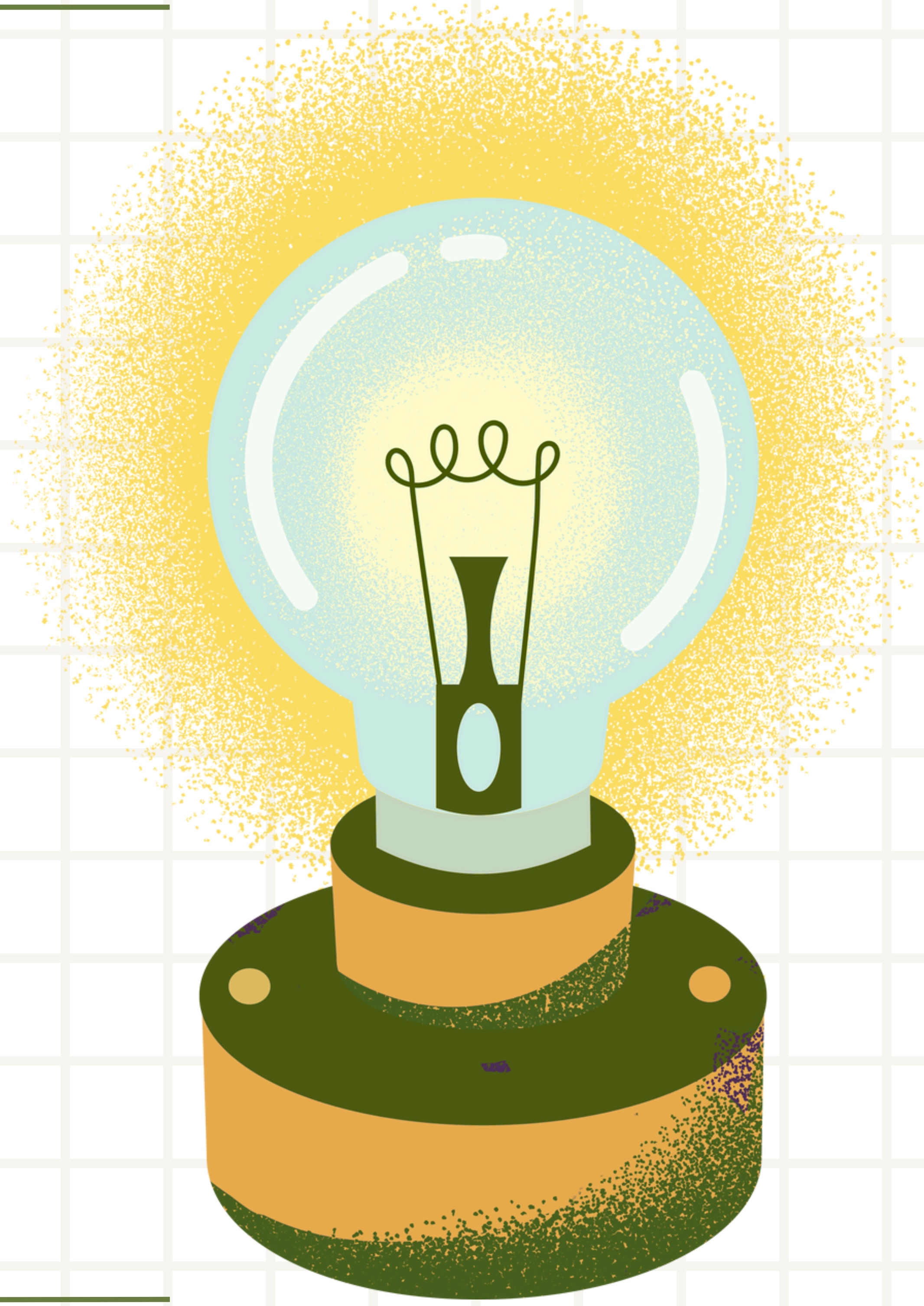
CentraleSupélec, 2025

Household energy consumption forecasting

in the UK

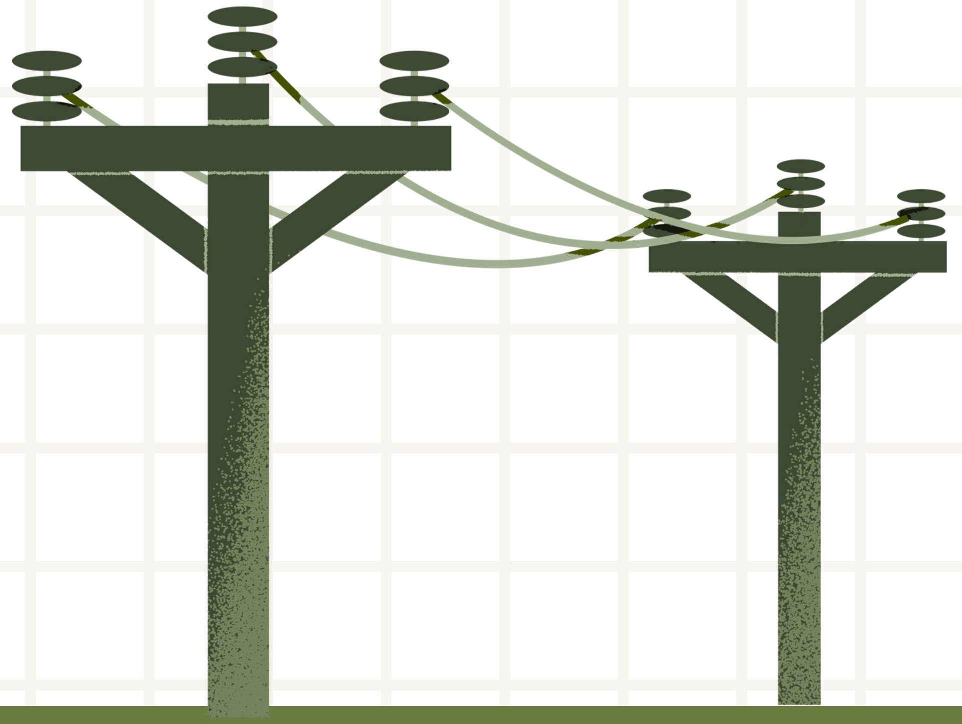
EI - ST4 - Énergie et Climat



Group 4

Gustavo C. NOVICKI, Mariana E. F. MENEZES, Thiago N. C. ARAUJO, Alexsandro S. ROSA, Martin SCHULZ.

SUMMARY



BACKGROUND

UK domestic energy sector faces increasing pressures

RISING DEMAND

CHANGING WEATHER PATTERNS

policy targets for EFFICENCY and DECARBONIZATION

ACORN

Segmenting households by socio-demographic groups allows differentiated modeling of consumption profiles.

DATASET OVERVIEW

ACORN - segmented time-series of eletricity consumption

Training: half-hourly electricity consumption - 2012 - 2013

Target Forecasting windows: January - February 2014

Synchronized meteorological variables and holidays

TEMPERATURE + HUMIDITY + WIND + HOLIDAYS

MOTIVATION

Provide utilities and grid operators with accurate resolutions for planning, demand response, and reliability.

OBJECTIVES

PHASE 1
ANALYSIS & CHARACTERIZATION

PHASE 3
MEDIUM -TERM FORECASTING

PHASE 2
SHORT -TERM FORECASTING

PHASE 4
INTERACTIVE DASHBOARD

1. Introduction | 2. Phase 1 | 3. Phase 2 | 4. Phase 3 | 5. Phase 4

PHASE 1

Exploratory data analysis

- Daily / weekly/ seasonal patterns
- Autocorrelation structures
- Quantify weather impacts via **correlation analysis.**

PHASE 2

Development of a model

- 30-minute consumption - 13 - 14/01/2014
- Train on pre-January 2014 data
- Validate and evaluate

PHASE 3

Development of a model

- Historical daily aggregates
- Weather covariates.
- Evaluate the model

PHASE 4

Implement a Streamlit dashboard

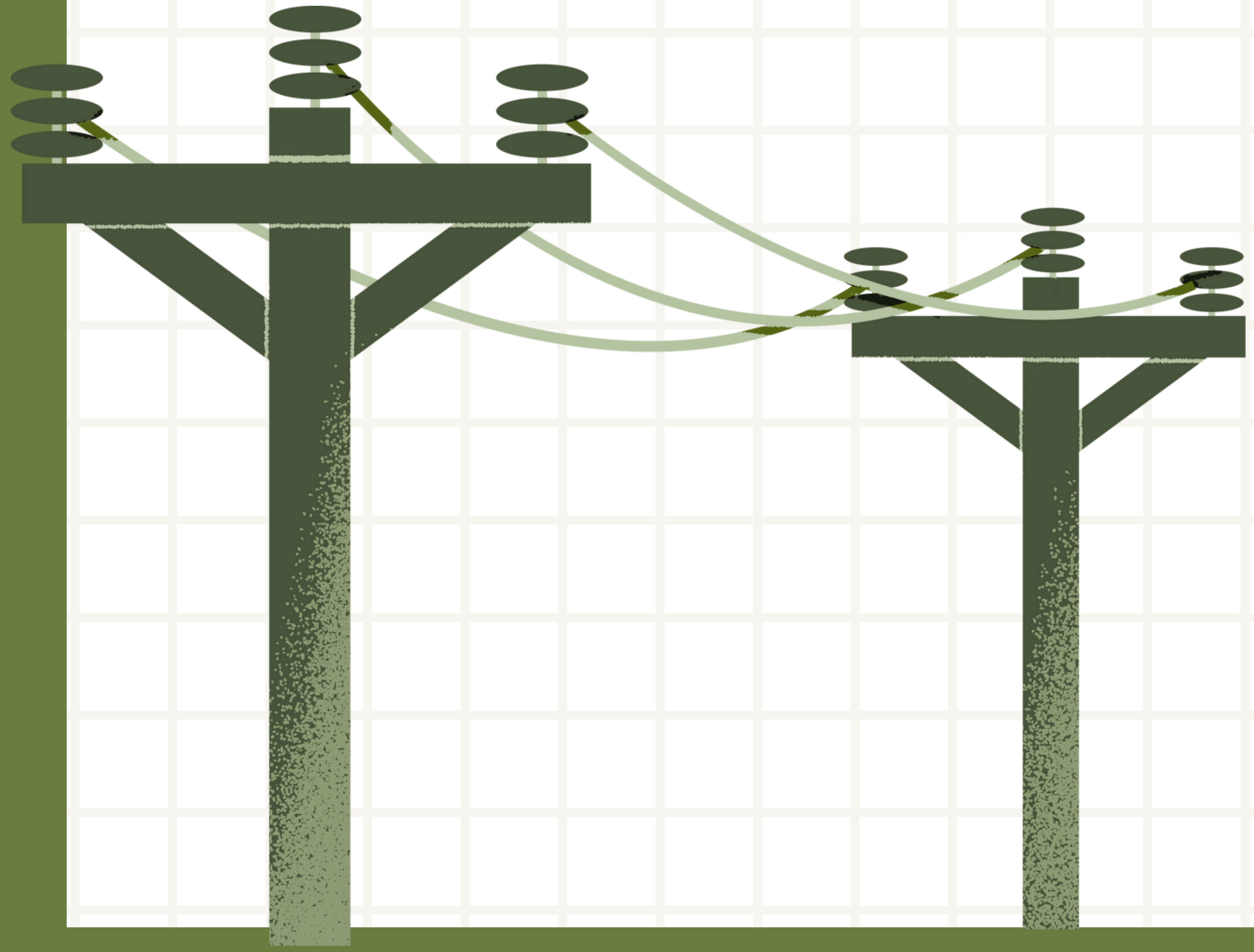
- EDA plots per segment
- Forecasts vs actual data
- Examine feature importance, error metrics, and Dashboard

METHODOLOGY

- Train / Test split for each target period
- Model training with default parameters
- Hyperparameter tuning - with cross-validation
- Model training with **best** parameters
- Model evaluation on the test set
- Prediction for the target period

PHASE 1

Exploratory Analysis & Data Characterization



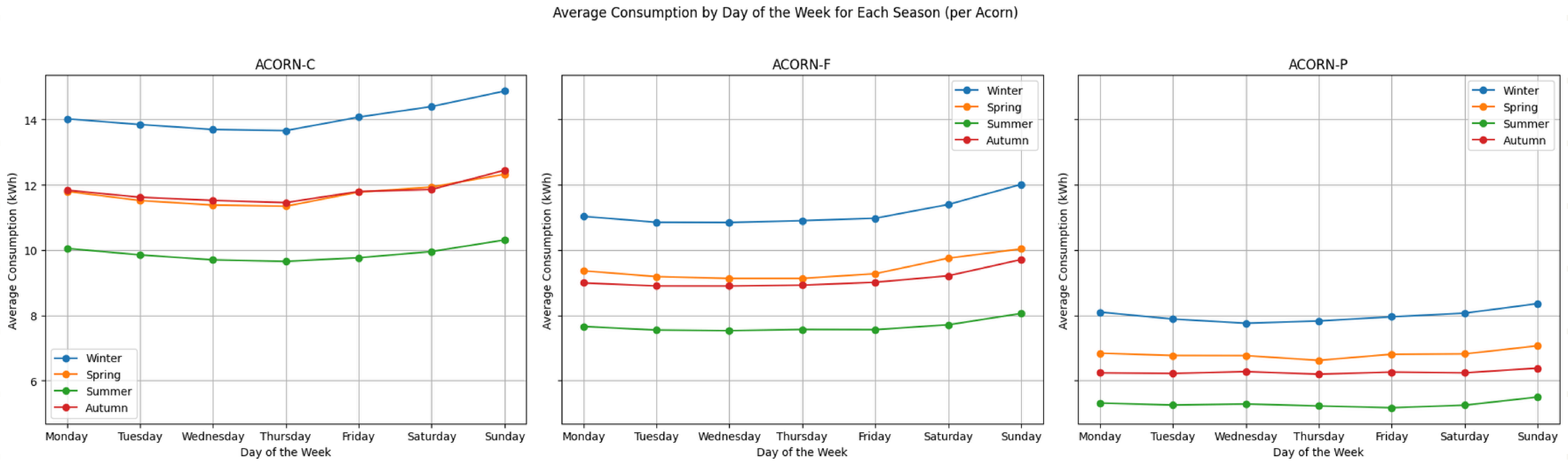
PHASE 1

EDA of consumption profiles

Verification of missing values, double lines, outliers

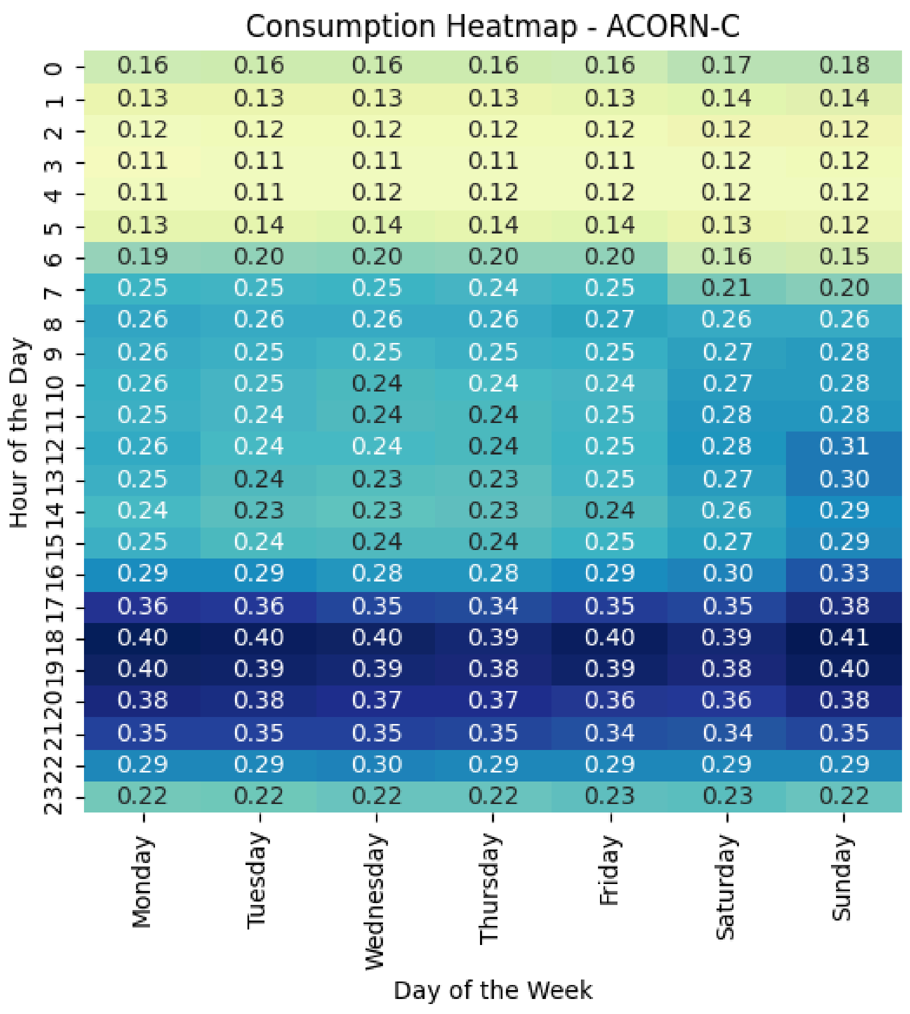
Weekly seasonality: weekdays X weekends

Analysis of holidays

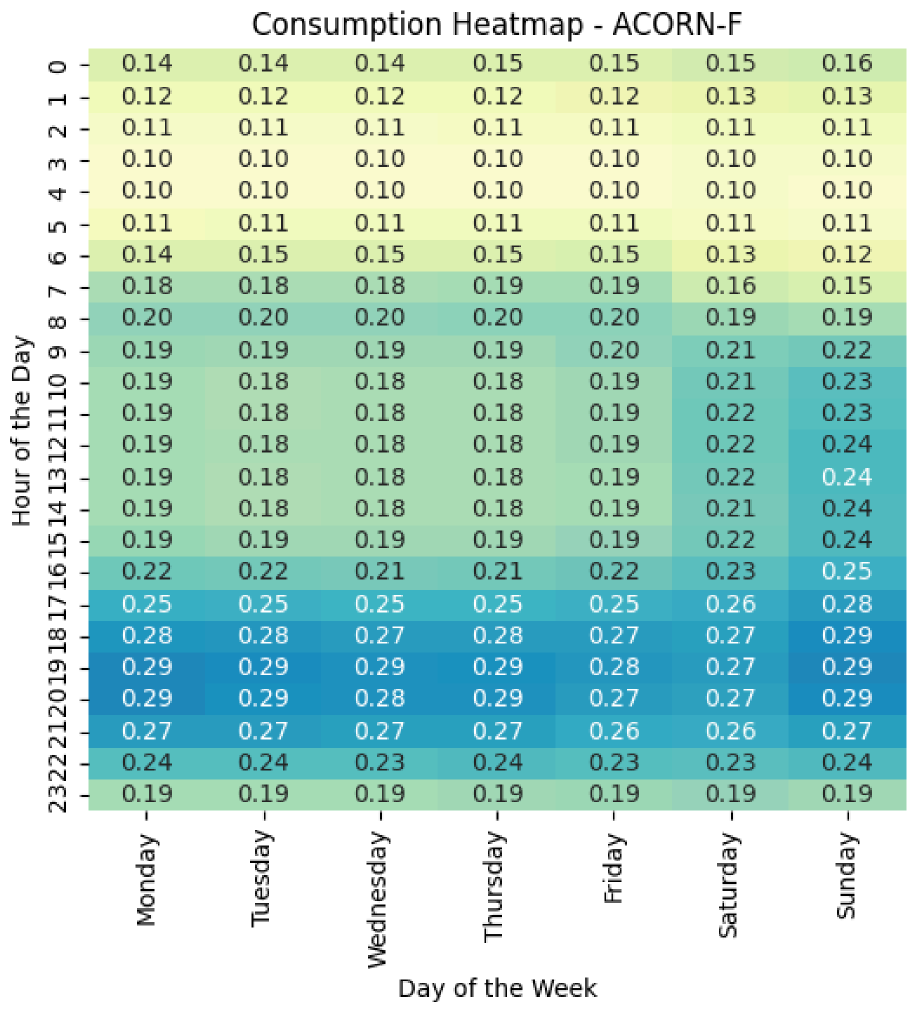


CONSUMPTION HEATMAPS

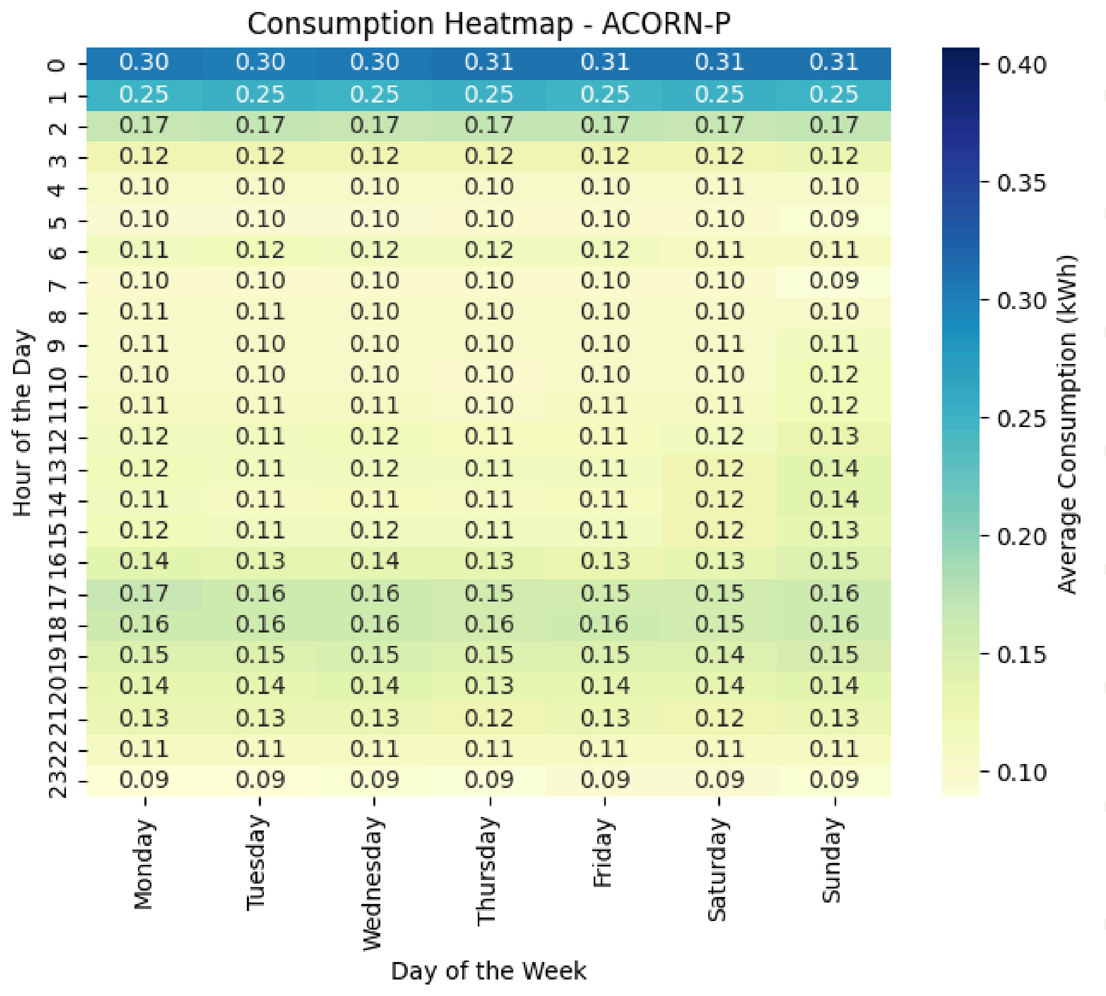
ACORN - C



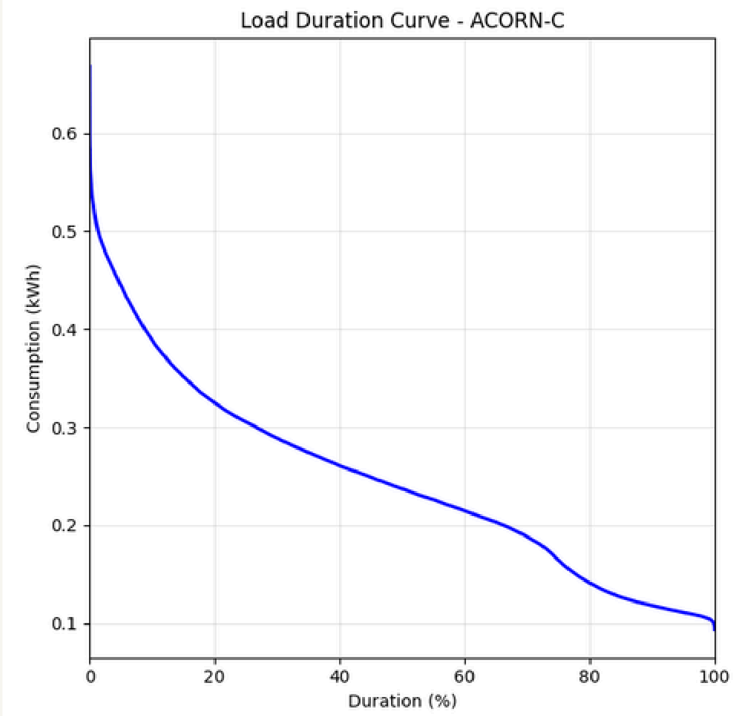
ACORN - F

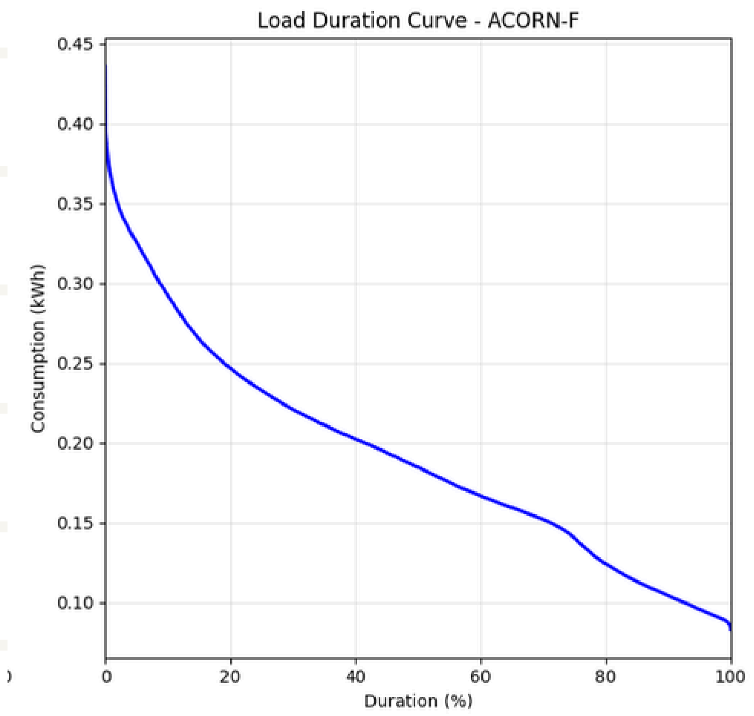


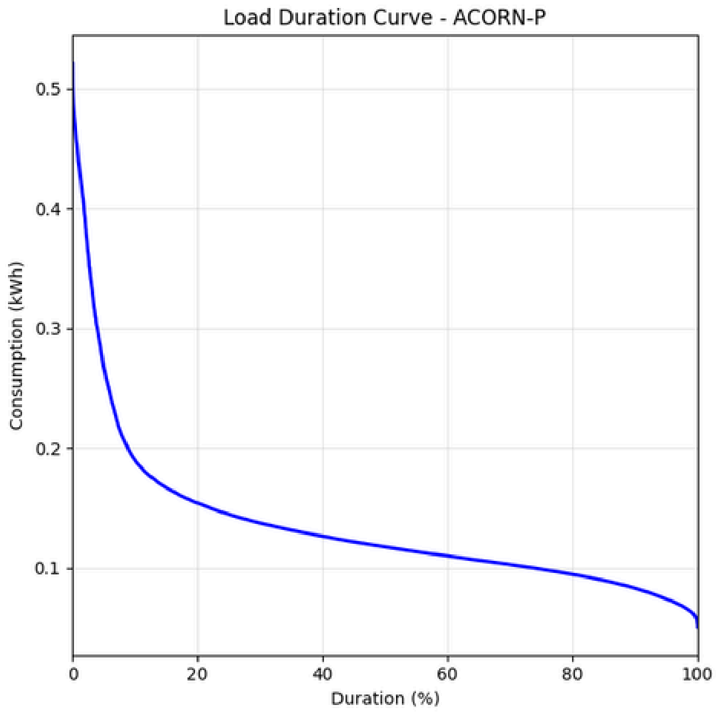
ACORN - P

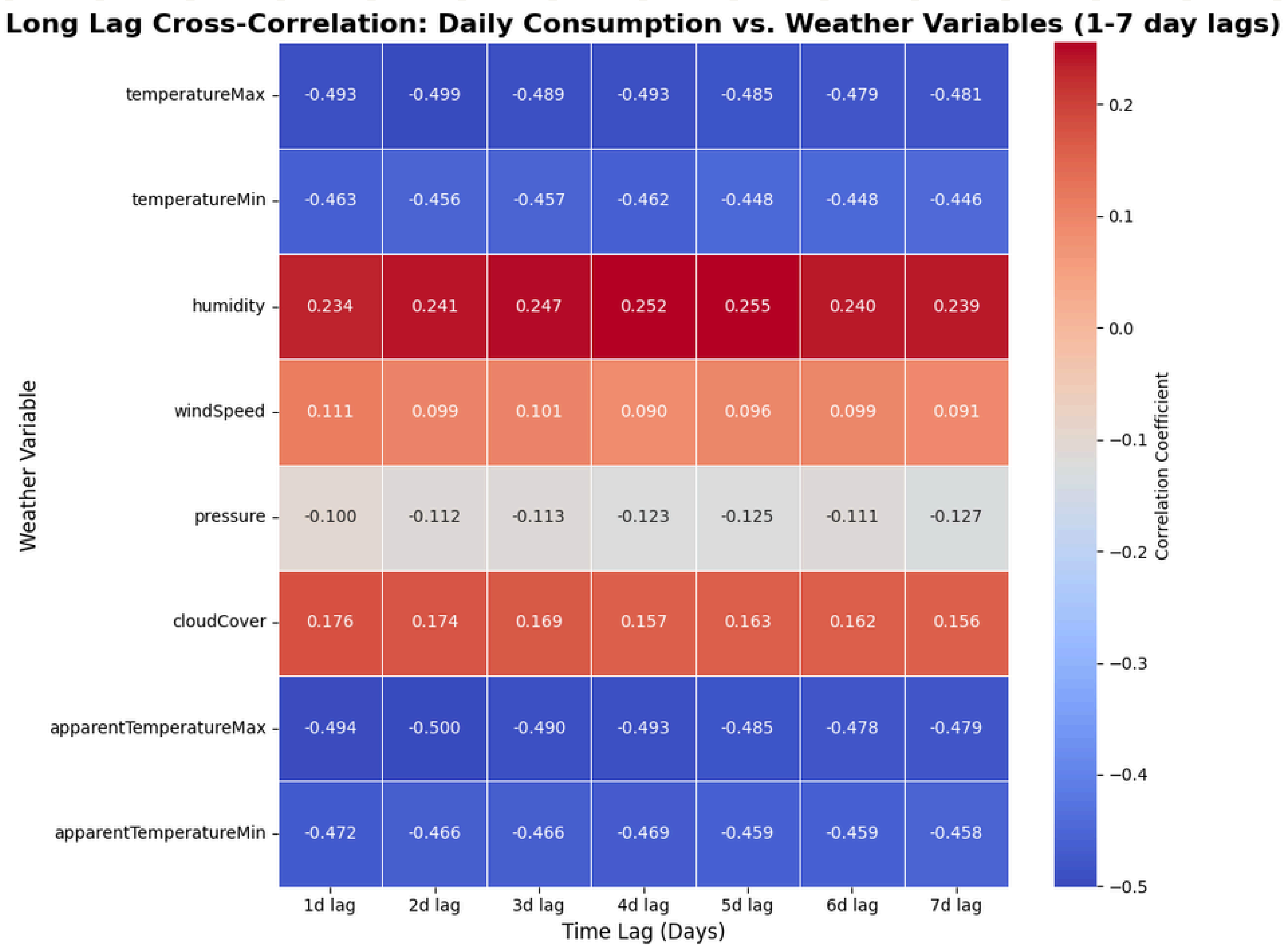


LOAD CURVES & LONG LAG CROSS CORRELATION







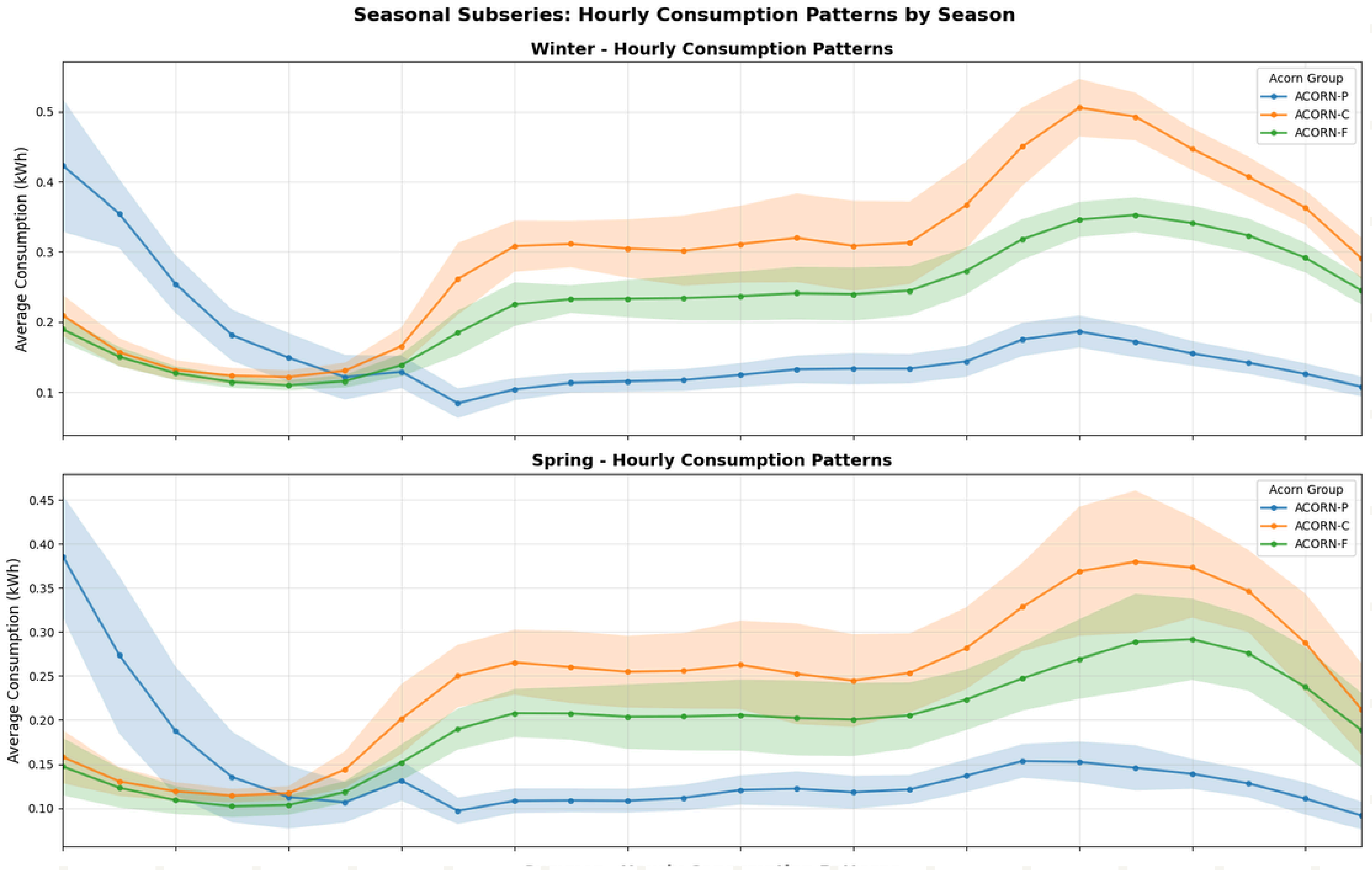


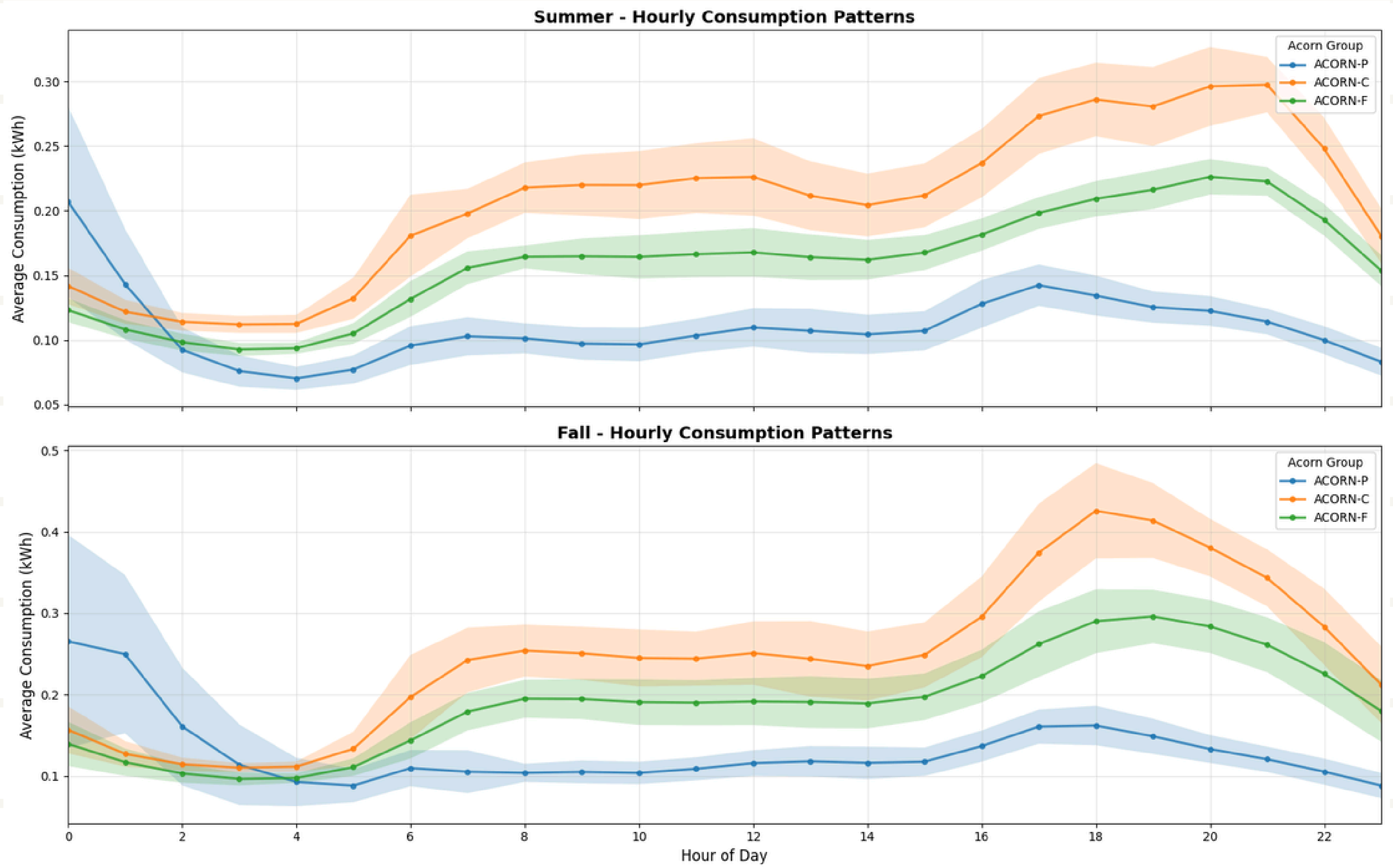
CONSUMPTION PATTERNS BY SEASON

Critical Season: Winter

very important factor in consumption volume

Daily Trends





DAILY CONSUMPTION - by the day of the week

Group C
Consumption increases throughout the week

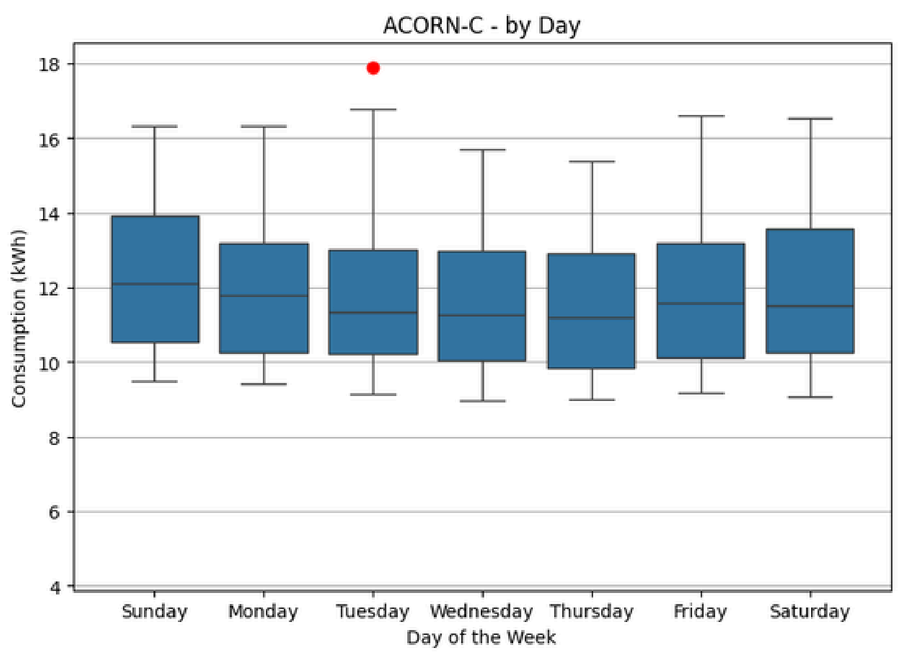
Peaks: friday and saturday

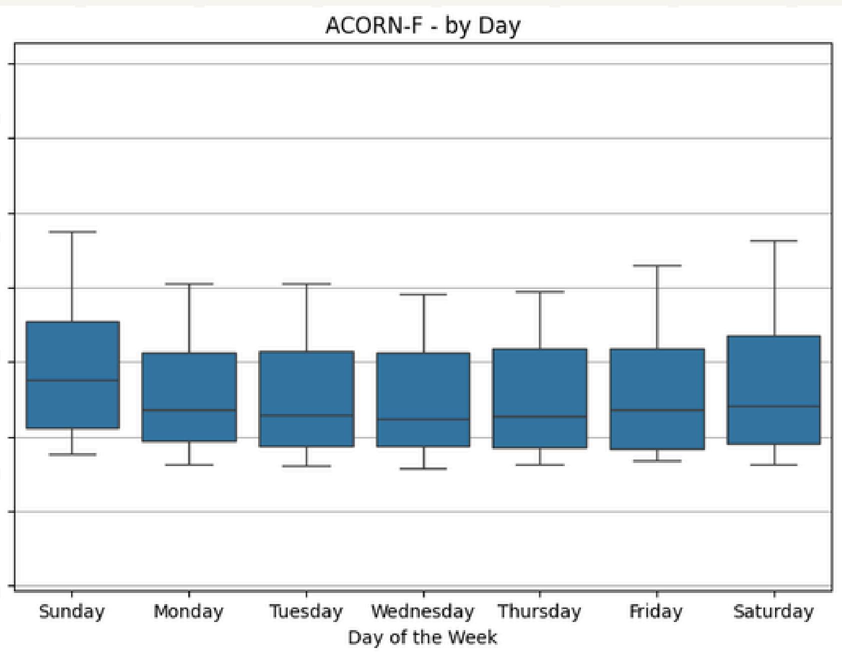
Group F
Intermediate level

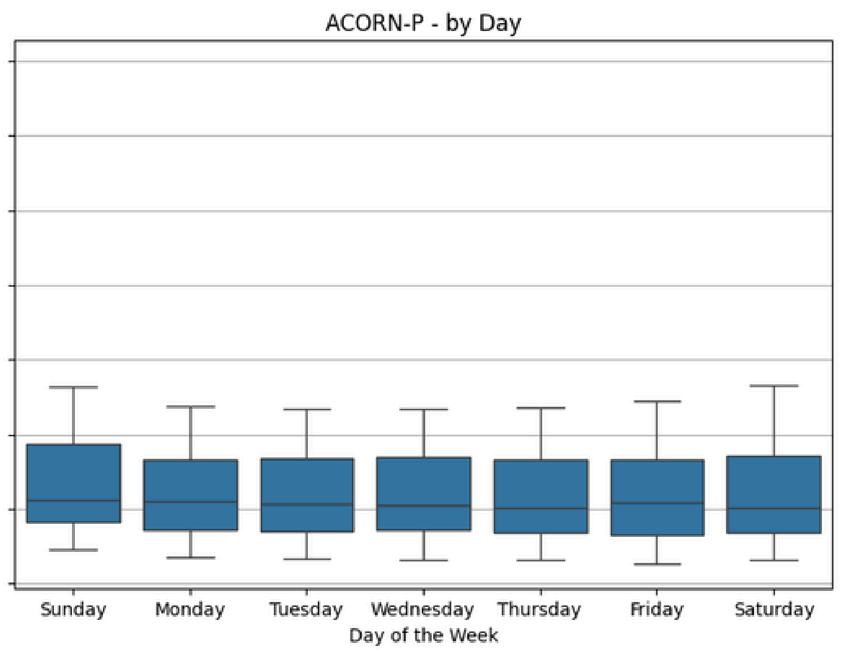
Peak: sunday

Group P
Much lower and more stable consumption

Slight Peak: sunday







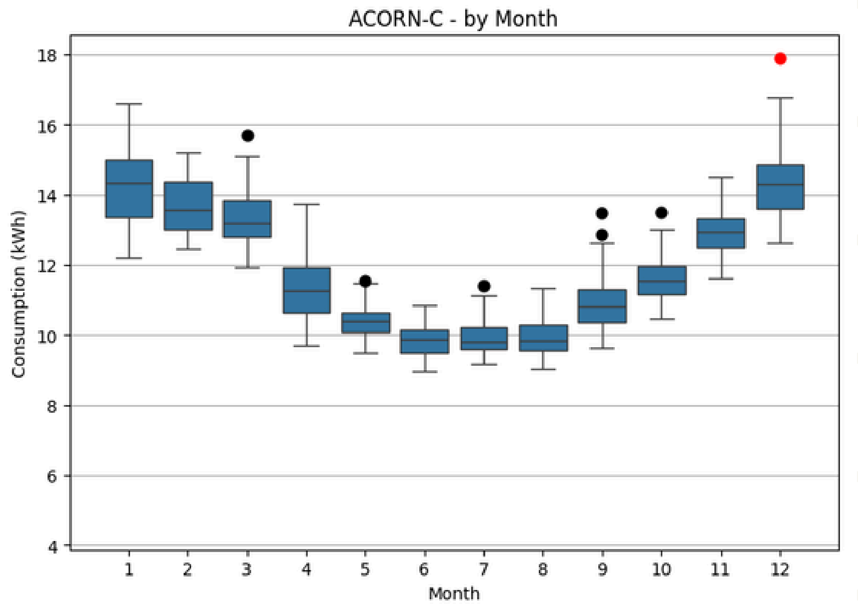
DAILY CONSUMPTION - by month

A Clear Seasonal "U" Pattern: All three groups show the same annual cycle

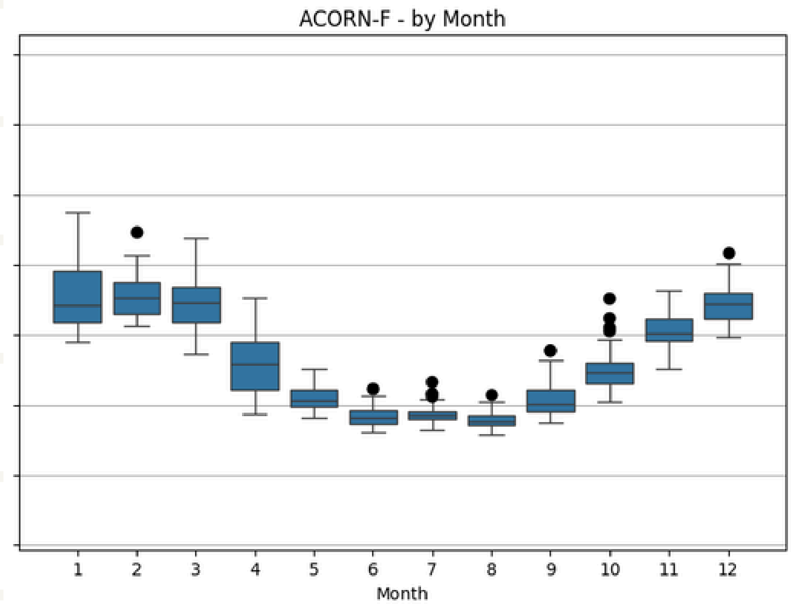
Higher consumption in the winter

More variability of consumption in the winter - for all groups

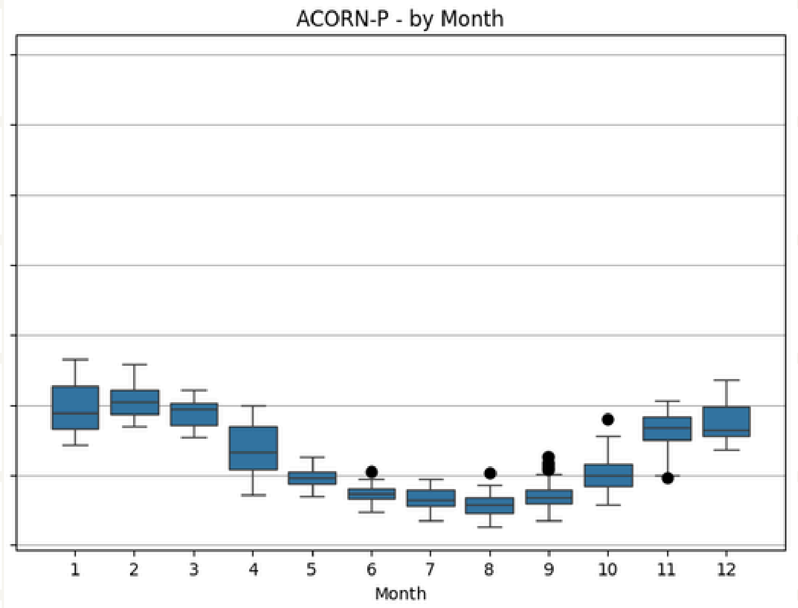
Higher consumption

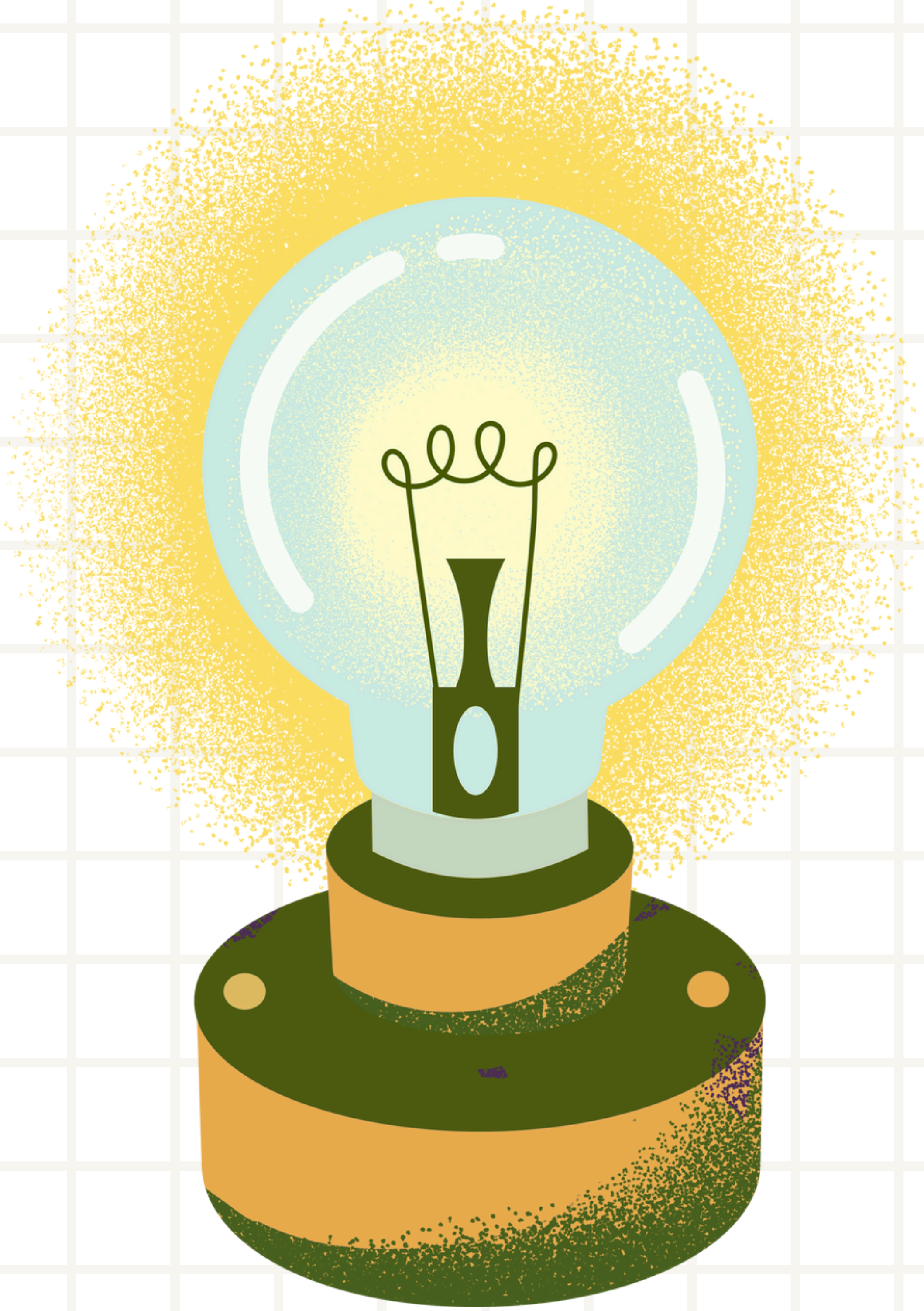


Medium consumption



Lower consumption





PHASE 2

Short-Term Forecasting

48-Hour, 30-Minute Resolution

phase 2

MODELS :

KNN & **RANDOM FOREST**

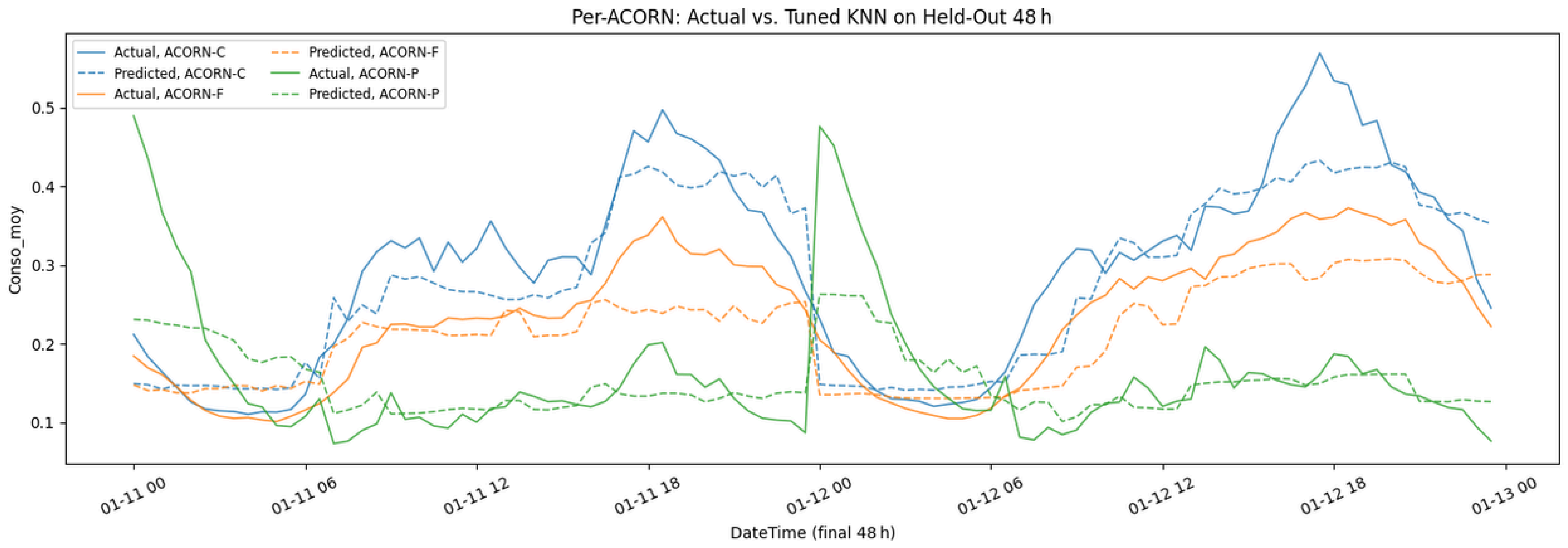
KNN

why?

Simple, fundamental machine learning model.
Great baseline to establish initial performance

process

Hyperparameters were individually tuned for each group using **Optuna** to find the **best settings.**



KNN

MODEL DISCARDED !

MAPE
over 30% ,
unacceptably high

not accurate forecasts

ACORN	MAE	MAPE (%)	RMSE
C	0.041	30.8348	0.0512
P	0.0337	30.8438	0.0596
F	0.037	30.8438	0.0448

random forest

why?

Powerful ensemble model - typically provides **higher accuracy** than single models.
Robust and **less** prone to **overfitting**.
It can effectively capture **complex, non-linear** relationships in the data.

ADVANTAGES

- High accuracy
- Strong predictive performance
- Can calculate feature importance

DISADVANTAGES

- Computationally slower
- More resource-intensive to train
- Less interpretable than simpler models

random forest

refining the evaluation

48-hour test window might be too short to calculate errors reliably. Thus, the test set was extended to a full week

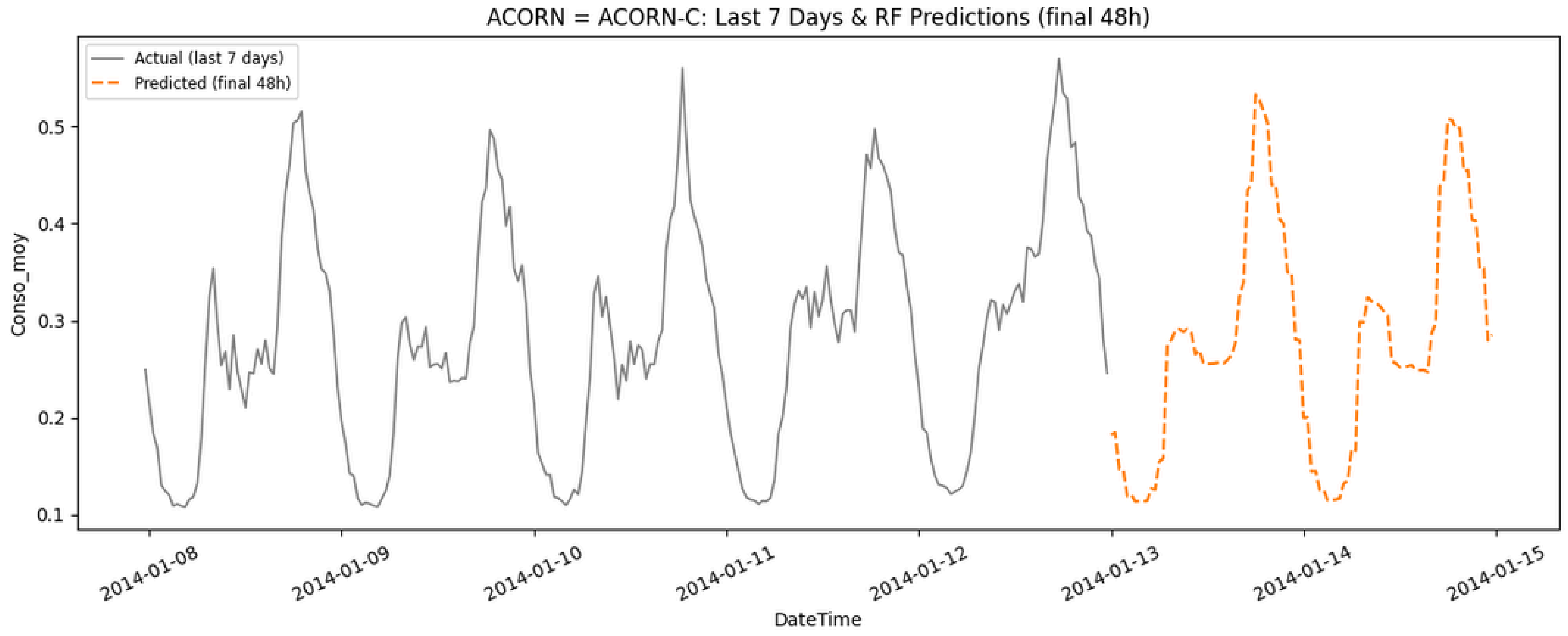
ACORN	MAE	MAPE (%)	RMSE
C	0.0185	1.8498	0.0274
P	0.0151	1.5087	0.0201
F	0.0127	1.2724	0.0164

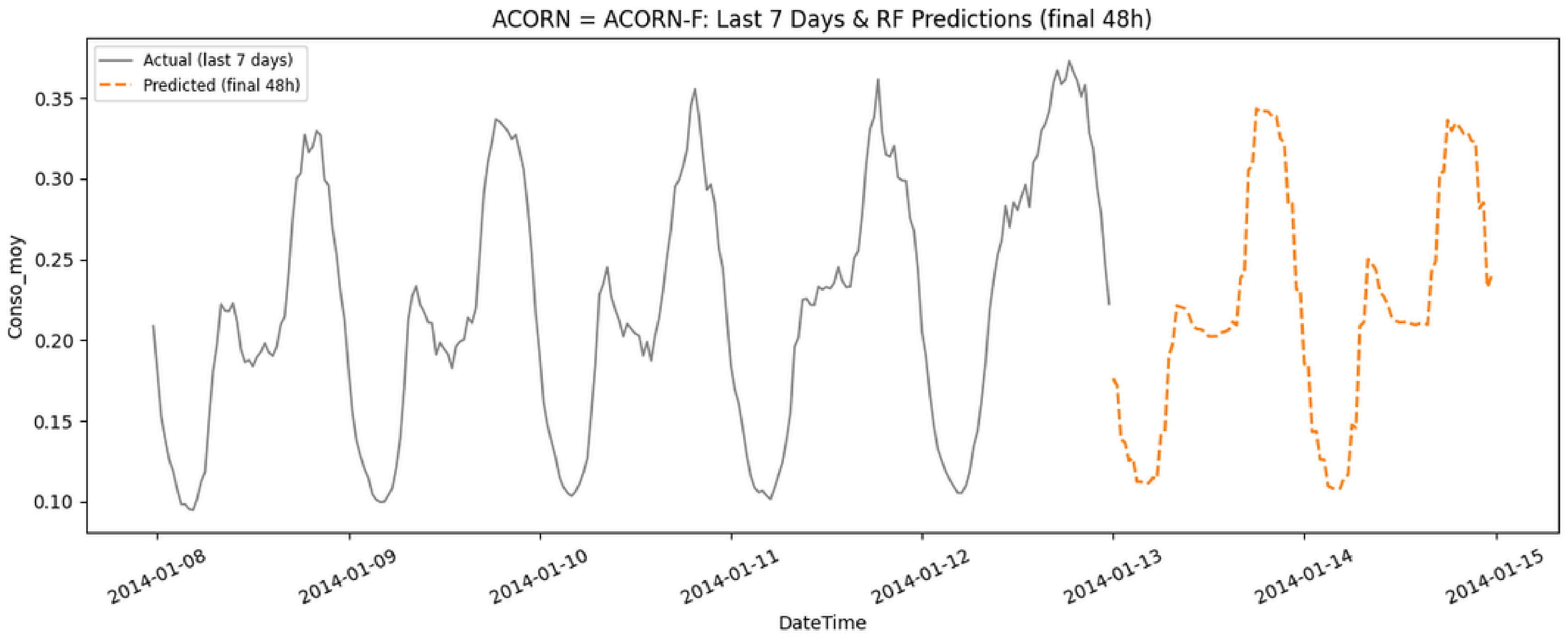
big improvement over KNN !

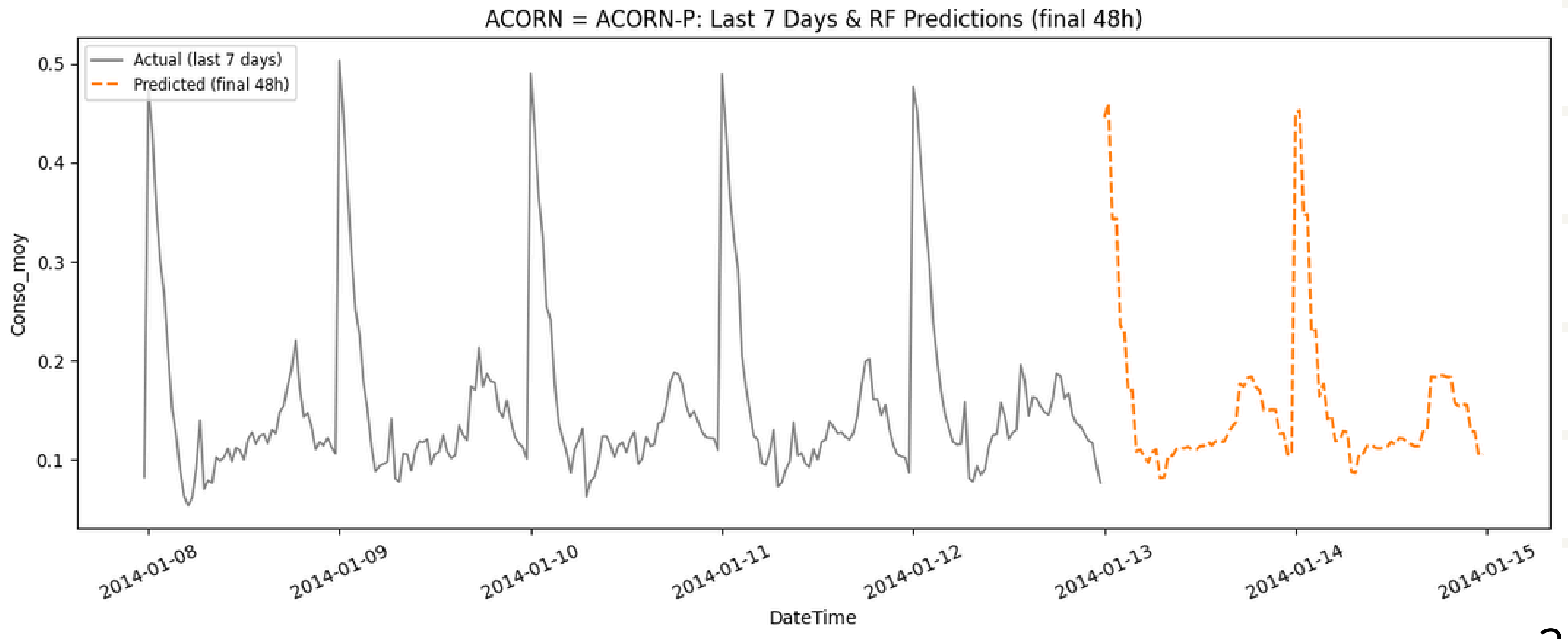
chosen model for the short-term forecast

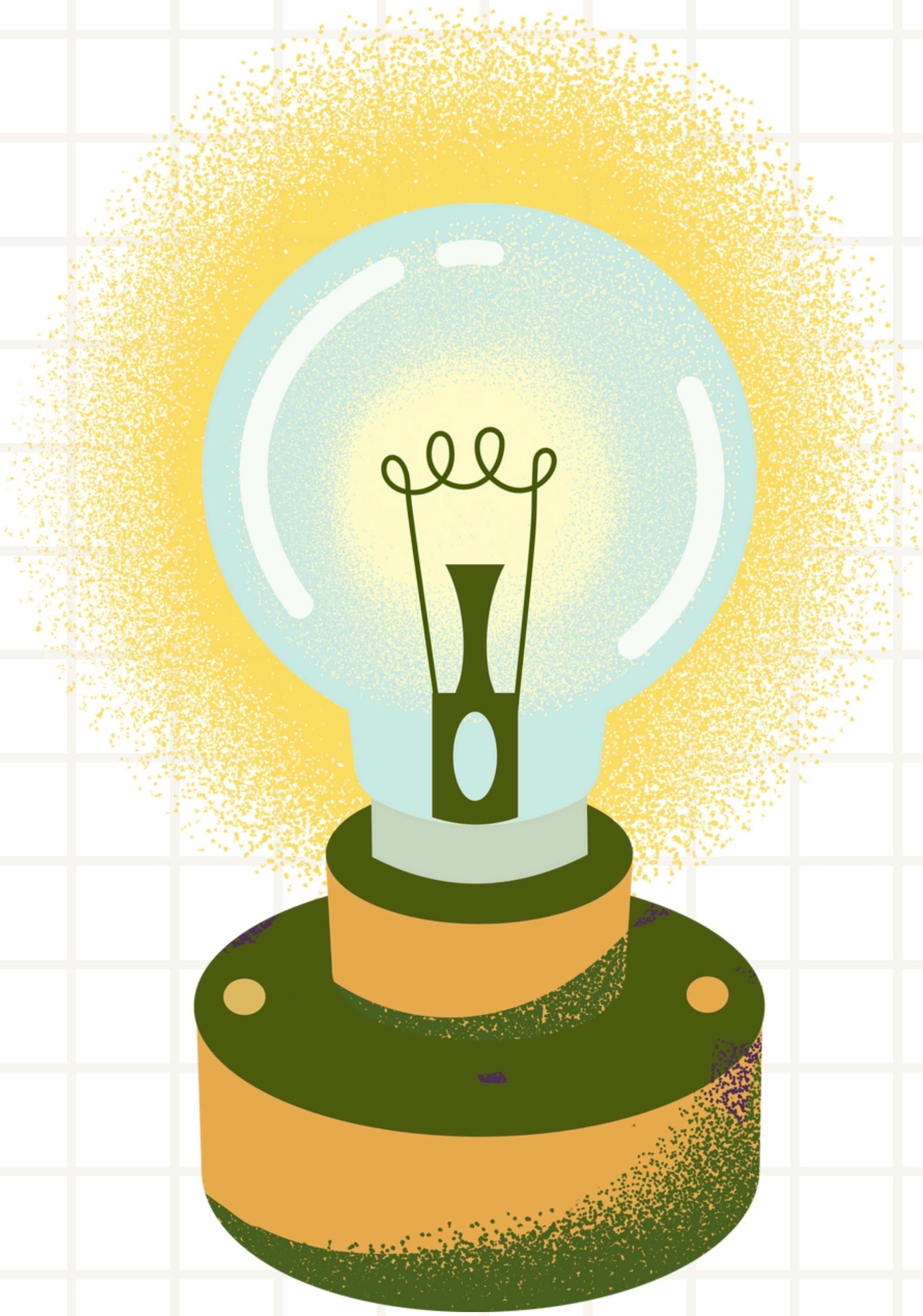
random forest

Prediction
for 13/01/2014 and 14/01/2014









PHASE 3

Medium-Term Forecasting

1-Month, Daily Resolution

phase 3

MODELS

- SARIMAX
- MLP
- SVM
- LightGBM
- LSTM

DISCARDED MODELS

	SARIMAX	MLP	SVM	LightGBM
MAPE range per ACORN	4-7%	3-7%	3-6%	1-2%
RMSE range per ACORN	0.6-1.1	0.4-1.2	0.48-1.4	0.1-0.4

CHOSEN MODEL : LSTM !

why?

It's a **recurrent neural network** that excels at **capturing long-term temporal patterns** in sequential data such as electricity usage, making it a great choice for medium-term consumption forecasting.

ADVANTAGES

- Captures long-term dependence
- Handles nonlinear patterns
- Robust to missing data

DISADVANTAGES

- Data and compute intensive
- Complex tuning & interpretation

LSTM

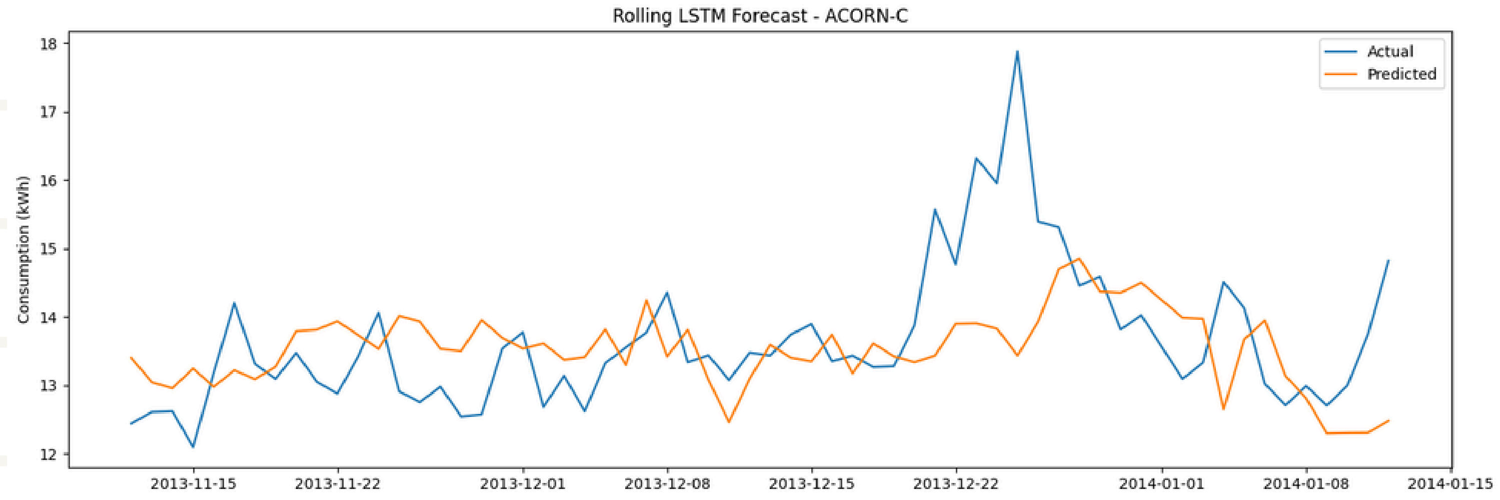
METHODOLOGY & RESULTS

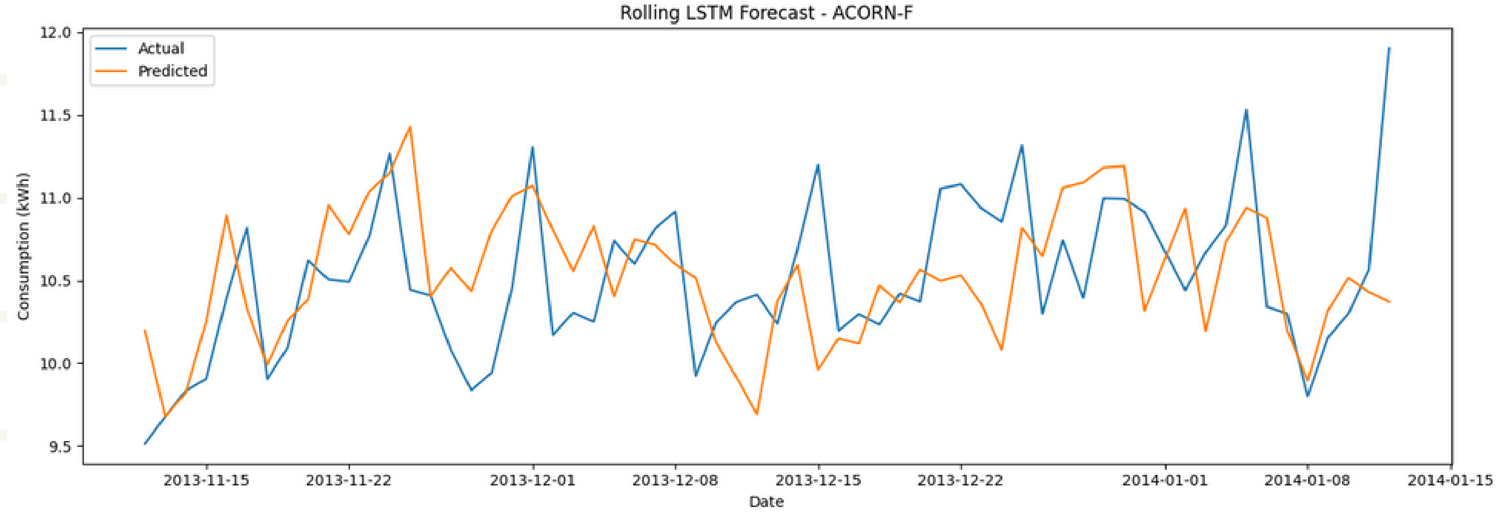
rolling forecast approach

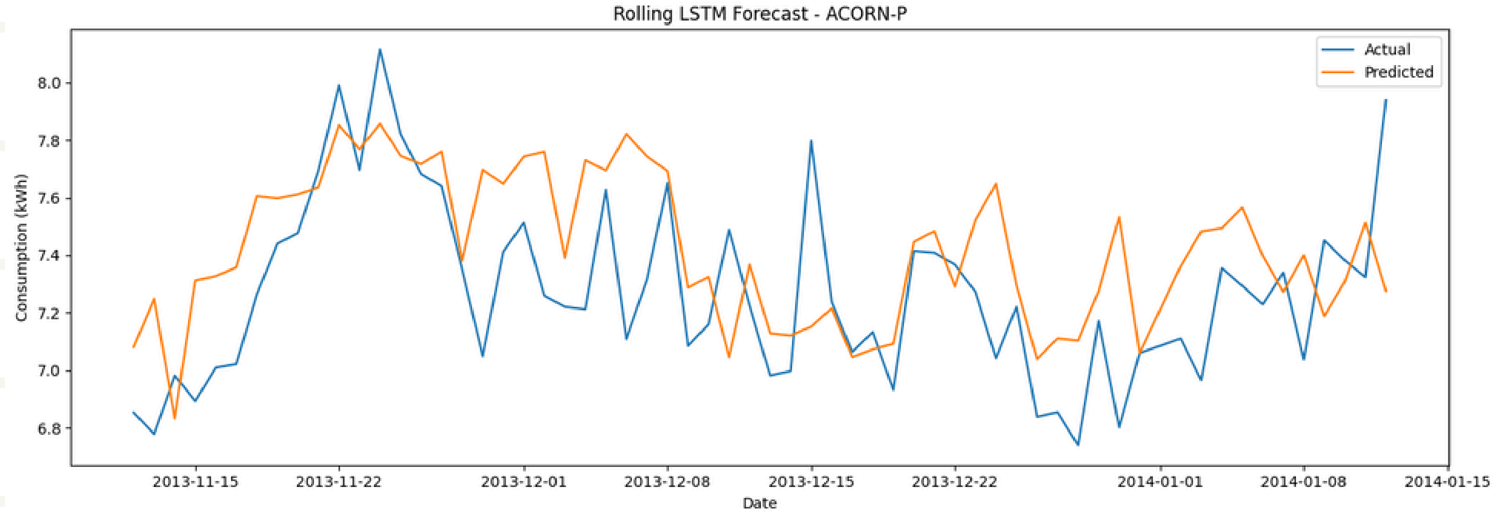
- ensure the highest possible accuracy
- mimics a real-world application.
- dynamically adapt to new trends

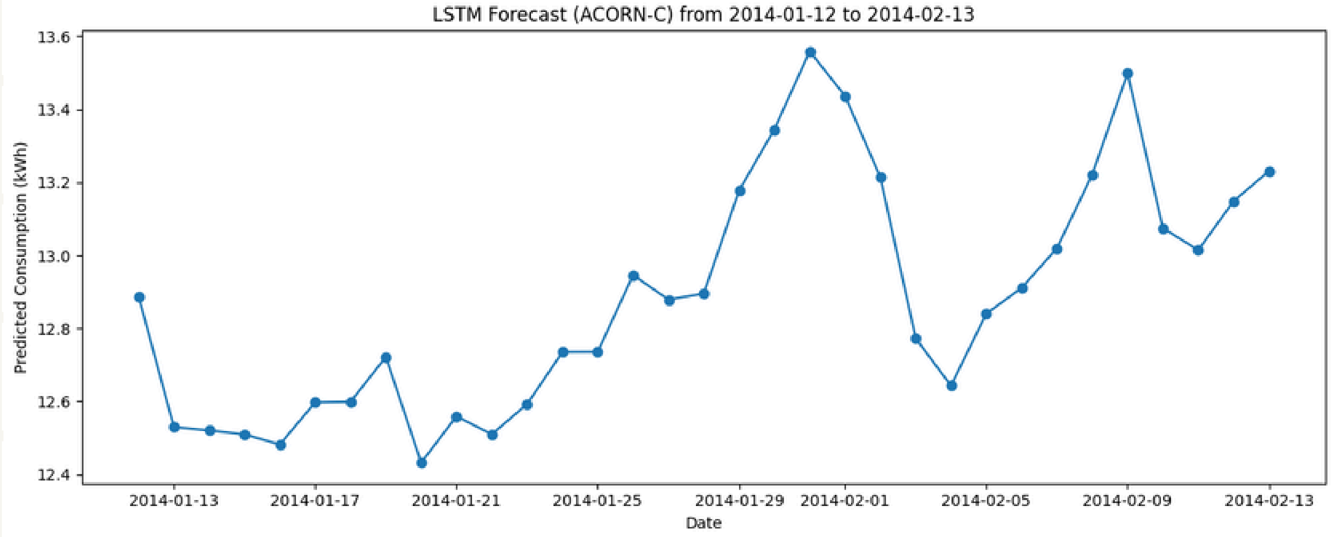
ACORN	MAE	MAPE (%)	RMSE
C	0.6392	4.50	0.9470
P	0.2938	4.07	0.5016
F	0.4093	3.88	0.3772

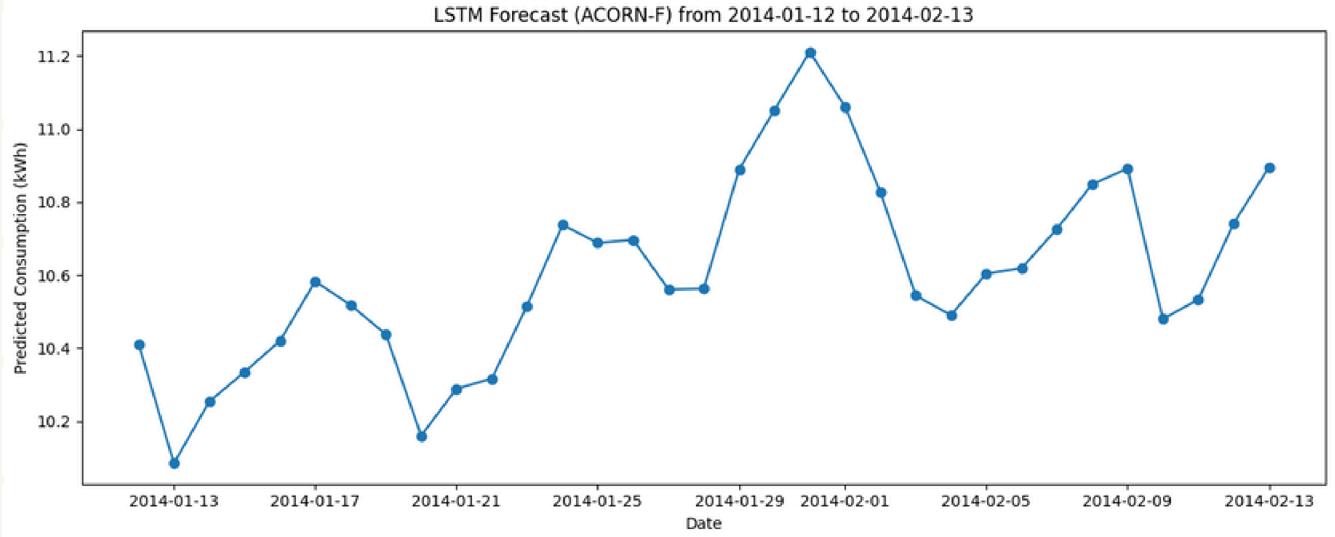
LSTM

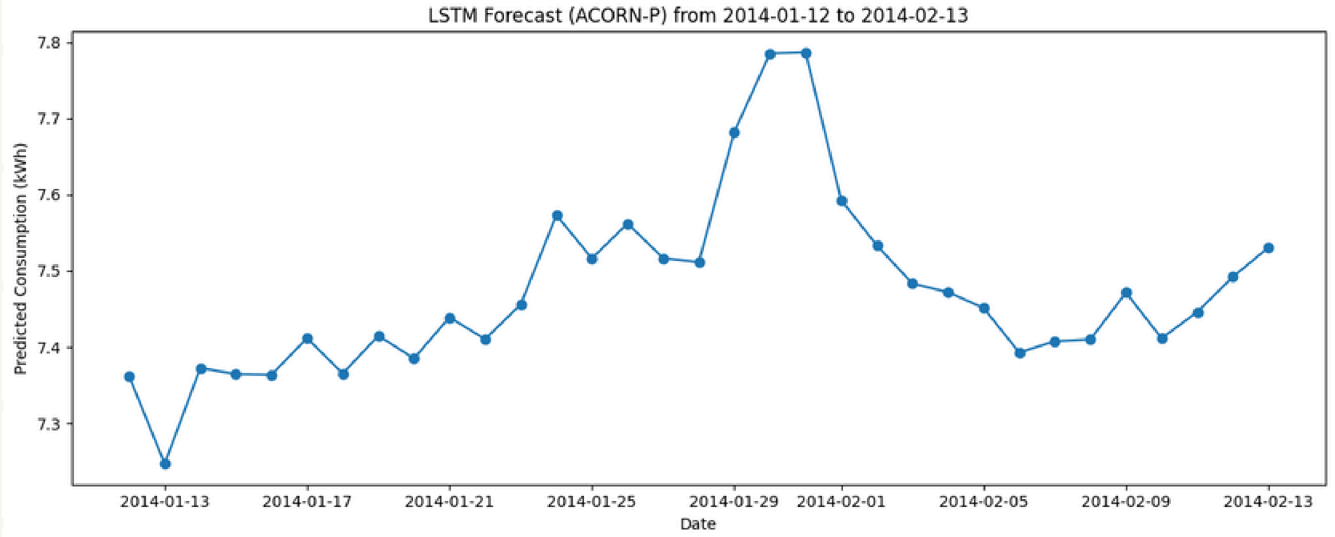

Rolling LSTM Forecast - ACORN-C
Actual
Predicted
Consumption (kWh)


Rolling LSTM Forecast - ACORN-F
Actual
Predicted
Consumption (kWh)
Date

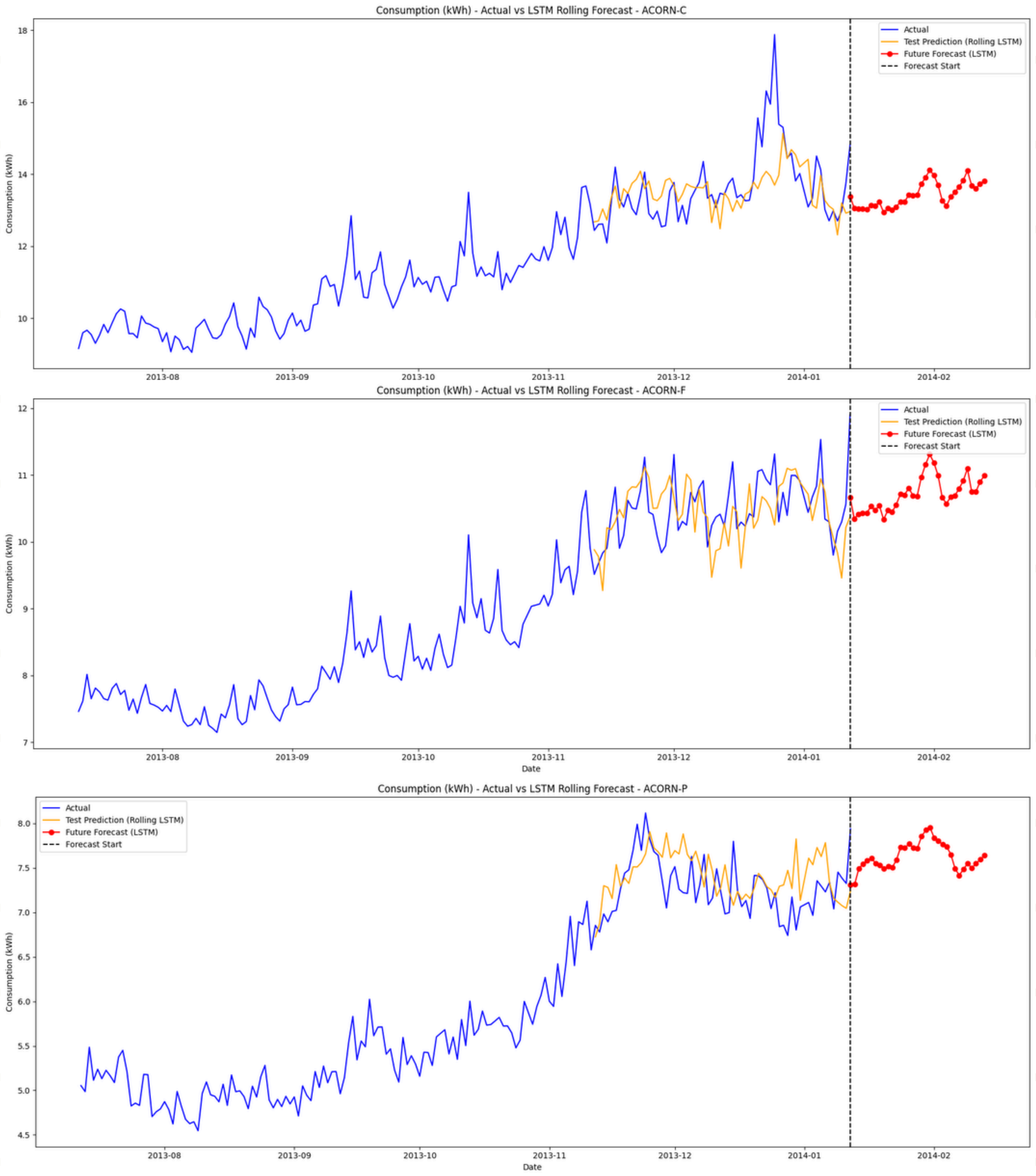

Rolling LSTM Forecast - ACORN-P
Actual
Predicted
Consumption (kWh)
Date

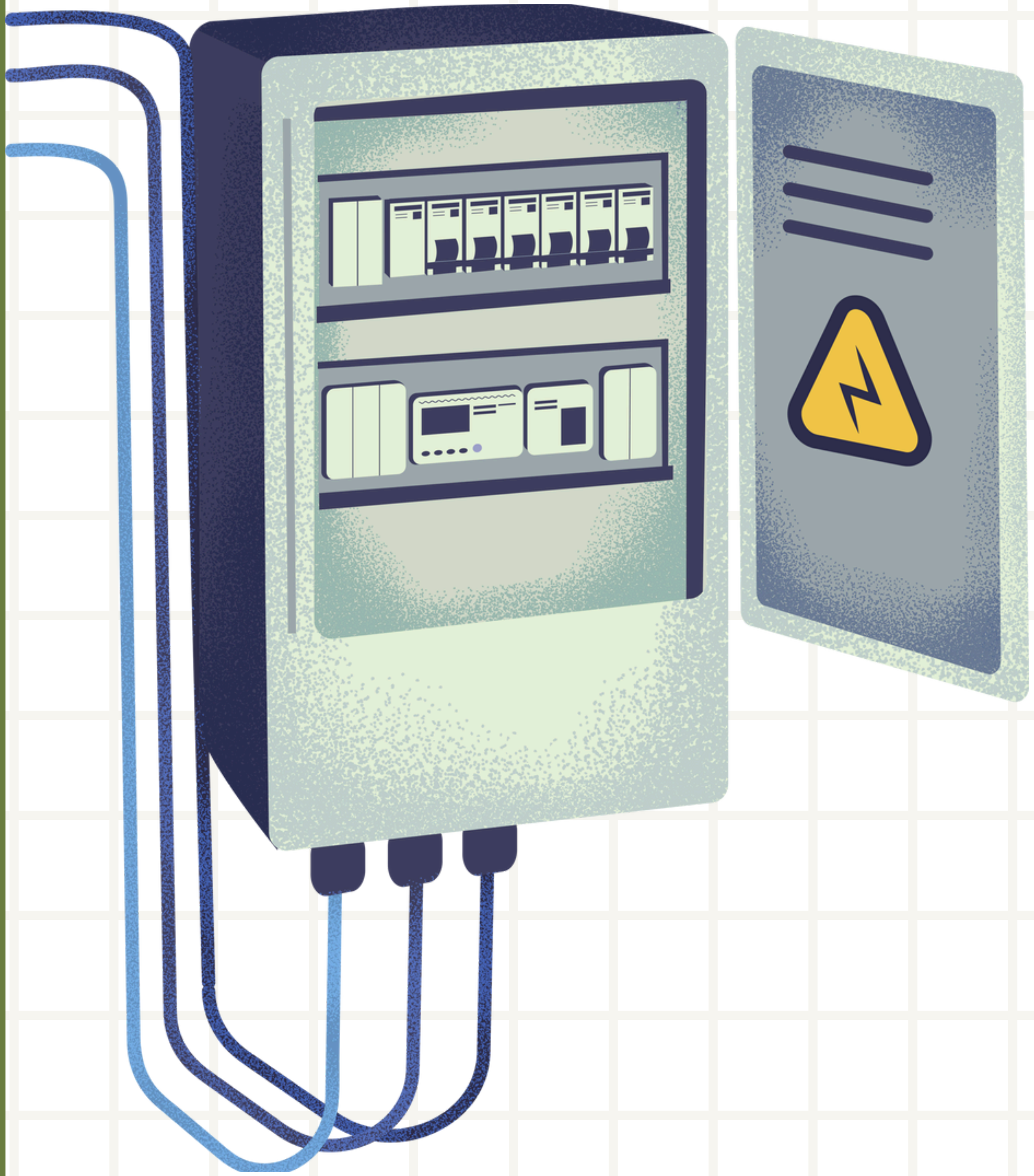

LSTM Forecast (ACORN-C) from 2014-01-12 to 2014-02-13
Predicted Consumption (kWh)
Date


LSTM Forecast (ACORN-F) from 2014-01-12 to 2014-02-13
Predicted Consumption (kWh)
Date


LSTM Forecast (ACORN-P) from 2014-01-12 to 2014-02-13
Predicted Consumption (kWh)
Date

LSTM

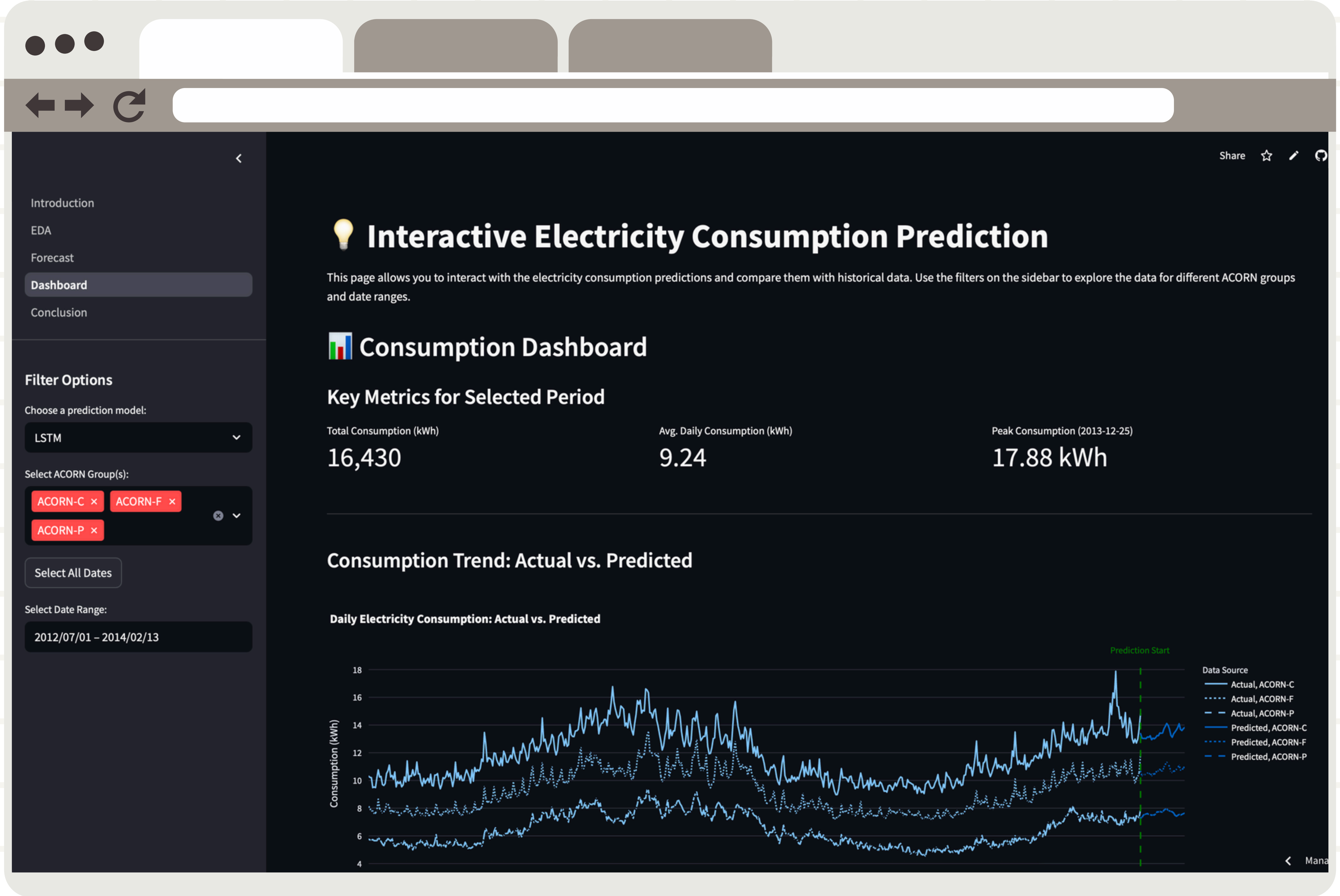

Consumption (kWh) - Actual vs LSTM Rolling Forecast - ACORN-C
Actual
Test Prediction (Rolling LSTM)
Future Forecast (LSTM)
Forecast Start
Consumption (kWh)
Consumption (kWh) - Actual vs LSTM Rolling Forecast - ACORN-F
Actual
Test Prediction (Rolling LSTM)
Future Forecast (LSTM)
Forecast Start
Consumption (kWh)
Date
Consumption (kWh) - Actual vs LSTM Rolling Forecast - ACORN-P
Actual
Test Prediction (Rolling LSTM)
Future Forecast (LSTM)
Forecast Start
Consumption (kWh)
Date



PHASE 4

Interactive Dashboard

Streamlit Link


Share
Introduction
EDA
Forecast
Dashboard
Conclusion
Filter Options
Choose a prediction model:
LSTM
Select ACORN Group(s):
ACORN-C
ACORN-F
ACORN-P
Select All Dates
Select Date Range:
2012/07/01 – 2014/02/13
Interactive Electricity Consumption Prediction
This page allows you to interact with the electricity consumption predictions and compare them with historical data. Use the filters on the sidebar to explore the data for different ACORN groups and date ranges.
Consumption Dashboard
Key Metrics for Selected Period
Total Consumption (kWh)
16,430
Avg. Daily Consumption (kWh)
9.24
Peak Consumption (2013-12-25)
17.88 kWh
Consumption Trend: Actual vs. Predicted
Daily Electricity Consumption: Actual vs. Predicted
Prediction Start
Consumption (kWh)
18
16
14
12
10
8
6
4
Data Source
Actual, ACORN-C
Actual, ACORN-F
Actual, ACORN-P
Predicted, ACORN-C
Predicted, ACORN-F
Predicted, ACORN-P

CONCLUSIONS

ACORN groups present **seasonal and behavioral patterns**

Short-term:

⟶ Random Forest outperformed KNN with tuned models per group

↳ MAPE : 1 -2 % - **robust** even with extended horizon

Medium-term:

⟶ Compared SARIMAX, MLP, SVM, LSTM and LightGBM (strange values)

⟶ Rolling LSTM best captured temporal patterns with MAPE < 4.5%, outperforming static models.

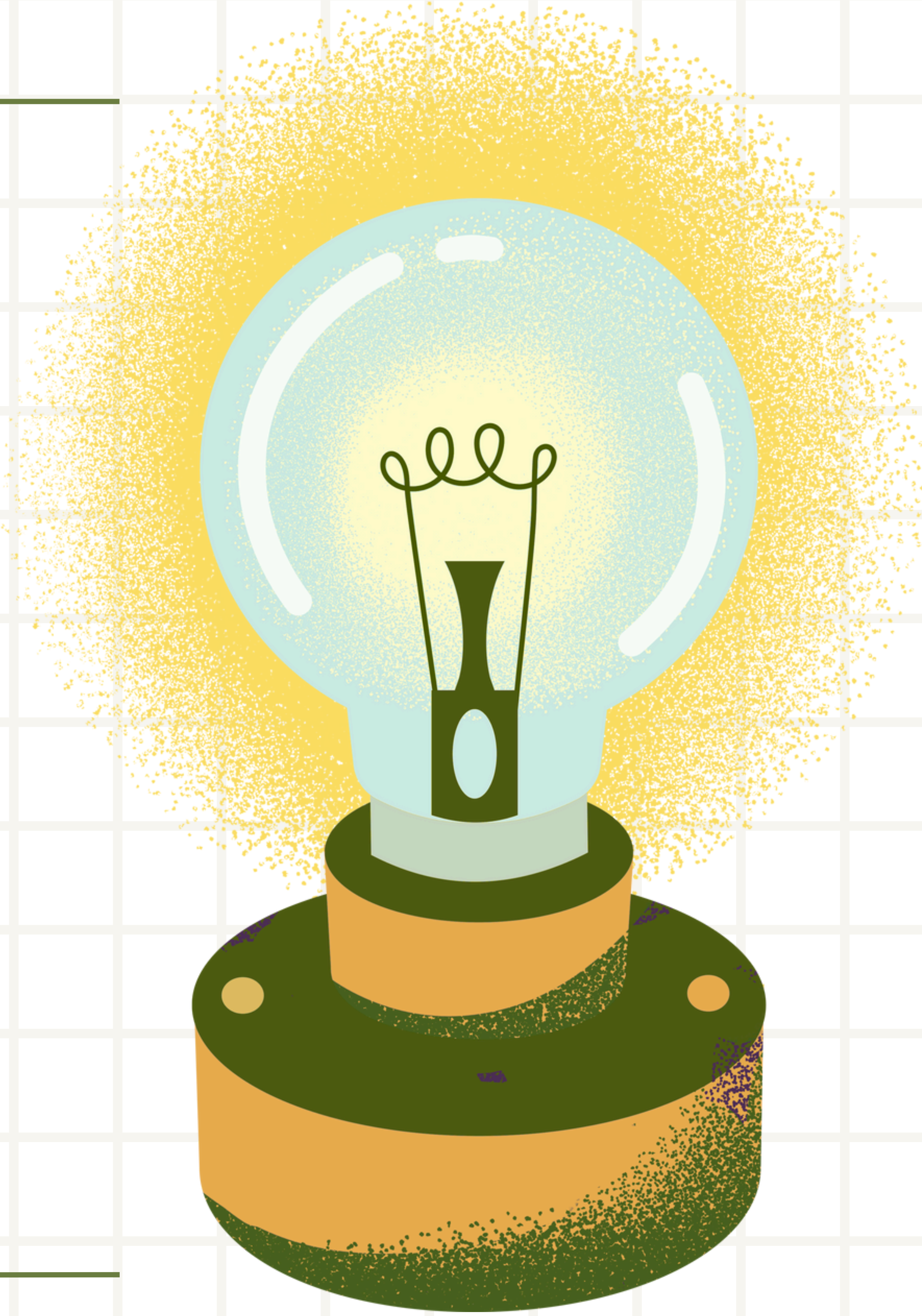
chosen models :

short-term	medium-term
Random Forest	**Rolling LSTM**

CentraleSupélec , 2025

THANK YOU

for your attention !



Group 4

Gustavo C. NOVICKI, Mariana E. F. MENEZES, Thiago N. C. ARAUJO, Alexsandro S. ROSA, Martin SCHULZ.