Hourly time-series prediction of power suppliesor for next week

Authors / Verfasser: Dipl-Ing Stefan Bratic

Overview

Project Scope

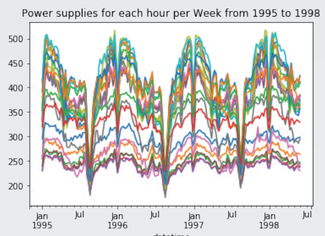
Project aims to predict future values of how much power supply is needed to be provided by plant, based on values in the future.

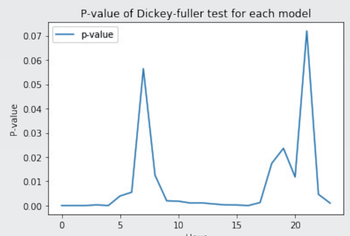
Approach

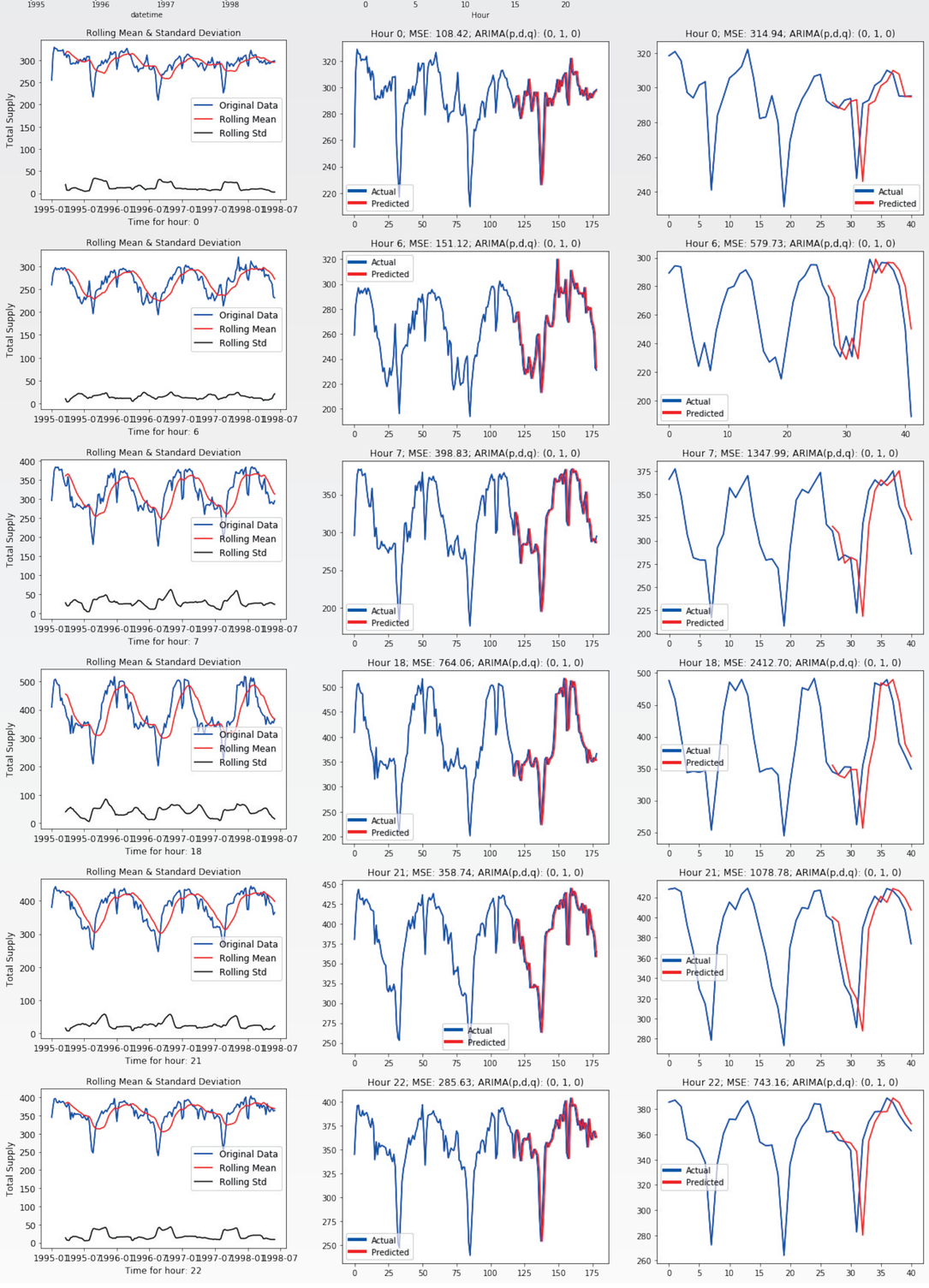
- The data was first separated in 24 models that will be used separately. Each model contained data for days between 1995 and 1998 for certain hour.
- Each model was analyzed for having stationary or non-stationary data.
- 6 models were extracted from 24 existing one and ARIMA with (0,1,0) and (1,0,0) parameters were applied
- After evaluating different forecasting times such as month, day or week. Optimal results were given by weekly forecasting.

Data Set

Extended Problem Statement







Details on the Approach

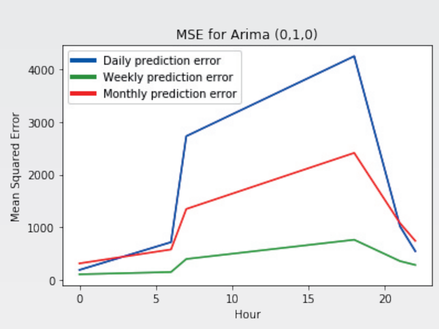
Approach

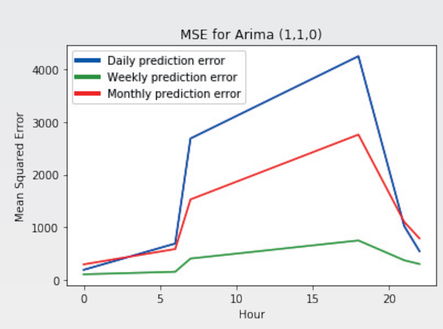
- Data separation was done in 24 different models because it contained data for each hour for each day from 1995 to 1998. We wanted to forecast on hourly basis because it was more likely that the data has pattern in repeating itself throughout the time for certain hour.
- After that, each model was tested if it contains stationarity or not. The reason for what is because if the data is non-stationary that means that we need to do some kind of transformation to make it stationary to be able to have lower error.
- The stationarity test was done via visualizations and statistical test called Dickey-Fuller test.
- Visual representation was used to check visually if rolling mean and rolling standard deviation had some trend and Dickey-Fuller test was to confirm what was seen on plot.
- After realizing that some hours hold identical values over the time, 6 models were extracted to represent 6 different types.
- To each model ARIMA was applied with parameters (0,1,0) and (1,0,0)
- To compare how good ARIMA was working three time frames were used: daily, weekly and monthly.
- Each model fitting was evaulated by comparing one-step ahead prediction with the actual value and at the end computing the total mean squared error

Detail Level

Start with an general, conceptual level and then more detailed level

Results





.

Discussion

- From the images above can be seen that the week-based predictions perform the best with given two Arima models. The biggest surprise was the error rate that day-based prediction has considering that the time frame is smaller than weekly one, which should induce that the prediction should contain lower error rate.
- Additionally, differencing technique was applied to the daily dataset to eliminate non-stationarity for particular models, but
- For future work what could be done is two compute the average for given hours that have almost equal values.

References

[1] Wastl, E.; "On the Feature Extraction of Features"; 2015; Journal on the Advanced Studies of Studies

TU Graz – Institute of Interactive Systems and Data Science
8010 Graz, Inffeldgasse 16c, Austria, Tel.: +43 (316) 873-5624
office.isds@tugraz.at, https://www.tugraz.at/institute/isds/home/