YAŞAR UNIVERSITY

GRADUATE SCHOOL

MASTER IN SCIENCE THESIS

A COMPARATIVE STUDY OF ARIMA AND LSTM FOR TWO-DAYS AHEAD FORECASTING OF ELECTRICITY DEMAND IN İZMİR-MANİSA REGION IN TURKEY

YİĞİT AHMET ARIKÖK

THESIS ADVISOR: ASSIST. PROF.(PHD) DR. GÖKHAN DEMİRKIRAN

ELECTRICAL AND ELECTRONICS ENGINEERING

BORNOVA / İZMİR
OCAK 2024

JURY APPROVAL PAGE

We certify that, as the jury, we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Jury Members:	**Signature:**
Assist./Assoc./Prof.(PhD) Xxx YYY … University	....
Assist./Assoc./Prof.(PhD) Xxx YYY … University	....
Assist./Assoc./Prof.(PhD) Xxx YYY … University	....
Assist./Assoc./Prof.(PhD) Xxx YYY … University	....
Assist./Assoc./Prof.(PhD) Xxx YYY … University	....
Assist./Assoc./Prof.(PhD) Xxx YYY … University	....

Prof. (PhD) Yucel Ozturkoglu

Director of the Graduate School

ABSTRACT

A COMPARATIVE STUDY OF ARIMA AND LSTM FOR TWO-DAYS AHEAD FORECASTING OF ELECTRICITY DEMAND IN İZMİR-MANİSA REGION IN TURKEY

Arıkök, Yiğit Ahmet
MSc, Electrical and Electronics Engineering
Advisor: Assist. Prof.(PhD) Dr. Gökhan DEMİRKIRAN
January 2024

Predicting electricity demand has become crucial for management of energy infrastructures and purchasing energy from energy markets. This research focus on on predicting the electric demand in Izmir and Manisa regions, using actual time series data from the Gdz Electric Distribution company. The dataset encompasses daily total energy consumption time series dating from January 2020 (post-COVID) to December 2022. Employing the conventional statistical technique ARIMA and comparing its performance with Long Short-Term Memory networks (LSTM), the study examines forecasting accuracy over a search space of hyperparameters. Results showcases that ARIMA can challenge computationally-intensive LSTM especially due to the evident patterns in data. Detrending the data using seasonal indicators further increases the performance of ARIMA and can provide insights to the statistically-inclined experts.

Keywords: electricity demand, two-days ahead, electric consumption, energy infrastructures, energy markets, time series, Gdz Electric Distribution company, ARIMA, Long Short-Term Memory networks (LSTM), deep learning

ÖZ

TÜRKİYE İZMİR-MANİSA BÖLGESİNDE ELEKTRİK TALEBİNİN İKİ GÜN SONRAKİ TAHMİNİ İÇİN ARIMA VE LSTM KARŞILAŞTIRILMASI

Arıkök, Yiğit Ahmet
Yüksek Lisans Tezi, Elektrik ve Elektronik Mühendisliği
Danışman: Dr. Öğr. Üyesi Gökhan DEMİRKIRAN
Ocak 2024

Elektrik talebini tahmin etmek, enerji altyapılarının yönetimi ve enerji piyasalarından enerji satın alma açısından hayati bir hale gelmiştir. Bu araştırma, Gdz Elektrik Dağıtım şirketinden elde edilen gerçek zaman serisi verilerini kullanarak İzmir ve Manisa bölgelerindeki elektrik talebini tahmin etmeye odaklanmaktadır. Veri seti, Ocak 2020 (COVID sonrası) ile Aralık 2022 tarihleri arasına uzanan günlük toplam enerji tüketim zaman serisini içermektedir. Geleneksel istatistiksel teknik olan ARIMA'yı kullanarak ve performansını Uzun Kısa Vadeli Hafıza ağları (LSTM) ile karşılaştırarak, çalışma hiparaparametrelerin arama alanı üzerinde tahmin doğruluğunu incelemektedir. Sonuçlar, ARIMA'nın verideki açık desenler nedeniyle özellikle hesaplama yoğun LSTM'ye meydan okuyabileceğini göstermektedir. Mevsimsel göstergeleri kullanarak veriyi detrend etmek, ARIMA'nın performansını artırır ve istatistiksel olarak eğilimli uzmanlara bilgi sağlayabilir.

Anahtar Kelimeler: elektrik talebi, iki gün önceden, elektrik tüketimi, enerji altyapıları, enerji sektörü, zaman serisi, Gdz Elektrik Dağıtım şirketi, ARIMA, Uzun-Kısa Vadeli Bellek(LSTM), derin öğrenme

ACKNOWLEDGEMENTS

I would like to thank to Gdz Electricity Distribution for providing data. Additionally, my sincere appreciation goes to my advisor Gökhan Demirkıran, and my dear Melis Bostan for their invaluable support and contributions.

Yiğit Ahmet Arıkök
İzmir, 2024

TEXT OF OATH

I declare and honestly confirm that my study, titled "A COMPARATIVE STUDY OF ARIMA AND LSTM FOR TWO-DAYS AHEAD FORECASTING OF ELECTRICITY DEMAND IN İZMİR-MANİSA REGION IN TURKEY" and presented as a Master's Thesis, has been written without applying any assistance inconsistent with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Yiğit Ahmet Arıkök
İzmir, 2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLS AND ABBREVIATIONS

SYMBOLS:

y_{tr} Output

μ_{tr} Deterministic Periodic Trend

s_{tr} Stochastic Process

x_t Time Series

ε_t Error

ϕ Weight Parameter of AR Term

θ Weight Parameter of MA Term

p Autoregressive Order

d Regular Differences Number

q Moving Average Order

ABBREVIATIONS:

ARIMA AutoRegressive Integrated Moving Average

LSTM Long-Short Term Memory

ACF Auto Correlation Function

PACF Partial Auto Correlation Function

MLP Multi-Layer Perceptron

SVR Support Vector Regression

KNNR K-Nearest Neighbor Regression

ANN Artificial Neural Network

CNN Convolutional Neural Network

FNN Feedforward Neural Network

RNN Recurrent Neural Network

GRU Gated Recurrent Unit

MAPE Mean Absolute Percentage Error

MSE Mean Squared Error

RMSE Root Mean Squared Error

ADF Augmented Dickey Fuller

SGD Stochastic Gradient Descent

1. CHAPTER: INTRODUCTION

As the technology advances, the reliance on electricity increases. For example, the transitioning of the industry towards Industry 4.0 concept relies on several automated machines and increase the demand for electricity (Tesch da Silva et al., 2020). Similarly, the shift towards electric vehicles will increase the electricity demand, even necessitating new technologies to maintain power quality (Çobanoğlu et al., 2021; Das et al., 2020; Justin et al., 2022; Rahman et al., 2022). Accurate prediction of electricity consumption is of great importance for sustainable electricity provider systems, and will be even more important in the near future with the introduction of demand response infrastructures (Alp & Demirkıran, 2023).

From the electricity distribution company's (specifically GDZ Electricity Distribution Company in this study) point of view, accurate demand forecasting is crucial for determining the appropriate size, location, and type of future power plants, as well as the operation and planning of electricity distribution to ensure uninterrupted energy exchange between various electrical areas in the region it covers (Lopez et al., 2019). For instance, maintenance of the system can be scheduled during lower demand periods in an electrical zone by transfering energy from nearby zones. Knowing the periods of lower demand beforehand via forecasting is crucial in this respect.

Another important aspect of short-term load forecasting lies in maintaining the balance between the generated electricity and the consumption. The generated electricity is bought within the day before the consumption day by GDZ company from the energy market. An important strategy for maintaining such balance is forecasting the total demand that will take place within tomorrow's day and purchase accordingly within today's day. At the day of forecasting, since the current day's total demand is not known yet, the electricity provider must develop two-days ahead forecasting strategies that forecasts tomorrow's demand from the time series whose last element is the previous day's consumption.

The deviations from the balance between the purchased electricity and the consumption has both environmental and financial consequences, thus the forecast of two-day ahead demand must be as accurate as possible. For example, in the case of scenario where the purchased electricity is less than the consumption, to ensure grid stability, backup electricity generation systems such as coal-powered plants, must be activated. Such backup systems emit higher CO_2 emissions, and produce electricity at higher costs, negatively affecting both the environment and the financial situation of the energy distributor company. Conversely, in the case of scenario where the purchased electricity is higher than the consumption, the excess of the purchased energy is wasted, again negatively impacting the environment and the financial situation of the distributor.

Regarding the considerations on balancing the demand and generation discussed above, this study aims to make a comparison between ARIMA -a linear forecasting method, (Box et al., 2013) and (Hochreiter & Schmidhuber, 1997) – a nonlinear artificial intelligence method- for two-day ahead forecasting. Importantly, this study uses the actual dataset constructed by the total electricity consumptions of İzmir and Manisa cities in Turkey spanning from January 2020 to December 2022.

The reason for our choices of these methods are that they both are widely used methods in forecasting. ARIMA method is being used for decades. Despite the introduction of several nonlinear methods, especially in the last decade, based on artificial intelligence and machine learning methods, ARIMA still remains an important choice especially in finance sector (such as X-12-ARIMA seasonal adjustment program of the U.S. Census Bureau (Findley et al., 1998). One of the main reasons for choosing ARIMA lies in its explainability(Erden, 2023; Mashadihasanli, 2022). That is, it can reveal the insight behind the forecasted outcomes. This is especially useful for experts that seeks to understand the daily, weekly or yearly patterns, intuitions and reasonings behind the results.

Regarding electricity load forecasting, several studies indicate the superior results of LSTM over ARIMA(Mpawenimana et al., 2020; Siami-Namini et al., 2018). However, it must be noted that such studies usually utilizes ARIMA without preprocessing Data. It must be noted that ARIMA application can be used with several data transforming methods (such as detrending, non-seasonal or seasonal differencing, filtering etc.) to both increase the performance and reveal insights about data. Herein, this study provides a compelling demonstration by showcasing how ARIMA can outperform basic LSTM, when combined with detrending technique. Moreover, we show that ARIMA with detrending technique provides a more robust technique to hyperparameter deviations than ARIMA without detrending and LSTM.

For the comparison of ARIMA and LSTM methods, a design of experiments was conducted for both methods. For ARIMA, the search space is created with the guidance ACF and PACF analyses with and without detrending. Subsequently, the obtained results were interpreted with the help of theoretical predictions. For LSTM, the search space is constructed based on hyperparameters, such as hidden units, optimizer, and the epoch number. An important point that can be emphasized here is that theoretically guided simple ARIMA models can challenge the empirical and computationally intensive LSTM models. Especially, we argue that the explainability capability of arima models can provide insights to experts for making further strategic moves, such as for making buy or sell decisions above or below the forecasted value.

2. CHAPTER: LITERATURE REVIEW

In the study by (Piotrowski et al., 2021), hourly data from a small wind turbine in southern Poland over almost 2 years was collected and utilized in various single and combined formats like LSTM, MLP, SVR, and KNNR to predict 48-hour production data. Their analysis show that LSTM, MLP, SVR, and KNNR were identified as the top-performing approaches in descending order. Additionally, incorporating lagged values of the time series, seasonal patterns, and daily variability indicators notably enhanced prediction accuracy.

(Chapagain et al., 2023) utilized deep artificial intelligence networks to model various scenarios for optimizing accurate electricity demand forecasting. One scenario excluded weekends, using an Artificial Neural Network (ANN) based Feed Forward Neural Network (FNN) with minimal prediction error. The second scenario, incorporating weekends, employed a Recurrent Neural Network (RNN) based Gated Recurrent Unit (GRU). Results favored the first scenario for better forecasting but recommended the second scenario for predicting demands on weekends and holidays.

(Khaled & Farouk Assar, 2022) emphasized the financial importance of estimating the electric load demand, by combining different machine learning algorithms and predicting the day ahead using hourly electric load usage data for Spain in the last quarter of 2018. Problem Temperature regression and load projection model (the model used in Spain) was compared separately with ARIMA, ANN, SVR and LSTM approaches. It was seen that the proposed hybrid model (LSTM&SVR) did not give a better prediction result than the model currently used in Spain.

(Torres et al., 2022) discussed the importance of the increasing demand for electricity day by day and carried out a study to predict the electrical load that will be needed. In this study, 4-hour predictions were made using the coronavirus optimization algorithm (CVOA), a meta-heuristic approach, and the LSTM algorithm, using electrical load data taken every 10 minutes for nine and a half years in Spain. Linear regression, decision tree, random forest and gradient-boosted trees approaches were also applied in the study, and it was seen that LSTM gave the best results.

(Islam & Ahmed, 2022) emphasized that as smart grids increase, the required electricity load increases and it is critical for electricity distribution companies to learn about this demand in advance. In their study, they compared the deep learning methods LSTM, ANN and RNN approaches to estimate short-term electricity demand using 2-year data of a Pakistani electricity distribution company, according to measurement criteria such as MAPE, MSE and RMSE. As a result of the study, it was seen that the LSTM method gave the best prediction result.

(Taylor et al., 2006) compared four different methods to predict the electricity load demand a day ahead, using hourly data from the city of Rio and half-hourly data from the UK. During the study, they did not consider holidays because of the variability of electricity load demand and reduced the possibility of obtaining reliable results, and they worked only on weekdays. In their studies, they compared Double seasonal ARMA modeling, Exponential smoothing for double seasonality, Artificial neural network and principal component analysis (PCA) methods.

Bilgili et al (2022), emphasizing the importance of electrical load demand forecasting in terms of the country's energy planning, used LSTM, ANFIS-SC, ANFIS-FCM and ANFIS-GP using 1460 electrical load demand data of Turkey between January 1, 2016 and December 31, 2019. compared to other approaches. As a result of the study, it was seen that the ANFIS-SC hybrid model gave the best result among ANFIS approaches, while the LSTM model gave the best prediction result among all models.

3. CHAPTER: DATASET

This research centers on forecasting the total electricity consumption in two cities, İzmir and Manisa, utilizing actual time series data obtained from the Gdz Electric Distribution company. The dataset comprises hourly energy consumption data post-COVID (06 January 2020, Monday) until 25 December 2022, Sunday (Figure 1).

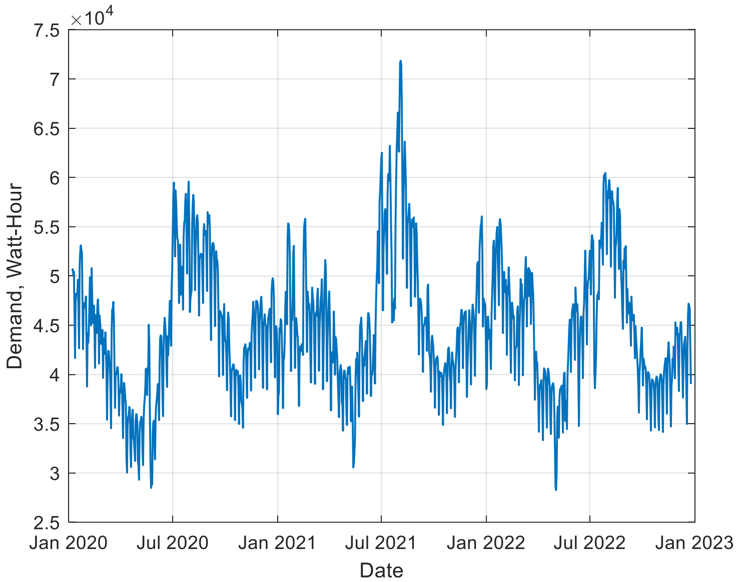

Figure 3.1. Demand (Watt-hour) vs. date of data used.

There are 155 weeks in the dataset. Last 30 weeks are chosen for testing which amounts to 19.35% of data whereas training part constitutes 80.65% (Figure 2).

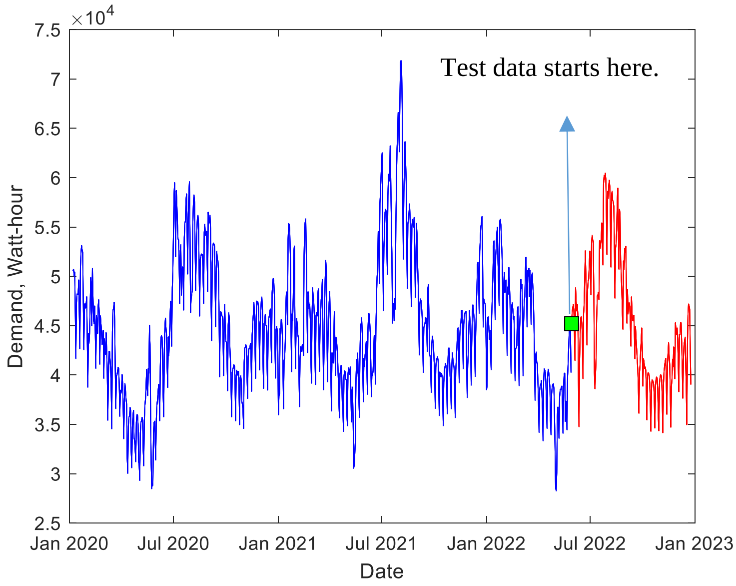

Figure 3.2. Train and test parts of data

3.1. Data Description

3.1.1. Augmented Dickey Fuller (ADF) Test

The Augmented Dickey-Fuller (ADF) test is used to determine whether a time series is stationary based on the presence of a unit root in the examined time series. Interpretations based on the p-values of null hypothesis of a unit root (i.e., unit root exists) in the time series obtained from the ADF test as follows. If the p-value is less than 5% confidence interval, then there is no unit root in the series (the series is stationary). If the p-value is higher than 5%, there is a unit root in the series (the series is non-stationary). Another crucial point is that if it is determined that the time series is non-stationary (there is a unit root in the series), differencing starting from the first order is necessary.

When ADF test (we use adftest function available in Econometrics Toolbox of MATLAB) is applied to the consumption series shown in Figure 1, it is seen that data is non-stationary with p-value equal to 20% (more than 5%). Thus, the first step is to make data stationary by differencing via $\Delta y_t = y_t - y_{t-1}$ (non-seasonal first order differencing). Applying first order differencing makes data stationary as can be seen from Fig.3. When we apply ADF test to the differenced data, ADF test also indicates that it is stationary with p-value equal to 0.001, less than 5%.

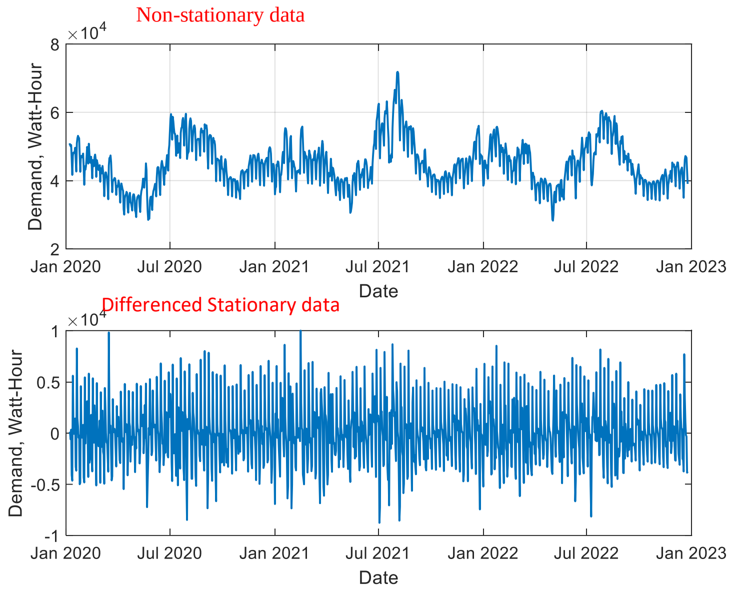

Figure 3.3. Normal time series data(upper) vs. differenced

3.2. Periodicity

Herein to see whether data has some periodicity, we apply sample autocorrelation functions in Fig. 4. Figure 4 indicates that there is a periodicity in the data. Therefore, when designing an ARIMA model, we will take this into account. An effective way to deal with periodicity is detrending the time series by decomposing it as the sum of a deterministic periodic trend using seasonal indicators and a stochastic process as in Equation (1).

$$y_{tr}(k) = \mu_{tr}(k) + s_{tr}(k) \quad (1)$$

Where $\mu_{tr}(k)$ is a deterministic periodic trend (i..e seasonal indicators) and $s_{tr}(k)$ is a stochastic process.

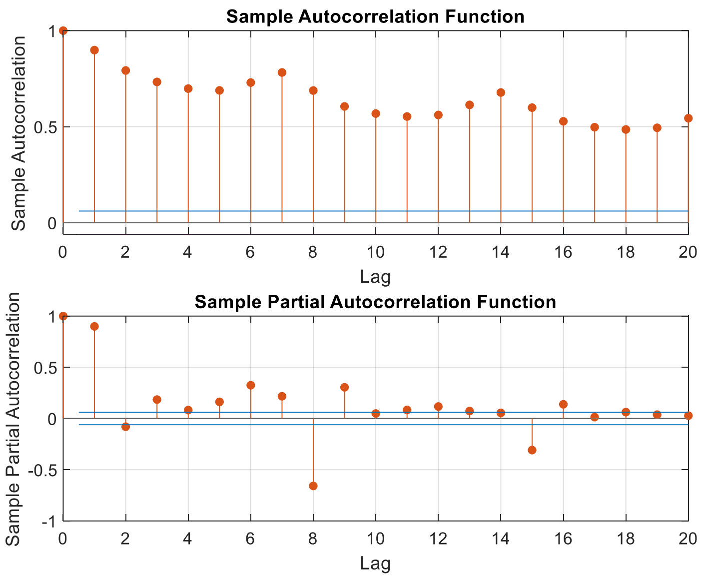

Figure 3.4. Autocorrelation nad Partial autocorrelation function of data.

We derive the seasonal indicators $\mu_{tr}(k)$ of the data through the following procedure: i) generate indicator (dummy) variables representing each day from 1 to 7. The initial indicator equals one for observations on Mondays and zero otherwise. Seven indicator variables are generated in total, each corresponding to a day of the week. Then, we conduct a regression of the detrended series against these seasonal indicators.

If we apply autocorrelation and partial autocorrelation to the detrended training data, $s_{tr}(k)$, we observe that periodicity is no more in the picture (Figure 5). Moreover, $s_{tr}(k)$ is stationary according to adftest with p-value being 1e-3. Thus, from now on since $s_{tr}(k)$ doesnot contain any periodicies, we can drop seasonality considerations.

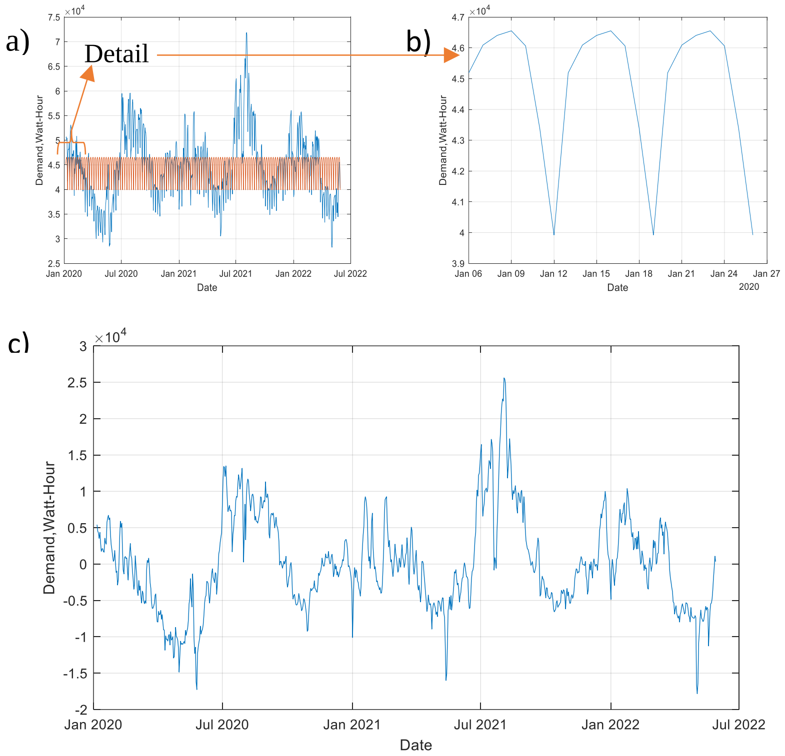

Figure 3.5. Training Demand data and its seasonal indicators. b) Seasonal indicators. c) Detrended training data: i.e. $s_{tr}(k) = y_{tr}(k) - \mu_{tr}(k)$.

A gradual decay of ACF and distinct cutoff PACF (Fig. 6) indicates that $s_{tr}(k)$ is an AR(3) process, the first and second lags being significant. Moreover, decay is slow, indicating an integration process, that is $D \geq 1$. In one-step ahead forecasting we would have used the non-zero lag indices of 1, 2 and 3. For two-step ahead (two-days ahead) forecasting we use the rolling ARIMA, that is we feed the one-step ahead forecasted data for the next one-step forecasting resulting in two-step ahead forecasting.

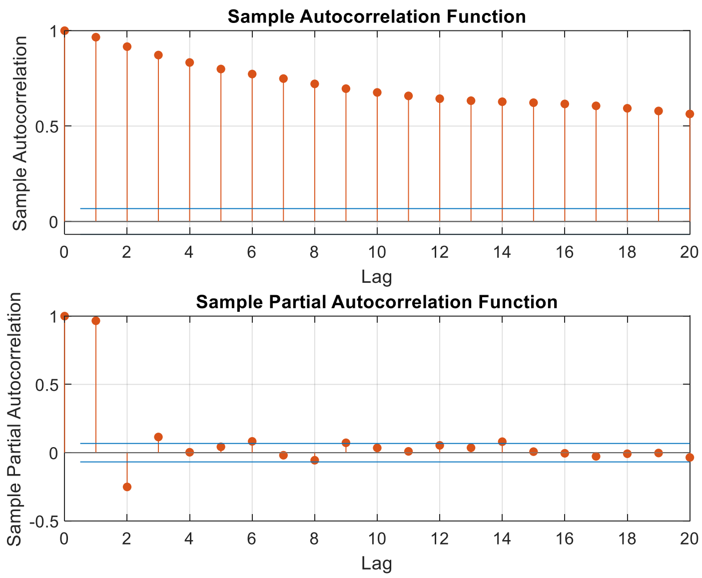

Figure 3.6. Autocorrelation and Partial Autocorrelation function of detrended training data, s_{tr} *(k)*, shows that this is AR(3) process.

4. CHAPTER: METHODOLOGY

4.1. ARIMA

The recursive form of the ARIMA(p,d,q) forecasting model for any time series x_t is as follows :

$$x_t = x_{t-1} + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Here, p denotes the autoregressive (AR) order, d represents the number of regular differences, and q indicates the moving average (MA) order. $\phi \epsilon \mathbb{R}^p$ signifies the weight parameters of the AR term, while θ $\epsilon \mathbb{R}^p$ represents the weight parameters of the MA term, and $\varepsilon(t) = x_t - \widehat{x_t}$ consists of the prediction errors of the series. For the case of $d = 0$ in the ARIMA(p,d,q) model, it corresponds to the model of the actual data's time series, whereas in other cases, it corresponds to the model of the new time series created by taking the d level differences of the actual data. The ARIMA(p,d,q) method requires selecting its parameters p, d, and q according to the characteristics of the time series. The degree of d represents an indicator of the stationarity of the time series. For stationary time series, $d = 0$ is chosen. One way to check whether a series is stationary is through the Dickey-Fuller Test.

4.2. LSTM

LSTM is a type of Recurrent Neural Network (RNN). One key difference in training RNNs compared to feedforward neural networks (such as MultiLayer Perceptron) is their training through backpropagation algorithms over time. One of the fundamental issues with RNNs is the vanishing gradient problem, where the gradient value intended to be obtained through backpropagation over time gets diminished after a certain depth due to the multiplication of partial derivatives according to the chain rule over the series or a specified range within the series. The LSTM structure provides a solution to this vanishing gradient problem at the mentioned depth in time. Due to the LSTM's architecture, during the computation of gradients via the backpropagation algorithm, along with the partial derivative multiplication terms (short-term), there are also terms involving the summation of partial derivatives (long-term), thereby mitigating the "vanishing gradient problem." Various LSTM or RNN variations exist that aggregate partial derivatives, preserving long-term information (Greff et al., 2017). In this report, the LSTM structure based on the article published by (Hochreiter & Schmidhuber, 1997) is used.

5. CHAPTER: COMPUTATIONAL EXPERIMENTS & RESULTS

5.1. Performances of ARIMA Strategies

Since according to Fig. 6, the detrended data is AR(3) process, we design a search space for the AR lags around 3 starting from 1 to 8. Moreover, we also search for the contribution of MA lags, as indicated in Table 1. We repeat the experiments for ARIMA with and without detrending method. Figure 7 below shows the Mean Absolute Percentage Error (MAPE) performances calculated according to Equation (2), where n correspond to the the length of the data, t is the index, A_t is the actual data, and F_t is the forecast data.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|A_t - F_t|}{A_t} \quad (2)$$

Table 5.1. Search Space for ARIMA strategies

Index	ARLags	MALags
1	{[NaN]}	{[1]}
2	{[NaN]}	{[1 2]}
3	{[NaN]}	{[1 2 3]}
4	{[NaN]}	{[1 2 3 4]}
5	{[NaN]}	{[1 2 3 4 5]}
:	:	:
76	{[1 2 3 4 5 6 7 8]}	{[1 2 3 4]}
77	{[1 2 3 4 5 6 7 8]}	{[1 2 3 4 5]}
78	{[1 2 3 4 5 6 7 8]}	{[1 2 3 4 5 6]}
79	{[1 2 3 4 5 6 7 8]}	{[1 2 3 4 5 6 7]}
80	{[1 2 3 4 5 6 7 8]}	{[1 2 3 4 5 6 7 8]}

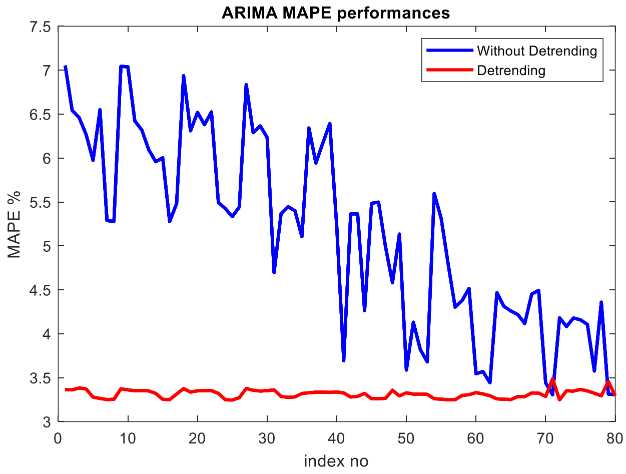

Figure 5.1. ARIMA MAPE performances with and without detrending

5.2. Performances of LSTM Strategies

We ran a number of experiments to find the best architecture and hyperparameters for our LSTM model. First, we used a single 50-unit LSTM layer with adam (adaptive moment estimation) optimizer and 500 training epochs with a learning rate of 0.005. Our model's Mean Absolute Percentage Error (MAPE) in this scenario was 4.80%. We experimented with changing the number of LSTM units, the number of training epochs, and the optimizer selection in order to further investigate different scenarios. We experimented with architectures containing 50-100-200 units and trained with 500-1000-2000 training epochs. Additionally, we compared the performance of the adam and Stochastic Gradient Descent (SGD) optimizers. Following several tests, we discovered that the top-performing setup included the adam optimizer and a single LSTM layer with 100 units and 500 training epochs. Our model was able to accurately recognize the underlying patterns in the data thanks to the selection of 500 epochs and 100 units when using the adam optimizer with MAPE of 3.4%.

Table 5.2. Search Space for LSTM

Optimizer	Learning Rate	Hidden Units	Epochs	MAPE (%)
Adam Optimizer	0.005	50	500	4.80
		50	1000	4.75
		50	2000	4.59
		100	500	**3.40**
		100	1000	5.78
		100	2000	5.30
		200	500	4.63
		200	1000	4.03
		200	2000	6.45
SGD	0.005	50	500	6.44
		50	1000	6.30
		50	2000	6.35
		100	500	6.41
		100	1000	6.37
		100	2000	6.34
		200	500	6.40
		200	1000	6.41
		200	2000	6.57

6. CHAPTER: DISCUSSION

The below compares the MAPE performances of ARIMA with detrending, ARIMA without detrending, and LSTM methods using boxplot (Figure 8). It can be seen that ARIMA with detrending provides the superior performances in average and in terms of the best performance than other methods. Moreover, the performances of ARIMA with detrending is relatively stable, meaning that it is more robust to hyperparameter variations in training.

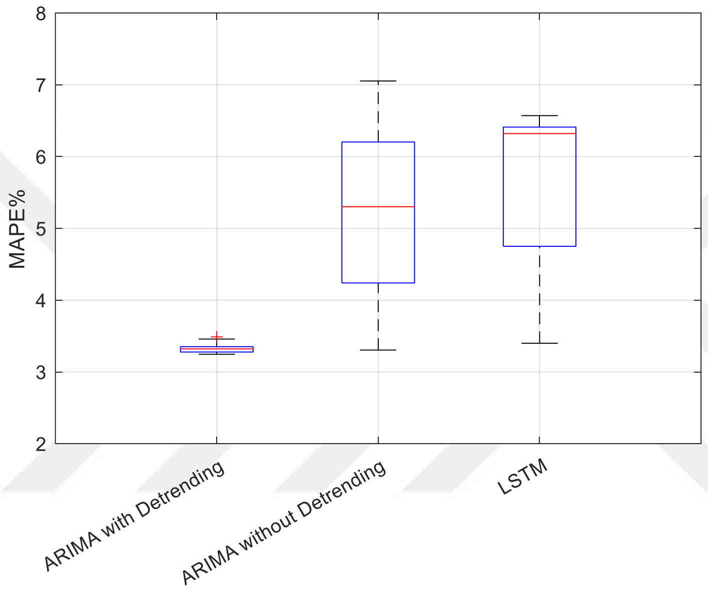

Figure 6.1. Boxplot Comparison of LSTM with ARIMA.

ARIMA without detrending challenges LSTM and has superior performance in average, and in terms of best performance. Although several studies show that LSTM is superior to ARIMA in general, this study showcases a different result. WE think there are two main reasons for these results:

1. Researchers that compare LSTM and ARIMA method should ensure thorough exploration of the space involving AR lags and MA lags of ARIMA. The referenced studies lack this comprehensive exploration, which is insufficient for such comparative studies.

2. The consumption dataset from the İzmir-Manisa region exhibits some patterns, as demonstrated in Figure 4 through ACF and PACF analyses. Such analysis of data before using ARIMA is valueable for experts.

Nonlinear Methods based on artificial intelligence, such as LSTM in this study, demand high computational resources to capture nonlinear relationships within data and can also integrate additional variables, such as temperature. However, this dataset only contains one-dimensional consumption data. If temperature data were incorporated, the outcomes might have favored LSTM.

Regarding the trend in data, the seasonal indicators depicted in Figure 5 provide valuable insights to experts, such as consumption patterns. It can be seen how the consumption decreases on Sunday and increases starting from Monday making a peak on Thursday.

7. CHAPTER: CONCLUSIONS AND FUTURE WORK

Pre-analysis of data, such as the ACF and PACF analyses, serve as crucial tools for decision-making by experts. Moreover, post analysis of ARIMA results, such as examining residual statistics and selecting information criteria, offer valuable insights to theoretically-inclined experts. Although a method can provide better forecasts than experts, it must be noted that forecasting methods of electricity consumption cannot incorporate everything, such as wars, natural diseases, economic collapses. The ultimate financial decisions regarding energy purchase are taken by experts. And we advise that such experts A expert should leverage pre and post-analysis of ARIMA results to gain a deeper understanding of its predictive insights. Such insights cannot be obtained from the study of nonlinear models. Even if one can use and obtain better results with artificial intelligence, the experts can better appreciate these results after applying and analyzing ARIMA. This is especially the case for the data that incorporates obvious patterns as in this study.

8. REFERENCES

Alp, M., & Demirkıran, G. (2023). A Novel Clustering-based Forecast Framework: The Clusters with Competing Configurations Approach. *Academic Platform Journal of Engineering and Smart Systems*, *11*(3), 151–162. https://doi.org/10.21541/apjess.1266610

Bilgili, M., Arslan, N., Şekertekin, A., & Yaşar, A. (2022). Application of long short-term memory (LSTM) neural network based on deep learning for electricity energy consumption forecasting. *Turkish Journal of Electrical Engineering and Computer Sciences*, *30*(1), 140–157. https://doi.org/10.3906/elk-2011-14

Box, G., Jenkins, G. M., C. Reinsel, G., & M. Ljung, G. (2013). Box and Jenkins: Time Series Analysis, Forecasting and Control. In *A Very British Affair* (pp. 161–215). Palgrave Macmillan UK. https://doi.org/10.1057/9781137291264_6

Chapagain, K., Gurung, S., Kulthanavit, P., & Kittipiyakul, S. (2023). Short-Term Electricity Demand Forecasting Using Deep Neural Networks: An Analysis for Thai Data. *Applied System Innovation*, *6*(6), 100. https://doi.org/10.3390/asi6060100

Çobanoğlu, A., Demirkıran, G., & Güneş, M. (2021). İzmir İlinde Elektrikli Kara Araçları için Güneş Enerjisi Destekli Bir Şarj İstasyonun Tasarlanması. *European Journal of Science and Technology*. https://doi.org/10.31590/ejosat.777874

Das, H. S., Rahman, M. M., Li, S., & Tan, C. W. (2020). Electric vehicles standards, charging infrastructure, and impact on grid integration: A technological review. In *Renewable and Sustainable Energy Reviews* (Vol. 120). Elsevier Ltd. https://doi.org/10.1016/j.rser.2019.109618

Erden, C. (2023). Derin Öğrenme ve ARIMA Yöntemlerinin Tahmin Performanslarının Kıyaslanması: Bir Borsa İstanbul Hissesi Örneği. *Yönetim ve Ekonomi Dergisi*, *30*(3), 419–438. https://doi.org/10.18657/yonveek.1208807

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., & Chen, B. C. (1998). New capabilities and methods of the x-12-arima seasonal-adjustment program. *Journal of*

Business and Economic Statistics, *16*(2), 127–152. https://doi.org/10.1080/07350015.1998.10524743

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Islam, B. ul, & Ahmed, S. F. (2022). Short-Term Electrical Load Demand Forecasting Based on LSTM and RNN Deep Neural Networks. *Mathematical Problems in Engineering*, *2022*, 1–10. https://doi.org/10.1155/2022/2316474

Justin, F., Peter, G., Stonier, A. A., & Ganji, V. (2022). Power Quality Improvement for Vehicle-to-Grid and Grid-to-Vehicle Technology in a Microgrid. *International Transactions on Electrical Energy Systems*, *2022*. https://doi.org/10.1155/2022/2409188

Assar, A. K. A. F. (2022). *Application of a Hybrid Machine Learning model on short term electricty demand prediction* [M.S. - Master of Science]. Middle East Technical University.

Lopez, J. C., Rider, M. J., & Wu, Q. (2019). Parsimonious Short-Term Load Forecasting for Optimal Operation Planning of Electrical Distribution Systems. *IEEE Transactions on Power Systems*, *34*(2), 1427–1437. https://doi.org/10.1109/TPWRS.2018.2872388

Mashadihasanli, T. (2022). Stock Market Price Forecasting Using the Arima Model: an Application to Istanbul, Turkiye. *Journal of Economic Policy Researches / İktisat Politikası Araştırmaları Dergisi*, *9*(2), 439–454. https://doi.org/10.26650/jepr1056771

Mpawenimana, I., Pegatoquet, A., Roy, V., Rodriguez, L., & Belleudy, C. (2020). A comparative study of LSTM and ARIMA for energy load prediction with enhanced data preprocessing. *2020 IEEE Sensors Applications Symposium (SAS)*, 1–6. https://doi.org/10.1109/SAS48726.2020.9220021

Piotrowski, P., Kopyt, M., Baczyński, D., Robak, S., & Gulczyński, T. (2021). Hybrid and Ensemble Methods of Two Days Ahead Forecasts of Electric Energy Production in a Small Wind Turbine. *Energies*, *14*(5), 1225. https://doi.org/10.3390/en14051225

Rahman, S., Khan, I. A., Khan, A. A., Mallik, A., & Nadeem, M. F. (2022). Comprehensive review & impact analysis of integrating projected electric vehicle charging load to the existing low voltage distribution system. *Renewable and Sustainable Energy Reviews*, *153*. https://doi.org/10.1016/j.rser.2021.111756

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, 1394–1401. https://doi.org/10.1109/ICMLA.2018.00227

Taylor, J. W., de Menezes, L. M., & McSharry, P. E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, *22*(1), 1–16. https://doi.org/10.1016/j.ijforecast.2005.06.006

Tesch da Silva, F. S., da Costa, C. A., Paredes Crovato, C. D., & da Rosa Righi, R. (2020). Looking at energy through the lens of Industry 4.0: A systematic literature review of concerns and challenges. In *Computers and Industrial Engineering* (Vol. 143). Elsevier Ltd. https://doi.org/10.1016/j.cie.2020.106426

Torres, J. F., Martínez-Álvarez, F., & Troncoso, A. (2022). A deep LSTM network for the Spanish electricity consumption forecasting. *Neural Computing and Applications*, *34*(13), 10533–10545. https://doi.org/10.1007/s00521-021-06773-2