IE-360 Term Project

Zeynep Talan - 2017402003

Problem Description

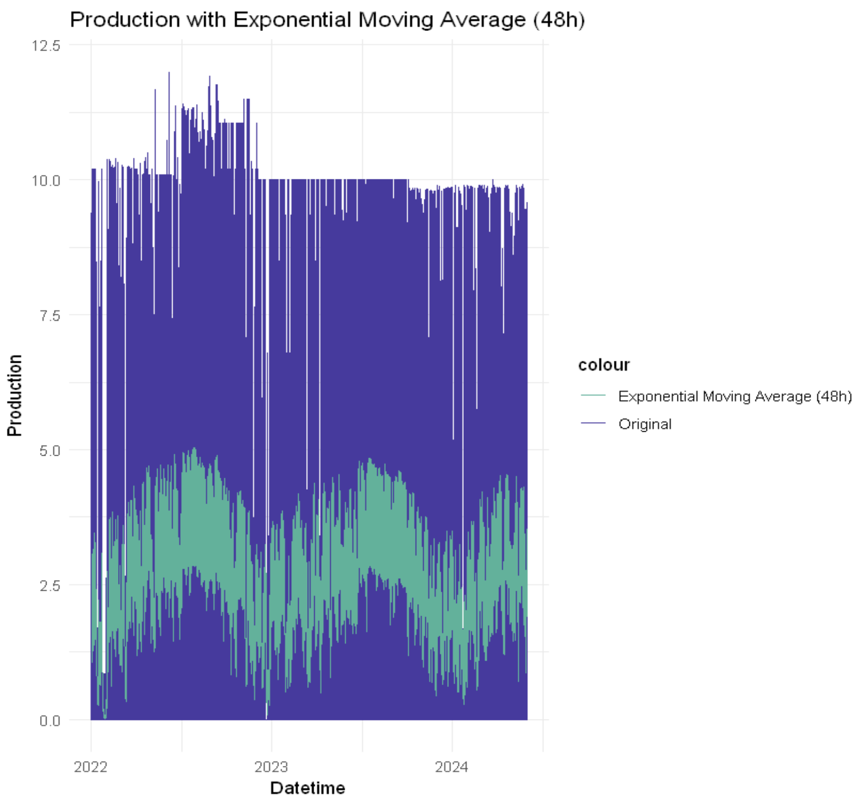
The aim of this project is to provide hourly solar power predictions for the Edikli GES (Güneş Enerjisi Santrali) located in Niğde, Turkey, for the next day. The predictions are essential for energy traders who need to submit their forecasts before noon on the preceding day to optimize their operations in the energy market. The data provided includes weather measurements from 25 grid points near the power plant, which are crucial for accurate solar power forecasting.

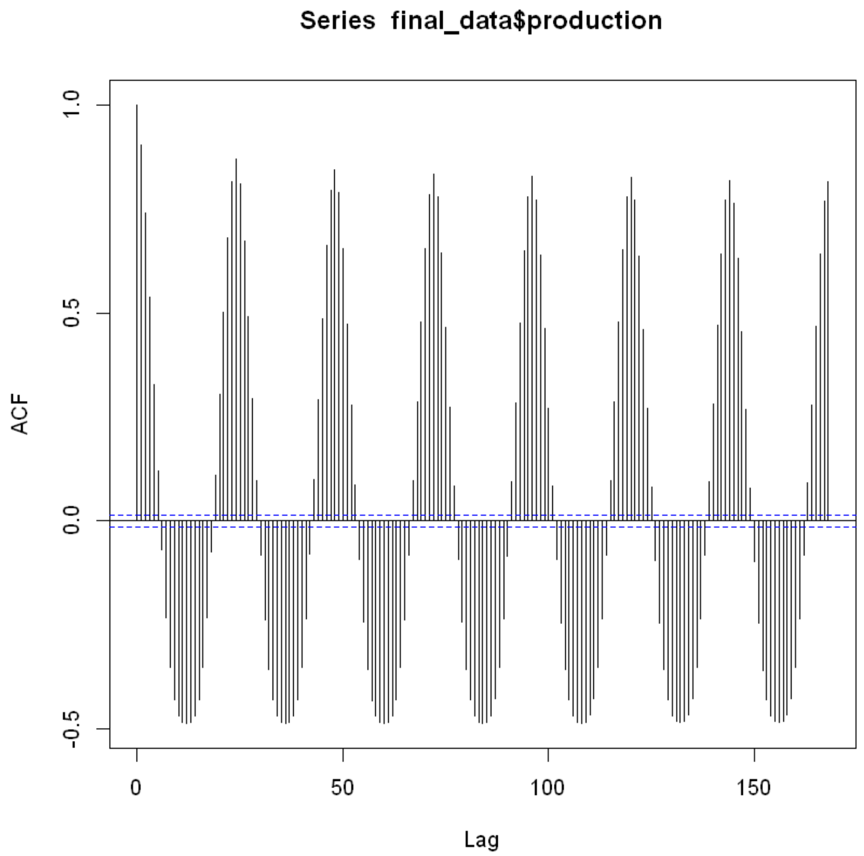
Data Preparation

Implemented necessary data preparation operations; handled missing values, to create a conclusive dataframe, filtered the hour data for the closest points and took the averages of these points' weather variables and use these as our data to make analysis on.

Data Visualization

Visualized data to observe patterns. Then use these information to build our models. Two of them is below, more in the code. For example, we've seen from these that our data may have yearly and daily seasonalities.





Feature Engineering

Extracted several features from the datetime information, such as hour and season, to capture the patterns in solar power production. Additionally, we calculated moving averages and lagged values of production to incorporate temporal dependencies.

Moving on to Build Models Required

Checked for multicollinearity before starting to select a base model for our lineer model approaches. Also decomposed the data to get trend, seasonal, remainder components to use in our models.

Built various combinations of the predictors for our models. All of the lineer models built, contain the same weather variablos and hour as predictors but other predictor variables they contain might differ in between them (lag1_production, lag2_production, rolling_mean_production, trend , seasonal). At the end the selected lm model that lastly created as model_with_lags resulted in %24.29 WMAPE.

Then moved to build various arima models and try to enhance our predictions. Also tried auto_arima at the end to see if it has any improving effect, yet it did not.

Conclusion

Since the WMAPE values of arima models are not better than the lm model choosed, concluded that it is the best model among the mentioned models.