All Sciences Proceedings
http://as-proceeding.com/

3rd International Conference on Innovative Academic Studies

September 26-28, 2023 : Konya, Turkey

ICIAS
https://www.icias.net/

Load Forecasting for Trakya University

Cihat Cagdas Uydur*[1], Firat Akin [2], Sureyya Guclu [1] and Bekir Dursun [1]

[1] *Technical Sciences Vocational School, Trakya University, Turkey*
[2] *Department of Electrical Engineering, Yildiz Technical University, Turkey*

**ccagdasuydur@trakya.edu.tr*

Abstract – Electricity demand prediction involves analyzing historical and current electricity usage, identifying factors influencing load estimation, and employing various methods and algorithms to project future demand. The primary objective of demand forecasting is to deliver cost-effective, reliable, and high-quality energy to consumers. To achieve this, electricity demand forecasting is segmented into different time periods, encompassing production, transmission, and distribution levels. Time-based classifications categorize load forecasting into very short-term, short-term, mid-term, and long-term forecasts, each serving specific planning purposes. Forecasts extending beyond a year are considered long-term, those spanning one week to one year are categorized as mid-term, forecasts covering one hour to one week are short-term, and estimates less than an hour fall under very short-term forecasts. The factors utilized to estimate electrical load vary based on the forecasting duration. Long-term load forecasting primarily relies on parameters like gross national product and population, while short-term load forecasting predominantly utilizes meteorological data. Electricity consumption fluctuates on an hourly basis throughout the day. Analyzing these consumption variations and forecasting power plant activity initiates with short-term load forecasting. An additional role of short-term load forecasting is optimizing energy usage economically. Addressing escalating energy requirements and ensuring sustainable energy supply commences with accurate load forecasting. Consequently, load forecasting applications are progressively gaining significance to meet the evolving energy demands. In this study, a short-term load forecast for Trakya University have been made. During the estimation and analysis processes, error rates have been revealed by using decision trees (DT), artificial neural networks (ANN) and least squares method (LSM). It is seen that the ANN algorithm comes to the forefront in predicting behavioral characteristics. The performance of the DT algorithm in behavior analysis is not unsuccessful. However, there is a higher margin of error in the prediction process using the DT algorithm compared to the ANN application. It has been revealed that LSM cannot be used as a load forecasting algorithm.

Keywords – Load Forecasting, Decision Trees, Artificial Neural Networks, Least Squares Method

I. INTRODUCTION

In today's economic conditions, electrical energy system planning is becoming more and more important in order to provide reliable, economical, high quality and uninterrupted supply of electrical energy with limited energy generation resources [1]. Accuracy of load forecast planning is of vital importance since load forecast planning affects all other planning, and an error here may cause errors in other planning as well. If the load forecast is underestimated from the actual real load, it will cause power outages and poor quality energy to prevent system collapse, and if it is overestimated from the actual real load, it will cause outages, failures and uneconomical electricity generation to power generation plants that do not operate at the appropriate capacity [2]. For this reason, the electricity load forecast should be estimated as close as possible to the actual real load.

It is known that load forecasting studies using statistical methods other than traditional methods have been conducted since 1962 [3]. In his study, New made long term load forecasting using time series analysis. G. Gross and F.D. Galiana in their 1987 paper "Short-term load forecasting" reviewed the important role of short-term forecasting in the online scheduling and security functions of the energy management system (EMS) and then discussed the different factors affecting the nature and behavior of load [4]. In the study by Papalexopoulos and Hesterberg using Pacific Gas and Electricity load data, the proposed hybrid model is compared to the classical models. The linear regression based hybrid model was found to provide more consistent forecasting results [5]. Bakırtzıs et al. applied a multilayer artificial neural network (ANN) model with 63 inputs and 24 outputs for daily load forecasting of the Greek power system. Previous load data and temperature data were used as independent variables as input parameters, while the next day was selected as the dependent variable [6].

In a study conducted by Liu using load data of a consumption region in the United States, fuzzy logic, neural networks and autoregressive model were compared and very short-term forecasting was performed. It was found that the model using fuzzy logic and neural network models gave higher accuracy [7]. In the long term load forecasting study conducted by Gürsoy in 2000, regression, time series and ANN were used to make long term forecasts for Antalya Kepez region and Çukurova Elektrik A.Ş (ÇEAŞ) region until 2020 [8]. In the study by Hippert et al. short-term load forecasting studies published between 1991 and 1999 using neural networks were analyzed and the effect of the input parameters used and the differences in the number of classes were examined [9]. In the study by Yalçınöz et al. electricity load forecasting of Niğde region was performed on a monthly basis. ANN and moving averages (MA) methods were used in the study. While the moving averages method gave more significant results for some months, ANN gave more consistent results for some months. When the overall error average is analyzed, it is concluded that the ANN algorithm is superior to the moving averages method [10].

In a study conducted by Huang in 2005, hourly short-term load forecasting was performed with an autoregressive moving average process (ARMA) model. Particle swarm optimization method was used for parameter selection and determination [11]. In the short-term load forecasting study conducted by Ceylan to predict the next day's load using the data of Ankara Gölbaşı region for the years 2002 and 2003, feed-forward multilayer perseptron network ANN and regression techniques were used as methodology. As a result of the study, ANN was found to give better and more consistent results than the regression model [12]. Autoregressive integrated moving average (ARIMA) and generalized regression ANN and SVM with simulated annealing in a study by Pai and Hong for forecasting electricity load in Taiwan, the SVM methodology with simulated annealing gave higher accuracy than the other two methods [13]. In the study by Demirel, Turkey's annual long term load data were estimated. GNP, energy consumed, energy produced, population and installed capacity are used as input parameters. Adaptive artificial neural network fuzzy logic inference system (ANFIS), ARMA, ANN and regression models were used as forecasting methods and the results were compared. It was concluded that the ANFIS method gave the best forecasting result [14].

In the study conducted by Akman, the hourly load data, maximum and minimum temperature data of the Central Anatolia Face Dispatch Operation Directorate between 2009-2015 were used as input parameters. As output, daily forecast was made and tested by selecting 10 different day types. ANN and hybrid methodologies including DT and ANN were used as forecasting methods. It is concluded that the hybrid model gives results closer to reality [15]. In the study by Khan, short-term load forecasting is performed for Faisalabad Electricity Supply Company. The input parameters are determined by Relief and Correlation technique. The forward propagation ANN method was used as the forecasting methodology [16]. Butekin performed load forecasting using ANN for Y-block of Yasar University campus buildings. In order to make the prediction results more meaningful, day types were separated, and load data were processed with some statistical methods. As a result, it was observed that the accuracy rate of the processed data divided into day types was higher [17]. Household data of Sakarya region were used by Doruk. For the forecasting method, the Dynamic Linear Model, which is a Bayesian approach of State-Space models, and the ARIMA model were compared and the performance of the process was demonstrated. It was revealed that the accuracy rate of the model created using seasonal factor, linear growth and ARMA was the highest [18].

In this study, a demand forecasting and analysis application is made for Trakya University considering regional conditions such as weather conditions and usage habits. The electricity consumption characteristics of Trakya University were determined. Along with the forecasting and analysis studies, the least squares method (LSM), decision trees (DT) and artificial neural networks (ANN) algorithms are evaluated comparatively and the algorithm with the best performance is determined. It was possible to reduce the electricity consumption of Trakya University.

II. MATERIALS AND METHOD

Within the scope of the study, the Department of Construction Affairs was contacted and requested information about the past records of electrical energy consumption values. However, it was determined that there were no detailed electrical energy consumption records for the past years. However, it was determined that electricity bills were listed and electricity consumption values were recorded.

After discussions on the existence of the OSOS (remote reading of electricity meters) system, it was agreed that Trakya University electricity consumption values could be obtained from the TREPAŞ-OSOSOS system. As a result, it was decided to use the 15-minute measurement values obtained from the TREPAŞ-OSOS remote reading system for monthly electricity consumption values. Electrical energy consumption values are periodically provided on a monthly basis in consultation with the Department of Construction Affairs of Trakya University.

Edirne province ambient conditions and weather data are obtained on a daily and hourly basis using the General Directorate of Meteorology information system (MEVBIS) and meteoblue databases. Fig. 1 shows the ambient conditions and weather data for October 2022 [19].

To be used in forecasting studies, Trakya University electricity consumption values were analyzed and reorganized on an hourly basis. In addition, they are categorized for each month depending on the days. Sample data for October 2022, representing the processed electrical energy consumption values, are shared in Table 1.

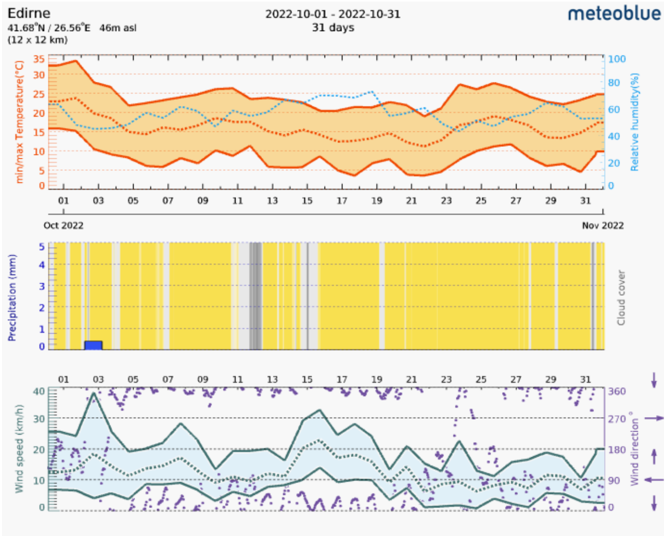

Fig. 1 Ambient Conditions and Weather Data for Edirne Province in October 2022 [19]

When the data in Table 1 are analyzed, it is seen that the days are classified as weekdays and weekends. Weekend days are marked with orange color while weekdays are colored with dark numbers.

In addition, the hours of the day are defined as day - peak - night in the electricity tariff. Daytime hours are highlighted in white, peak hours in dark blue and nighttime hours in dark gray.

Finally, the total electricity consumed per day is calculated in the bottom row and the total monthly value of one hour of consumption is calculated in the rightmost column.

As a result of the preliminary studies, the conditions determined for each estimation algorithm are presented in Table 2. The parameters indicated in the table were selected among the variations that gave the best results after 5 different trials.

Table 2. Optimum Conditions for Forecasting Algorithms

Method	**LSM**	**DT**	**ANN**
Program	MATLAB	Python	MATLAB
Type	LSM	Gradient Boosting	Levenberg–Marquardt
Factor	Poly: 6th	Tree: 500	Hidden Layer: 12 Hidden Layer: 3

Table 1. Trakya University October 2022 Electric Energy Hourly Consumption Values (kWh)

OCTOBER		1	2	3	4	5	6	7	8	9
00:00	01:00	1.656,94	1.578,86	1.482,92	1.576,35	1.636,49	1.591,35	1.595,41	1.569,28	1.474,72
01:00	02:00	1.617,68	1.603,14	1.455,49	1.585,12	1.621,96	1.579,37	1.630,75	1.555,52	1.507,16
02:00	03:00	1.556,83	1.543,23	1.414,87	1.568,42	1.576,58	1.559,52	1.576,92	1.580,10	1.507,28
03:00	04:00	1.536,29	1.491,27	1.368,94	1.547,94	1.556,23	1.532,45	1.566,67	1.553,53	1.485,51
04:00	05:00	1.512,54	1.428,03	1.426,70	1.556,07	1.526,11	1.521,11	1.587,84	1.532,23	1.472,94
05:00	06:00	1.478,38	1.423,36	1.391,46	1.490,88	1.523,46	1.494,10	1.526,63	1.510,14	1.454,06
06:00	07:00	1.444,83	1.378,30	1.373,92	1.459,55	1.480,08	1.482,35	1.508,66	1.463,08	1.428,89
07:00	08:00	1.449,25	1.402,24	1.454,15	1.568,01	1.625,70	1.604,97	1.591,00	1.520,50	1.469,77
08:00	09:00	1.477,24	1.478,31	1.763,94	1.826,77	1.891,01	1.875,54	1.879,19	1.525,53	1.507,34
09:00	10:00	1.535,13	1.566,70	2.103,09	2.234,36	2.254,38	2.241,89	2.232,29	1.507,47	1.502,68
10:00	11:00	1.539,85	1.611,35	2.255,17	2.355,43	2.409,19	2.424,44	2.305,95	1.459,20	1.491,58
11:00	12:00	1.603,75	1.621,91	2.350,26	2.418,38	2.386,87	2.482,96	2.329,92	1.478,65	1.506,44
12:00	13:00	1.661,52	1.589,77	2.326,18	2.510,25	2.438,64	2.423,44	2.303,45	1.542,59	1.512,88
13:00	14:00	1.799,65	1.642,01	2.386,74	2.464,47	2.432,10	2.455,14	2.300,72	1.539,90	1.530,70
14:00	15:00	1.853,60	1.696,69	2.415,39	2.467,69	2.430,07	2.462,20	2.287,65	1.513,28	1.519,69
15:00	16:00	1.874,40	1.688,37	2.414,66	2.384,14	2.375,28	2.437,75	2.294,52	1.534,11	1.527,54
16:00	17:00	1.876,65	1.652,79	2.255,92	2.270,39	2.222,40	2.224,78	2.081,64	1.533,57	1.510,55
17:00	18:00	1.826,64	1.601,53	1.979,16	2.015,48	2.003,45	1.929,20	1.872,83	1.500,97	1.519,92
18:00	19:00	1.826,73	1.562,35	1.821,26	1.815,13	1.830,19	1.825,60	1.726,87	1.541,40	1.533,53
19:00	20:00	1.804,81	1.580,27	1.792,18	1.785,74	1.803,05	1.755,76	1.713,15	1.554,18	1.537,82
20:00	21:00	1.723,49	1.573,98	1.779,86	1.772,03	1.781,81	1.720,98	1.642,71	1.557,67	1.557,14
21:00	22:00	1.703,77	1.544,96	1.662,36	1.707,70	1.649,95	1.694,29	1.595,78	1.521,19	1.514,57
22:00	23:00	1.632,06	1.535,61	1.621,17	1.663,36	1.638,89	1.646,78	1.569,19	1.486,81	1.520,88
23:00	00:00	1.589,30	1.508,67	1.624,95	1.620,59	1.598,25	1.640,09	1.534,19	1.500,05	1.522,07
	TOP	**39.581,32**	**37.303,68**	**43.920,73**	**45.664,26**	**45.692,14**	**45.606,07**	**44.253,94**	**36.580,96**	**36.115,65**

Analyzing the data, we observe fluctuations in electricity consumption throughout the day. The consumption tends to peak during specific hours, typically during midday and early evening, and drops during the night and early morning.

The highest daily consumption is on the 5th of October, reaching 45,692.14 megawatt-hours, while the lowest daily consumption is on the 9th of October, amounting to 36,115.65 megawatt-hours.

Understanding these consumption patterns is crucial for effective energy management and distribution, enabling appropriate resource allocation and planning to meet the energy demands efficiently. Furthermore, such data can aid in developing strategies to optimize energy usage and enhance the sustainability of the power grid.

III. Forecasting and Analysis

In this study, 5 different models were used for forecasting. For each model, weekdays are separated into Monday, Tuesday, Wednesday, Wednesday, Thursday, Thursday, Friday, weekdays and weekends and a dummy variable is used by adding an independent variable.

In Model-1, a single training and test data set was used by using the distinction of weekdays as a dummy variable in the input parameters. Weekday models are defined as 0, while weekend models are assigned a dummy variable of 1. In Model 1, humidity, temperature and consumption load one hour ago are used as independent variables, while consumption load at time t is estimated.

In Model 2, the load at time t is estimated using humidity, temperature and the load one day ago as the independent variable set. In this model, the effect of using one hour ago load and 24 hours ago load on the forecasting result is compared by using a dummy variable.

In Model-3, as in Model-1, weekdays were separated. Humidity, temperature, previous hour's load and 24 hours ago's load were selected as independent variables or in other words as input parameters and the electric load at that hour was estimated.

In Model-4, the distinction between weekdays and weekends was made using a dummy variable and the estimation process was performed using a single training and test data. In Model-4, the humidity parameter was removed from the independent variables. Temperature, the previous electricity consumption value and the electricity consumption value one day, i.e. 24 hours ago were used as independent variables.

In Model-5, weekday-weekend separation is provided with dummy variables. In the forecasting algorithms used in Model-5, meteorological data were not used for the load to be forecast. Season, one hour ago load parameter and 24 hours ago load parameter were used as independent variables and forecasts were made. The purpose of this model is to investigate the effect of the forecasting process performed only with the previous load parameters on the forecast accuracy when meteorological data are removed from the independent variables.

As a result of the analysis, the result data of all algorithms were transferred to a separate file and the success rate of the methods was compared with the mean absolute percentage error (MAPE). The formula for MAPE is given in Equation 1.

$$MAPE = \frac{100}{n}\sum_{j=1}^{n}\frac{|A_j - P_j|}{|A_j|} \qquad (1)$$

In Equation 1, A_j is the measured value, P_j is the predicted value and n is the number of data. The MAPE statistic eliminates the disadvantages that may arise when comparing models with different unit values.

Comparative prediction results for all models are presented in Fig. 2 to Fig. 6.

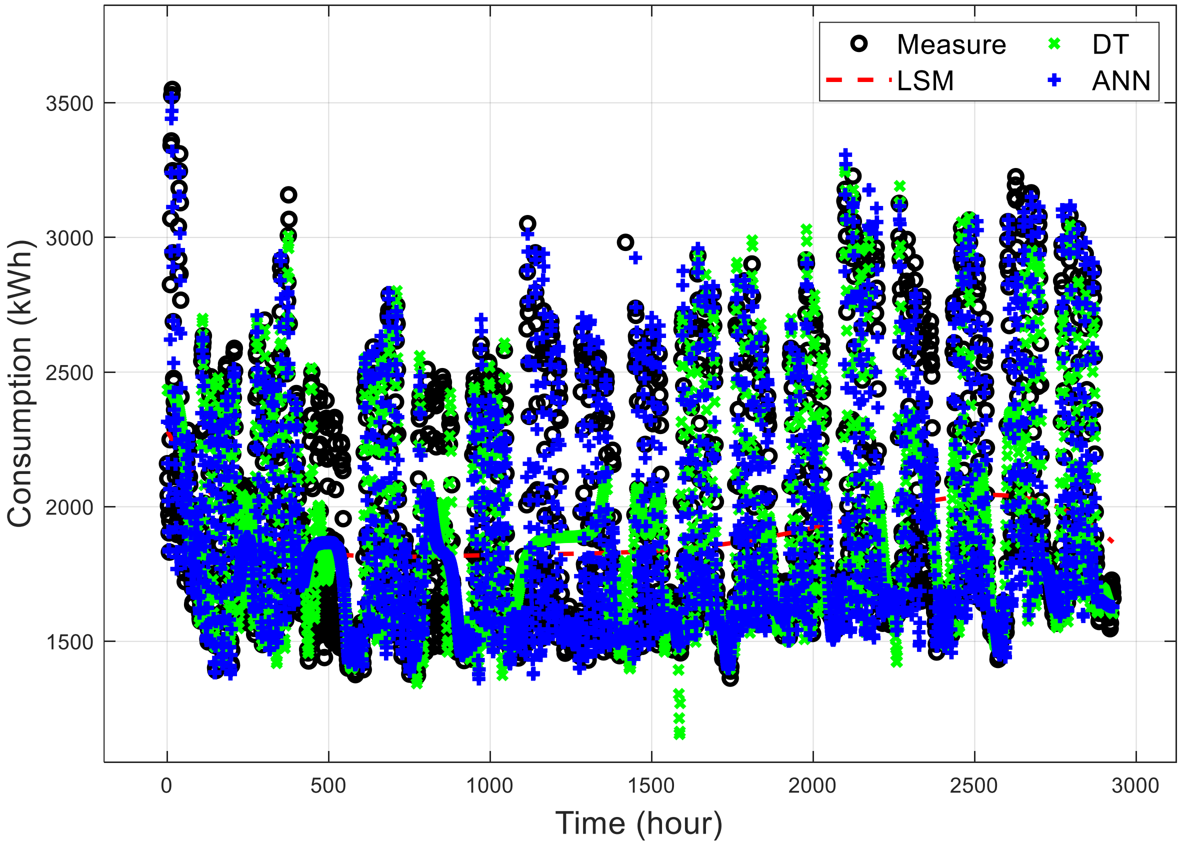

Fig.2 Comparison of prediction results of Model-1

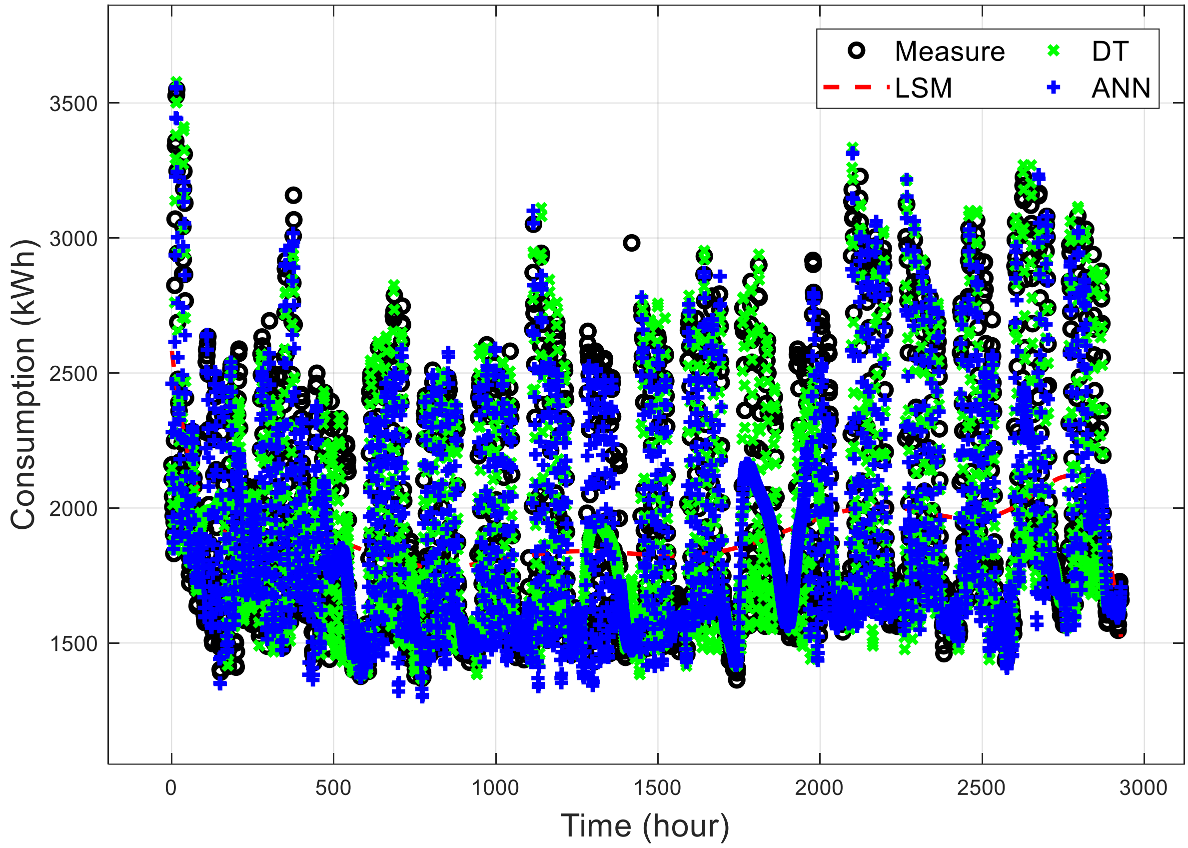

Fig.3 Comparison of prediction results of Model-2

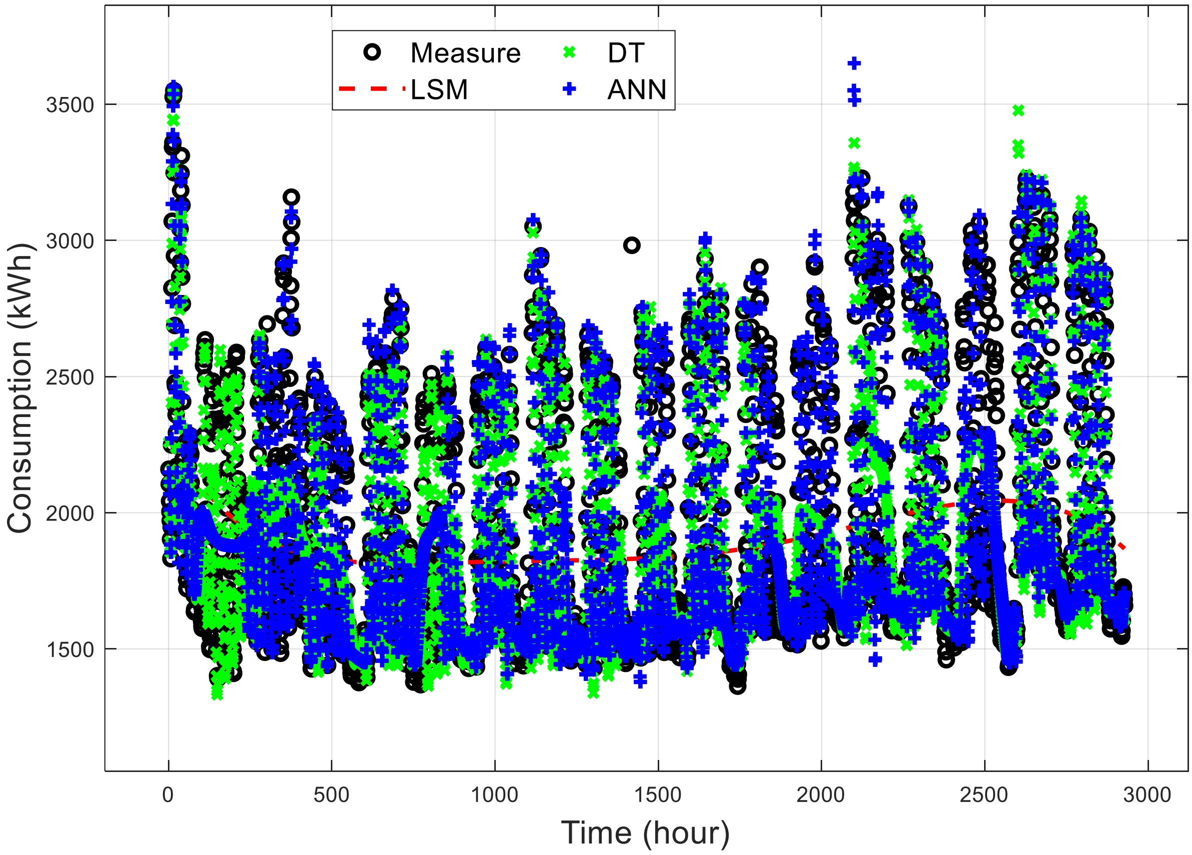

Fig.4 Comparison of prediction results of Model-3

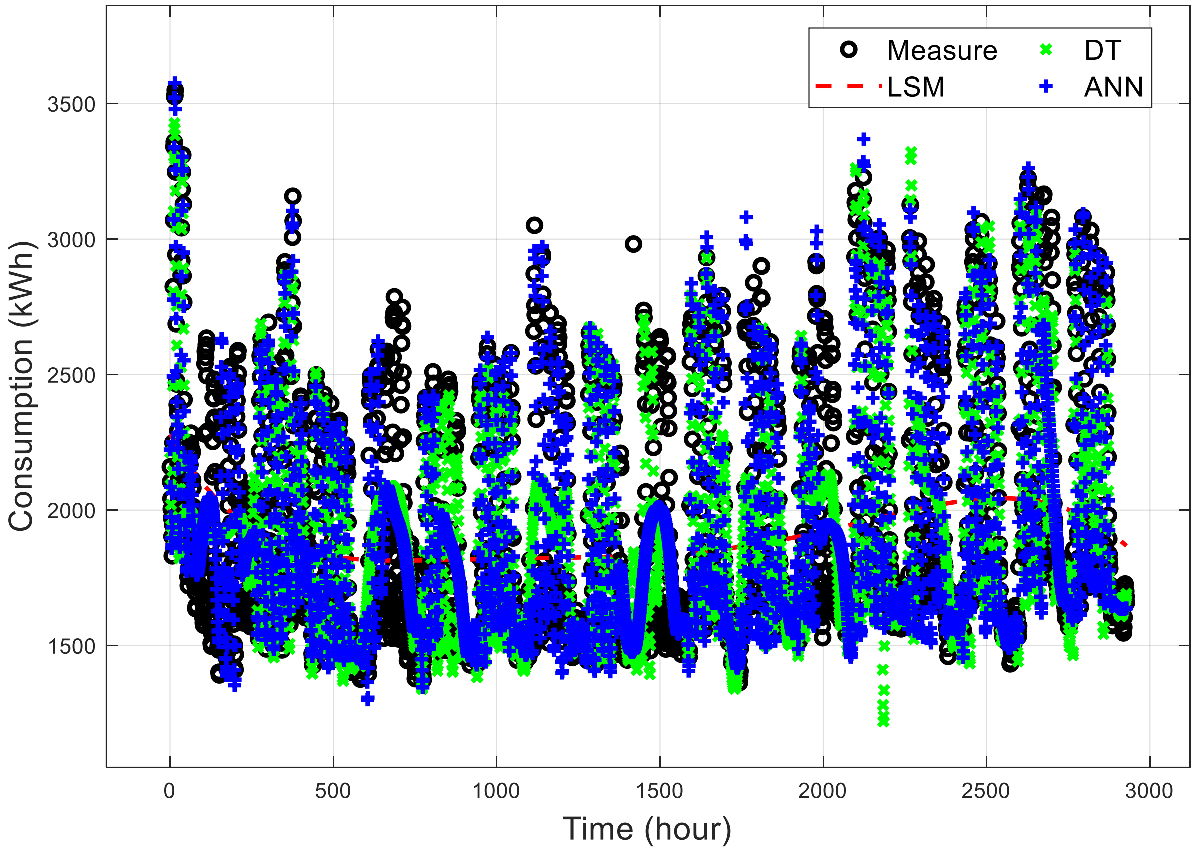

Fig.5 Comparison of prediction results of Model-4

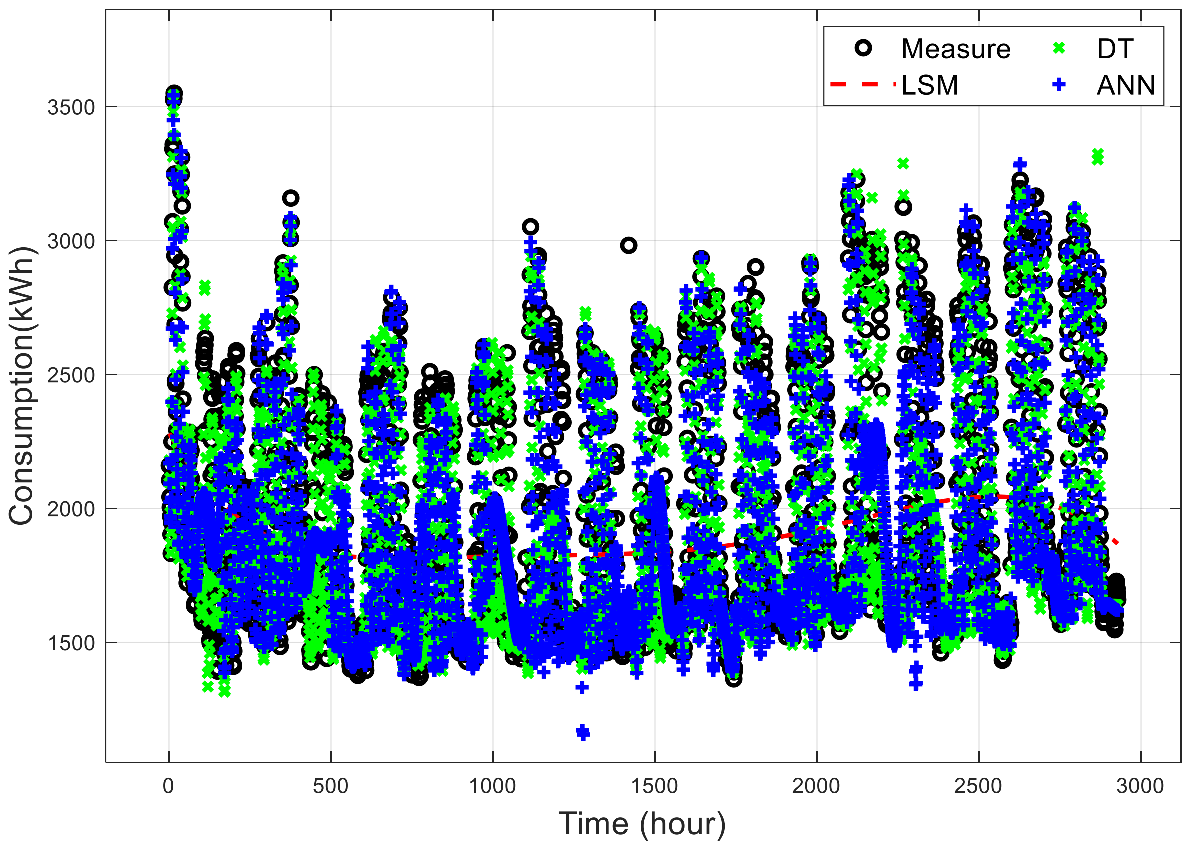

Fig.6 Comparison of prediction results of Model-5

IV. DISCUSSION

In this study, hourly electricity load forecasting is performed with various methods using Trakya University electricity consumption data. Hourly electricity data for the last 4 months of 2022 is used as forecast data. Hourly electricity consumption data for the first 8 months of 2022 were used as training data. As independent variables, 5 different data sets were created and studies were carried out with 5 different prediction models. The forecasting process was carried out by using ANN, DT and LSM forecasting algorithms in 5 different forecasting models.

The MAPE error rates obtained as a result of the forecasting studies are presented in Table 3.

Table 3. MAPE Values Calculated for Forecasting Algorithms

Scenario	ANN	DT	LSM
Model-1	%4.96	%6.36	%17.44
Model-2	%5.64	%7.41	%17.44
Model-3	%4.88	%6.25	%17.44
Model-4	%4.21	%6.16	%17.44
Model-5	%4.08	%5.04	%17.44

From the MAPE values, it's evident that across all models, the ANN consistently performs the best in terms of forecasting accuracy, achieving the lowest MAPE values compared to DT and LSM. The LSM, on the other hand, consistently shows the highest MAPE, indicating lower accuracy compared to ANN and DT. These results provide valuable insights into the effectiveness of each modeling technique for the given scenario, highlighting ANN as a particularly effective choice for accurate load forecasting.

It is seen that the ANN algorithm comes to the forefront in predicting behavioral characteristics. The performance of the DT algorithm in behavior analysis is not unsuccessful. However, there is a higher margin of error in the prediction process using the DT algorithm compared to the ANN application. The LSM application cannot be used as a load forecasting algorithm.

The effect of day separation does not have a definite result in all models. For this reason, weekdays should be separated according to the characteristics of the region to be forecasted and the algorithm to be used in forecasting.

Online and remote monitoring systems should be preferred during periodic measurements so that the system can be kept under constant control. The data obtained from measurements made with the help of remote monitoring systems should be analyzed using forecasting algorithms.

V. CONCLUSION

In this study, a demand forecasting and analysis application is made for Trakya University considering regional conditions such as weather conditions and usage habits. The electricity consumption characteristics of Trakya University were determined. Along with the forecasting and analysis studies, the least squares method (LSM), decision trees (DT) and artificial neural networks (ANN) algorithms are evaluated comparatively and the algorithm with the best performance is determined. It was possible to reduce the electricity consumption of Trakya University. In this context, the results obtained from the study are presented below.

When comparing Model 1 and Model 2, the model utilizing the load data from one hour prior demonstrates superior accuracy compared to the model utilizing load data from one day earlier. The load consumption from one hour ago exerts a more significant influence on the forecasted load value.

In the assessment of Model 3, Model 4, and Model 5 against other models, the model incorporating both the consumption load from one hour ago and one day ago achieves higher accuracy than models using them individually. Merely relying on the load from either one hour ago or 24 hours ago for the specific region to be forecasted does not suffice for achieving a high level of accuracy.

In the evaluation of Model 4 and Model 5 in comparison to Model 3, the inclusion of humidity data from meteorological records adversely impacts prediction accuracy. Hence, a thorough understanding of the region's characteristics is essential in choosing appropriate meteorological data, ensuring accurate predictions.

Model-5 is an algorithm that conducts forecasts without meteorological data, utilizing solely the load data from one hour and one day ago. Notably, Model-5 demonstrates the highest forecast accuracy in this scenario. Consequently, it has been determined that incorporating a dummy variable for seasonal effects and employing an algorithm trained exclusively on consumption data can enhance the success rate.

ACKNOWLEDGMENT

This study was supported by Trakya University Scientific Research Projects Coordination Unit. Project Number: 2022/124. The author would like to thank Trakya University for their support.

REFERENCES

[1] Yoldaş, U. C. (2016) “Elektrik enerjisinde yük tahmini yöntemleri ve Türkiye’nin 2005-2020 yılları arasındaki elektrik enerjisi talep gelişimi ve arz planlaması”, Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara,

[2] Aksel, F. (2000). Regresyon Analizi ve Yapay Sinir Ağı Yöntemleri ile Uzun Dönem Yük Tahmini, Yüksek Lisans Tezi, İ.T.Ü. Fen Bilimleri Enstitüsü, İstanbul

[3] New W. R. (1962). Load Forecasting on the TVA System Par I-Substation Loads, Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems, 81(3), 101-105.

[4] Gross G., Galiana F. D. (1987). Short-term load forecasting, Proceedings of the IEEE, 75(12), 1558-1573.

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

[6] Bakırtzıs, A. G., Petrıdıs, V., Klartzıs, S. J., Alexıadıs, M.C. (1996). A Neural Network Short Term Load Forecasting Model For Greek Power System, IEEE Transactions on Power Systems, 11(2), 858-863.

[7] Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. I., & Naccarino, J. (1996). Comparison of very short-term load forecasting techniques. IEEE Transactions on power systems, 11(2), 877-882.

[8] Gürsoy, E. (2000). Yük Tahmini Yöntemleri ve Çukurova Elektrik AŞ, Kepez Elektrik T.A.Ş. Bölgelerine Uygulanması. (Yüksek Lisans Tezi). İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

[9] Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. IEEE Transactions on power systems, 16(1), 44-55.

[10] Yalçınöz, T., Herdem, S. ve Eminoğlu, U. (2002). Yapay Sinir Ağları ile Niğde Bölgesinin Elektrik Yük Tahmini, ELECO’2002 Elektrik-Elektronik ve Bilgisayar Mühendisliği Sempozyumu ve Fuarı, sayfa 25-29, Bursa, 18-22 Aralık 2002.

[11] Huang, C. M., Huang, C. J., & Wang, M. L. (2005). A particle swarm optimization to identifying the ARMAX model for short-term load forecasting. IEEE Transactions on Power Systems, 20(2), 1126-1133.

[12] Ceylan, G. (2004). Yapay Sinir Ağları ile Kısa Dönem Yük Tahmini. (Yüksek Lisans Tezi). İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

[13] Pai, P. F., & Hong, W. C. (2005). Support vector machines with simulated annealing algorithms in electricity load forecasting, Energy Conversion and Management, 46(17), 2669-2688.

[14] Demirel, Ö. (2009). Anfıs ve Arma Modelleri ile Elektrik Enerjisi Yük Tahmini. (Yüksek Lisans Tezi). Marmara Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

[15] Akman, T. (2018). Yapay Zekâ Modelleri Kullanarak Ankara Bölgesinin Kısa Dönem Elektrik Enerjisi Yük Tahmini. (Yüksek Lisans Tezi). Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

[16] Khan, N. U. (2018). Short-Term Load Forecasting by Using Artificial Neural Networks. (Yüksek Lisans Tezi). Bahçeşehir Üniversitesi, FBE, İstanbul.

[17] Butekin Ç, T. (2019) Ann Based Electrıcıty Consumption Forecasting in Yaşar University. (Yüksek Lisans Tezi). Yaşar Üniversitesi, Fen Bilimleri Enstitüsü, İzmir.

[18] Doruk, E. (2019). Sakarya Bölgesi Hanehalkı Elektrik Tüketiminin Dinamik Lineer Model ile Tahmini. (Yüksek Lisans Tezi). Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.

[19] https://www.meteoblue.com/en/weather/historyclimate/weatherarchive/edirne_turkey_747712?fcstlength=1m&year=2022&month=10