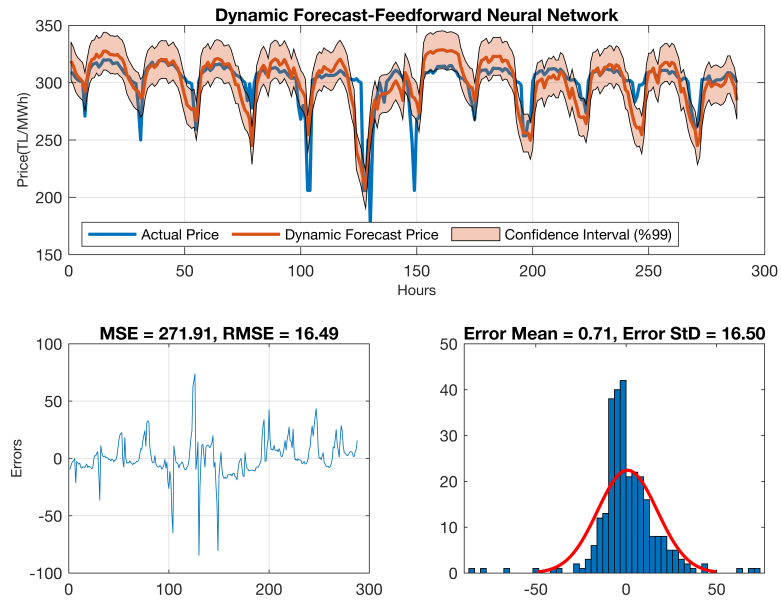


Figure 25: Dynamic Rolling Forecasting FFNN Prices vs. Actual Prices

CHAPTER V

CONCLUSION

Time-series prediction and classification methods are implemented into day-ahead electricity prices for the selected European countries. Although the main study was to find out the best prediction methods for those countries in the first place, the situation on electricity prices began to change right during the COVID-19 pandemic around the world. Therefore, we compared the changes in the best methods of electricity price predictions.

The whole family of exponential smoothing methods which are called simple, double, and triple exponential methods are initially used in hourly electricity price data sets, as well as Naïve forecasting methods. The best time-series prediction methods are selected for the European countries using RMSE. At the same time, May. 27th, 2019 was estimated using the ARIMA method for each of 12 European countries. Also, five look-back experiments are analyzed in the LSTM model and the error performance is measured by the RMSE. The 168 look-back sizes in Slovenia are reached that the lowest 11.55 RMSE. While the number of look-back sizes increases from 24 to 168, the error performance of most European countries decreased, excluding Greece. In chapter 3, 33 features of those countries are used to classify 12 European countries by nine different classification methods using the best parameters. In chapter 4, the studies on the electricity market and price estimation are evaluated by using multiple linear regression. The test results demonstrate that commodities that correlated such as electricity, NG, oil, and coal affect each other because of supply-demand changes. Also, there is possible how much training set we would use to reduce the MAPE when estimating electricity prices. Historical electricity prices which are previous one-day

prices, previous one-week prices, and previous moving average one-day prices allow the error rate to decrease in the forecasting model. Also, the historical electricity prices have an important role in this forecasting model. Whereas the estimation results of this model are generally successful, it is observed how much training data can be utilized to decrease error rates. It is precise that using a training set of one year has the minimum error rate than other training sets in this model. Furthermore, when it is particularly used natural gas and oil prices, it is indicated that the impacts of these commodities reduce the MAPE and RMSE, except coal prices. The dynamic forecast method and regular forecast method are closed to each other according to their error performances. At the same time, feedforward network is implemented on the model with 21 independent variables. The regular forecast FFNN is also a good choice to predict hourly prices. The regular forecast FFNN on Sep. 5th, 2019 is reached that the lowest 1.83% MAPE. The MAPE values of Dynamic Forecast FFNN illustrate dramatically poor than the MAPE results of Regular Forecast FFNN.

Bibliography

- [1] N. C. D. Ph, Z. Li, “Market operations in electric power systems: forecasting, scheduling and risk management,” vol. 40, pp. 40–1574–1540–1574, 2002.
- [2] V. Kaminski, *Energy Markets*. Risk Books, 2012.
- [3] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.
- [4] A. Eydeland and K. Wolyniec, “Energy and power risk management,” 01 2003.
- [5] D. Savitski, “Pricing in competitive electricity markets, a. faruqui and k. eakin, editors. forward by j. robert malko,” *Review of Industrial Organization*, vol. 21, pp. 329–333, 02 2002.
- [6] R. Weron, “Modeling and forecasting electricity loads and prices: A statistical approach,” 2006.
- [7] E. S. Gardner, “Exponential smoothing: The state of the art—part ii,” *International Journal of Forecasting*, vol. 22, no. 4, pp. 637 – 666, 2006.
- [8] L.-E. Öller and P. Stockhammar *International Journal of Forecasting*, vol. 26, no. 1, pp. 204 – 205, 2010. Special Section: European Election Forecasting.
- [9] T. O. Benli, “A comparison of various electricity tariff price forecasting techniques in turkey and identifying the impact of time series periods,” 2016.
- [10] J. S. McMenamin and F. A. Monforte, “Statistical approaches to electricity price forecasting,” pp. 249–263, 2000.
- [11] H. Raj, V. Jha, and P. Rajib, “Short term forecasting of day-ahead electricity prices: Holt winters double seasonality model vs. artificial neural network model 1 short term forecasting of day-ahead electricity prices: Holt winters double seasonality model vs. artificial neural network model,” 08 2016.
- [12] T. Jonsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, “Forecasting electricity spot prices accounting for wind power predictions,” *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 210–218, 2013.
- [13] J. Crespo Cuaresma, J. Hlouskova, S. Kossmeier, and M. Obersteiner, “Forecasting electricity spot-prices using linear univariate time-series models,” *Applied Energy*, vol. 77, pp. 87–106, 01 2004.
- [14] G. Zhang, “Time series forecasting using a hybrid arima and neural network model,” *Neurocomputing*, vol. 50, pp. 159 – 175, 2003.

- [15] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "Arima models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, 2003.
- [16] R. Weron and A. Misiorek, "Forecasting spot electricity prices with time series models," *Proceedings of the European Electricity Market EEM-05 Conference*, 05 2005.
- [17] M. A. Kölmek and Navruz, "Forecasting the day-ahead price in electricity balancing and settlement market of turkey by using artificial neural networks," *Turkish Journal of Electrical Engineering and Computer Science*, vol. 23, pp. 841 – 852, 2015.
- [18] A. Azadeh, M. Moghaddam, M. Mahdi, and S. H. Seyedmahmoudi, "Optimum long-term electricity price forecasting in noisy and complex environments," *Energy Sources, Part B: Economics, Planning, and Policy*, vol. 8, no. 3, pp. 235–244, 2013.
- [19] O. Güngör, B. Akşanlı, and R. Aydoğan, "Algorithm selection and combining multiple learners for residential energy prediction," *Future Generation Computer Systems*, vol. 99, pp. 391 – 400, 2019.
- [20] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
- [21] R. Gomathi, "A new approach for classification of prices in the electricity market using core vector machine," 2017.
- [22] D. Huang, H. Zareipour, W. D. Rosehart, and N. Amjady, "Data mining for electricity price classification and the application to demand-side management," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 808–817, 2012.
- [23] X. Yan and N. A. Chowdhury, "Mid-term electricity market clearing price forecasting using multiple least squares support vector machines," *IET Generation, Transmission Distribution*, vol. 8, no. 9, pp. 1572–1582, 2014.
- [24] S. Fan, J. R. Liao, K. Kaneko, and L. Chen, "An integrated machine learning model for day-ahead electricity price forecasting," in *2006 IEEE PES Power Systems Conference and Exposition*, pp. 1643–1649, 2006.
- [25] M. Hambali, Akinyemi, M. Oladunjoye, and Y. N., "Electric power load forecast using decision tree algorithms," vol. 7, pp. 29–42, 01 2017.
- [26] S. Obadi, "What are the causes of high crude oil price? causality investigation," *International Journal of Energy Economics and Policy*, vol. 3, pp. 80–92, 01 2013.

- [27] D. Dickey and M. P. Muñoz, “Are electricity prices affected by the us dollar to euro exchange rate? the spanish case,” *Energy Economics*, vol. 31, pp. 857–866, 11 2009.
- [28] F. Zhang and H. Fleyeh, “A review of single artificial neural network models for electricity spot price forecasting,” in *2019 16th International Conference on the European Energy Market (EEM)*, pp. 1–6, 2019.
- [29] A. Cervone, E. Santini, S. Teodori, and D. Romito, “Electricity price forecast: a comparison of different models to evaluate the single national price in the italian energy exchange market,” *International Journal of Energy Economics and Policy*, vol. 4, pp. 744–758, 12 2014.
- [30] Z. Tan, J. Zhang, J. Wang, and J. Xu, “Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models,” *Applied Energy*, vol. 87, pp. 3606–3610, 11 2010.
- [31] I. P. Panapakidis and M. N. Moschakis, “Comparison of machine learning models for the prediction of system marginal price of greek energy market,” *International Journal of Energy and Environmental Engineering*, vol. 13, no. 3, pp. 148 – 152, 2019.
- [32] M. Ranjbar, S. Soleymani, N. Sadati, and A. Ranjbar, “Electricity price forecasting using artificial neural network,” pp. 1 – 5, 01 2007.
- [33] L. Jiang and G. Hu, “Day-ahead price forecasting for electricity market using long-short term memory recurrent neural network,” 11 2018.
- [34] J. Lago, F. De Ridder, and B. De Schutter, “Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms,” *Applied Energy*, vol. 221, pp. 386 – 405, 2018.
- [35] A. Shiri, M. Afshar, A. Rahimi-Kian, and B. Maham, “Electricity price forecasting using support vector machines by considering oil and natural gas price impacts,” in *2015 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*, pp. 1–5, 2015.
- [36] C. Chatfield, “The Holt-Winters Forecasting Procedure,” *Journal of the Royal Statistical Society Series C*, vol. 27, pp. 264–279, November 1978.
- [37] P. Maçaira, R. Souza, and F. C. Oliveira, “Modelling and forecasting the residential electricity consumption in brazil with pegels exponential smoothing techniques,” *Procedia Computer Science*, vol. 55, pp. 328 – 335, 2015. 3rd International Conference on Information Technology and Quantitative Management, ITQM 2015.
- [38] P. Goodwin, “Using naave forecasts to assess limits to forecast accuracy and the quality of fit of forecasts to time series data,” *SSRN Electronic Journal*, 01 2014.

- [39] J. M. Keller, M. R. Gray, and J. A. Givens, “A fuzzy k-nearest neighbor algorithm,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 4, pp. 580–585, 1985.
- [40] B. Kaminski, M. Jakubczyk, and P. Szufel, “A framework for sensitivity analysis of decision trees,” *Central European Journal of Operations Research*, vol. 26, 03 2018.
- [41] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, “A practical guide to support vector classification chih-wei hsu, chih-chung chang, and chih-jen lin,” 2003.
- [42] R. K. Agrawal, F. Muchahary, and M. M. Tripathi, “Ensemble of relevance vector machines and boosted trees for electricity price forecasting,” *Applied Energy*, vol. 250, pp. 540–548, 2019.
- [43] S. Z. Li and A. Jain, “Adaboost,” in *Encyclopedia of Biometrics*, pp. 9–9, Springer US, 2009.
- [44] S. Shalev-Shwartz and S. Ben-David, *Stochastic Gradient Descent*, p. 150–166. Cambridge University Press, 2014.
- [45] Y. Liu, Y. Wang, and J. Zhang, “New machine learning algorithm: Random forest,” in *Information Computing and Applications*, pp. 246–252, Springer Berlin Heidelberg, 2012.
- [46] I. I. Baskin, G. Marcou, D. Horvath, and A. Varnek, *Bagging and Boosting of Classification Models*, ch. 15, pp. 241–247. John Wiley Sons, Ltd, 2017.
- [47] M. Grüning and S. Kropf, “A ridge classification method for high-dimensional observations,” in *From Data and Information Analysis to Knowledge Engineering*, pp. 684–691, Springer Berlin Heidelberg, 2006.
- [48] *MathWorks, I. MATLAB : the language of technical computing : computation, visualization, programming : installation guide for UNIX version 5.* 1996.
- [49] T. Ulgen, “Github repository.” <https://github.com/toygarulgen?tab=repositories>. Available online: 2021-03-01.
- [50] ENTSO-E, “Day-ahead electricity prices transparency platform.” <https://transparency.entsoe.eu/transmission-domain/r2/dayAheadPrices/show>. Available online: 2020-05-05.
- [51] EXIST, “Transparency platforms.” <https://seffaflik.epias.com.tr/transparency>. Available online: 2019-12-22.
- [52] O. Karabiber and G. Xydis, “Electricity price forecasting in the danish day-ahead market using the tbats, ann and arima methods,” *Energies*, vol. 12, p. 928, 03 2019.

- [53] S. Shastri, A. Sharma, V. Mansotra, A. Sharma, A. S. Bhadwal, and M. Kumari, “A study on exponential smoothing method for forecasting,” *International Journal of Computer Sciences and Engineering*, vol. 6, pp. 482–485, 4 2018.
- [54] D. Risteski, A. Kulakov, and D. Davcev, “Single exponential smoothing method and neural network in one method for time series prediction,” pp. 741 – 745, 02 2004.
- [55] R. Anggrainingsih, G. R. Aprianto, and S. W. Sihwi, “Time series forecasting using exponential smoothing to predict the number of website visitor of sebelas maret university,” in *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, pp. 14–19, 2015.
- [56] S. Makridakis, S. Wheelwright, and R. Hyndman, *Forecasting: Methods and Applications*, vol. 35. 01 1984.
- [57] C. R. Mege, I. Nashirul Haq, E. Leksono, and F. X. Nugroho Soelami, “Battery discharging temperature prediction using holt’s double exponential smoothing,” in *2019 6th International Conference on Electric Vehicular Technology (ICEVT)*, pp. 378–381, 2019.
- [58] C. Chatfield, “The holt-winters forecasting procedure,” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 27, no. 3, pp. 264–279, 1978.
- [59] D. Bissing, M. T. Klein, R. A. Chinnathambi, D. F. Selvaraj, and P. Ranganathan, “A hybrid regression model for day-ahead energy price forecasting,” *IEEE Access*, vol. 7, pp. 36833–36842, 2019.
- [60] EuropeanCommission, “Eu energy statistical pocketbook and country datasheets.” <https://ec.europa.eu/energy/data-analysis/energy-statistical-pocketbook>. Accessed: 2020-05-05.
- [61] O. G. Poyrazoglu and G. Poyrazoglu, “Impact of natural gas price on electricity price forecasting in turkish day-ahead market,” in *2019 1st Global Power, Energy and Communication Conference (GPECOM)*, pp. 435–439, 2019.
- [62] T. Ashfaq and N. Javaid, “Short-term electricity load and price forecasting using enhanced knn,” in *2019 International Conference on Frontiers of Information Technology (FIT)*, pp. 266–2665, 2019.
- [63] J. R. Quinlan, “Decision trees and decision-making,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 339–346, 1990.
- [64] J. C. R. Filho, C. M. Affonso, and R. C. L. de Oliveira, “Energy price classification in north brazilian market using decision tree,” in *2015 12th International Conference on the European Energy Market (EEM)*, pp. 1–5, 2015.

- [65] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.
- [66] R. A. Swief, Y. G. Hegazy, T. S. Abdel-Salam, and M. A. Bader, "Support vector machines (svm) based short term electricity load-price forecasting," in *2009 IEEE Bucharest PowerTech*, pp. 1–5, 2009.
- [67] M. Gumus and M. S. Kiran, "Crude oil price forecasting using xgboost," in *2017 International Conference on Computer Science and Engineering (UBMK)*, pp. 1100–1103, 2017.
- [68] D. Barrow, "A comparison of adaboost algorithms for time series forecast combination," *International Journal of Forecasting*, vol. 32, pp. 1103–1119, 06 2016.
- [69] B. Hsueh, W. Li, and I. Wu, "Stochastic gradient descent with hyperbolic-tangent decay on classification," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 435–442, 2019.
- [70] S. Nallathambi and K. Ramasamy, "Prediction of electricity consumption based on dt and rf: An application on usa country power consumption," in *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, pp. 1–7, 2017.
- [71] T. Ulgen, A. E. Sayed, and G. Poyrazoglu, "Prediction algorithm learner selection for european day-ahead electricity prices," in *2020 2nd Global Power, Energy and Communication Conference (GPECOM)*, pp. 285–286, 2020.
- [72] A. Singh, B. S. Prakash, and K. Chandrasekaran, "A comparison of linear discriminant analysis and ridge classifier on twitter data," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 133–138, 2016.
- [73] Bloomberg, "Currencies markets." <https://www.bloomberg.com/quote/EURTRY:CUR>. Available online: 2019-12-26.
- [74] P. Wang, B. Liu, and T. Hong, "Electric load forecasting with recency effect: A big data approach," *International Journal of Forecasting*, vol. 32, pp. 585–597, 07 2016.
- [75] H. Zhang and D. Li, "Naïve bayes text classifier," in *2007 IEEE International Conference on Granular Computing (GRC 2007)*, pp. 708–708, 2007.
- [76] X. Liu, Z. Liu, G. Wang, Z. Cai, and H. Zhang, "Ensemble transfer learning algorithm," *IEEE Access*, vol. 6, pp. 2389–2396, 2018.