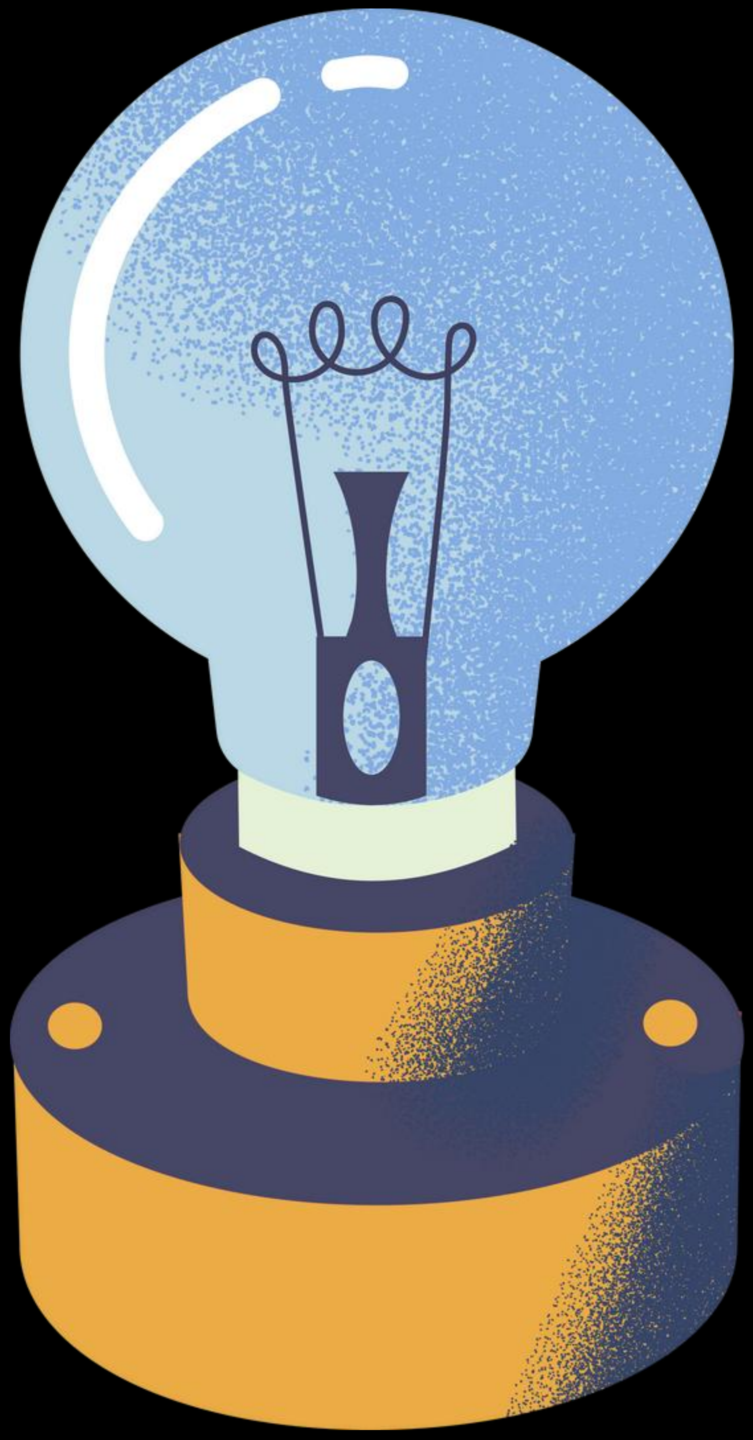
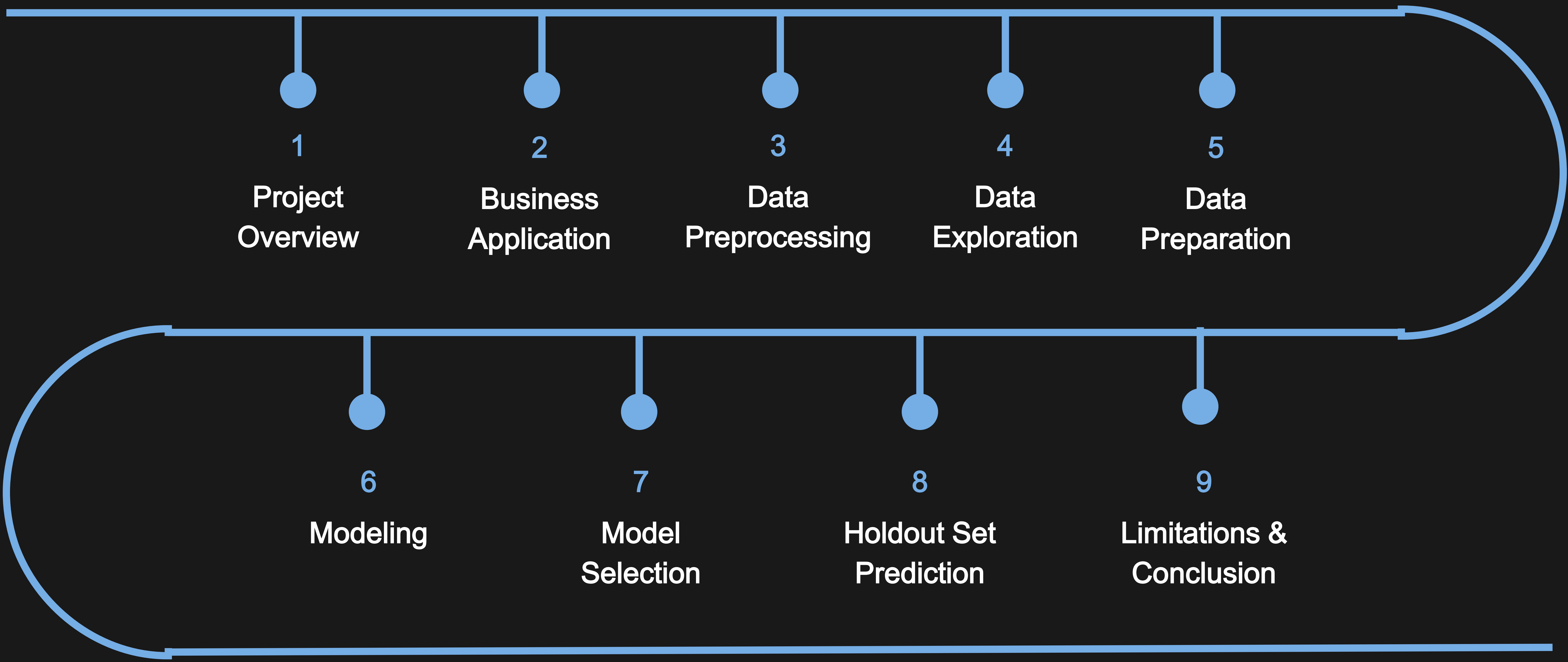
Residential Electricity Usage Prediction

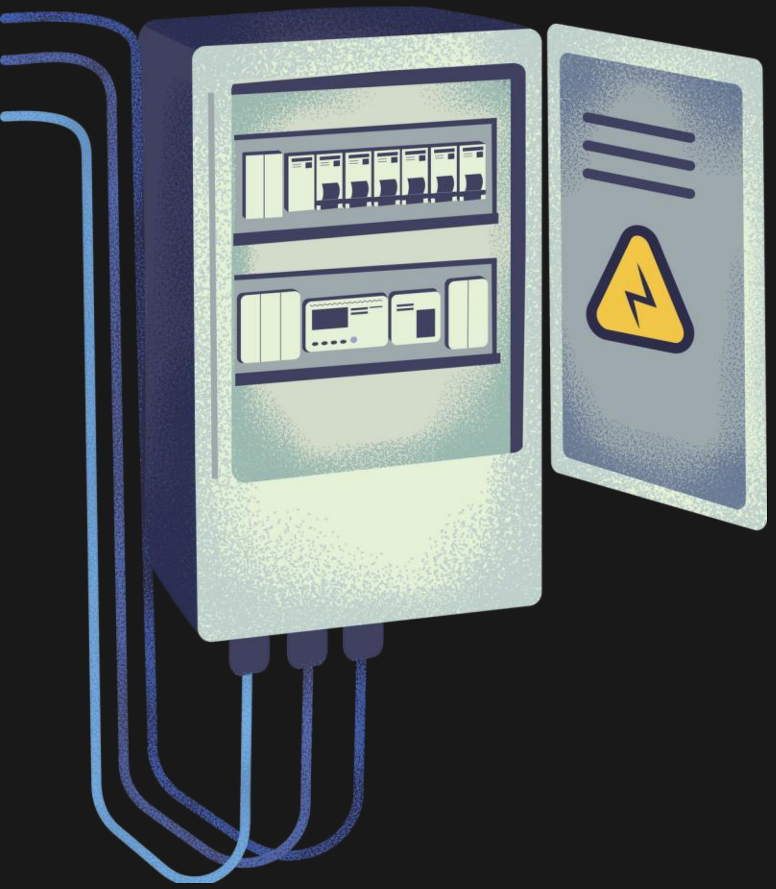


Maralmaa Batnasan

DATA SCIENCE LIFECYCLE



1 PROJECT OVERVIEW



This project focuses on forecasting household electricity usage. The approach combines feature engineering, time series modeling, and an ensemble machine learning strategy to deliver accurate and actionable predictions.

2

BUSINESS APPLICATION

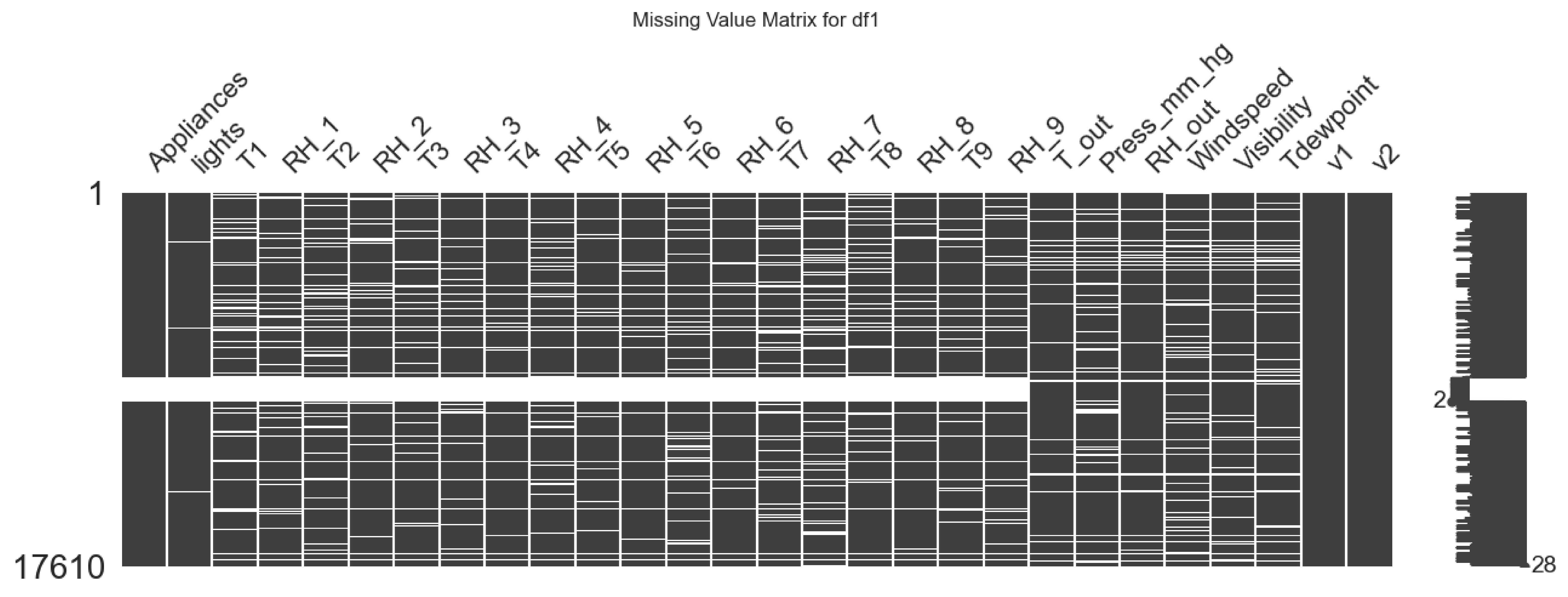
Business Applications:

- Demand Response & Grid Optimization
- Investment Risk Assessment
- Renewable Energy Integration
- Dynamic Pricing
- Personalized Energy Report
- Policy & Planning

Risks of Applying Machine Learning to Time Series Data:

- Temporal Leakage
- Error Accumulation
- Overfitting

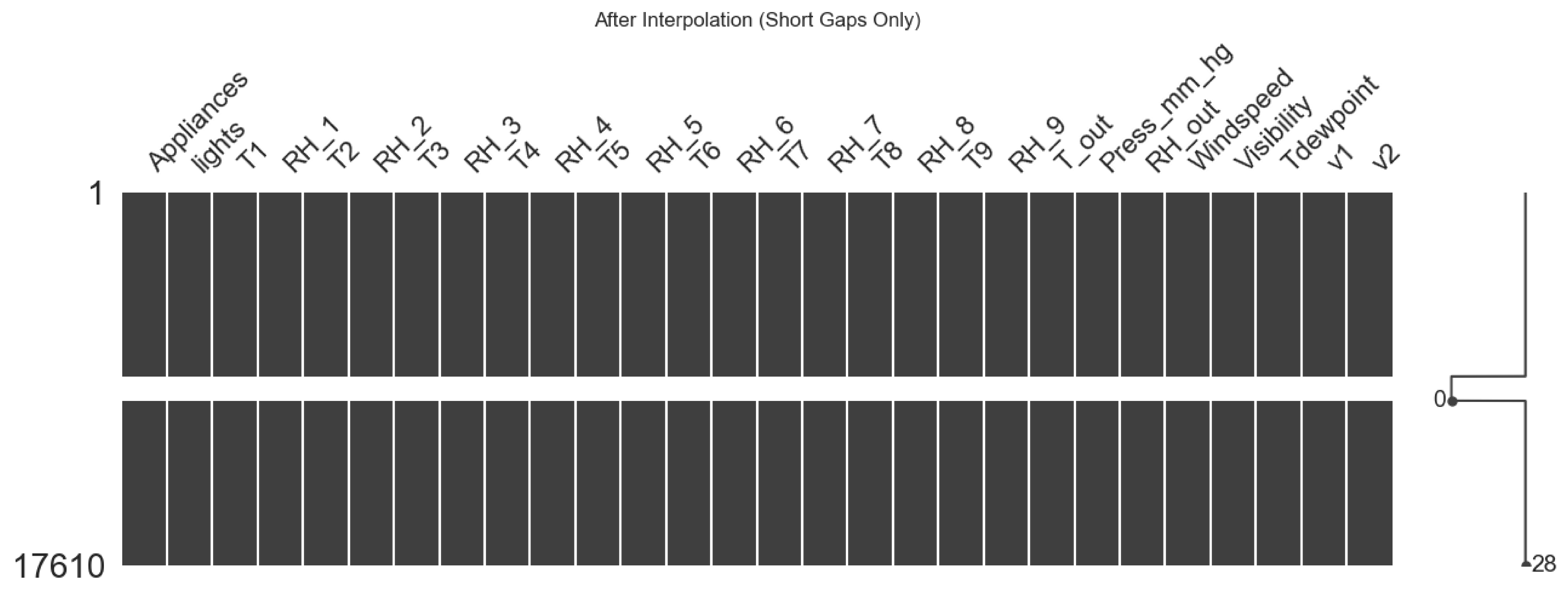
3 DATA PREPROCESSING



- Mark March 12–19 as fully missing to prevent inaccurate lags
- Interpolate short gaps (≤ 1 hour) using linear interpolation.

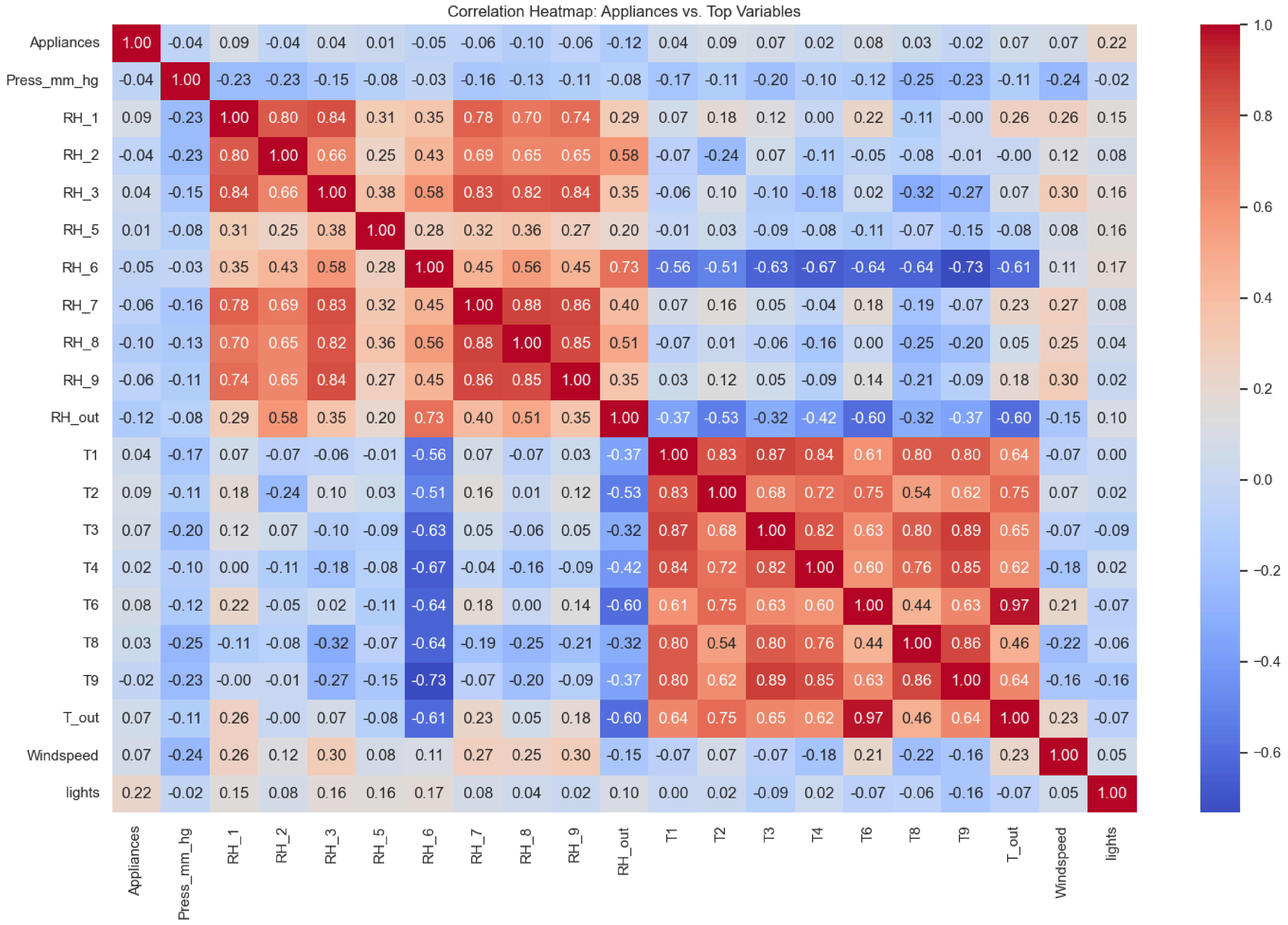
3

DATA PREPROCESSING

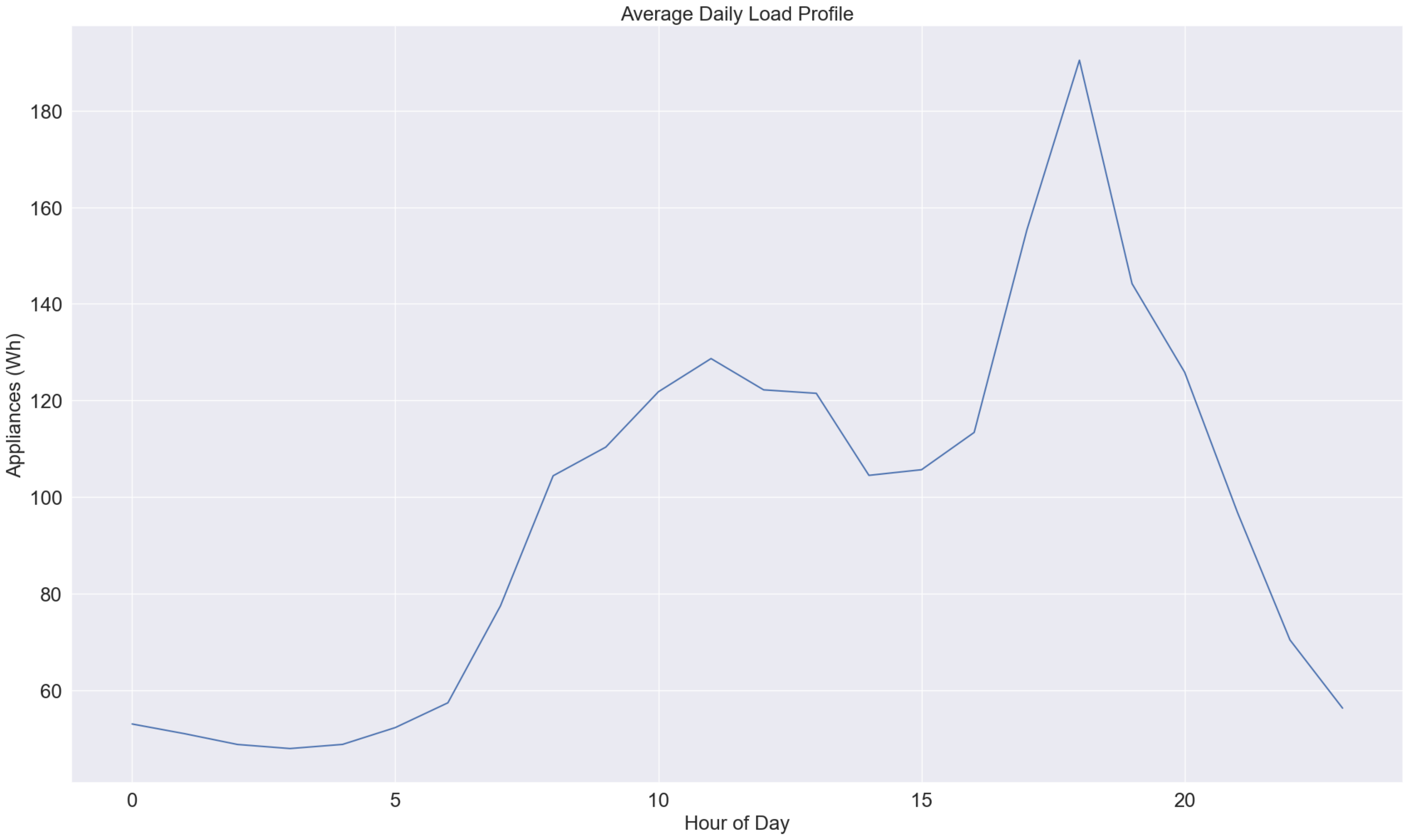


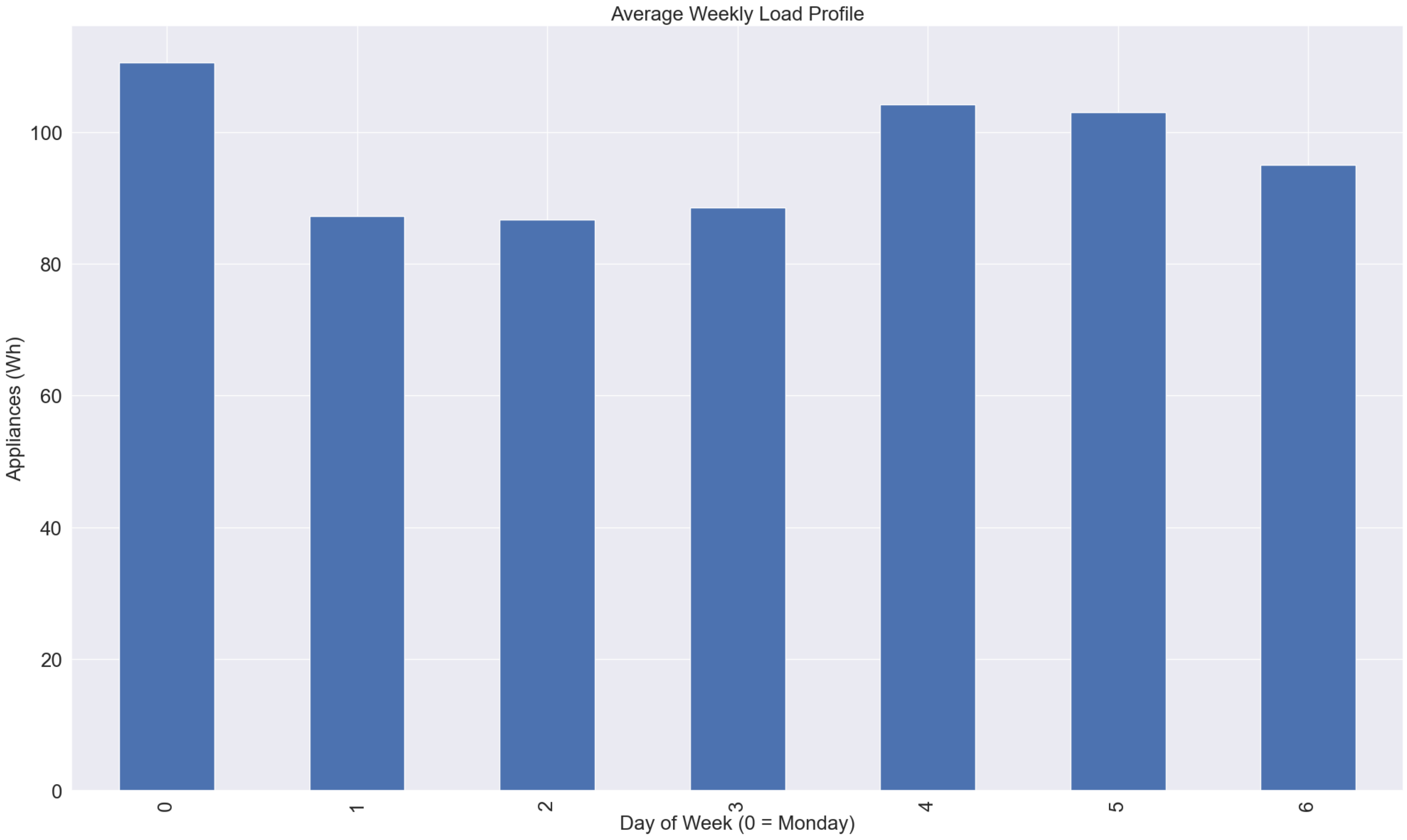
4

DATA EXPLORATION

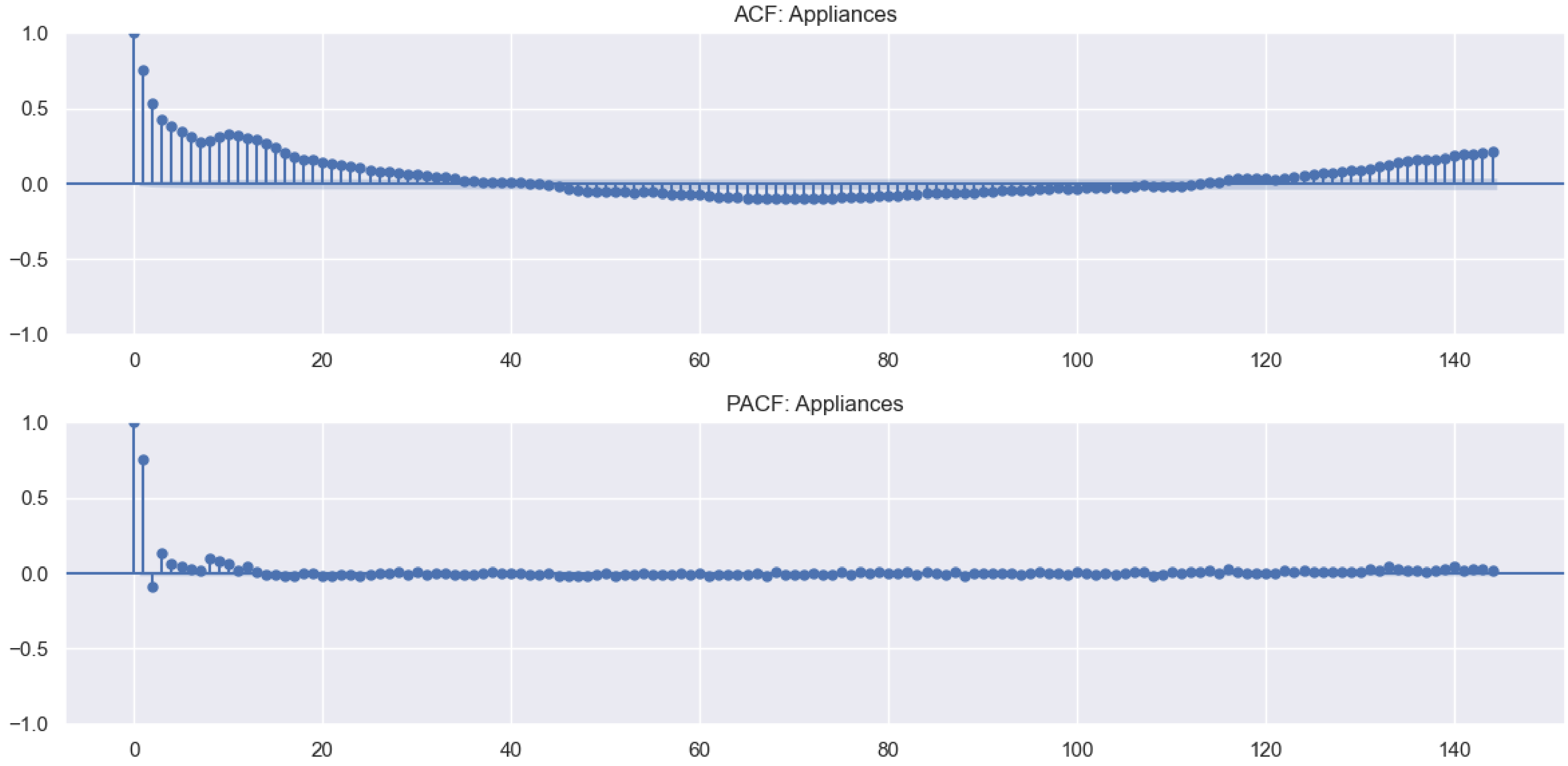


4 DATA EXPLORATION



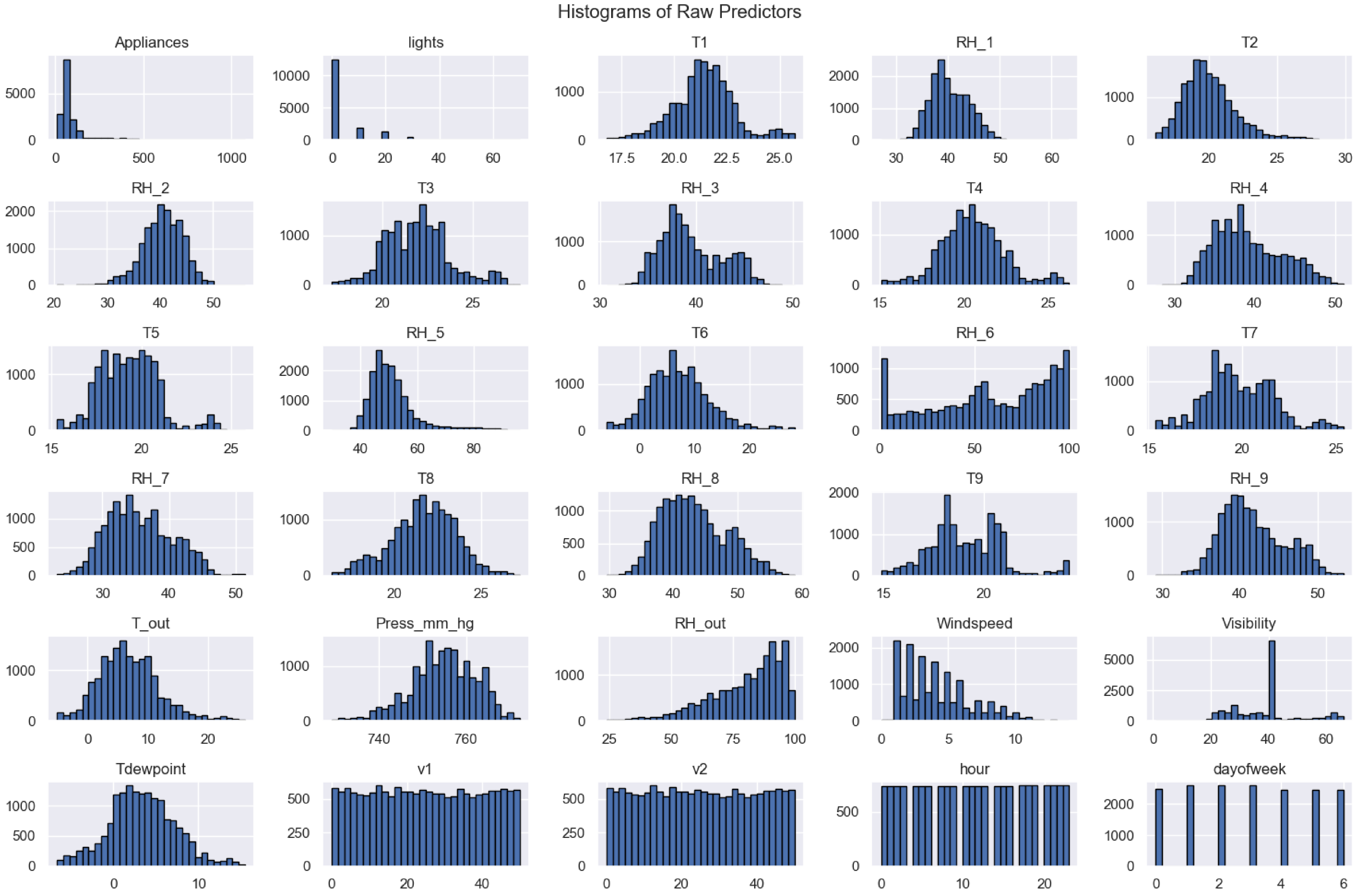


4 DATA EXPLORATION



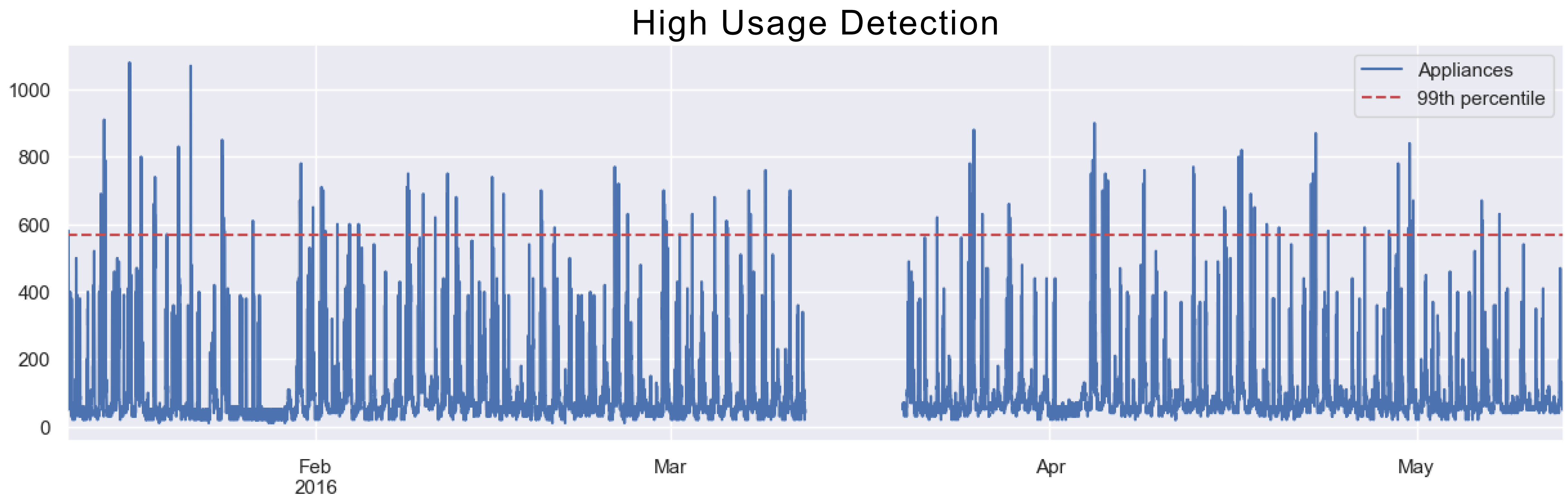
4

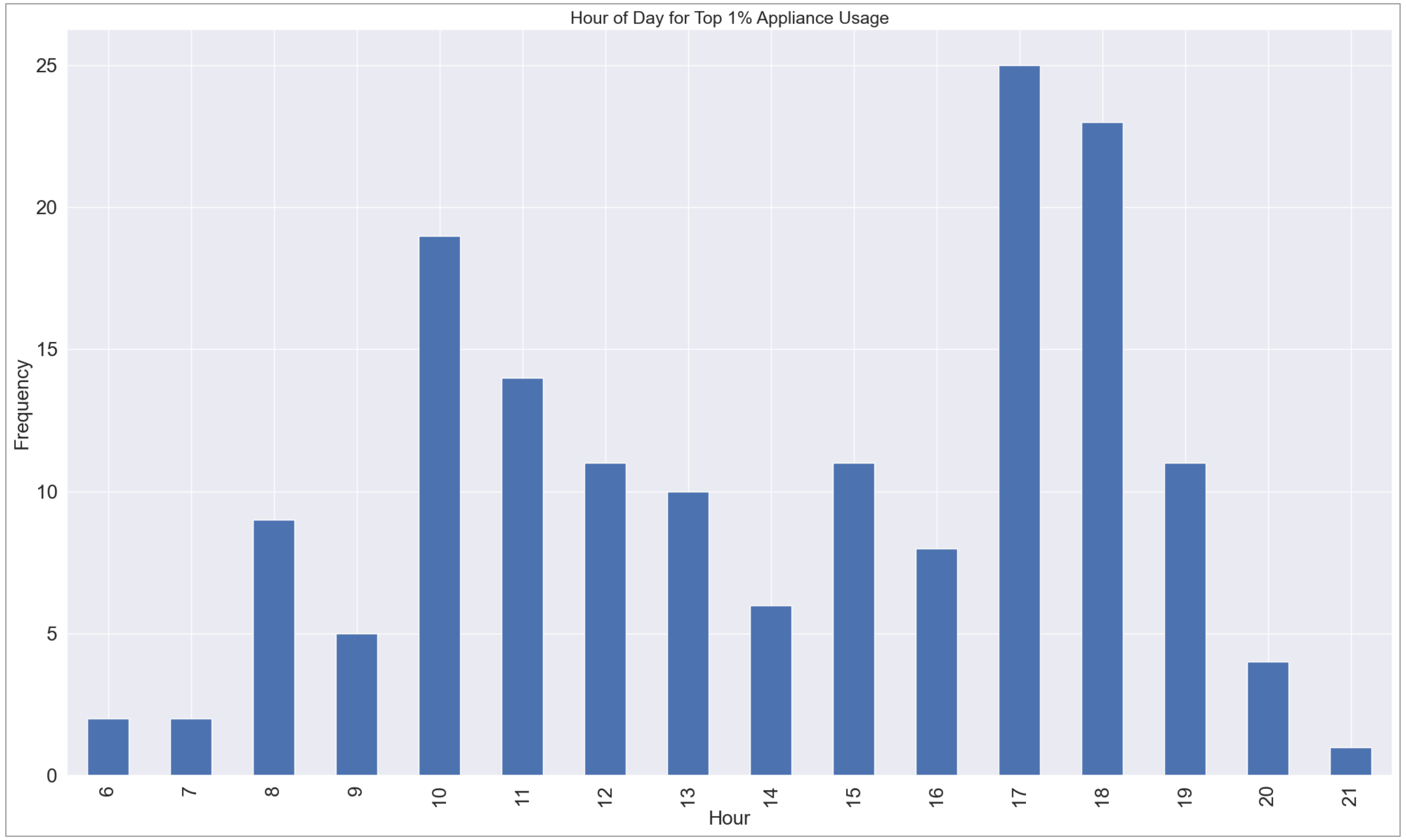
DATA EXPLORATION



4

DATA EXPLORATION - OUTLIER





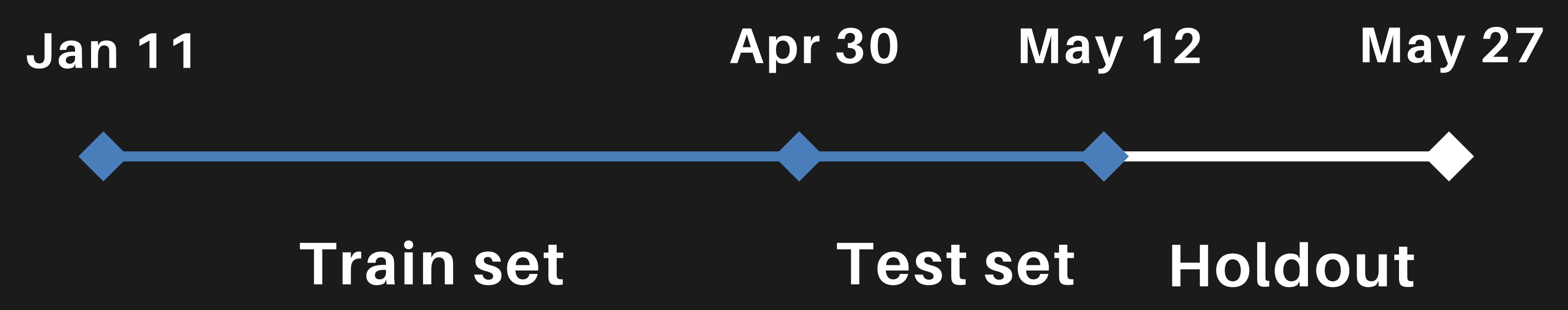
5 DATA PREPARATION

KEY FEATURES
• **External variables** (temperature, humidity, windspeed, etc.) • **Time-based features** (hour, day of week, weekend indicator, peak hour indicator, hour_sin, hour_cos) • **Interaction features** (e.g., temperature × hour, windspeed × weekend) • **Rolling and lag features** on key predictors and target

	FEATURE TYPE	TIME WINDOWS
Appliances	Lag features (to capture autoregressive effects)	1 to 144 steps, i.e., 10 mins to 1 day
	Rolling statistics (shifted to avoid data leakage)	over 1-hour and 3-hour windows
(T1, RH_1, T_out, RH_out etc.)	Lag features	1, 6 steps, i.e., 10-60 mins
	Rolling statistics	over 1-hour and 3-hour windows

6 MODELLING

LightGBM | **TimesFM** | **LSTM**



LightGBM

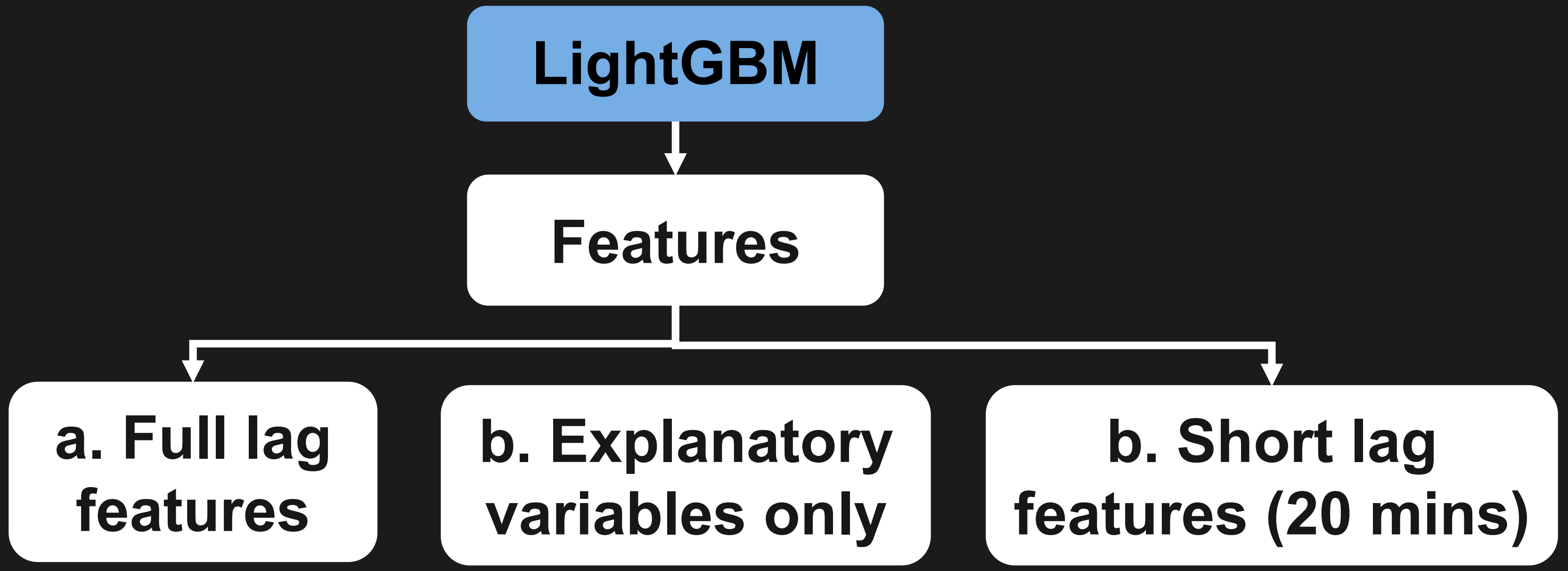
↓

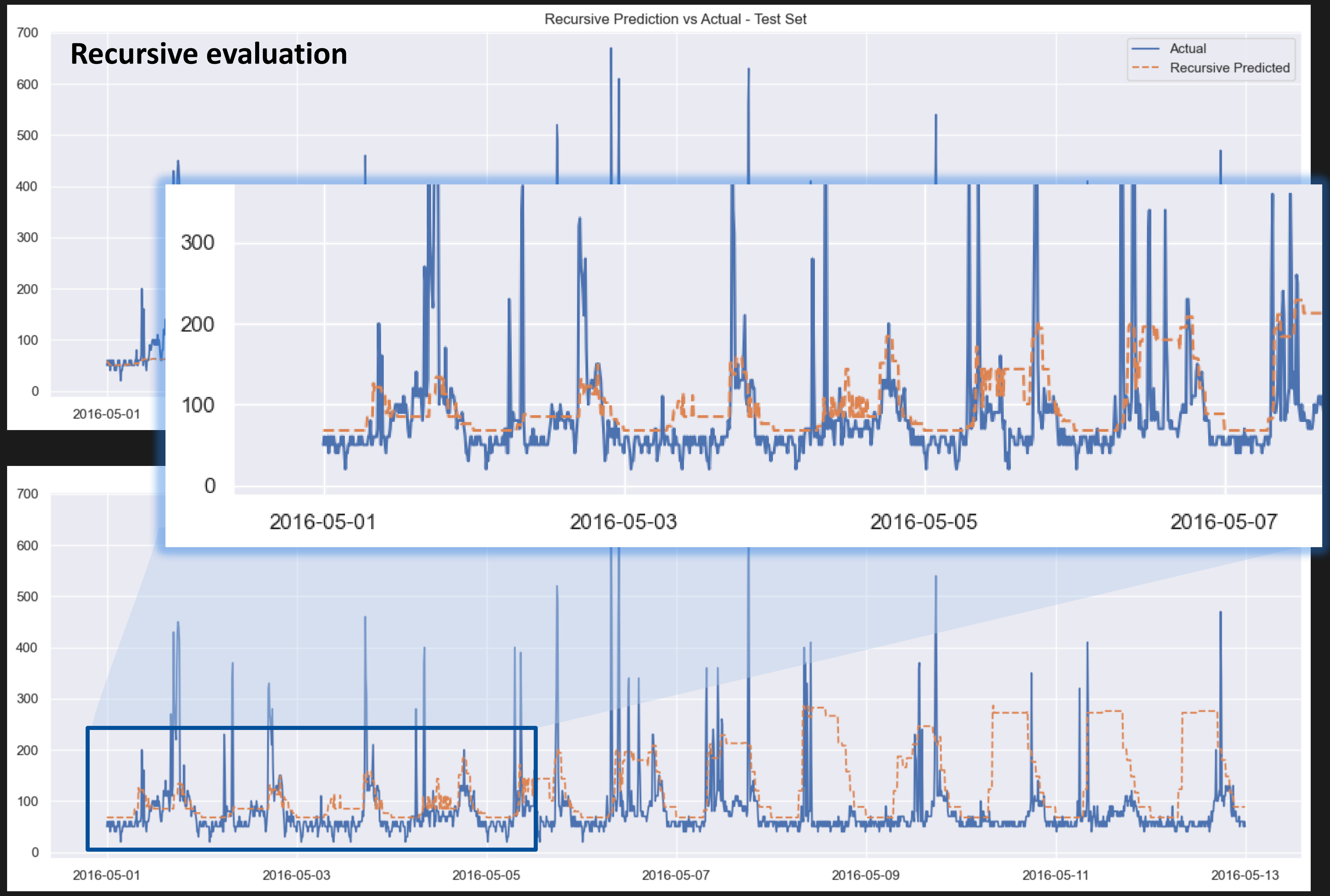
Features

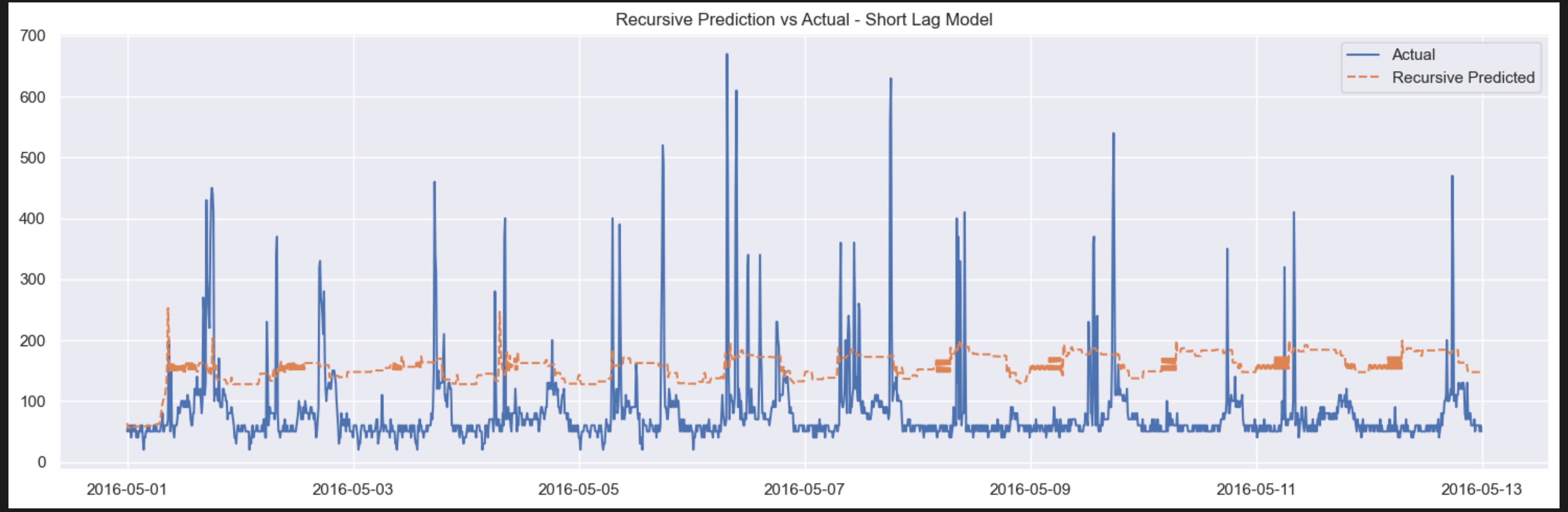
- a. Full lag features
- b. Explanatory variables only
- c. Short lag features (20 mins)

6 MODELLING

	Test MAE	Test RMSE
a. Full lag features	41.195	72.869
b. Explanatory variables only	65.805	100.258
b. Short lag features (20 mins)	89.43	98.62







7 MODEL SELECTION

Ensemble approach (only on explanatory variables)

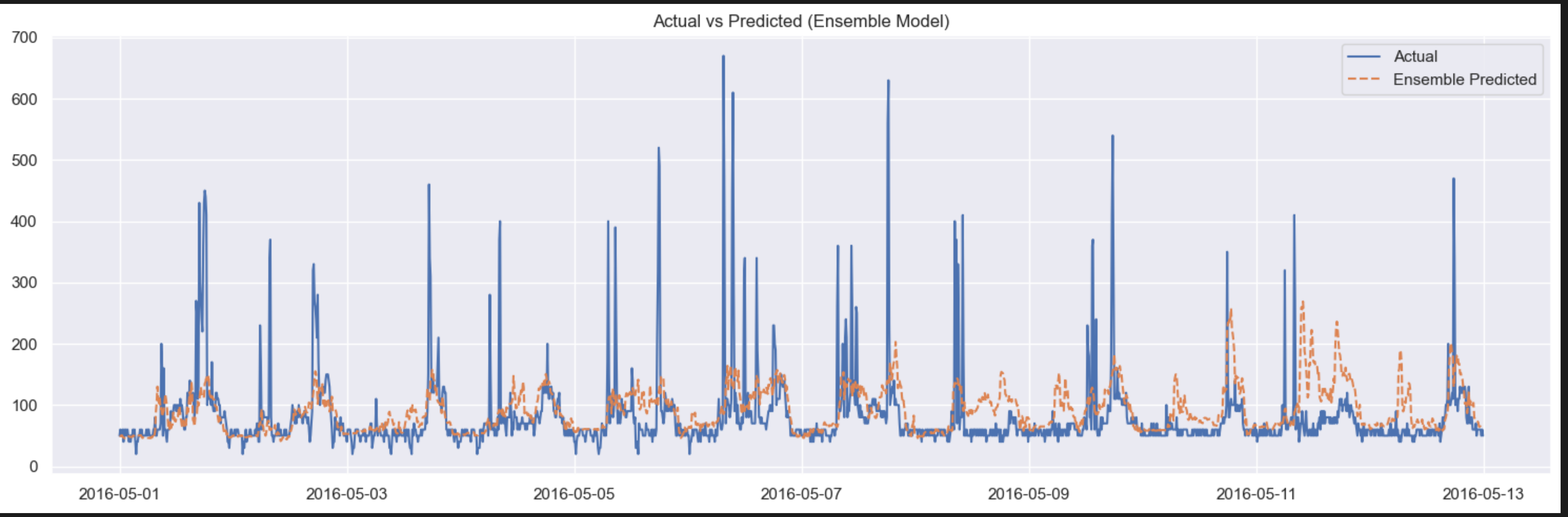
Peak Classifier

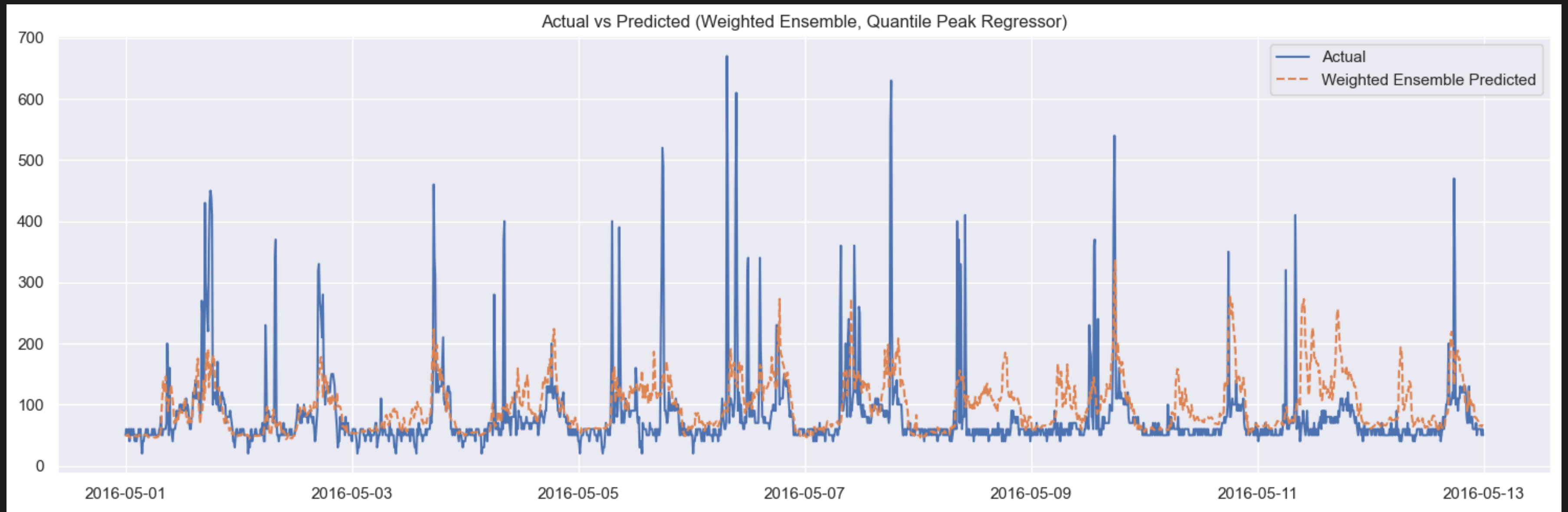
"Non-Peak"
Regressor for non-peak periods

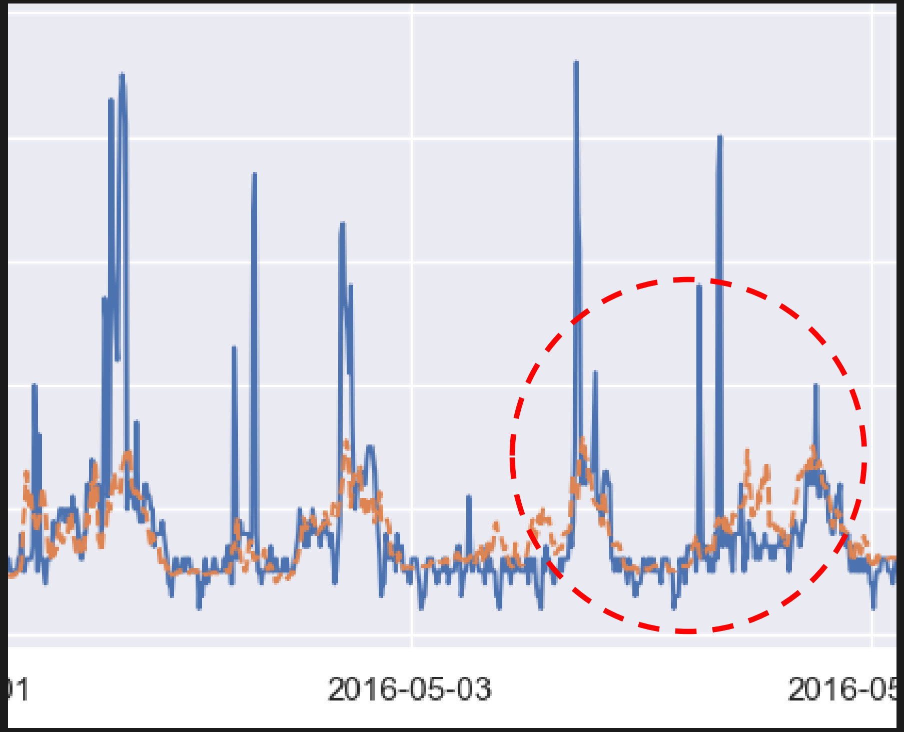
"Peak"
Regressor for peak periods

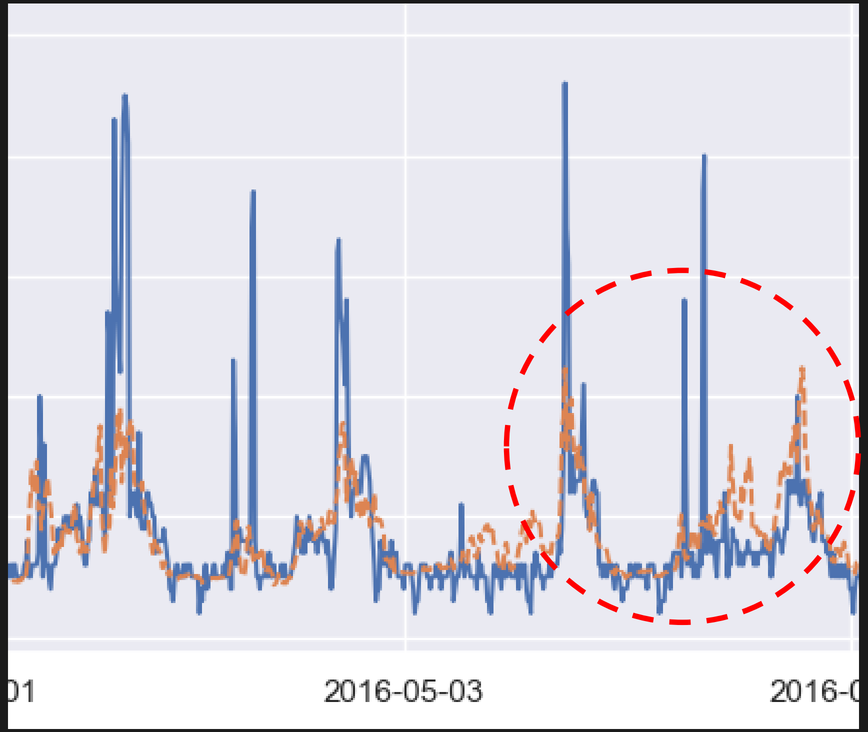
Final Prediction
Weighted average of the two regressors

	Test MAE	Test RMSE
a. Default	34.42	63.50
b. Quantile Regression	37.34	64.22
b. Higher quantile (alpha=0.98)	39.91	64.86

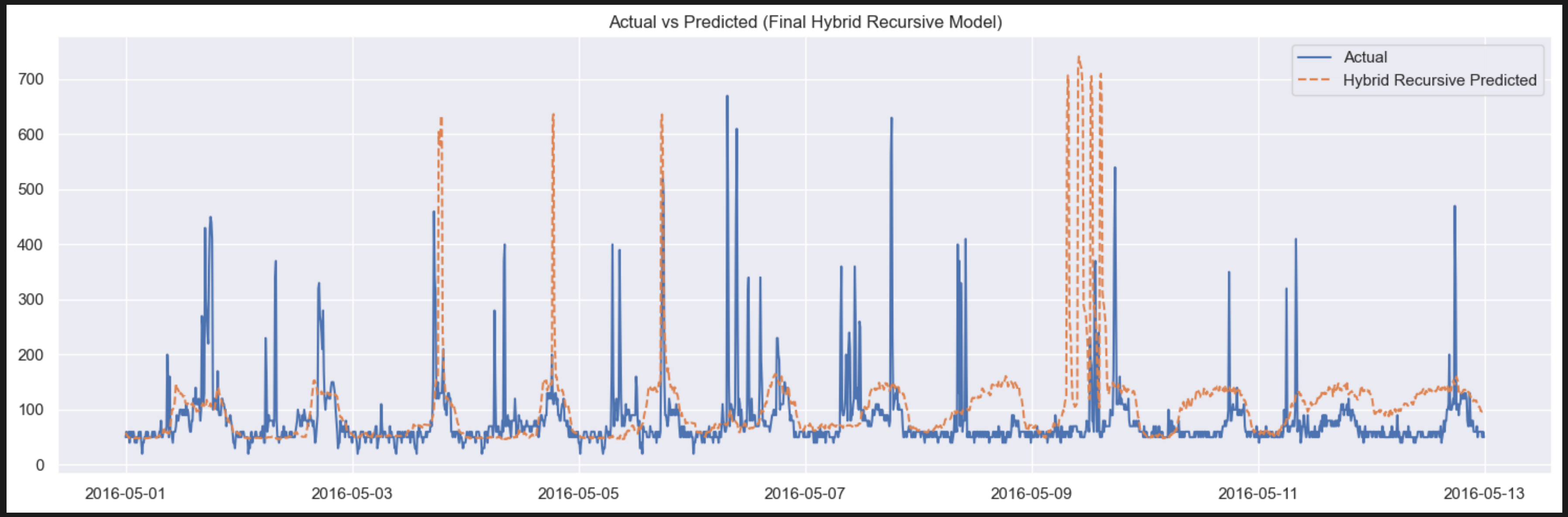


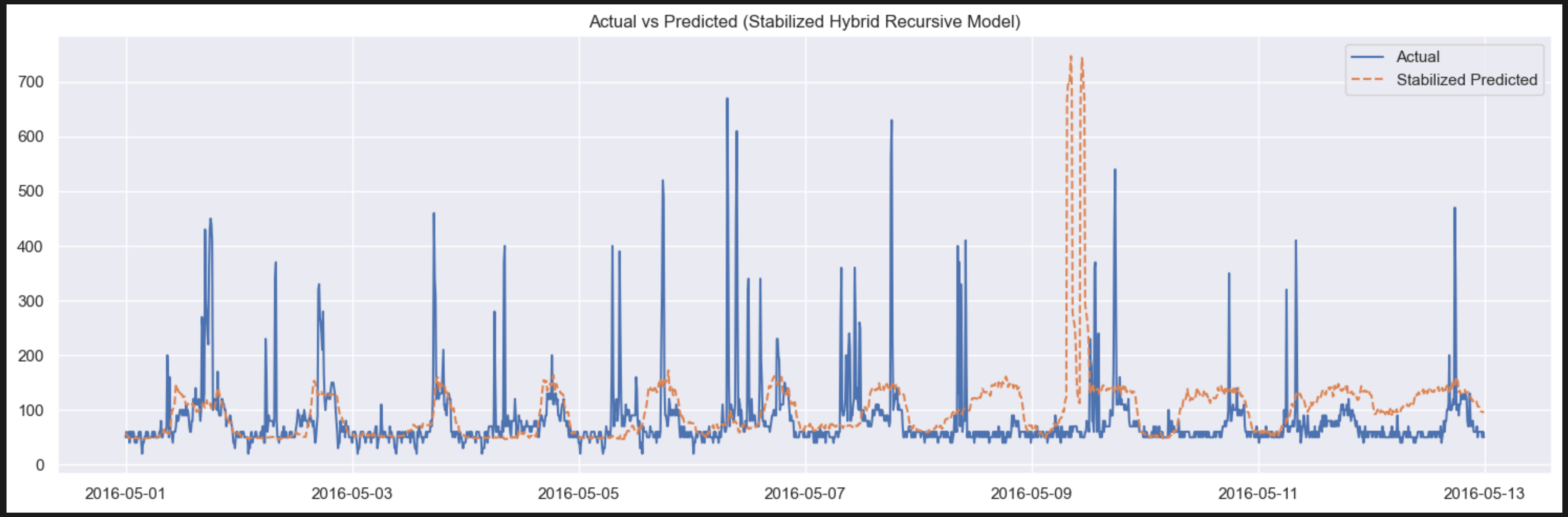




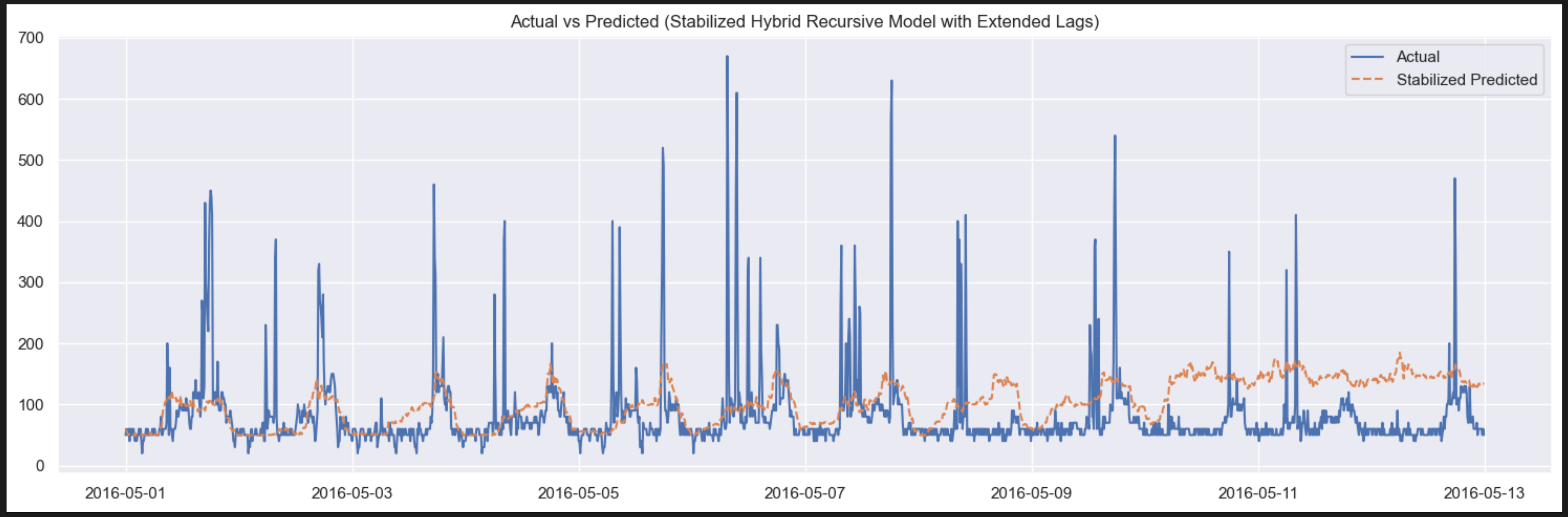


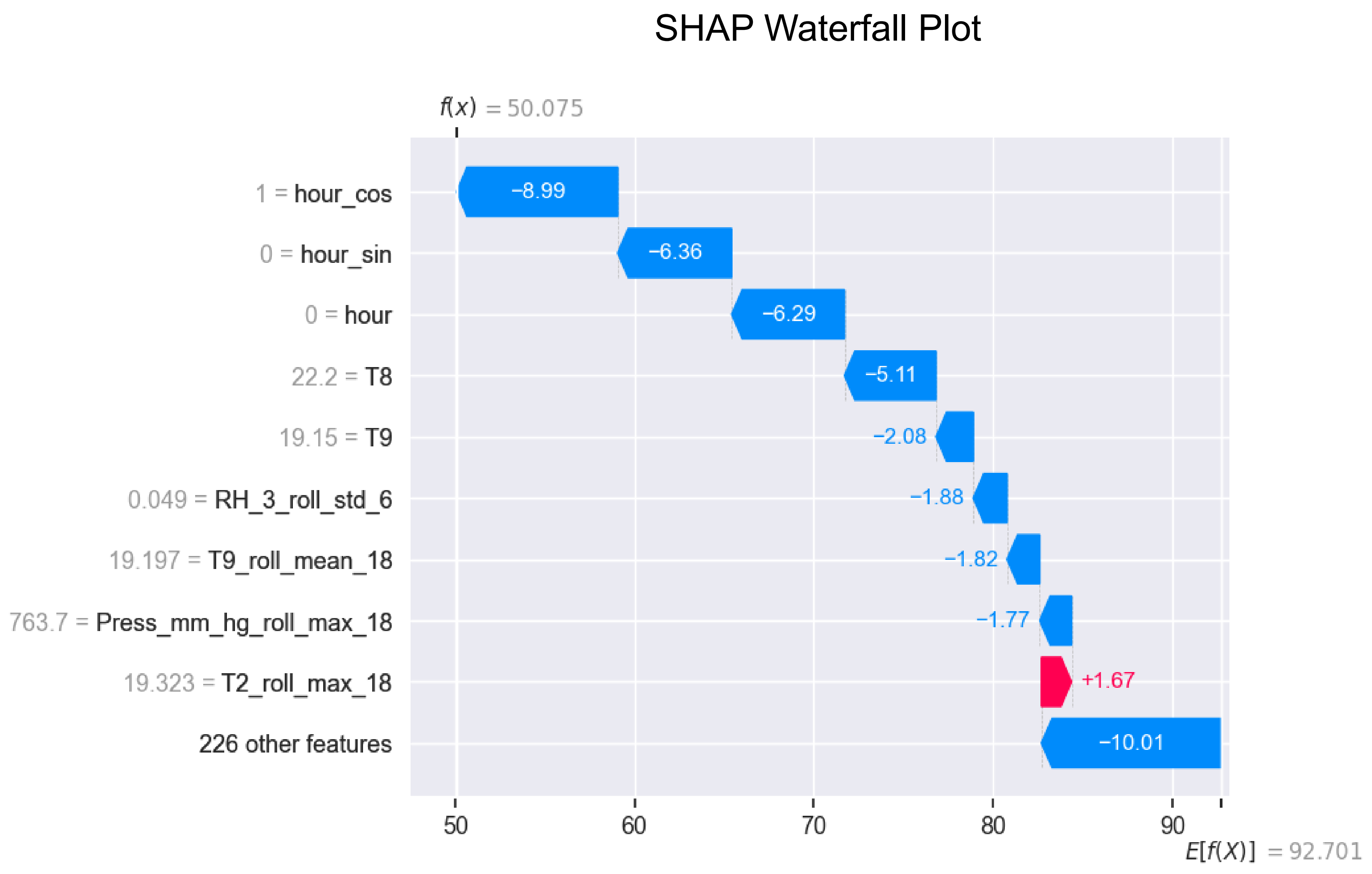
7 MODEL SELECTION

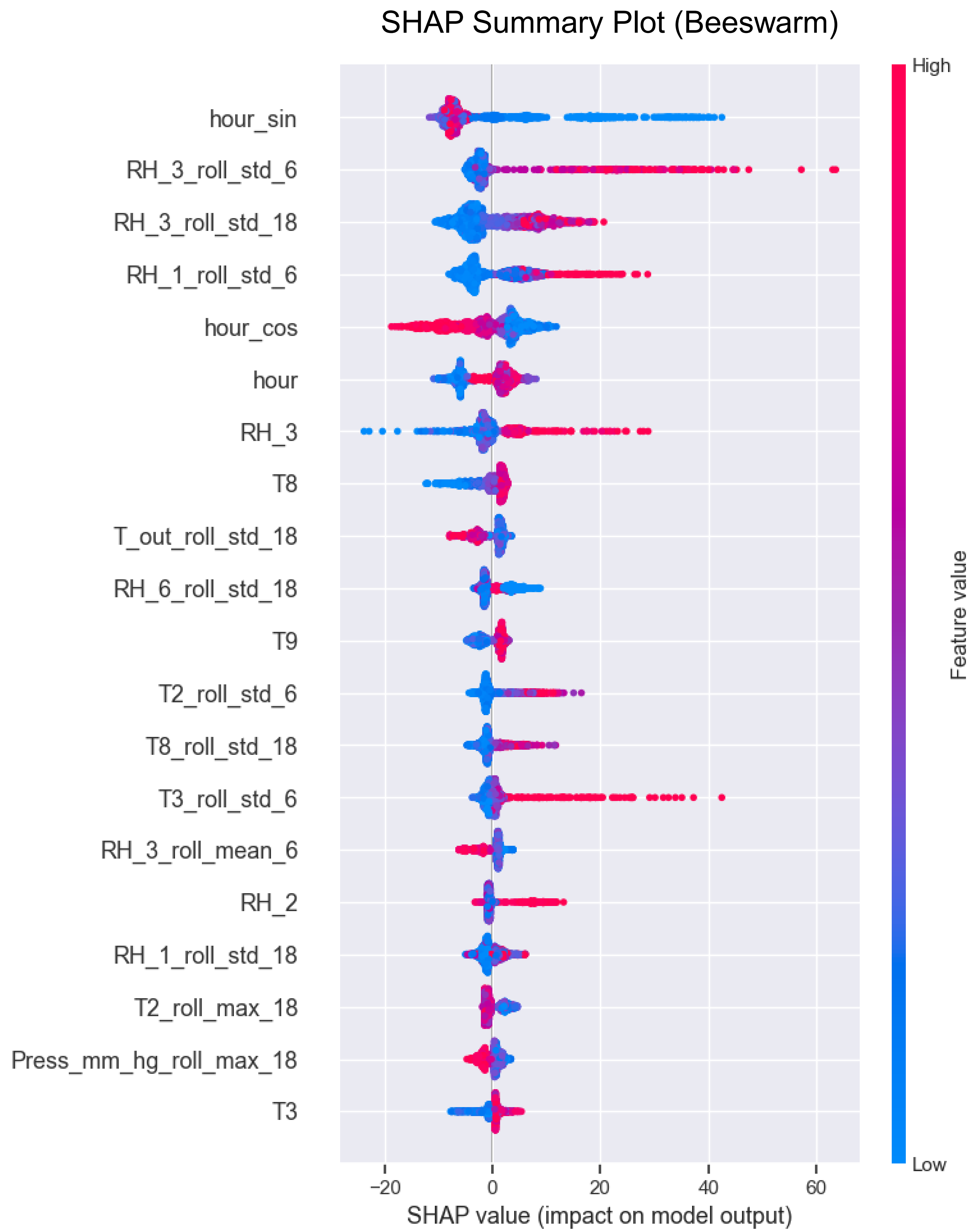




Trials: Appliance lags added back

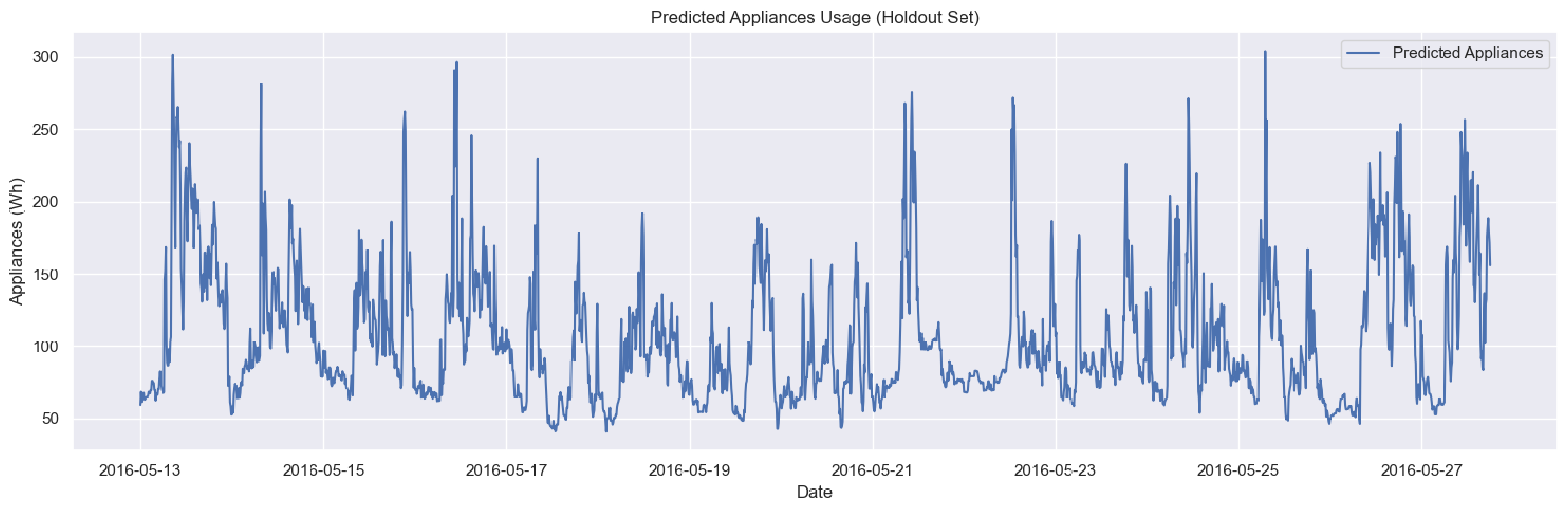



SHAP Waterfall Plot
f(x) = 50.075
1 = hour_cos
0 = hour_sin
0 = hour
22.2 = T8
19.15 = T9
0.049 = RH_3_roll_std_6
19.197 = T9_roll_mean_18
763.7 = Press_mm_hg_roll_max_18
19.323 = T2_roll_max_18
226 other features
−8.99
−6.36
−6.29
−5.11
−2.08
−1.88
−1.82
−1.77
+1.67
−10.01
50
60
70
80
90
E[f(X)] = 92.701


SHAP Summary Plot (Beeswarm)
hour_sin
RH_3_roll_std_6
RH_3_roll_std_18
RH_1_roll_std_6
hour_cos
hour
RH_3
T8
T_out_roll_std_18
RH_6_roll_std_18
T9
T2_roll_std_6
T8_roll_std_18
T3_roll_std_6
RH_3_roll_mean_6
RH_2
RH_1_roll_std_18
T2_roll_max_18
Press_mm_hg_roll_max_18
T3
−20
0
20
40
60
SHAP value (impact on model output)
High
Low
Feature value

8

HOLDOUT SET PREDICTION



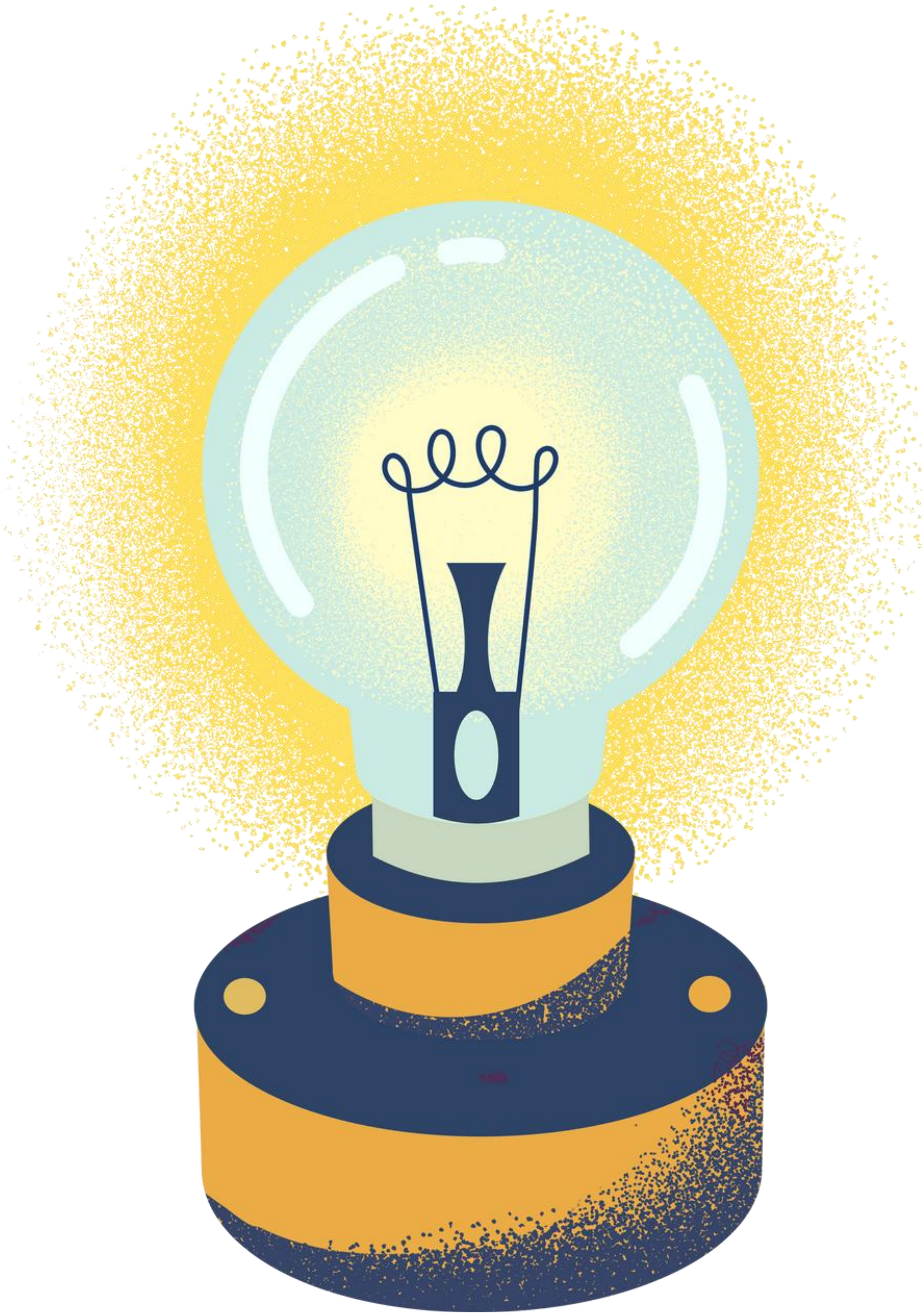
9 LIMITATIONS & CONCLUSION

- For high-demand periods where actual usage exceeded 500–700 Wh, model predictions remained significantly lower.
- Post-prediction adjustment could be done
- Alternative architectures (e.g., sequence models – further improvement on TimesFM and LSTM)

KEY TAKEAWAYS:

- Recursive evaluation is critical for honest assessment of time series models using lagged targets.
- Ensemble approaches (with peak-aware logic) can improve both accuracy and robustness.
- Consistent feature engineering and careful handling of missing data are essential for reliable deployment.

Thank you



Maralmaa Batnasan