TÜRKİYE CUMHURİYET MERKEZ BANKASI

İTÜ

ISTANBUL ELECTRICITY DEMAND FORECAST WITH ARTIFICIAL NEURAL NETWORKS

Hayriye Yasak OZKAL, Buse KAYLAN, Meltem SIPAHI

Advisor: Prof. Dr. Altan ÇAKIR

19.02.2025

Rome, Italy

Disclaimer

The views are those of the authors and do not necessarily reflect the views of the Central Bank of the Republic of Türkiye or its board members

Content

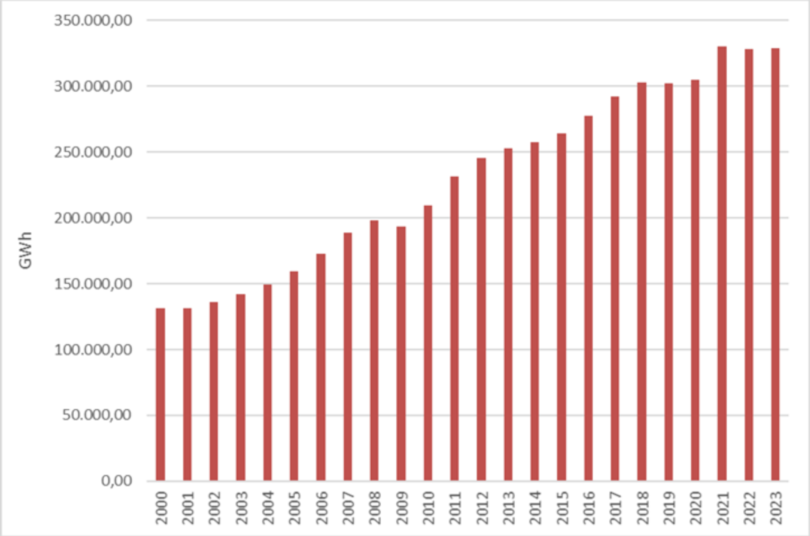
- **Introduction**
- **Main Objectives**
- **Methodology**
  - Data Collection
  - Generating Synthetic Data
  - Exploratory Data Analysis (EDA)
  - Time Series and Variable Selection
  - Data Preprocessing
  - Model Development
  - Model Evaluation
- **Key Findings & Future Work**

Introduction

Number of Consumers

Years	2021	2022	2023	2022-2023 Change
Istanbul	8.346.307	8.482.702	8.594.757	1.32 (%)
Türkiye	47.311.986	48.563.459	49.726.481	2,39 (%)

Development of Actual Consumption by Years (GWh)



As examined in the Energy Market Regulatory Authority (EMRA) 2023 sectoral report, the demand for electricity consumption and the number of consumers have increased over the years, ignoring minor deviations.

TÜRKİYE CUMHURİYET MERKEZ BANKASI

Main Objectives

- Ensuring System Stability
- Preventing Energy Shortages
- Minimizing Cost Increases
- Reducing Environmental Impact
- Preventing Resource Waste

The main objective of energy demand instability projects is to understand, manage and solve the problems caused by inaccurate demand forecasts and the resulting imbalances.

Methodology

Data Collection

- Gathering hourly humidity and temperature data for Istanbul
- Gathering hourly electricity consumption data for Turkey from January 2021 to December 2023
- Generation of Synthetic Data for Istanbul Electricity Consumption (Failed)
- Ratio of Istanbul's monthly electricity consumption to Turkey.

Exploratory Data Analysis (EDA)

- Identify patterns, trends, and correlations between electricity consumption and external factors.
- Examine historical data to understand consumption variations over time.

Model Development

- NARX RNN MODEL
- NARX LSTM MODEL
- PROPHET – LSTM HYBRID MODEL

Performance Evaluation

- Performance Metrics (MAE, RMSE)
- Model Loss Function
- Comparison of Actual and Estimated Values

Methodology

Data Collection

EPİAŞ

Actual Electricity Consumption

Temperature Humidity

OpenMeteo

In addition to these data collected on an hourly basis, information such as religious and public holidays, weekends and working days have also been added to the data set.

Methodology

Generating Synthetic Data For Istanbul Electricity Consumption

Data Source

- Total electricity consumption data of Turkey (2021-2023) from Energy Exchange Istanbul (EPIAS)
- Data includes date, hour, and consumption amount in MWh
- Monthy consumption percentages for Istanbul were used to scale the generated data

Challenge

- Lack of city-specific electricity consumption data
- Need for synthetic data to represent Istanbul's consumption

Methodology

GAN Model Architecture

GENERATOR

- **Dense Layer (256 units)** LeakyReLU Activation
- Batch Normalization
- **Dense Layer (512 units)** LeakyReLU Activation
- Batch Normalization
- **Dense Layer (1024 units)** LeakyReLU Activation
- Batch Normalization
- **Output Layer: Dense Layer (1 unit)** Linear Activation

DISCRIMINATOR

- **Dense Layer (1024 units)** LeakyReLU Activation Dropout (0.3)
- **Dense Layer (512 units)** LeakyReLU Activation Dropout (0.3)
- **Dense Layer (256 units)** LeakyReLU Activation Dropout (0.3)
- **Output Layer: Dense Layer (1 unit)** Sigmoid Activation

TRAINING PROCESS

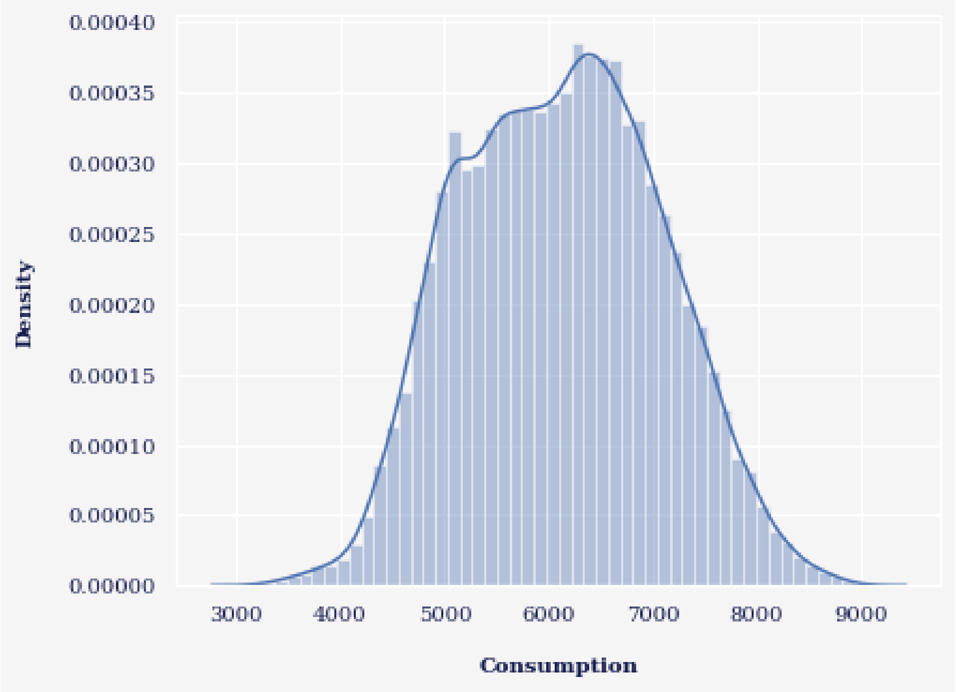
- **Epochs:** 10,000
- **Adversarial Training:** Generator and Discriminator iteratively trained to improve quality of synthetic data.

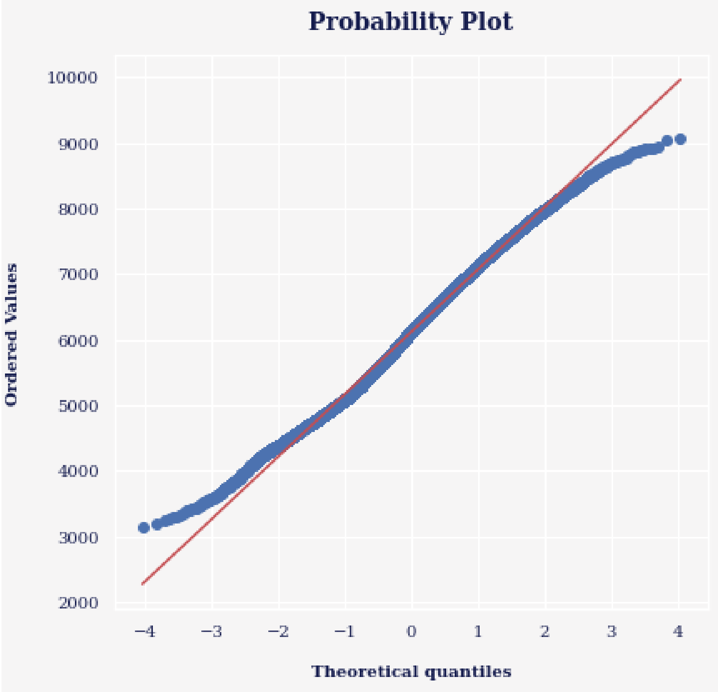
When the synthetic data generated by the GAN model is compared with the real electricity consumption data, it is observed that the synthetic data cannot simulate the real data pattern.

Methodology

Exploratory Data Analysis (EDA)

- The hourly consumption data of Istanbul province used in the analyses and modelling were obtained by scaling with the province-based consumption percentages in Turkey in the monthly reports of EMRA.
- Electricity consumption is usually concentrated between 5000-7000 MWh and extremely low/high consumption is rare.
- The data are approximately normally distributed but deviations are observed especially at the endpoints.



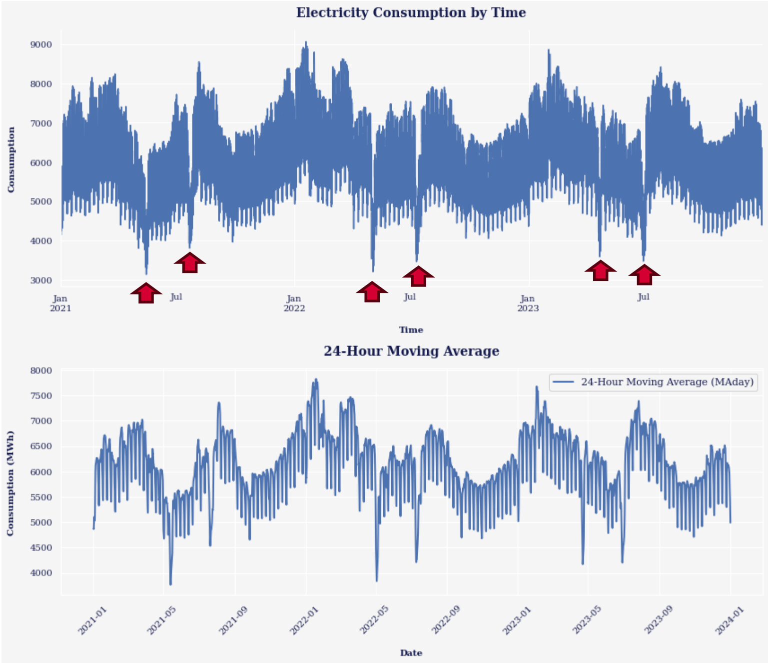


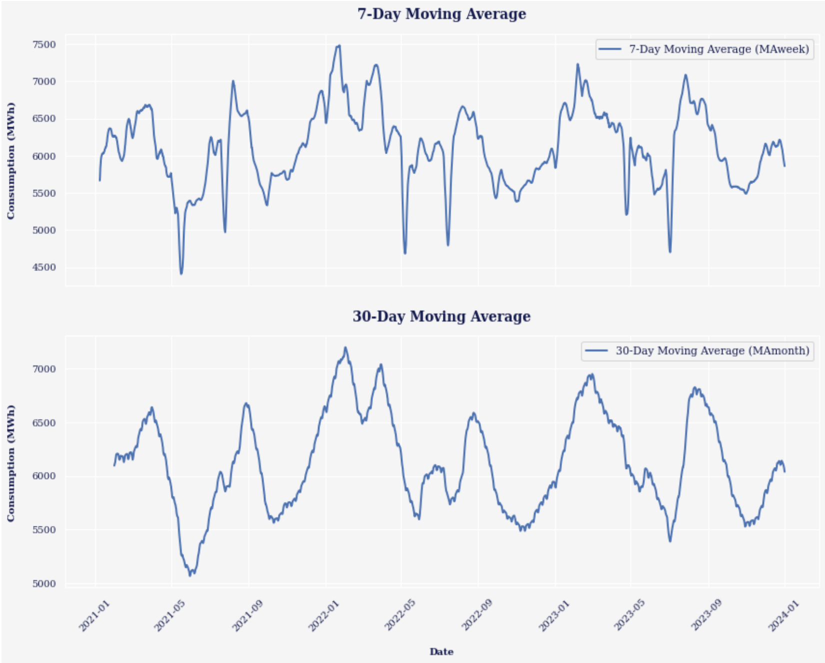
TÜRKİYE CUMHURİYET MERKEZ BANKASI

Methodology

Exploratory Data Analysis (EDA)

- When the hourly electricity consumption data of Istanbul between January 2021 and December 2023 are analysed, it is seen that consumption increases in winter months and there are fluctuations in summer months.
- It is seen that significant decreases occur in the data during religious holidays.
- The 30-Day Moving Average graph shows that the increases and decreases in overall energy consumption are seasonally significant



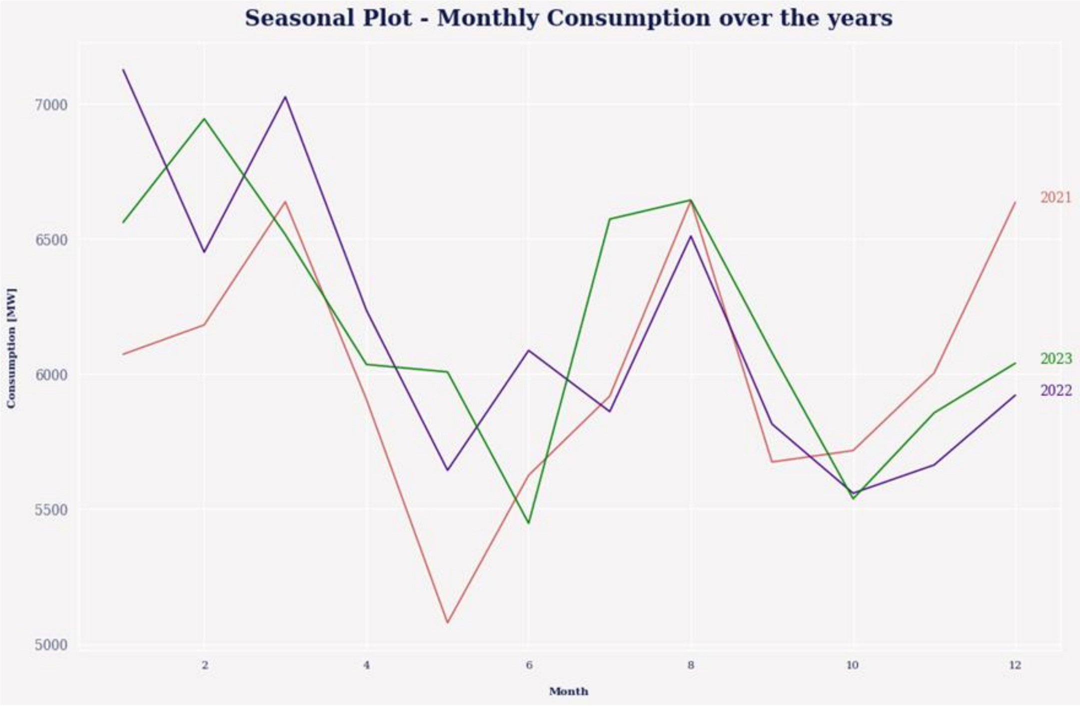


TÜRKİYE CUMHURİYET MERKEZ BANKASI

Exploratory Data Analysis (EDA)

Monthly Consumption

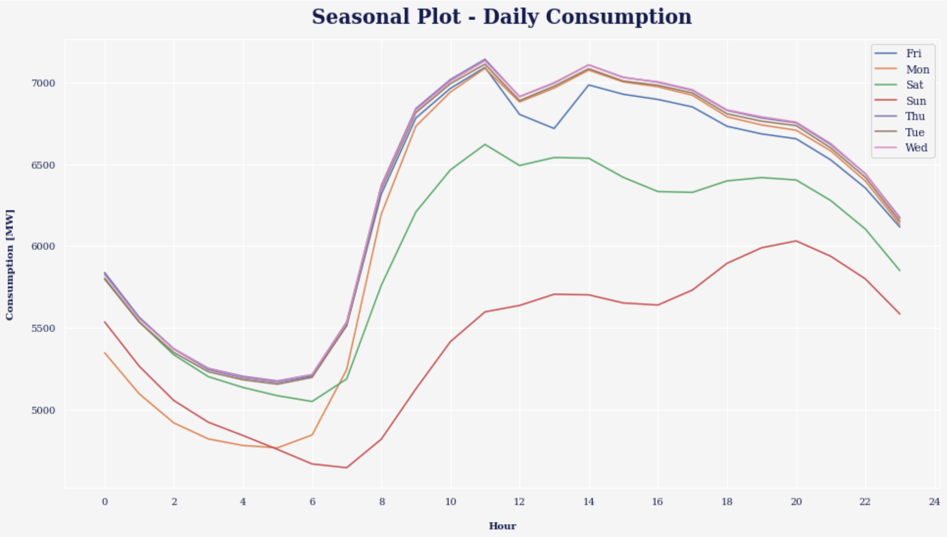
- When the monthly electricity consumption graph is analyzed, it is observed that consumption increases in summer and winter months and decreases in spring months.
- 2021 shows a different trend in terms of consumption, which may be due to the pandemic and the mild winter compared to other years

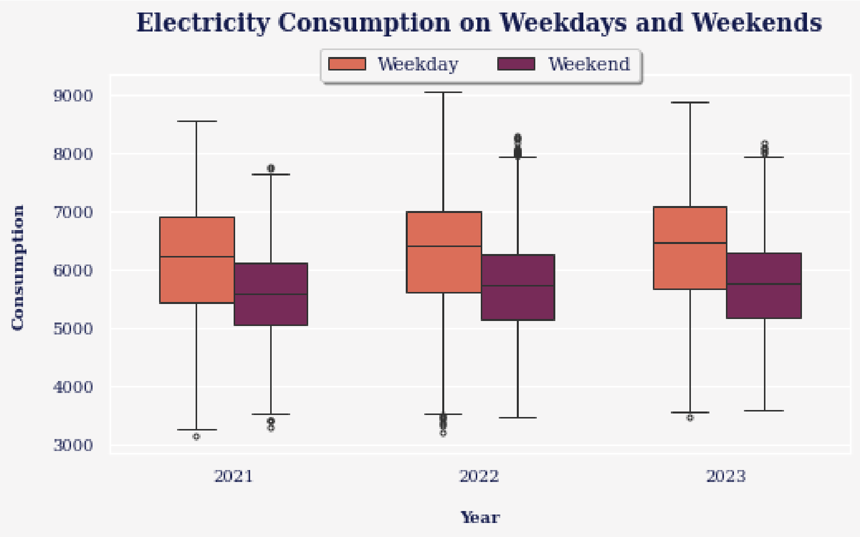


Exploratory Data Analysis (EDA)

Hourly & Daily Consumption

- When the electricity consumption graph is analysed in terms of days of the week there is a clear distinction between weekday and weekend consumption.
- For all days, there is a sharp increase in consumption starting around 5-6 AM, peaking nearly 11 AM.
- There is a gradual decline in consumption starting from late afternoon, around 5-6 PM, continuing through the evening and night.





Methodology

Time Series Analysis &

Variable Selection

- **Stationary:** The stationarity of the independent variables to be used in the model was measured by Augmented Dickey Fuller, Phillips-Perron and KPSS tests and those that failed were eliminated.
- **Correlation Matrix:** The correlation values of the dependent variable and all independent variables are observed and the independent variables to be eliminated are decided with the help of the threshold value in the decision mechanism.
- **Multicollinearity:** The VIF value used to examine whether there is multicollinearity between the explanatory variables in the data set.
- **PACF (Lag Analysis):** The appropriate number of lags to include in autoregressive (AR) models are determined by this analysis.

Selected Factors
Sensed Temperature
Perceived Humidity
Holiday
Weekend
Related Electricity Consumption Hour
Month of the year

Methodology

Data Preprocessing

- Based on the results of data exploration and analysis, a final data sets were created for electricity consumption data and related factors, with the time interval from 00:00 on January 1, 2021 to 23:00 on December 31.
- The dataset is divided into 60% training data, 20% validation data to tune model hypermeters and 20% test data to measure the model performance in chronological order.
- Next, all the time series recorded values are normalized using standart MinMaxScaler.
- Lags of the normalized series up to 24 time steps determined by PACF were added to the feature series.

Methodology

Model Development

Narx RNN Model : The Nonlinear AutoRegressive model with exogenous inputs (NARX) is a powerful modeling technique for time series forecasting.

Narx LSTM: LSTM (Long Short Term Memory) is a specialized type of RNN (Recurrent Neural Network) widely used in time series due to its capacity to learn long-term relationships within consecutive time steps.

Prophet - LSTM Hybrid Model : Combines the strengths of two popular time series forecasting techniques. The idea behind this hybrid approach is to leverage the advantages of both methods to improve the accuracy and robustness of time series forecasts.

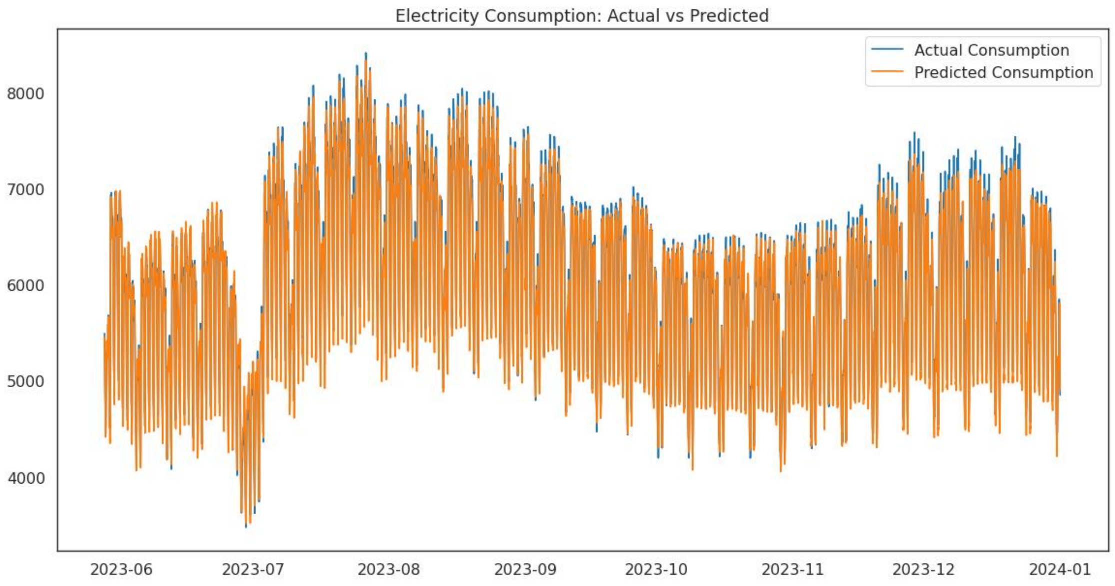
Prophet: Prophet is an open source library for forecasting time series with complex features such as trends, seasonality and holidays.

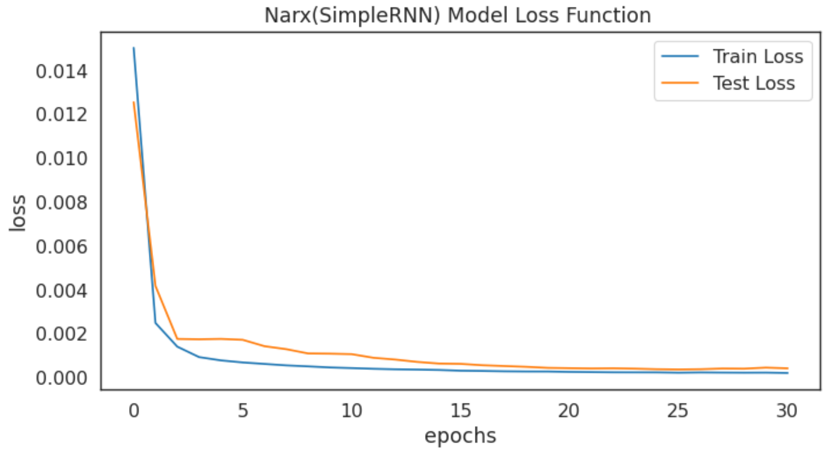
LSTM: Can learn more complex and short-term patterns that Prophet cannot model.

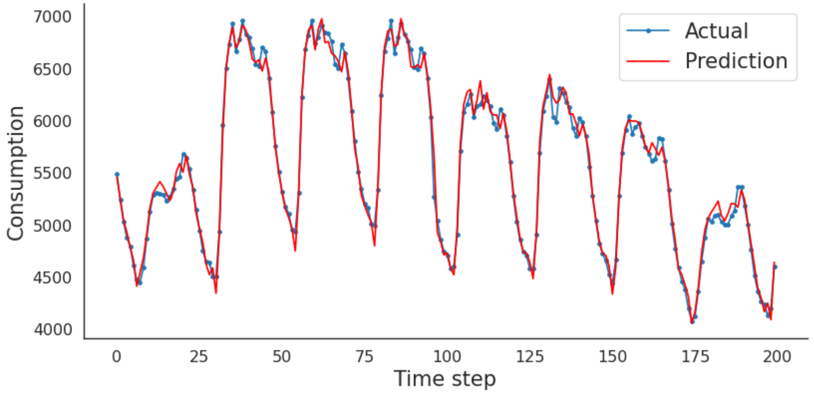
Model Name	Layer (type)	Output Shape	Parameters	Total Parameters
NARX RNN Model	simple_rnn (SimpleRNN)	(None, 50)	2,900	2,951
	dense (Dense)	(None, 1)	51	
LSTM Based NARX Model	lstm (LSTM)	(None, 24, 50)	11,600	31,851
	dropout (Dropout)	(None, 24, 50)	0	
	lstm_1 (LSTM)	(None, 50)	20,200	
	dense (Dense)	(None, 1)	51	
Prophet LSTM Hybrid Model	Prophet Model	(26.256, 32)	0	31,851
	lstm (LSTM)	(None, 24, 50)	11,600	
	lstm_1 (LSTM)	(None, 50)	20,200	
	dense (Dense)	(None, 1)	51	

Model Development

NARX RNN Model





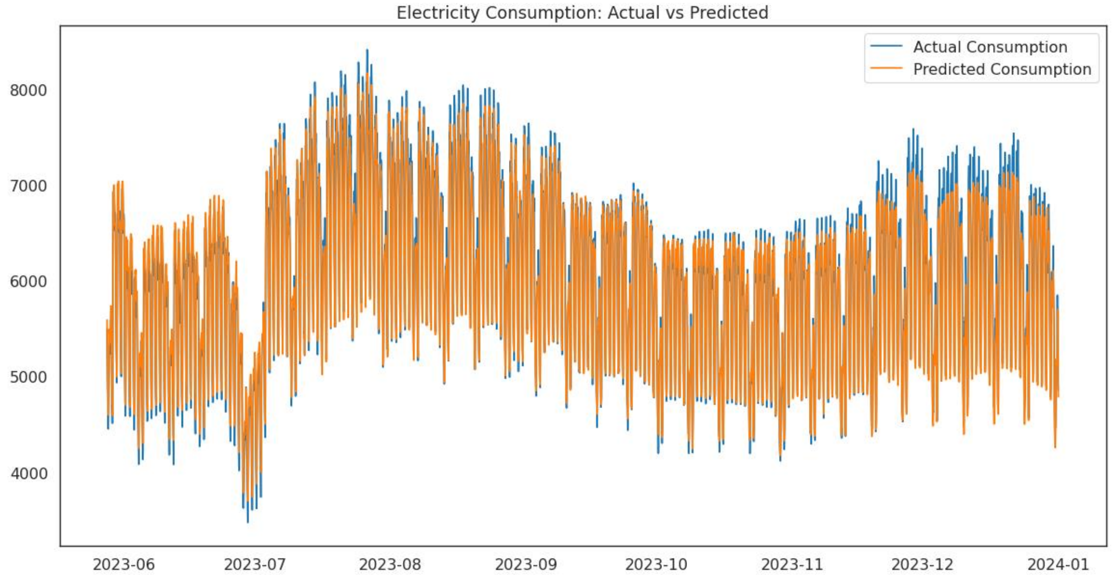


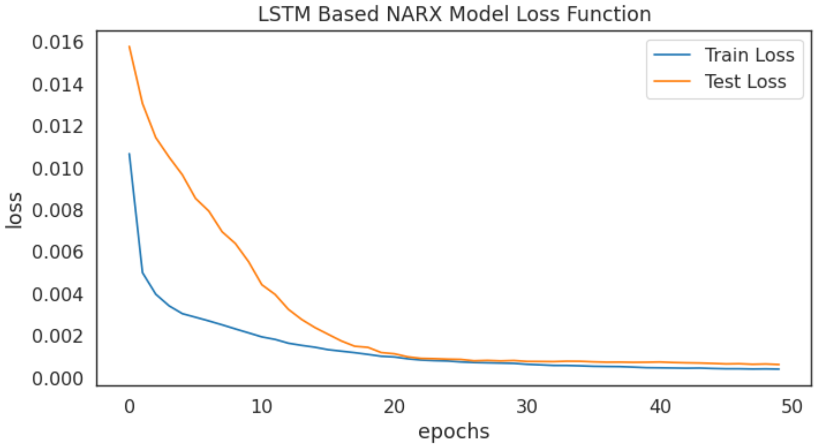
- Lag: 24
- Model: SimpleRNN (50 units, ReLU)

TÜRKİYE CUMHURİYET MERKEZ BANKASI

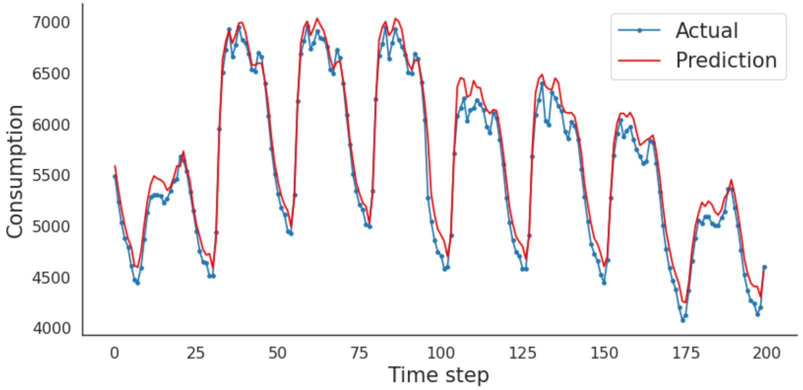
Model Development

LSTM Based NARX Model





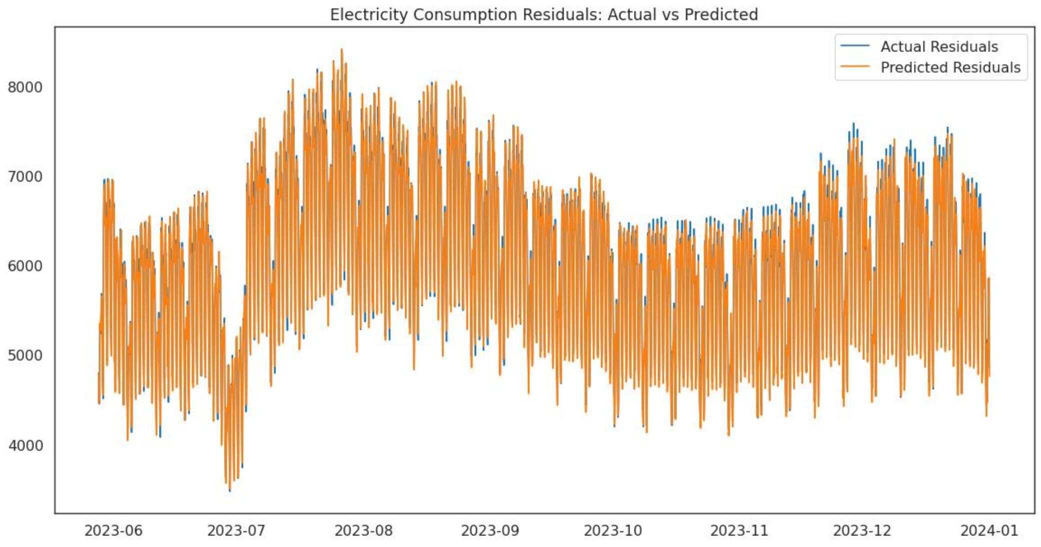
- Lag: 24
- Model: LSTM(50 units)



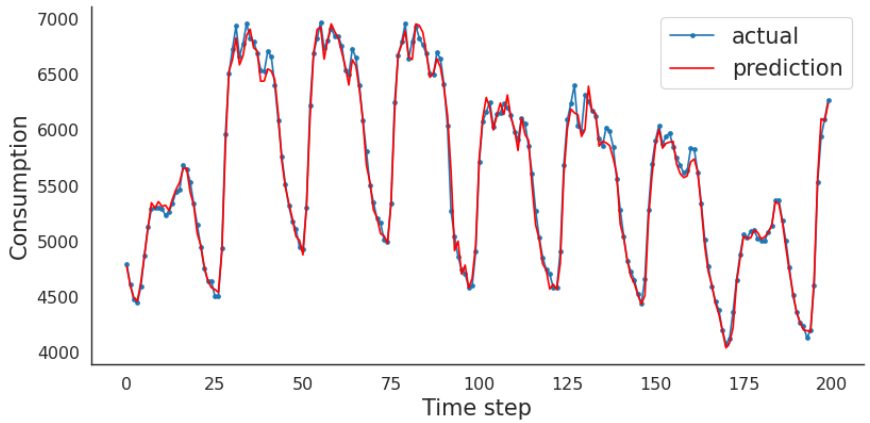
TÜRKİYE CUMHURİYET MERKEZ BANKASI

Model Development

Prophet-LSTM Hybrid Model




```python
[21]: predicted_values=pd.DataFrame()

predicted_values[['ds', 'trend', 'daily', 'weekly', 'yearly', 'extra_regressors_additive']] = prophet_predictions[['ds', 'trend', 'daily', 'weekly', 'y

y_pred_flatten = np.full(len(predicted_values), np.nan)

# y_pred değerlerini prophet_predictions ile hizalayarak ekleyin
y_pred_flatten[-len(y_pred):] = y_pred.flatten()

y_pred_flatten[-len(y_pred):] = y_pred.flatten()

y_pred_flatten[:-len(y_pred)] = y_pred_train.flatten()

predicted_values['residual'] = y_pred_flatten

# Sezonluk ve diğer bileşenleri toplayarak y_pred sütununu oluşturma
predicted_values['y_pred'] = (
    predicted_values['trend'] +
    predicted_values['daily'] +
    predicted_values['weekly'] +
    predicted_values['yearly'] +
    predicted_values['extra_regressors_additive'] +
    predicted_values['residual']
)
```



- **Model:** Using Prophet initial predictions generated and these predictions used as input to train an LSTM model. The LSTM model learns from the residuals (the differences between the actual values and the Prophet predictions) and generates refined forecasts.

Model Development

Model Evaluation

- MAE: Measures the mean of error magnitudes, ignores direction, is easy to interpret and has low sensitivity to outliers.
- RMSE: Expresses errors in original data units, gives more weight to large errors and is more sensitive to outliers.

MODELS	PERFORMANCE METRICS			
	Train		Test	
	MAE	RMSE	MAE	RMSE
NARX (SimpleRNN)	86.65	115.68	90.39	116.87
LSTM Based NARX	133.41	168.11	117.48	148.66
Prophet	82.93	111.95	85.88	112.22
Prophet-LSTM Hybrid Model	62.17	80.48	65.58	84.73

When the performances of the models are evaluated with MAE and RMSE metrics over training and test data, it is seen that the models can generalise the data and there is no risk of overfitting in the models, except for the LSTM-based NARX model.

Findings & Feature Work

Key Findings:

- Synthetic datasets could failed to replicate time series data.
- When the performance metrics, loss function graphs are evaluated and the actual predicted consumption values are compared, it is seen that the Prophet-LSTM hybrid model is more successful in the prediction of hourly electricity consumption.

Future Work:

- Optimizing performance metrics
- Evaluation of short-term and long-term forecast performances
- Improving the performance of the model by selecting shorter and more significant periods.
- Updating the model regularly with new data and adapting to changing conditions.

THANK YOU!

Hayriye Yasak Ozkal

hayriye.yasakozkal@tcmb.gov.tr

Big Data Technologies Directorate

Central Bank of the Republic of Turkey