THE UNIVERSITY OF CHICAGO

Power Consumption of Tetouan City

Martin Copello, Kartik Garg
Maya Zhao, David Wen

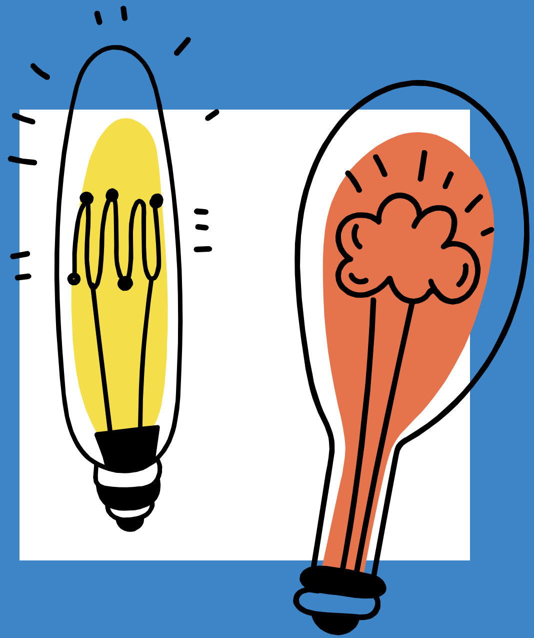
Source: https://github.com/martincopello/power_consumption_time_series

Table of Contents

THE UNIVERSITY OF CHICAGO

01

Problem Statement

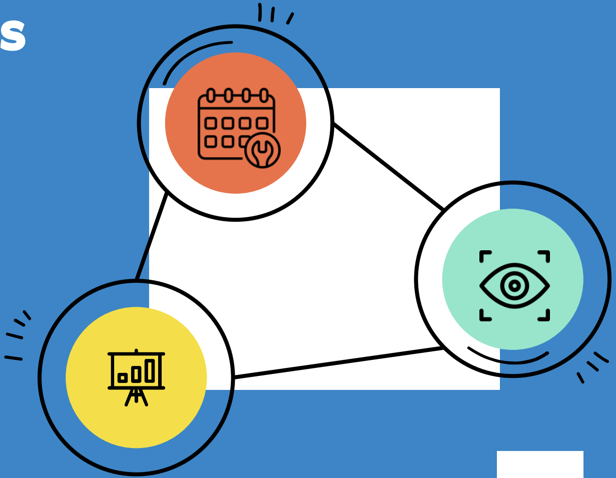


Data Problem

THE UNIVERSITY OF CHICAGO

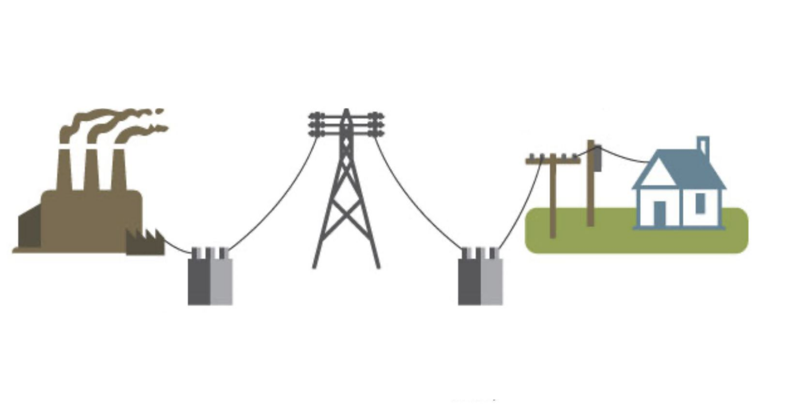
Accurately forecasting load on the electric grid can have a massive impact on electric reliability, especially as we move towards electrification. Some benefits include:

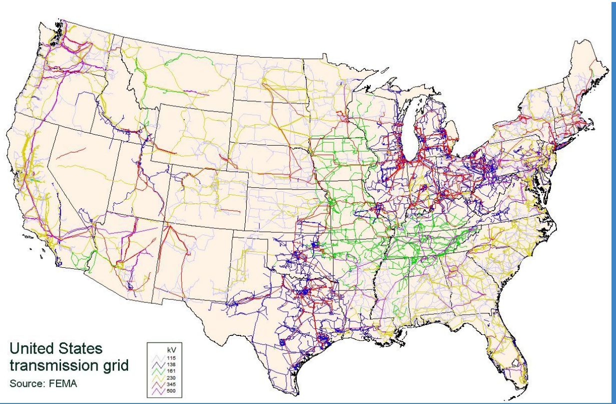
- Proactive preventive maintenance schedules
- Grid operations to lessen load
- Prediction of equipment misoperations

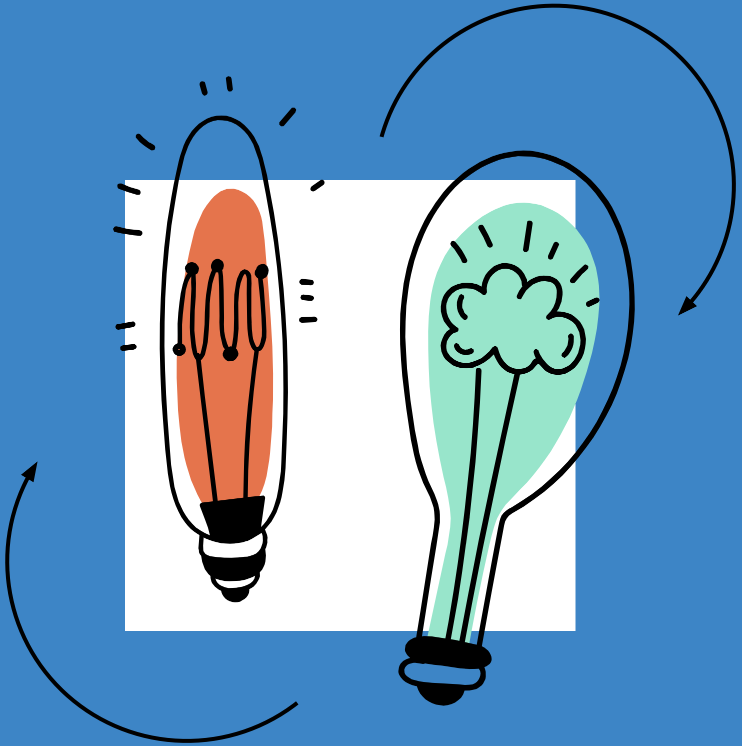


The Electric Grid

- Storage of electricity
- Adjust generation to forecasted usage of electricity
- Interconnected networks
- Reduce carbon emissions/footprint





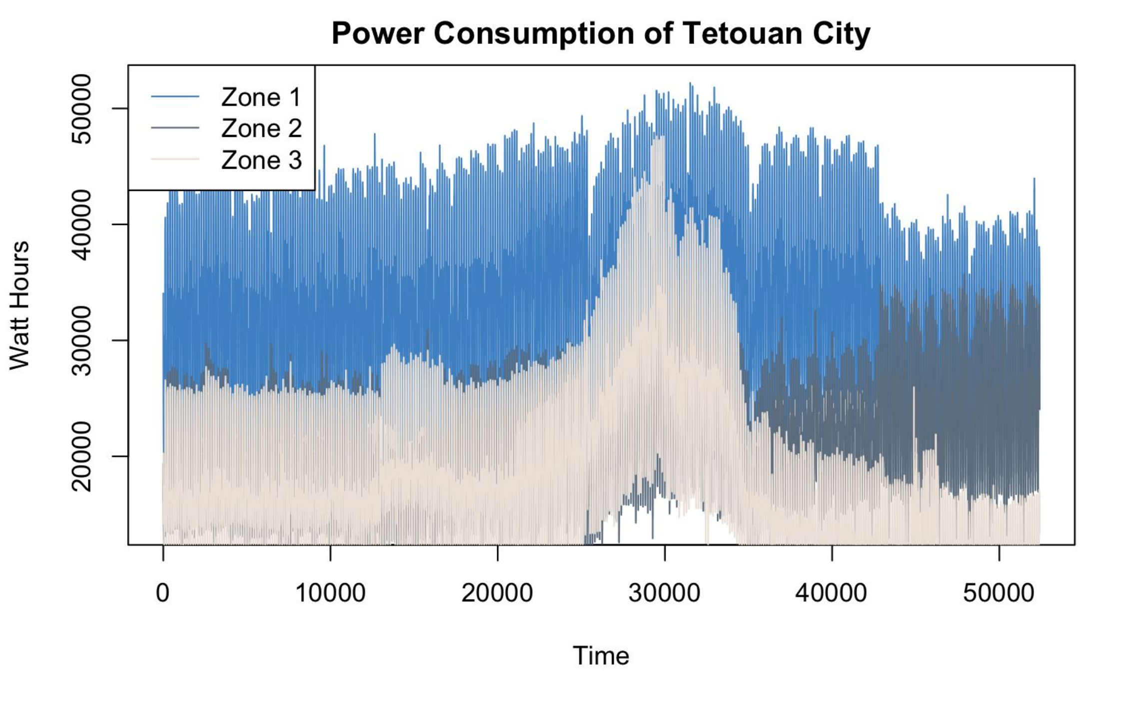


02

Data

Check the data

THE UNIVERSITY OF CHICAGO



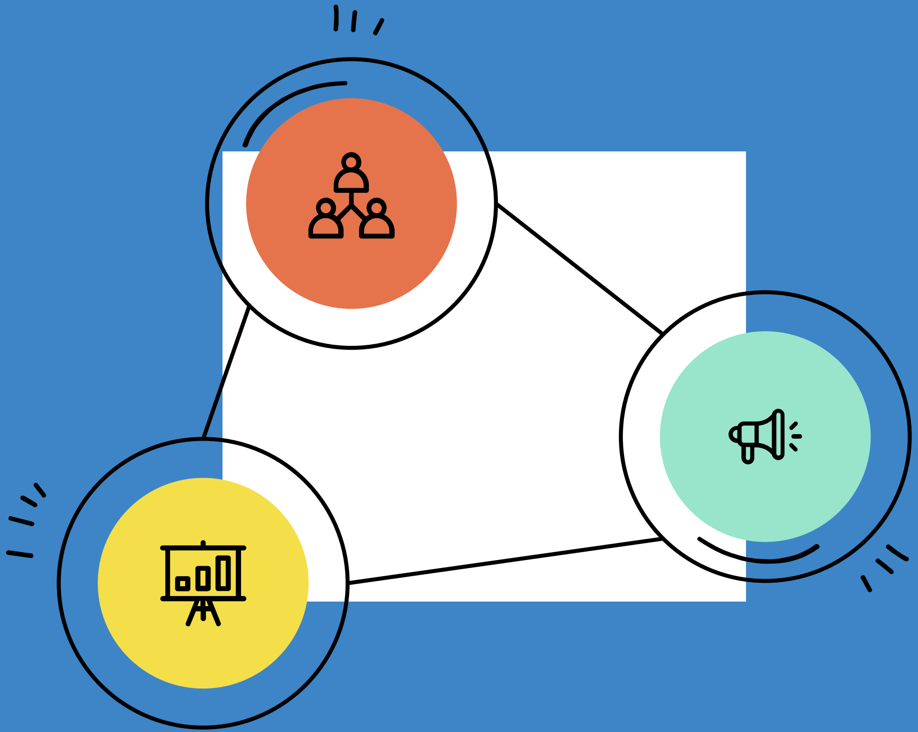
Power consumption of Tetouan City, Morocco

- UCI Machine Learning Repository
- Power consumption of distribution networks of Tetouan city in north Morocco
- 52416 obs. of 9 variables (no missing values)
- Frequency: 10 mins
- 01/01/2017 00:00 - 12/30/2017 23:50
- Focus: Zone 1 (blue line)

Source: https://archive.ics.uci.edu/ml/datasets/Power+consumption+of+Tetouan+city#

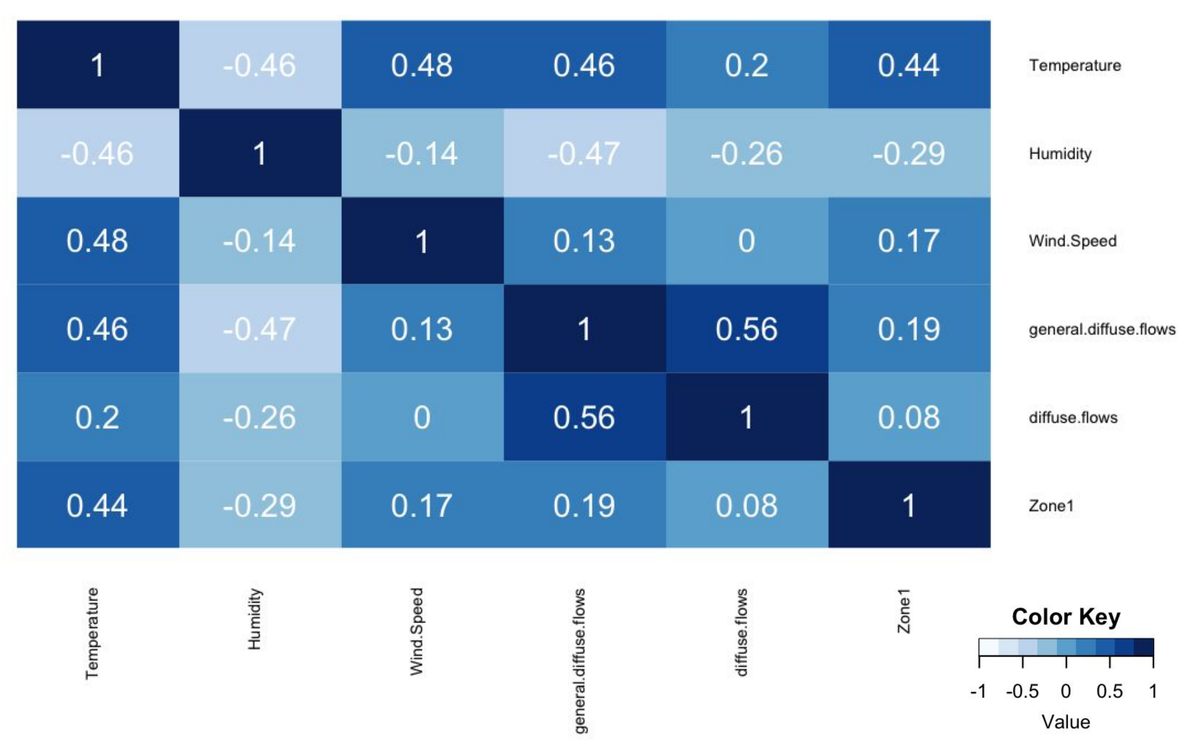
03

Exploratory Data Analysis



Exploratory data analysis

THE UNIVERSITY OF CHICAGO



Data Dictionary

- Temperature

Weather Temperature of Tetouan city.

- Humidity

Weather Humidity of Tetouan city.

- Wind Speed

Wind Speed of Tetouan city.

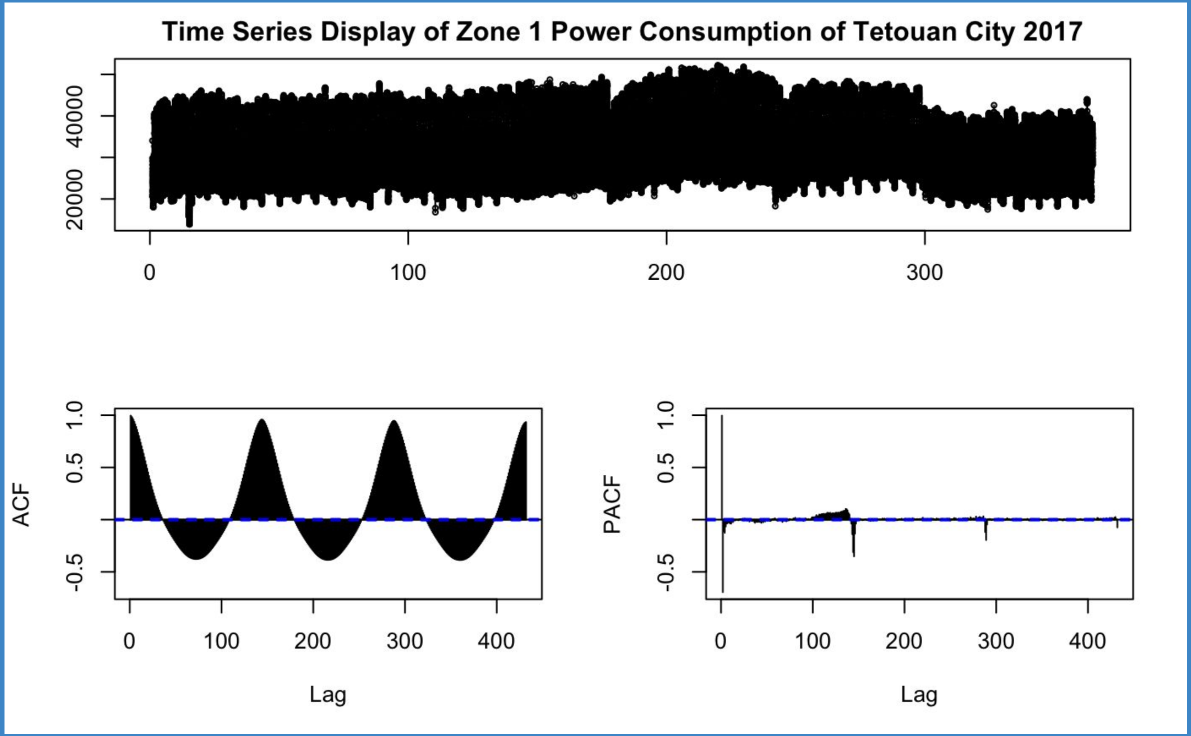
- General diffuse flows

diffuse flows

- **Zone I**

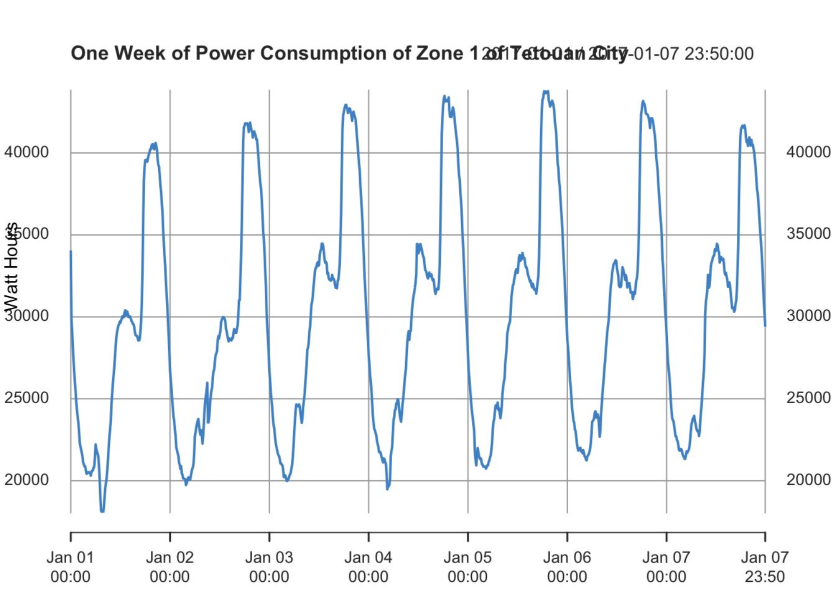
power consumption of zone I of Tetouan city

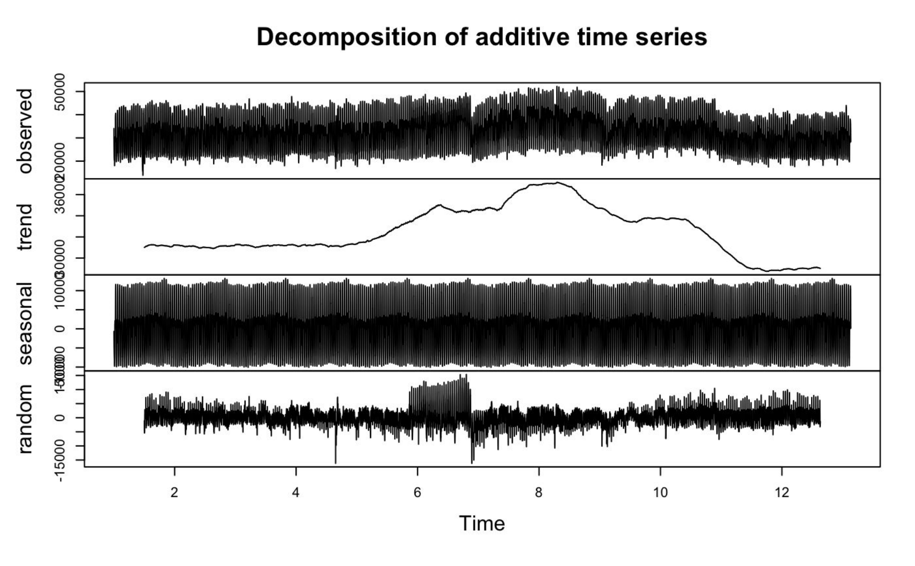
Zone 1 Power Consumption Overview



- Up & down patterns
- Seasonality
- Slow decay in ACF shows non-stationarity

Over different periods of time





Power consumption in a week

- Daily pattern
- Peak in evening

Decomposition using monthly frequency (6*24*30)

- Additive seasonality
- Peak in summer

Stationarity

THE UNIVERSITY OF CHICAGO

Tests

```
        Augmented Dickey-Fuller Test

data:  df$Zone1
Dickey-Fuller = -32.63, Lag order = 37, p-value = 0.01
alternative hypothesis: stationary
```

```
        KPSS Test for Level Stationarity

data:  df$Zone1
KPSS Level = 6.6276, Truncation lag parameter = 19, p-value = 0.01
```

Tests for stationarity

Augmented Dickey-Fuller (ADF) Test

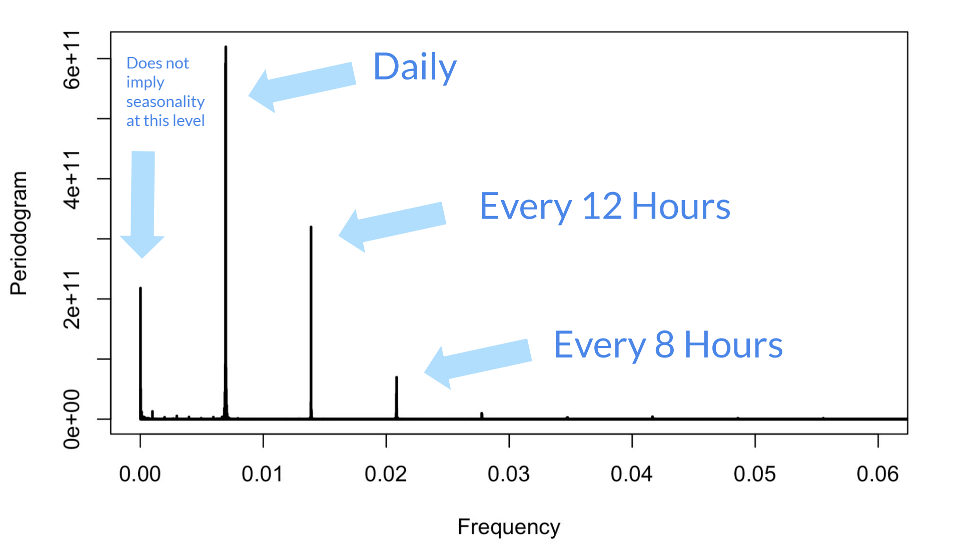
- Null hypothesis: non-stationary
- P-value: 0.01, reject null hypothesis

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test

- Null hypothesis: stationary
- P-value: 0.01, reject null hypothesis

Conclusion: The series is difference stationary.

Seasonality Decomposition in Frequency Domain



Processing technique

- Differencing
- Decomposition in daily & weekly frequencies

Feature Engineering & Selection

THE UNIVERSITY OF CHICAGO

Initial Features

- Temperature
- Humidity
- Wind Speed
- Diffuse Flows

+

Feature Engineering

- Hour
- Minute
- Day of the week

→

Feature Selection

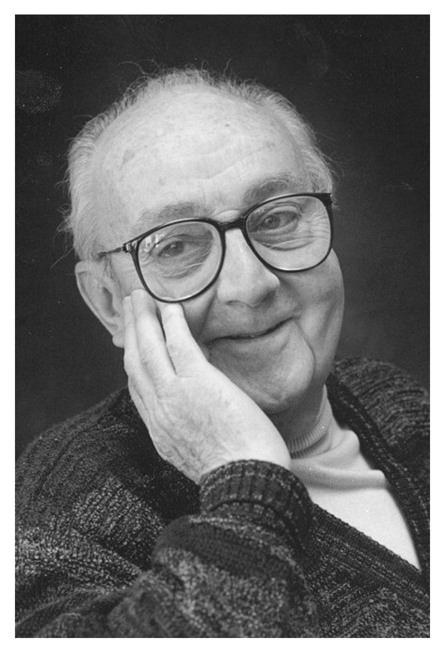
Based on correlation wrt Power Consumption

- Temperature
- Hour
- Day of the week

04 Forecasting Models

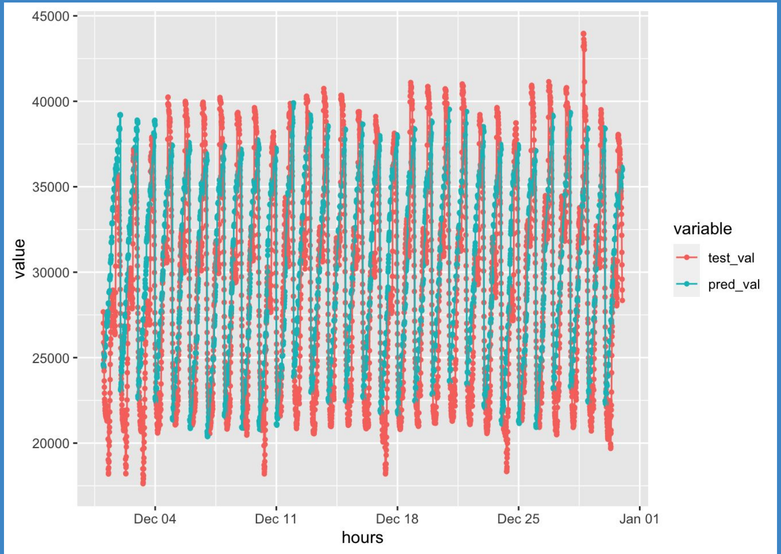
"All models are wrong, some models are useful."

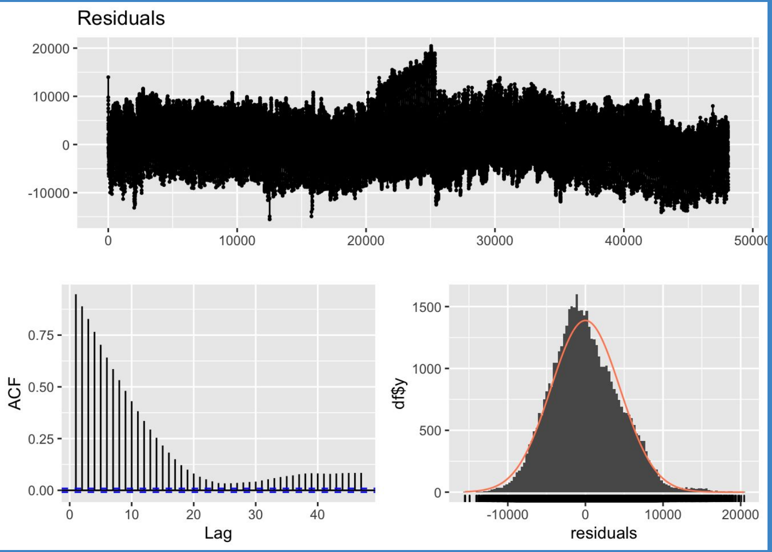
— George Box



Baseline Model: Linear Regression

Model selection: Temperature | **Temperature + hour** | Temperature + hour + dayOfWeek





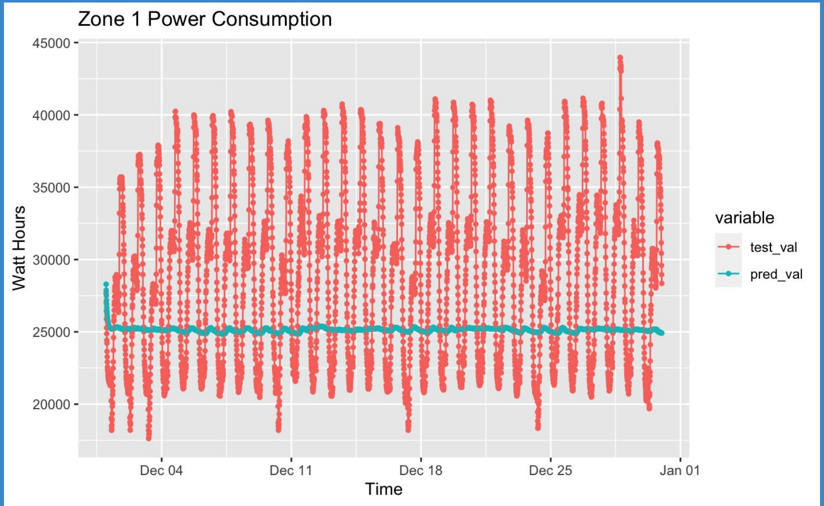
Best Performing: Temperature + hour
sMAPE: 0.2028148
Training time: 27.51 secs

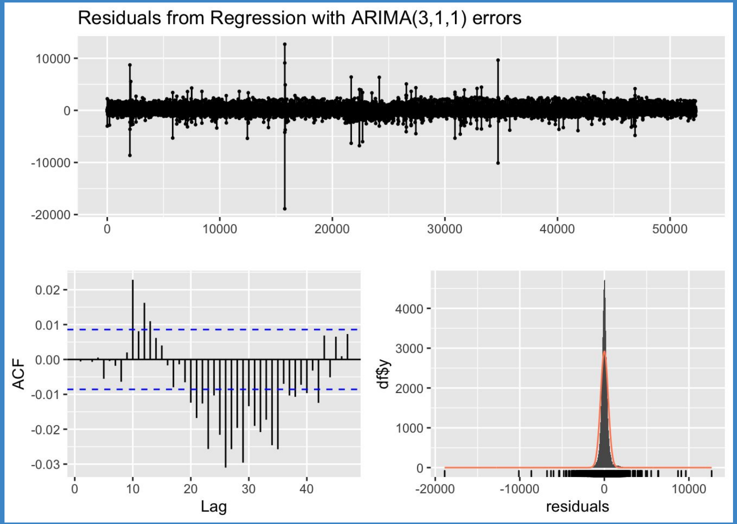
- Residuals are slightly right skewed
- ACF: Slow decay, spikes at lag 1...

Baseline Model: Regression with ARIMA errors

THE UNIVERSITY OF CHICAGO

Model selection: Temperature | Temperature + hour | Temperature + hour + dayOfWeek





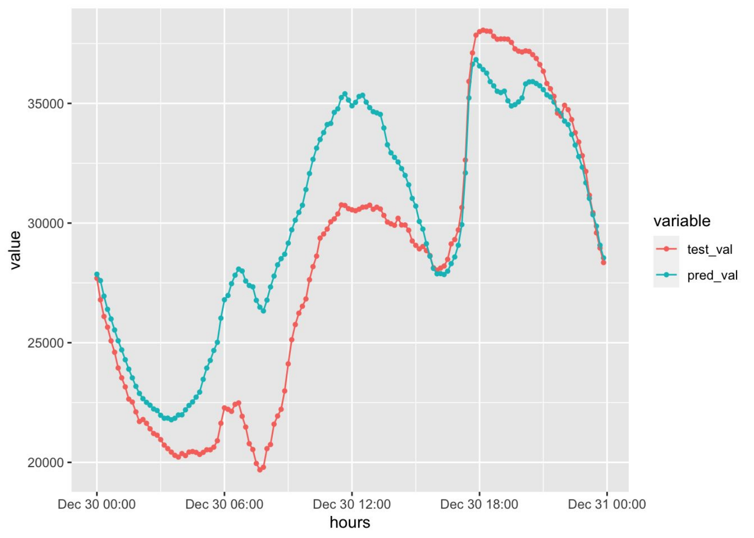
Best Performing: Temperature
sMAPE: 0.2082446
Training time: 30.76 secs

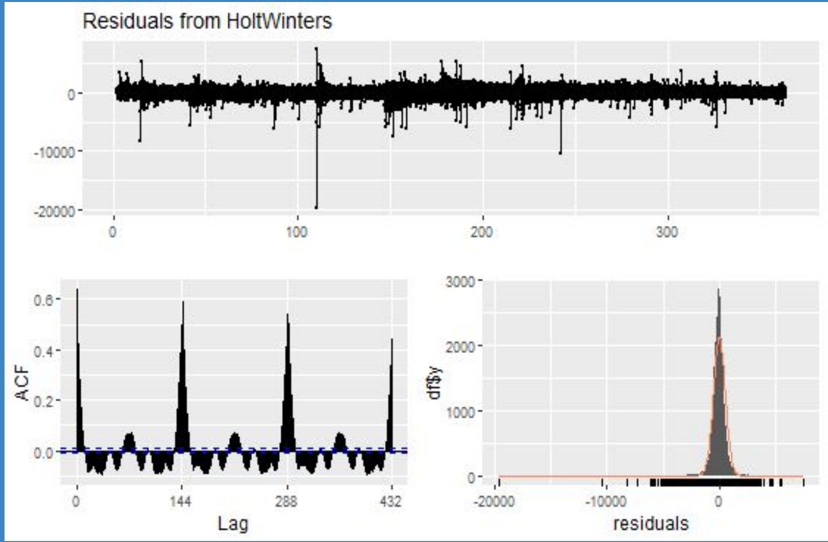
- Residuals are close to normal distribution with a smaller standard deviation
- ACF: Slow decay, spikes at lag 10, 26...

Holt Winters with Exponential Smoothing

THE UNIVERSITY OF CHICAGO

Daily Frequency | Weekly Frequency





Best Performing: Daily Frequency
sMAPE: 0.09130168
Training time: 3.44 secs

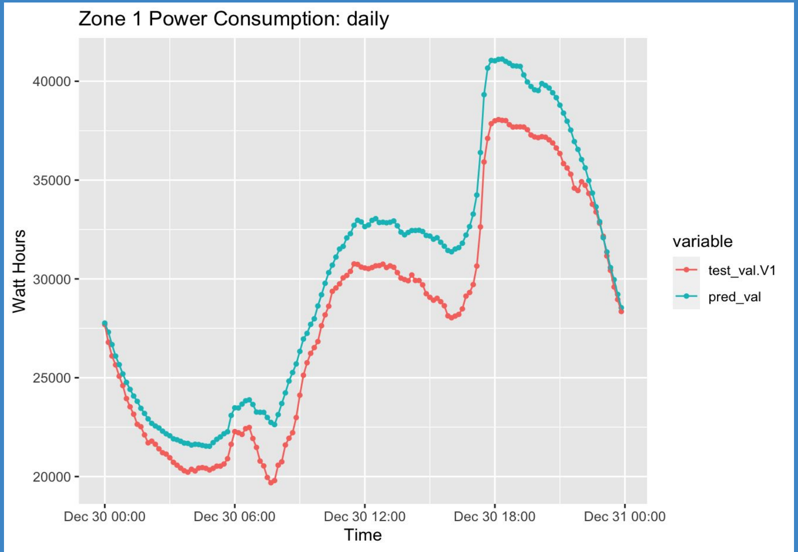
- Residuals are close to normal distribution with a smaller standard deviation
- ACF: Slow decay, spikes at lag 0, 144,288...

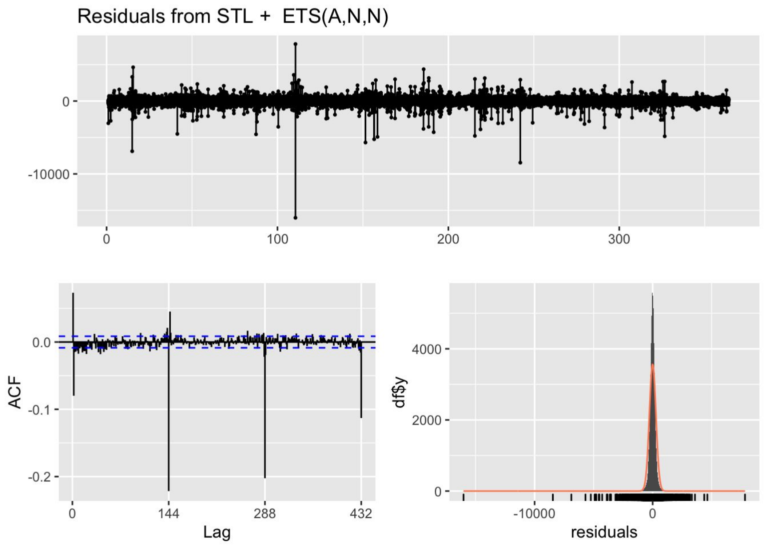
STLF

THE UNIVERSITY OF CHICAGO

Seasonal and Trend decomposition using Loess Forecasting model

Daily Frequency | Weekly Frequency



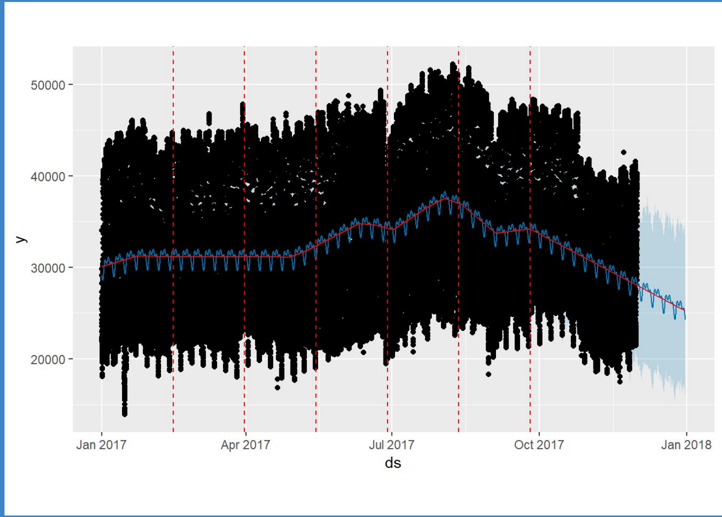


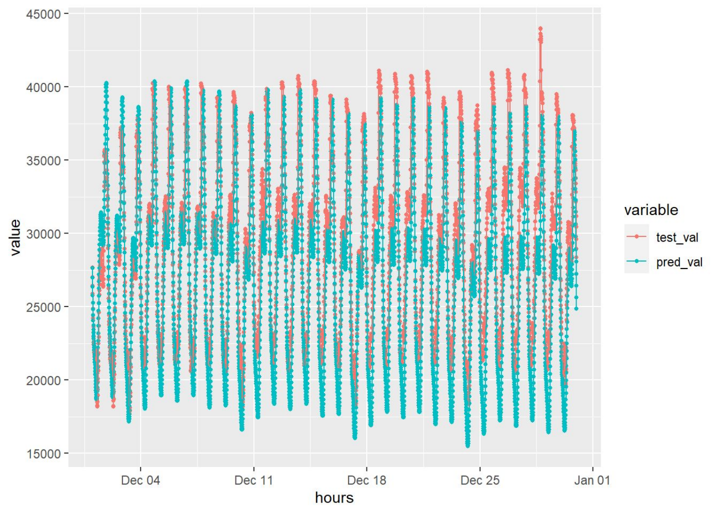
Best Performing: Daily Frequency
sMAPE: 0.06732304
Training time: 3.53 secs

- Residuals are close to normal distribution with a smaller standard deviation
- ACF: spikes at lag 0, 144, 288,432…

Prophet

Daily Frequency | Weekly Frequency | **Daily Frequency with modified number of change points**



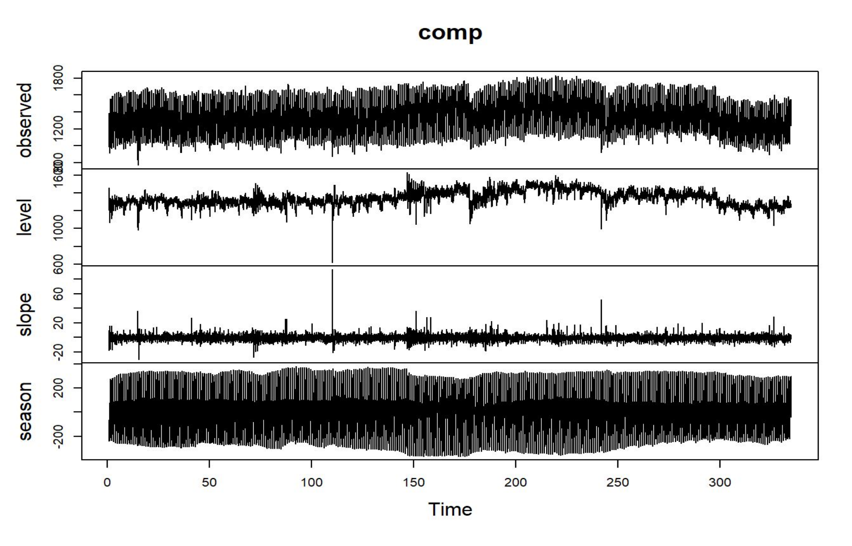


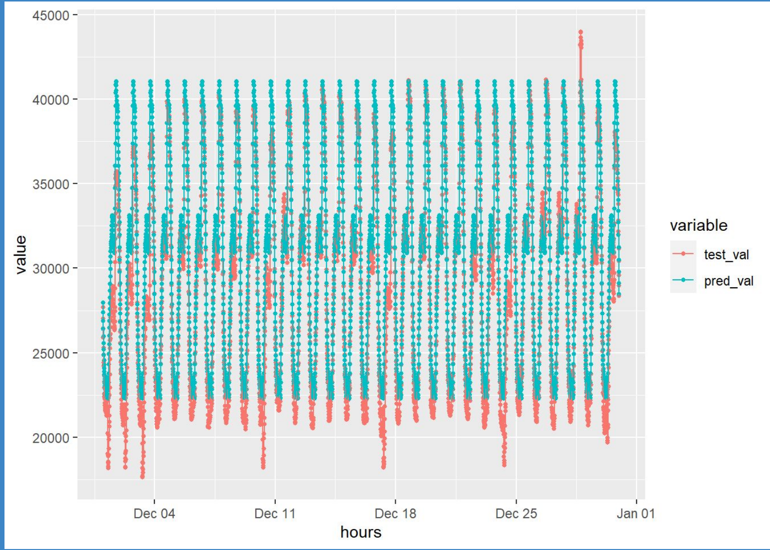
Best Performing: Daily Frequency / Change Points
sMAPE: 0.089074
Training time: 36.04 secs

##	Models	sMAPE	RMSE
## 1	Model daily	0.11210268	3694.453
## 2	Model Weekly	0.19683392	6700.145
## 3	Model daily with modified changepoints	0.08907442	3112.064

Tbats

Daily Frequency | Weekly Frequency



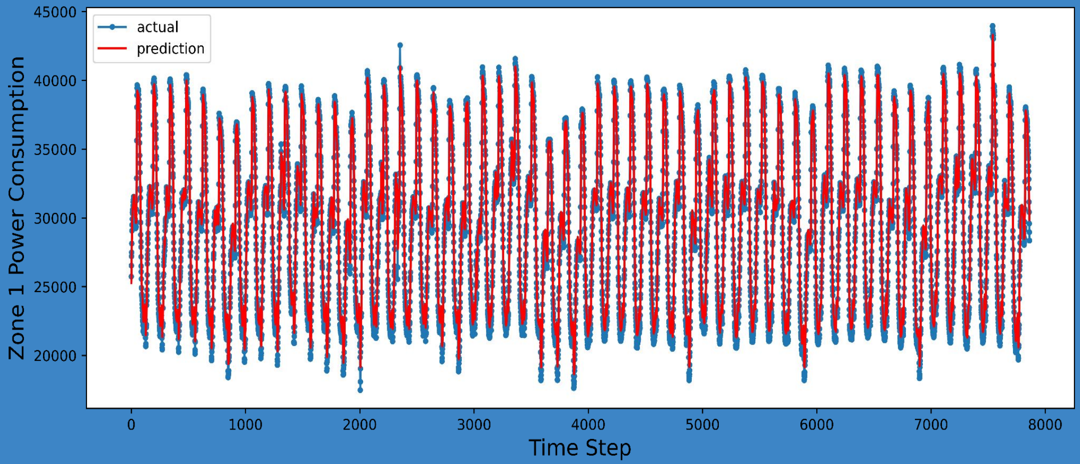


Best Performing: Daily Frequency / Change Points
sMAPE: 0.0465972
Training time: 421.14 secs

```
##          Models      sMAPE     RMSE
## 1  Model daily 0.04659725 3106.174
## 2 Model weekly 0.09198016 3106.174
```

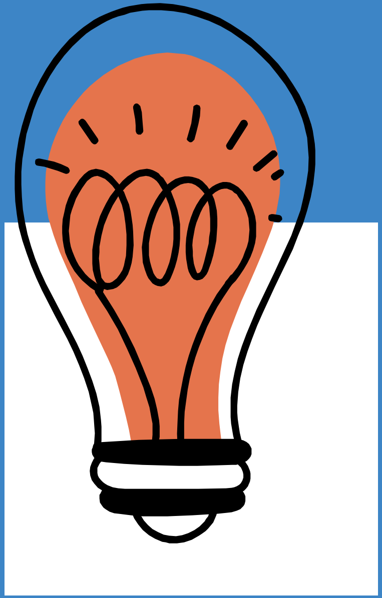
LSTM

1. Use MinMaxScaler()
2. Test-train split (lag=3/12/24 , test_ratio=0.15)
3. Implement LSTM model
4. Invert the transformation



	RMSE	sMAPE
Lag = 3	979.0666	0.02402
Lag = 12	1595.1529	0.03810
Lag = 24	1490.0596	0.03416

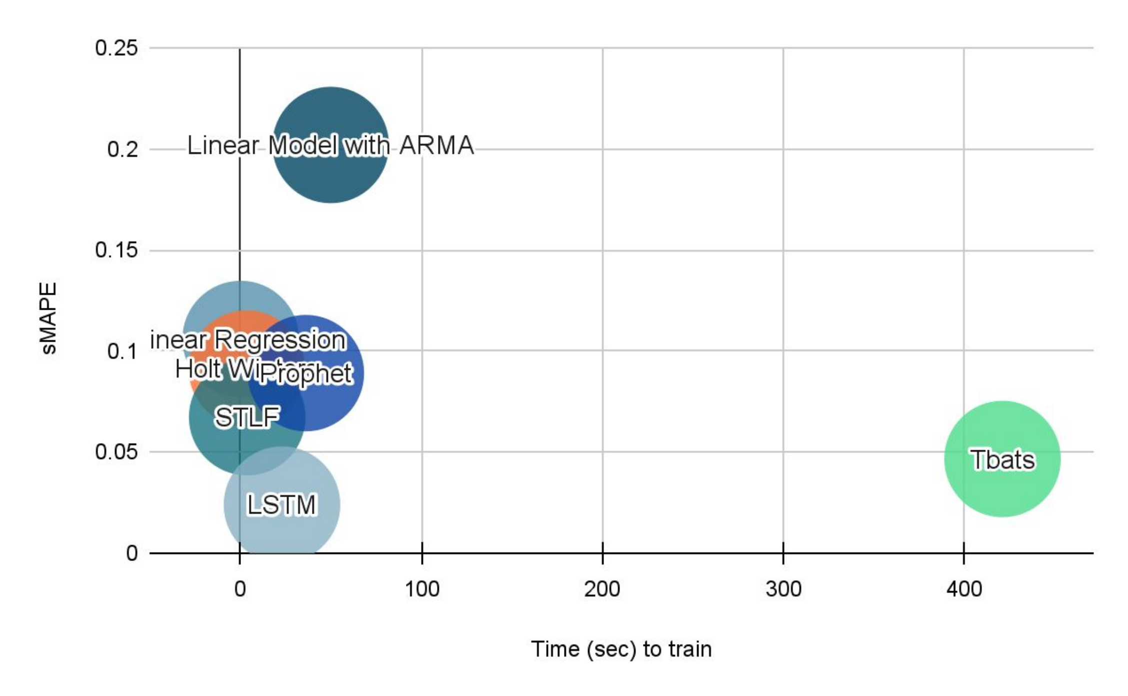
Best Performing: Lag = 3
sMAPE: 0.02402
Training time: 22.88 secs



05

Model Evaluation

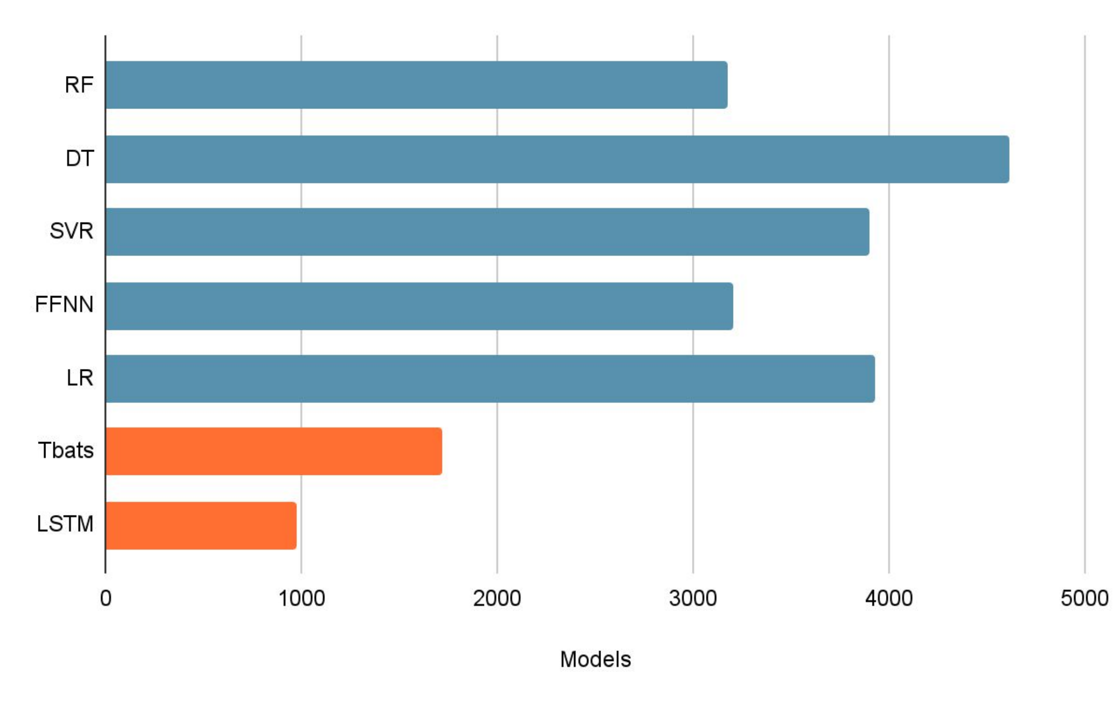
Model Comparison



	sMAPE	Time(sec)
Linear Regression	0.1060	0.04
Linear Model with ARIMA	0.2080	30.76
Holt Winters	0.0913	3.44
STLF	0.0673	3.53
Prophet	0.0891	36.04
Tbats	0.0466	421.14
LSTM	0.024	22.88

Forecasting vs supervised learning models

THE UNIVERSITY OF CHICAGO

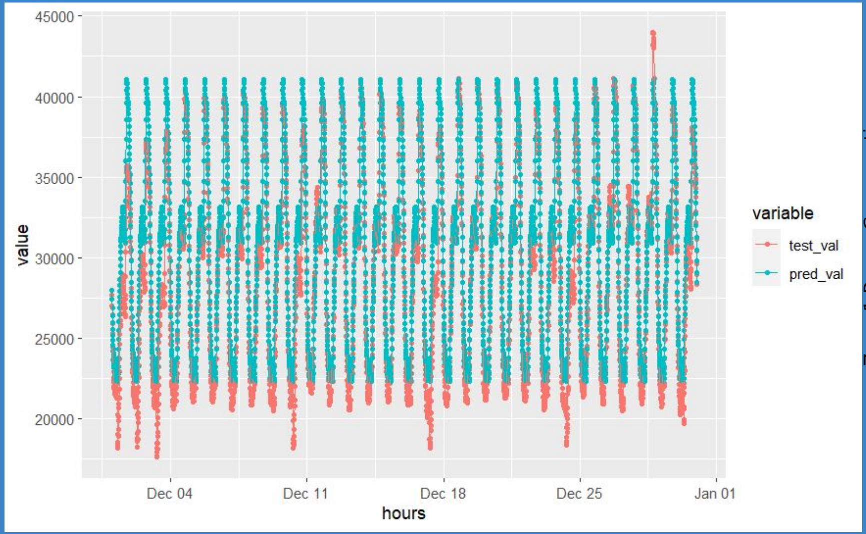


	RMSE
RF*	3174.7
DT*	4613.9
SVR*	3898.7
FFNN*	3203.6
LR*	3925.5
Tbats	1722.6
LSTM	979.1

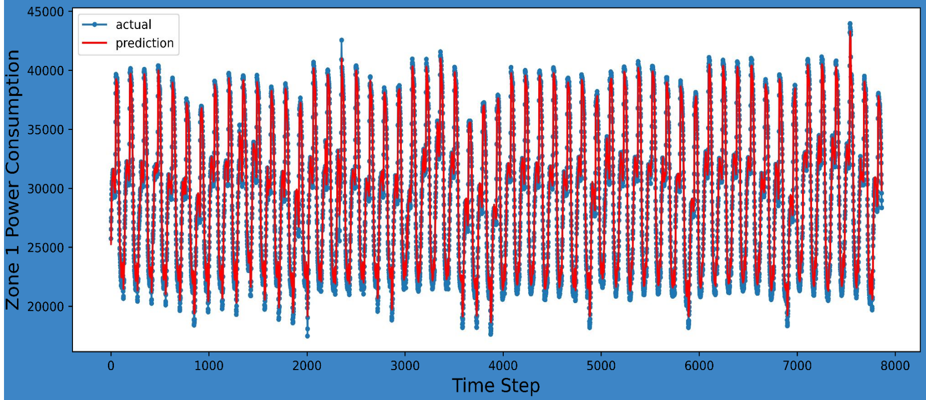
*source: https://ieeexplore.ieee.org/document/8703007

Model Forecasts

Actual vs predicted electricity power consumption of comparative models every 10 minutes



TBATS Model



LSTM Model

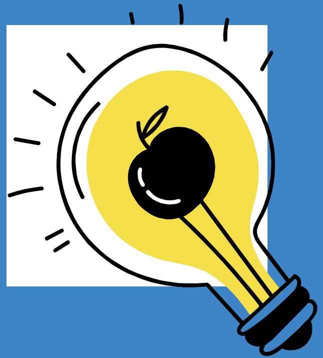
06

Potential Improvements

Potential next steps and improvements

THE UNIVERSITY OF CHICAGO

- Apply the same study to different power grids
- Financial study to measure the economic impact to configure power supply based on forecasts.
- Real time model performance and improvement



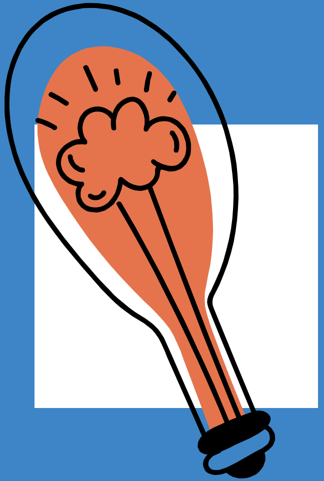
References

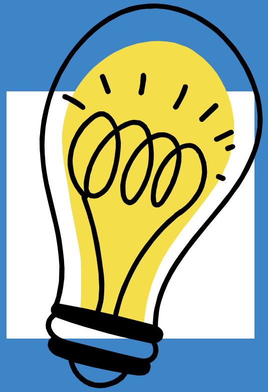
- https://archive.ics.uci.edu/ml/datasets/Power+consumption+of+Tetouan+city
- https://ieeexplore.ieee.org/document/8703007
- Prophet | Forecasting at scale. (facebook.github.io)
- Time Series Forecasting using TBATS Model | by Nadeem | Analytics Vidhya | Medium

Link to our Github repository:

- https://github.com/martincopello/power_consumption_time_series

Thank you!
Q & A





Appendix

Correlation Plot with added features