DSC 425
Time Series Analysis of Power Consumption

Dr. John McDonald

Group – 3 (Power)
Akanksha Singh
Yashkumar Prajapati

OBJECTIVE

Rising Electricity Demand

Weather Influence

High Energy Dependence

Rapid Urban Expansion

DATASET OVERVIEW

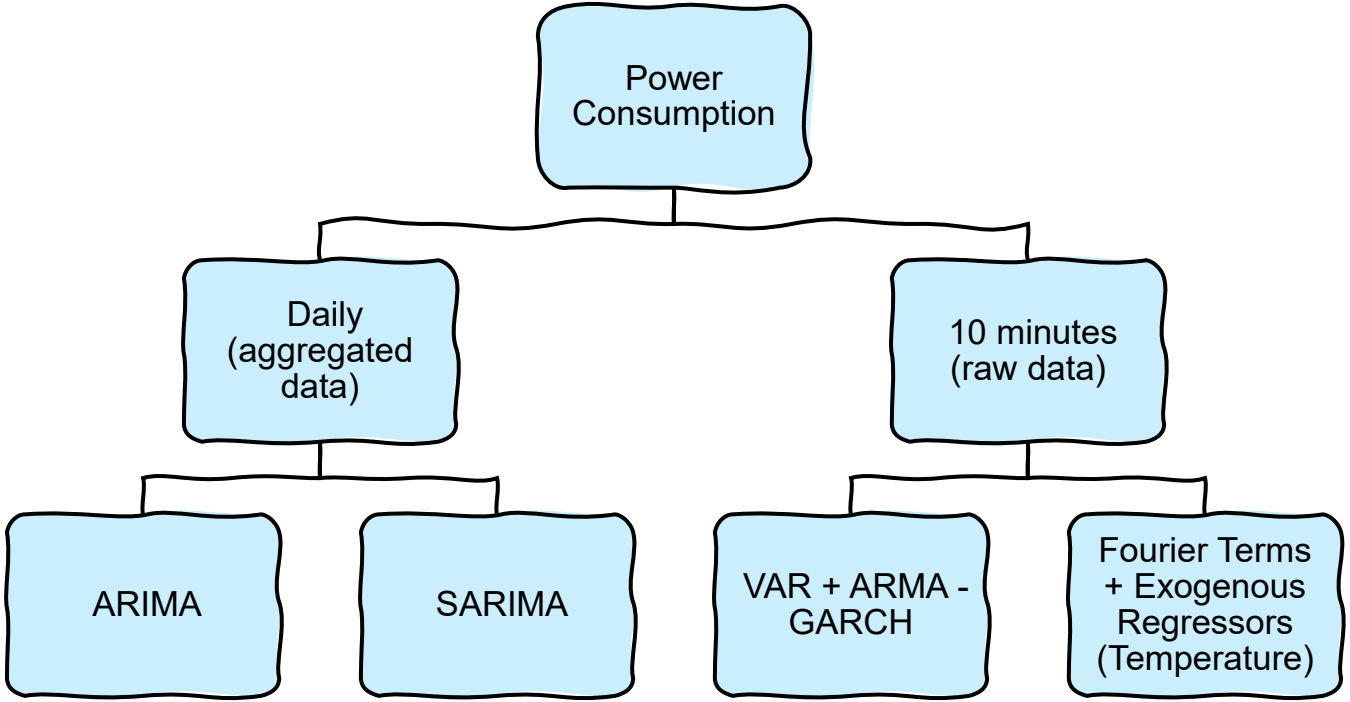
Observations
52416

9
Features

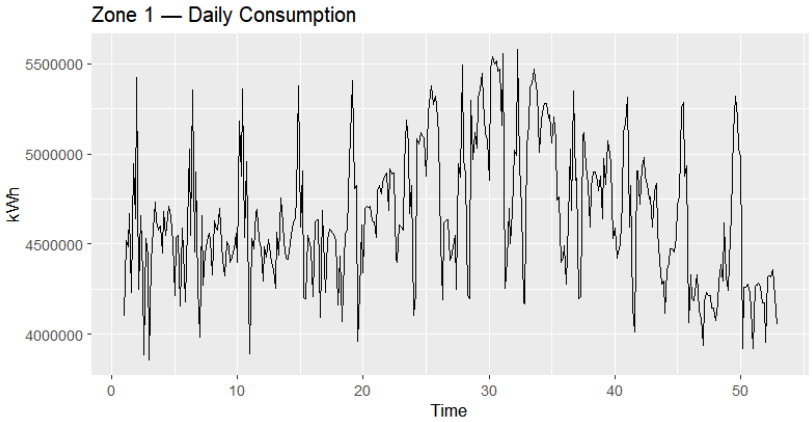
Frequency
10 Minutes
Intervals

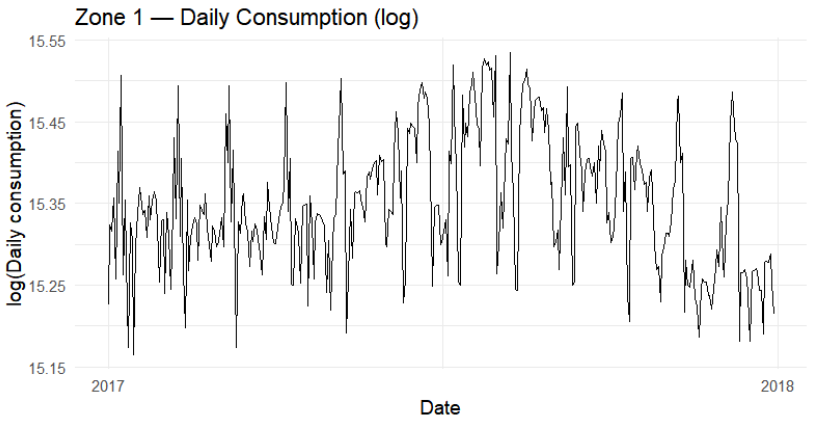
Data Range
1 Jan 2017 -
30 Dec 2017

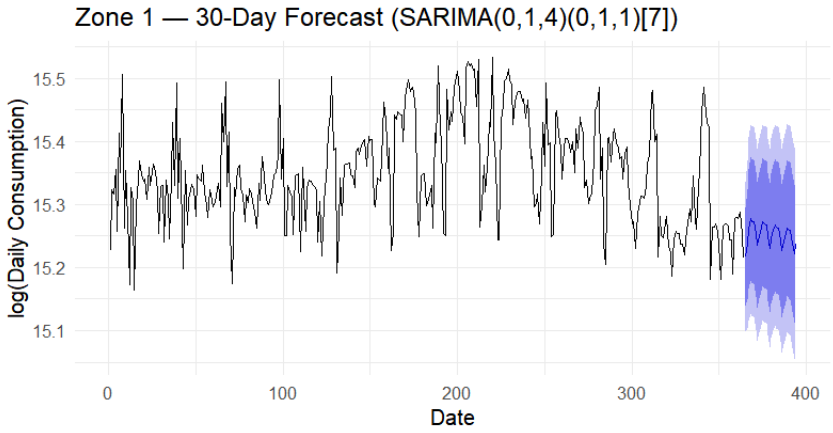
Modelling Strategy Overview



Baseline Modelling Using Daily Aggregation - SARIMA

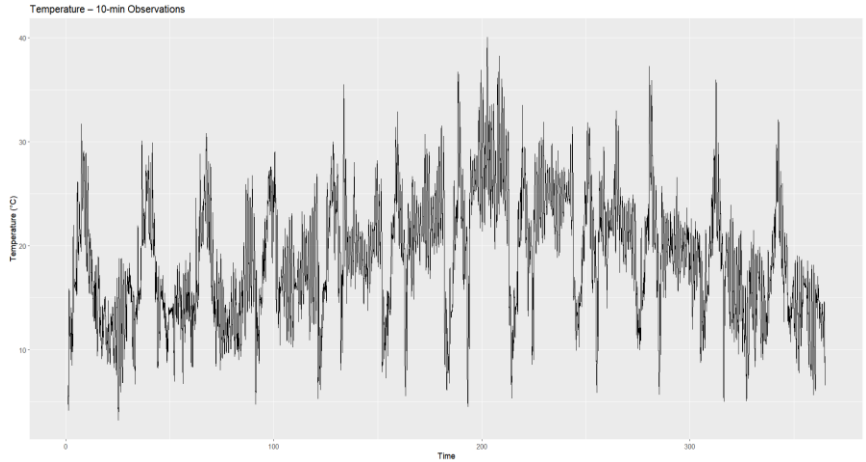


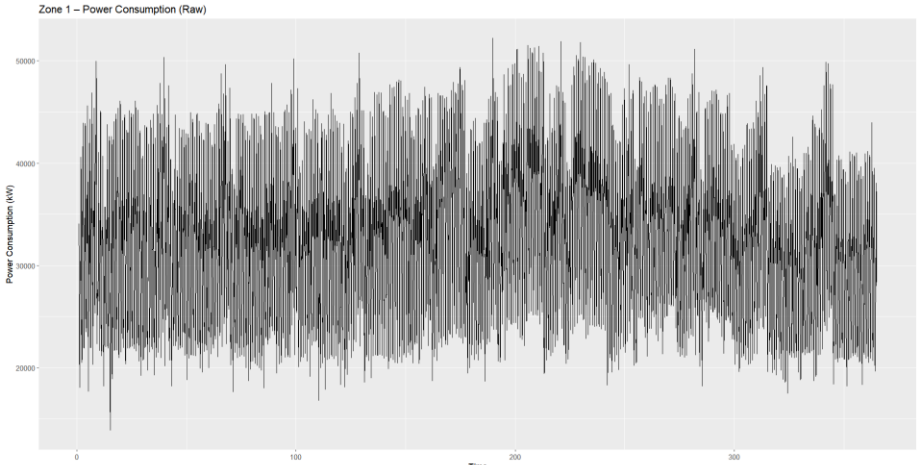


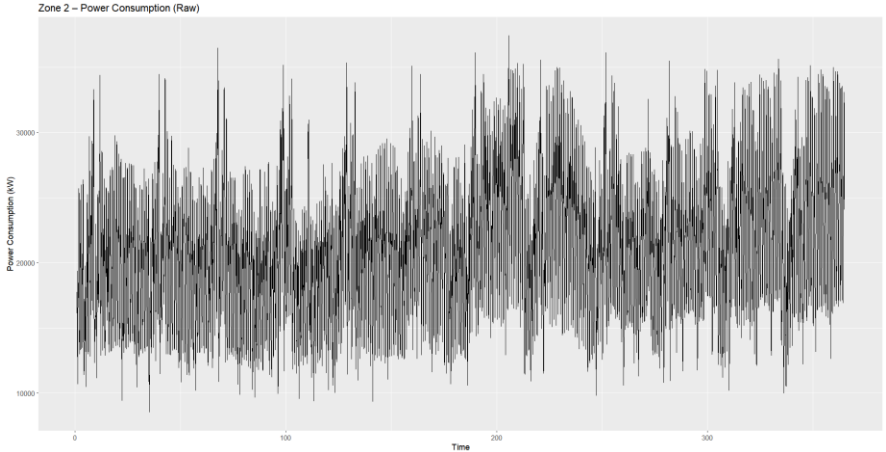


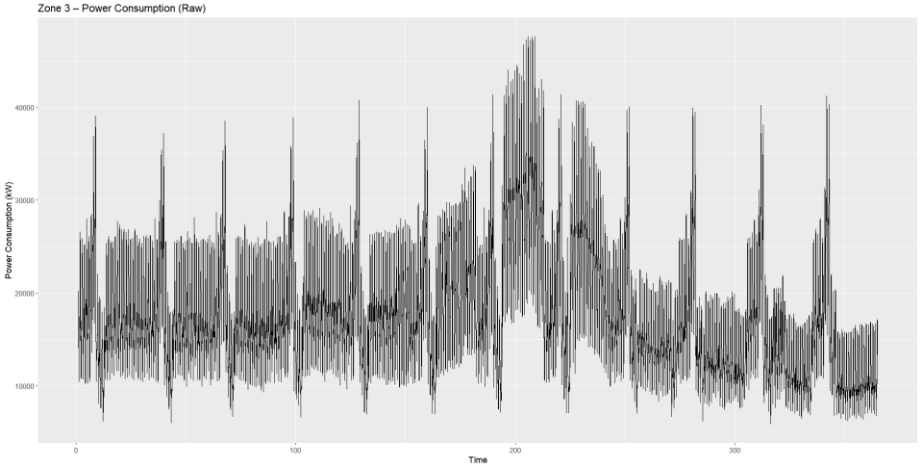
Model	AICc (approx.)	BIC (approx)	Residual Diagnostics (Ljung-Box p-value)
ARIMA(0,1,1)	−949	−941	9.4e-07 (fails)
ARIMA(1,1,1)	−983	−972	0.0105 (fails)
ARIMA(0,1,2)	−954	−942	2.2e-06 (fails)
ARIMA(0,1,1) + drift	−947	−935	9.4e-07 (fails)
Best ARIMA via AICc/BIC (0,1,4)	−999	−980	0.83 (passes)
SARIMA(0,1,4)(0,1,1)[7]	−963	−940	0.91 (passes)

Raw Time Series Overview



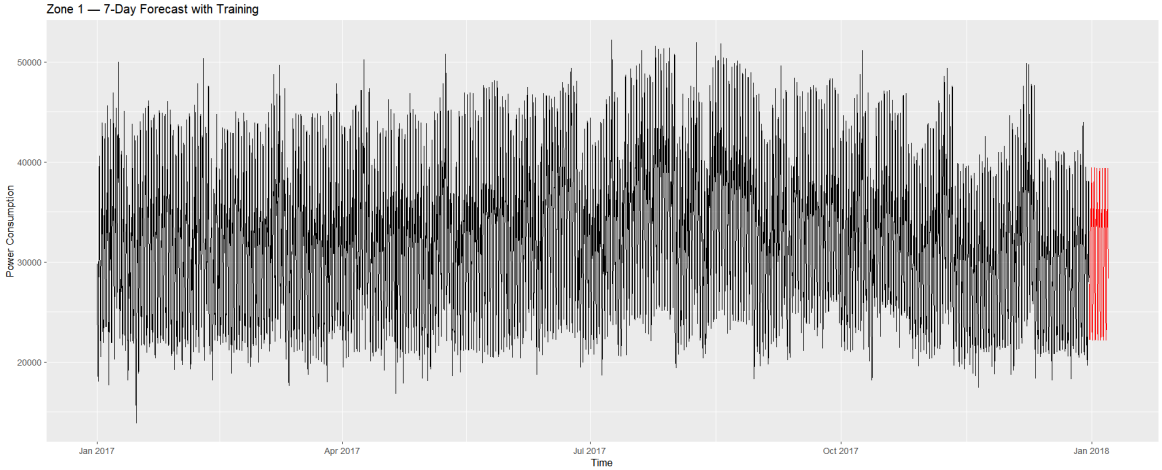




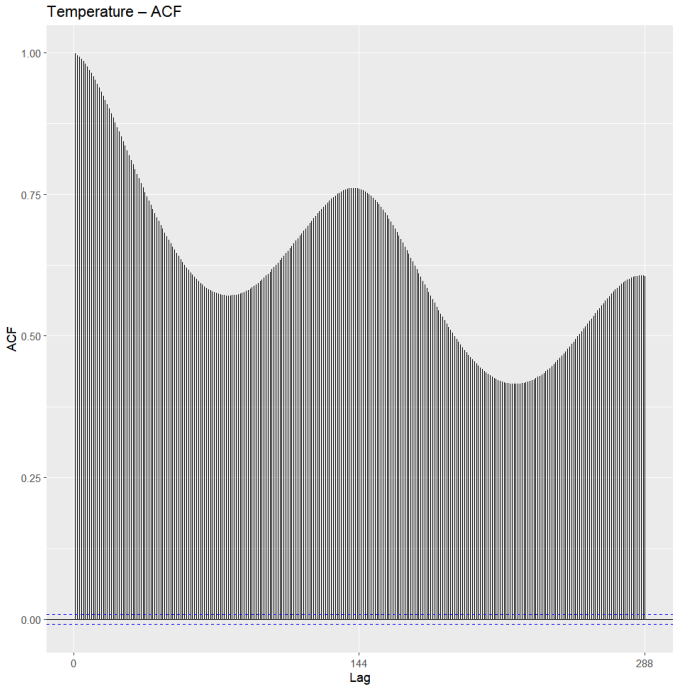


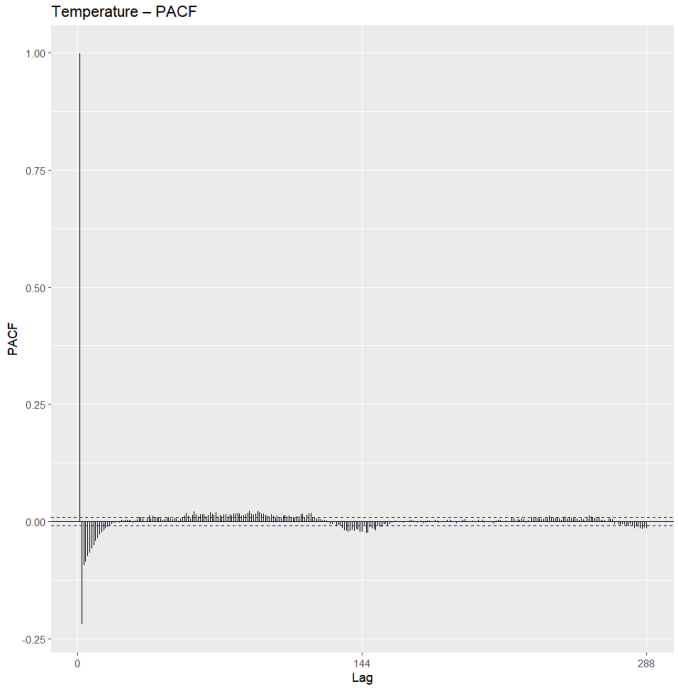
High-Frequency VAR + ARMA–GARCH (10-Minute Data)

Model	Information Criteria / IC Selections	Residual Serial Correlation	Heteroskedasticity (ARCH)	Key Interpretation
VAR(10)	**SC selected p = 10** AIC → 24 HQ → 23 SC → **10**	**Fail** — Portemanteau $\chi^2 \approx 50$, df =8, $p \ll 0.001$	**Fail** — ARCH $\chi^2 \approx 7022$, $p \ll 0.001$	VAR(10) still leaves autocorrelation and strong ARCH patterns.
VAR(24)	**AIC selected p = 24** HQ → 23 SC → 10 FPE → 24	**Fail** — Portemanteau $\chi^2 \approx 7677$, df =480, $p \ll 0.001$	**Fail** — ARCH $\chi^2 \approx 6913$, $p \ll 0.001$	Longer-lag VAR does not resolve autocorrelation or volatility.
ARMA(1,1)–GARCH(1,1) (Baseline, skew-t, Fourier K=6)	**AIC = −6.037 BIC = −6.0336**	**Fail** — LB on z: $\chi^2 \approx 758$, $p \ll 0.001$	**Pass** — LB on z^2: $p \approx 1$	Good variance modeling, mean autocorrelation remains.
ARMA(3,3)–GARCH(1,1) (Final, skew-t, Fourier K=6)	**AIC = −6.0387 BIC = −6.0346**	**Fail** — LB on z: $\chi^2 \approx 733$, $p \ll 0.001$	**Pass** — LB on z^2: $p \approx 1$	Lowest AIC; strong volatility fit; mean still autocorrelated.



Temperature - Dependence & Stationarity Checks



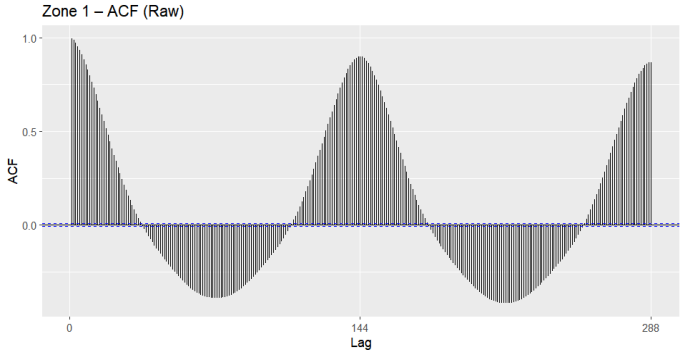


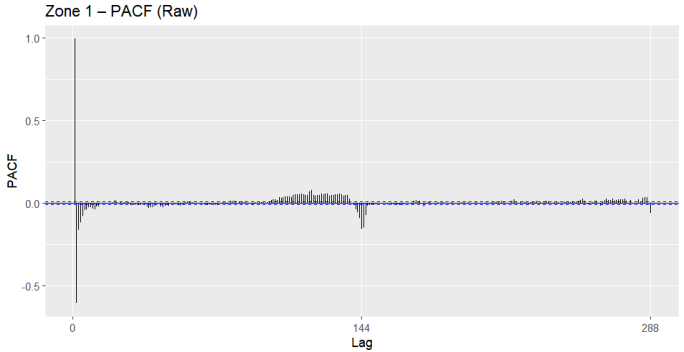

```
Extended ACF (EACF) – Temperature
> print(eacf(temp))
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x x x x x x  x  x  x
1 x x x x x x x x x x  x  x  x
2 x x o o o o o o o o  o  o  o
3 x x x o o o o o o o  o  o  o
4 x x x x o o o o o o  o  o  o
5 x x x x o o o o o o  o  o  o
6 x x x x o o o o o o  o  o  o
7 x x x o x o o o o o  o  o  o
```

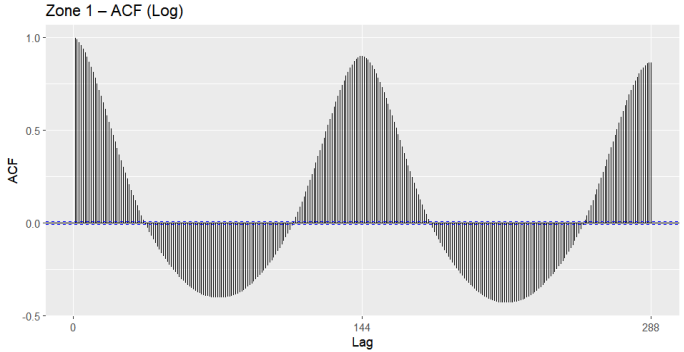
ADF: pass (stationary in differences)
KPSS: fails (not level-stationary)

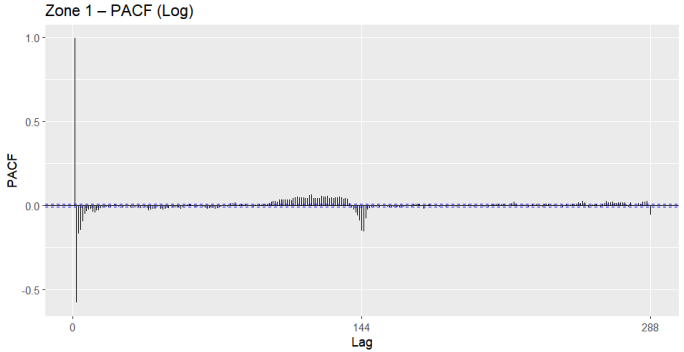
- Strong daily cycle
- Smooth, persistent autocorrelation
- Good candidate as exogenous regressor

Zone 1 – Dependence & Stationarity Checks







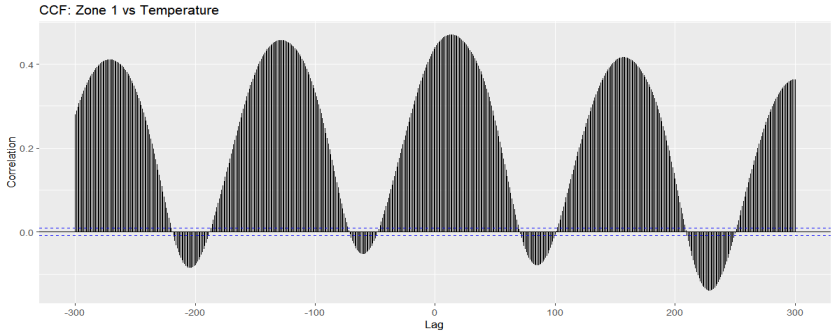


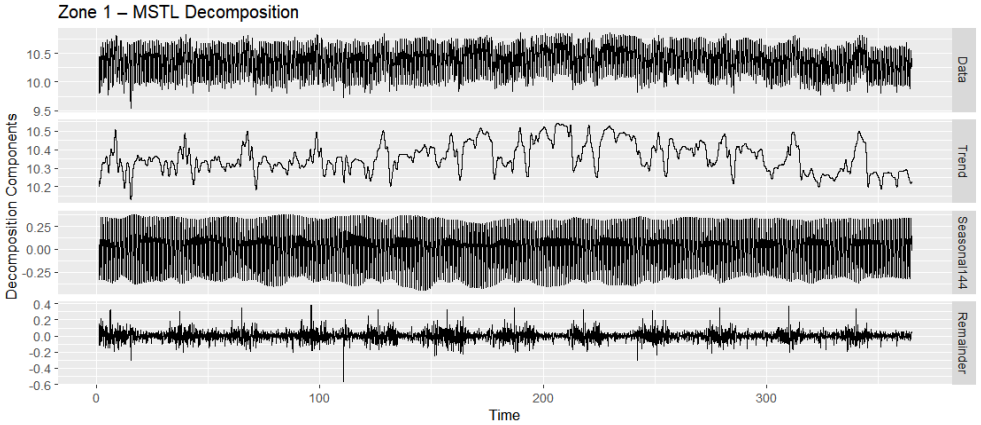

```
Extended ACF (EACF) – Log Zone 1
> print(eacf(lz1))
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x x x x x x x  x  x  x
1 x x x x x x x x x x x  x  x  x
2 x x o x o o o x o x o  o  o  o
3 x x x x o o o x x o o  o  o  o
4 x x x x o o o x x x o  o  o  o
5 x x x x o o o o x x o  o  o  o
6 x x x x o x o o x x x  x  o  o
7 x x o x x x o o x o o  o  o  o
```

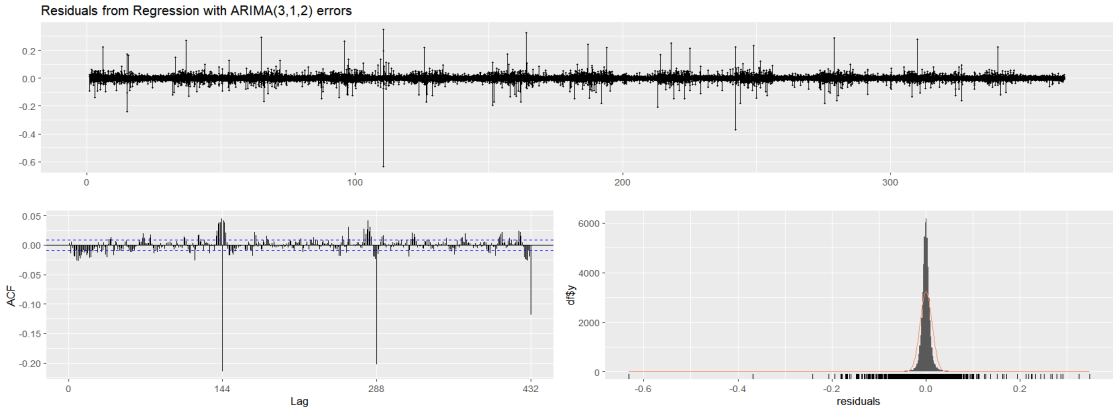
ADF: pass (stationary in differences)
KPSS: fails (not level-stationary)

- Strong intraday seasonality
- Slow decay autocorrelation
- Needs seasonal adjustment

Zone 1: CCF, MSTL, and Model Fit Diagnostics



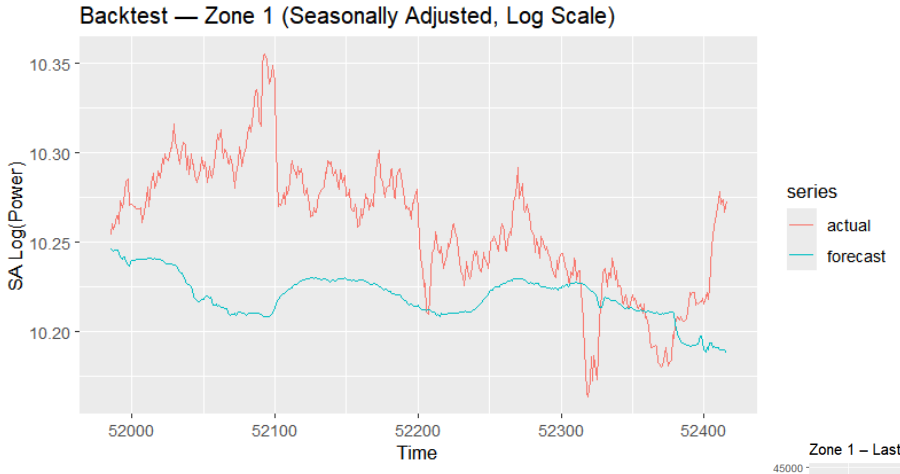




CCF with Temperature: clear repeating co-movement
MSTL: dominant intraday cycle; small trend
Model Form: ARIMA(3,1,2) + Fourier + Temperature
Residuals: centred; variance stable; mild seasonal autocorrelation remains

Model Selection
- AICc lowest among tested structures
- Fourier terms capture daily shape
- Temperature improves short-horizon alignment

Zone 1 — Back-test Accuracy and 3-Day Forecast

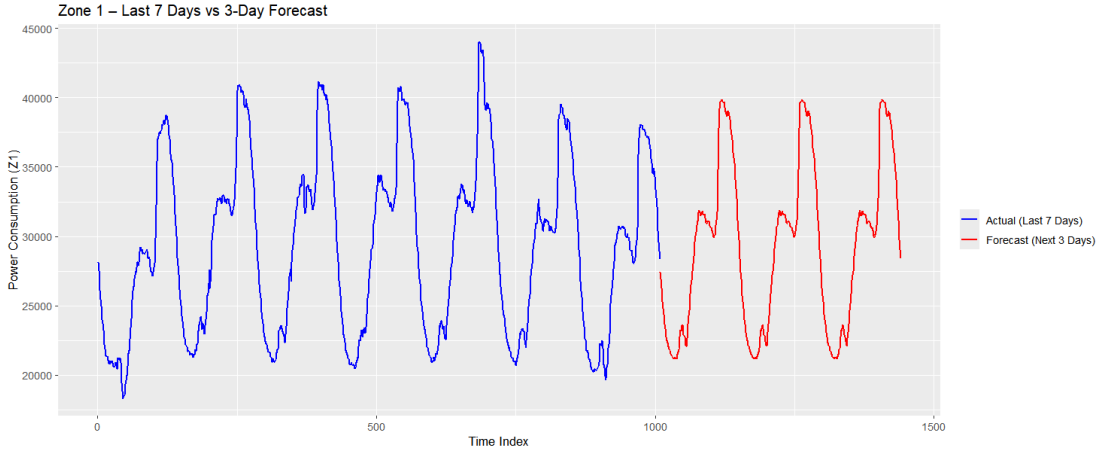


Back-test Accuracy (Seasonally Adjusted, Log Scale):

- **RMSE:** 0.05299
- **MAE:** 0.04404
- **MAPE:** 0.43%

Model Behavior:

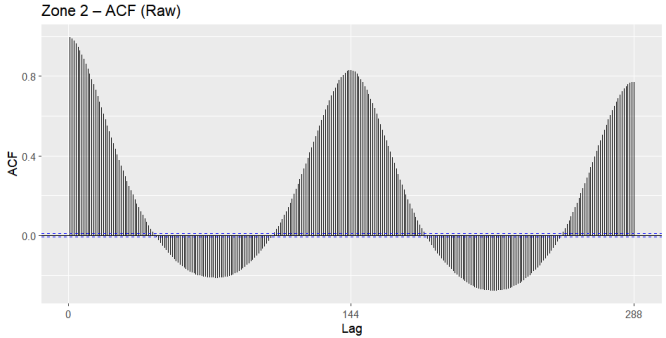
- Captures intraday cycle reliably
- Forecast follows overall shape but smooths sharp drops
- Daily pattern recovered; level shifts under-tracked

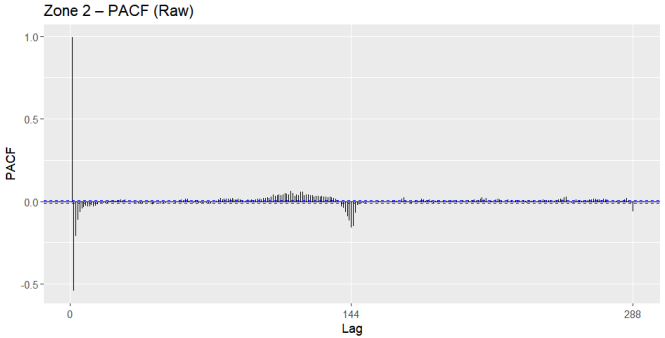


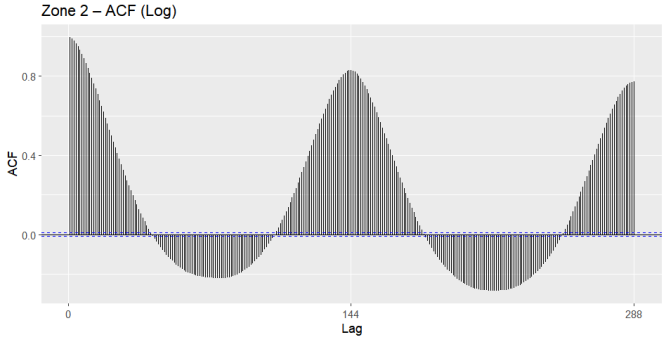
Final Takeaway:

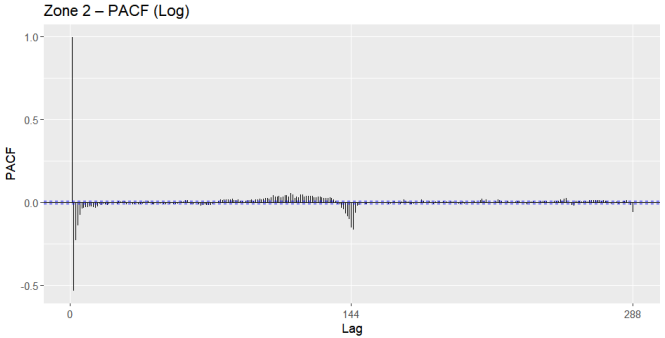
- Strong short-horizon pattern accuracy
- Limited capability for sudden load changes
- Sufficient as a **pattern forecaster**, not a **level-change** predictor

Zone 2 – Dependence & Stationarity Checks







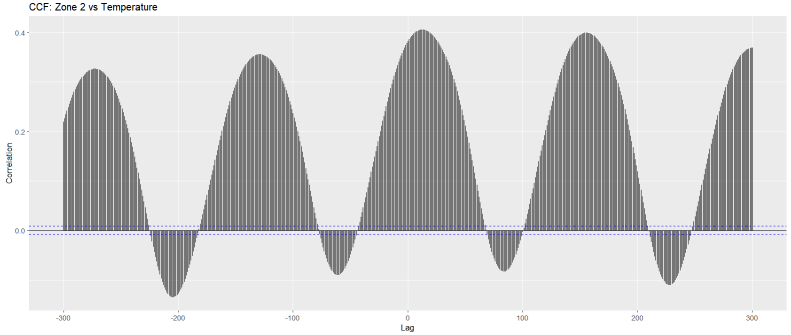


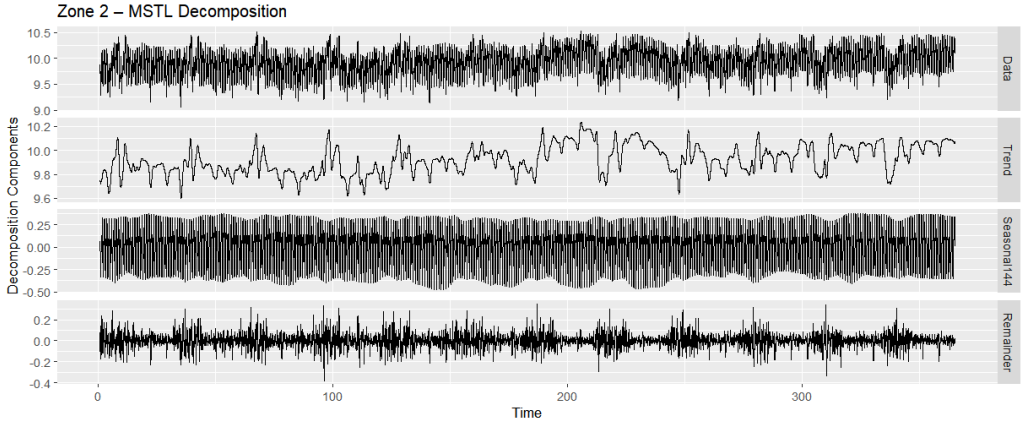

```
Extended ACF (EACF) – Log Zone 2
> print(eacf(lz2))
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x x x x x x x  x  x  x
1 x x x x x x x x x x x  x  x  x
2 x x o o x o o o o o o  o  o  o
3 x x x o x o o o o o o  o  x  o
4 x x x x x o o o o o o  o  x  o
5 x x x x x x o o o o o  o  x  o
6 x x x x o o o o o o o  o  o  o
7 x x x o o x o o o o o  x  o  o
```

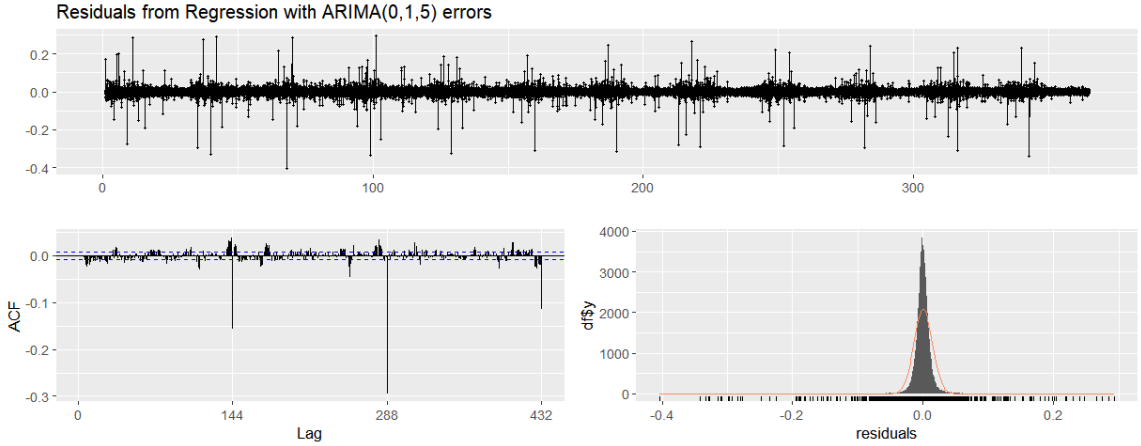
ADF: pass (stationary in differences)
KPSS: fails (not level-stationary)

- High seasonal persistence
- Long memory behavior
- Requires seasonal removal

Zone 2: CCF, MSTL, and Model Fit Diagnostics





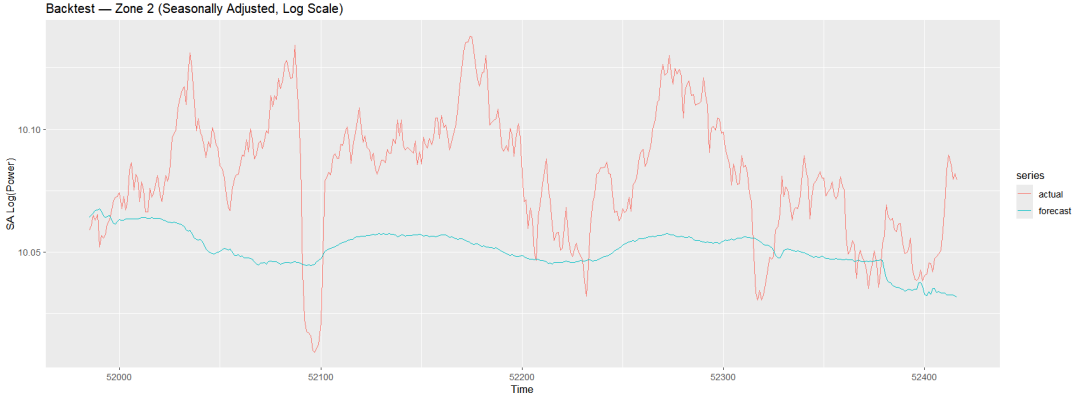


CCF with Temperature: clear repeating cycle; moderate co movement
MSTL: strong 144-point intraday seasonality; mild trend drift
Model Form: ARIMA(0,1,5) + Fourier + Temperature
Residuals: centred; variance stable; small seasonal autocorrelation remains

Model Selection:

- Best AICc among candidate Zone 2 structures
- MA-dominant pattern aligns with EACF shape
- Fourier terms handle daily cycle; temperature improves short-term fit

Zone 2 — Back-test Accuracy and 3-Day Forecast



Back-test Accuracy (Seasonally Adjusted, Log Scale):

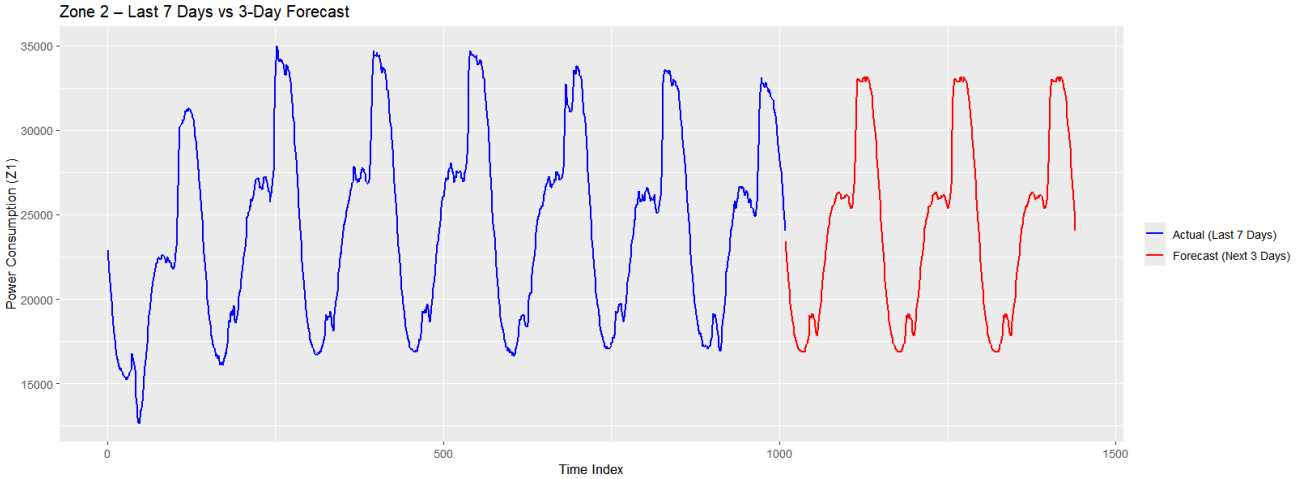
- **RMSE:** 0.03831
- **MAE:** 0.03247
- **MAPE:** 0.32%

Model Behavior:

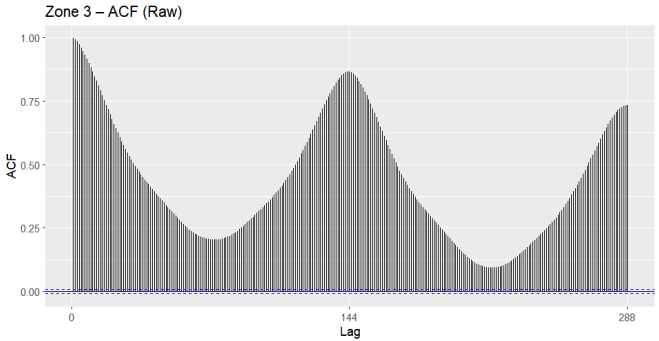
- Recovers intraday cycle consistently
- Forecast tracks the general pattern but stays smoother
- Underestimates sharp rises and dips in demand

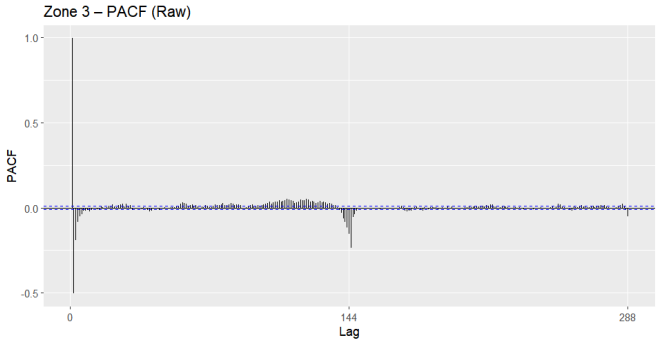
Final Takeaway:

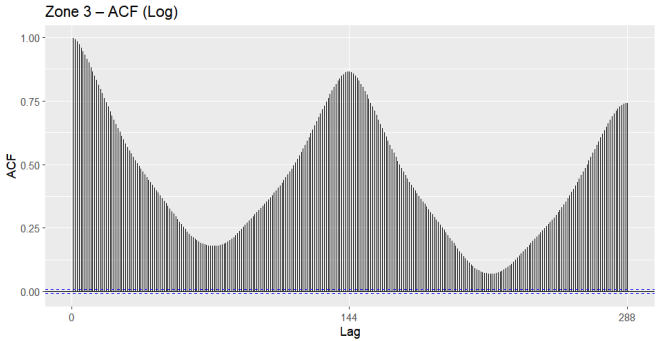
- Reliable short-horizon cycle prediction
- Weak on sudden consumption changes
- Performs well as a **pattern-focused forecaster**, not a **peak-level predictor**

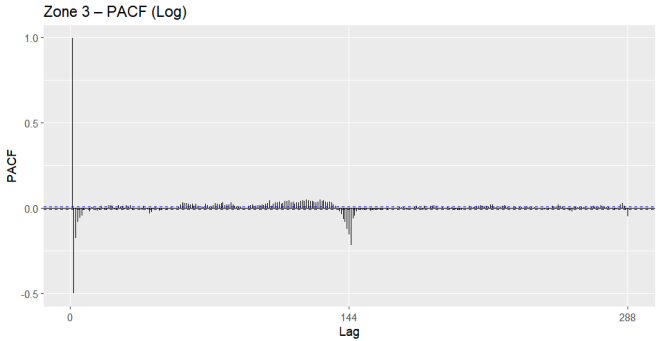


Zone 3 – Dependence & Stationarity Checks







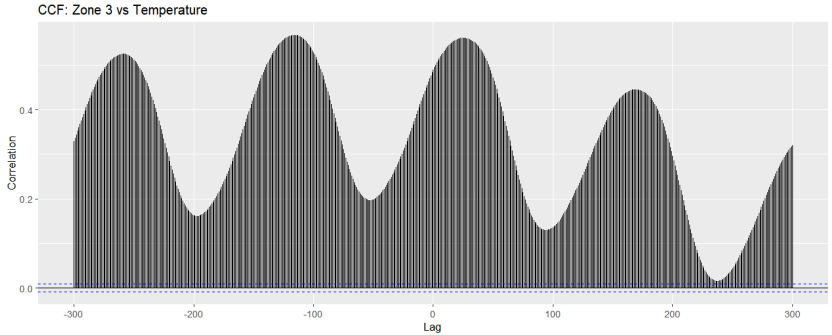


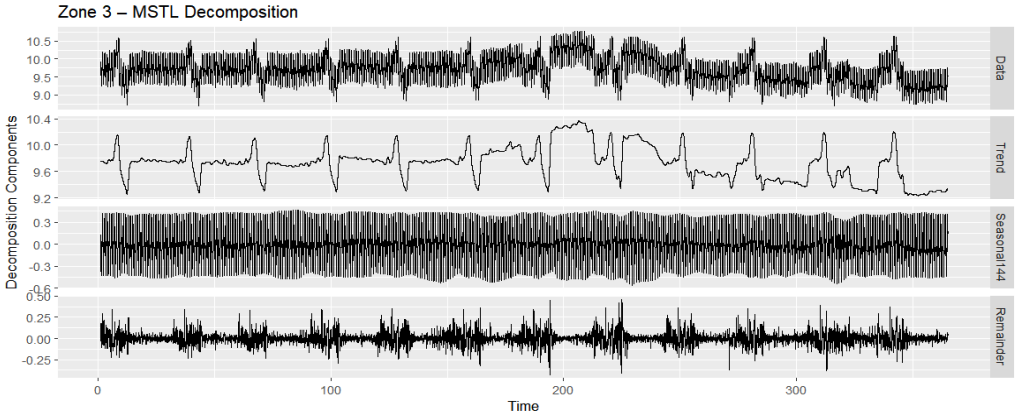

```
Extended ACF (EACF) – Log Zone 3
> print(eacf(lz3))
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x x x x x x x  x  x  x
1 x x x x x x x x x x x  x  x  x
2 x x x o x o o x x o o  x  o  x
3 x x x x x o o x x o o  x  o  x
4 x x x x x o o x x o o  o  o  x
5 x x x x x x o x o x o  o  o  o
6 x x x x x o o x o x o  x  o  o
7 x x o x x x x o o x o  o  o  o
```

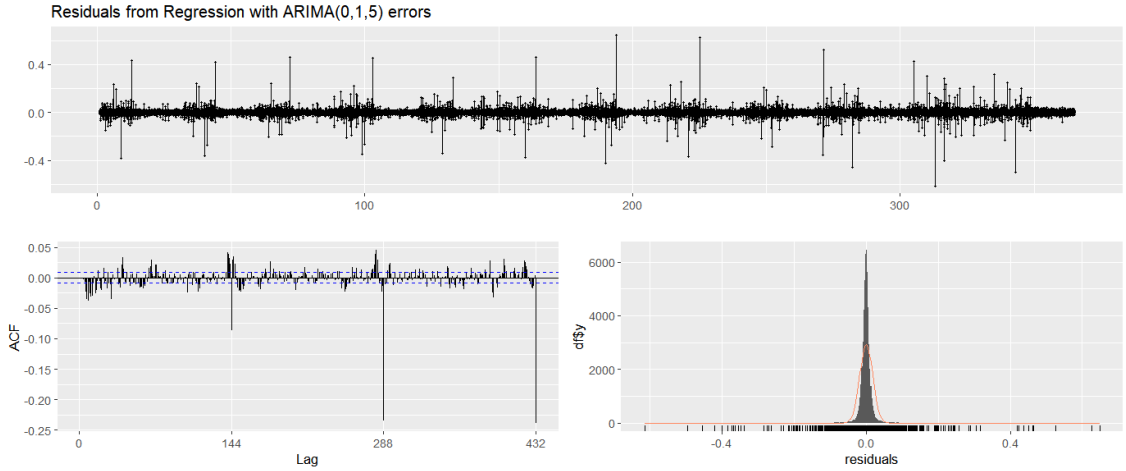
ADF: pass (stationary in differences)
KPSS: fails (not level-stationary)

- Most volatile zone
- Deep seasonal structure
- MA-dominant pattern

Zone 3: CCF, MSTL, and Model Fit Diagnostics





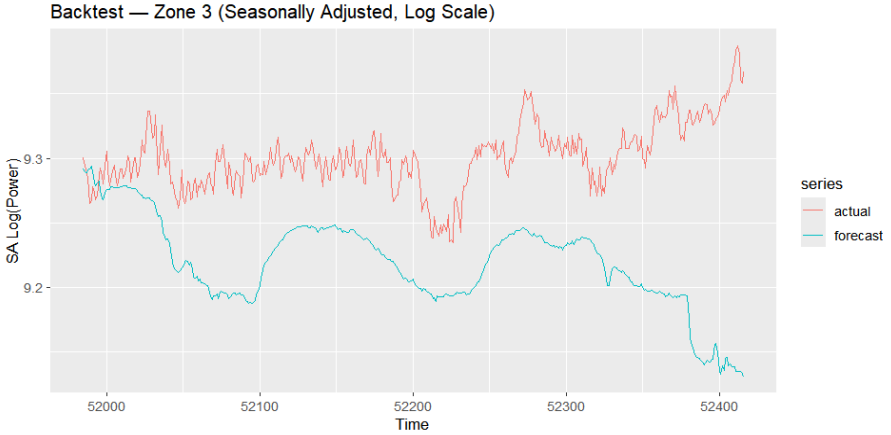


CCF with Temperature: repeating high-amplitude cycle; strongest temperature link
MSTL: clear intraday seasonality; trend more irregular than Z1/Z2
Model Form: ARIMA(0,1,5) + Fourier + Temperature
Residuals: centred; variance stable; seasonal lags still visible

Model Selection

- Lowest AICc among Zone 3 candidates
- MA-dominant remainder matches EACF pattern
- Fourier captures daily structure; temperature improves short-horizon response

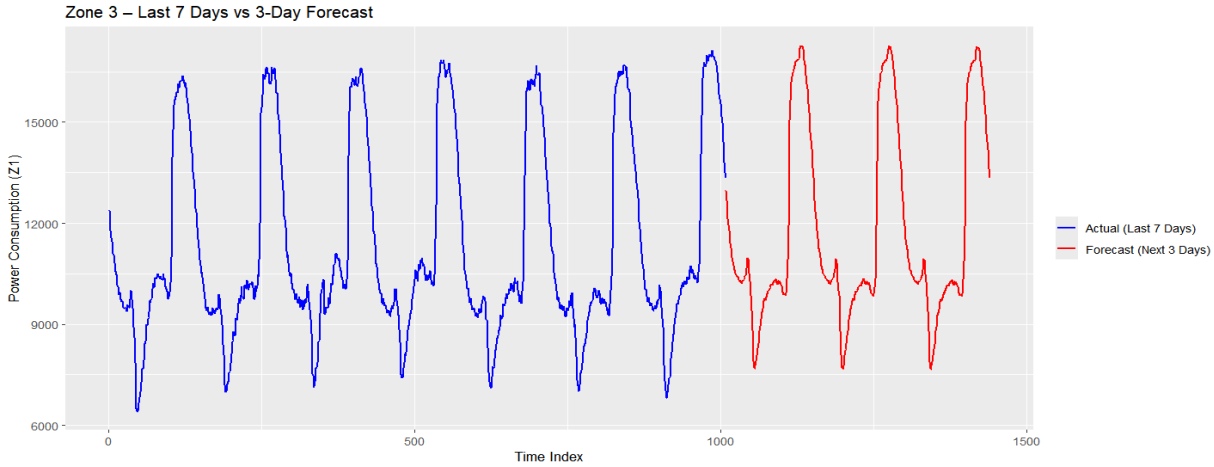
Zone 3 — Back-test Accuracy and 3-Day Forecast



Back-test Accuracy (Seasonally Adjusted, Log Scale):
- **RMSE:** 0.04122
- **MAE:** 0.03490
- **MAPE:** 0.37%

Model Behavior:
- Daily cycle captured with consistent timing
- Forecast stays smoother than true consumption
- Underestimates sharp peaks and sudden load changes

Final Takeaway:
- Strong short-term seasonal pattern recovery
- Limited handling of abrupt demand spikes
- Effective for **cycle prediction**, not for **level-shift or volatility forecasting**



Challenges Encountered

- Structural limitations of the dataset
- Time and frequency structure problems
- Daily aggregation wiped out essential signal
- Autocorrelation far stronger than standard models assume
- Multivariate modeling instability
- Exogenous and seasonal components were difficult to extract cleanly

Key Learnings

- ✓ High-frequency data behaves fundamentally differently
- ✓ Strong autocorrelation dominates model behavior
- ✓ Seasonality is multi-layered, not single-cycle
- ✓ Aggregation can destroy critical information
- ✓ Exogenous signals may be weak even if correlated
- ✓ Classical models struggle with sudden level shifts
- ✓ Multi-step forecasting amplifies uncertainty

Conclusion

Our analysis showed that reliable forecasting for this dataset is only achievable when the model directly captures the strong intraday seasonal structure. Daily ARIMA/SARIMA and VAR-GARCH approaches struggled with the extreme autocorrelation and high-frequency volatility. The Fourier + Exogenous Regressor model produced the most stable short-term results by learning the 10-minute cycle, though it still smoothed sharp level shifts. Overall, the project established a solid baseline forecaster and clarified the limits of classical time-series methods for dense, operational load data.

Future Work

- Extend dataset to multiple years to capture annual and long-term seasonal structure
- Incorporate additional exogenous drivers (holidays, operational schedules, regional demand)
- Explore machine-learning and hybrid models (LSTM, TFT, Prophet + ARMA residuals) for regime shifts

Thank You