Technical Report

Predicting Daily Electricity Consumption in Norway

By Charles Aleksander Ravndal

1. Executive summary

This capstone examined whether daily electricity consumption in Norway can be predicted using weather, seasonality, and lagged demand features. Accurate demand forecasting supports sustainable hydropower reservoir management, grid stability, and reduced reliance on imports.

The initial hypothesis assumed that consumption could be predicted primarily from weather and seasonality. Results disproved this: lagged demand features (yesterday's and last week's consumption) were far more important.

The best-performing model was Ridge Regression (α=1000), which achieved a test R^2 of 0.92. Weather and seasonality contributed little predictive value, reflecting recent climate shifts (colder summers, milder winters) and Norway's high electrification, particularly from widespread electric vehicle adoption.

2. Introduction

Norway's electricity system is unique: nearly 100% of generation comes from hydropower, making it one of the world's cleanest grids. At the same time, demand forecasting is becoming more complex due to structural factors:

- Electrification: Over 80% of new cars sold in Norway are electric, contributing to high but stable demand.
- Climate shifts: Colder summers and milder winters weaken traditional seasonal peaks.

Research question:
Can we predict daily electricity consumption in Norway using weather and seasonality, and what insights can we gain about the main drivers of demand?

Hypothesis:
Consumption would be driven primarily by temperature and seasonal cycles.

3. Data and methodology

3.1 Data Sources

- ENTSO-E Transparency Platform – Daily total load for Norway, 2015–2025.
- Norwegian Meteorological Institute (MET Norway) – Daily temperature and precipitation data.

3.2 Preprocessing

- Merged consumption and weather data by date.
- Engineered features: Lag features (lag1, lag7).
    - Heating Degree Days (HDD),
    - Cooling Degree Days (CDD).
    - Calendar features (day of week, month, weekend flag).
- Standardized inputs for Ridge regression.

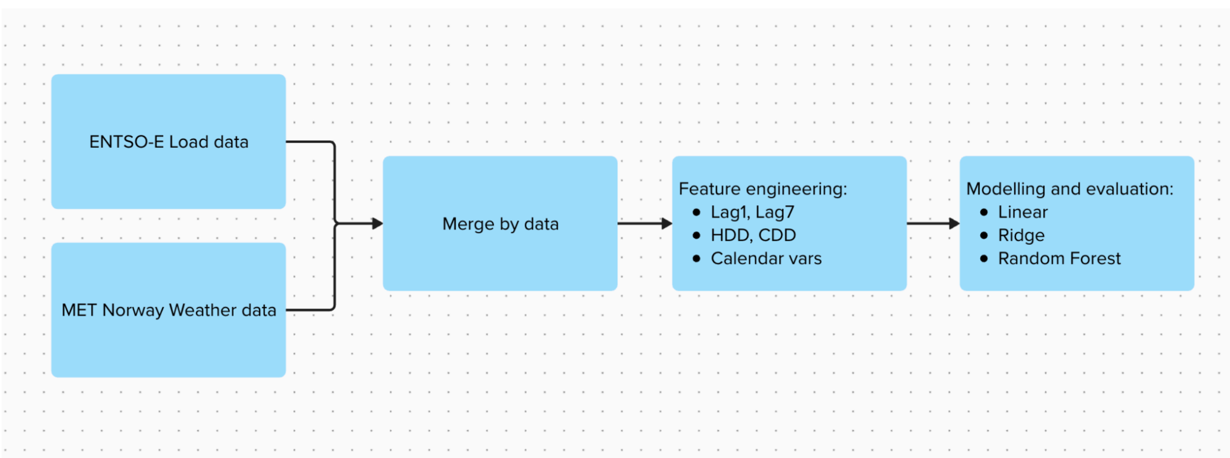

Figure 1. Workflow of data preparation and feature engineering

4. Exploratory data analysis

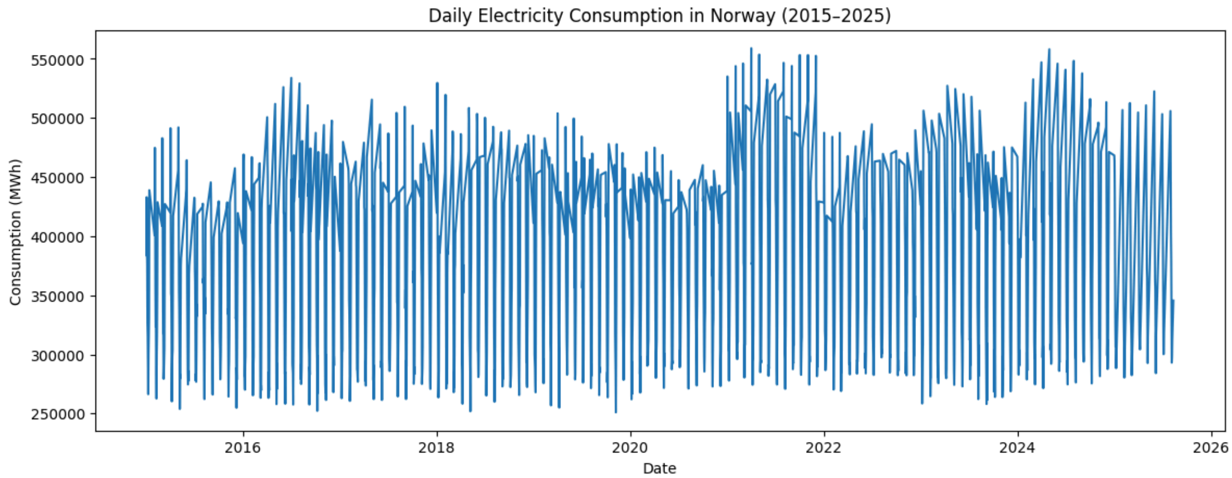

Figure 2. Daily electricity consumption in Norway (2015 - 2025)

Daily electricity consumption in Norway between 2015–2025 is shown in **Figure 2**. The series is relatively stable, with daily fluctuations but no major structural shifts. Average daily consumption typically falls between 300,000 and 500,000 MWh.

Seasonal decomposition (**Figure 3**) highlights a recurring yearly pattern. Demand is moderately higher in spring and early summer (April–June) and somewhat lower in winter (December–February). However, the peaks and troughs are less pronounced than expected, consistent with recent trends of colder summers and milder winters.

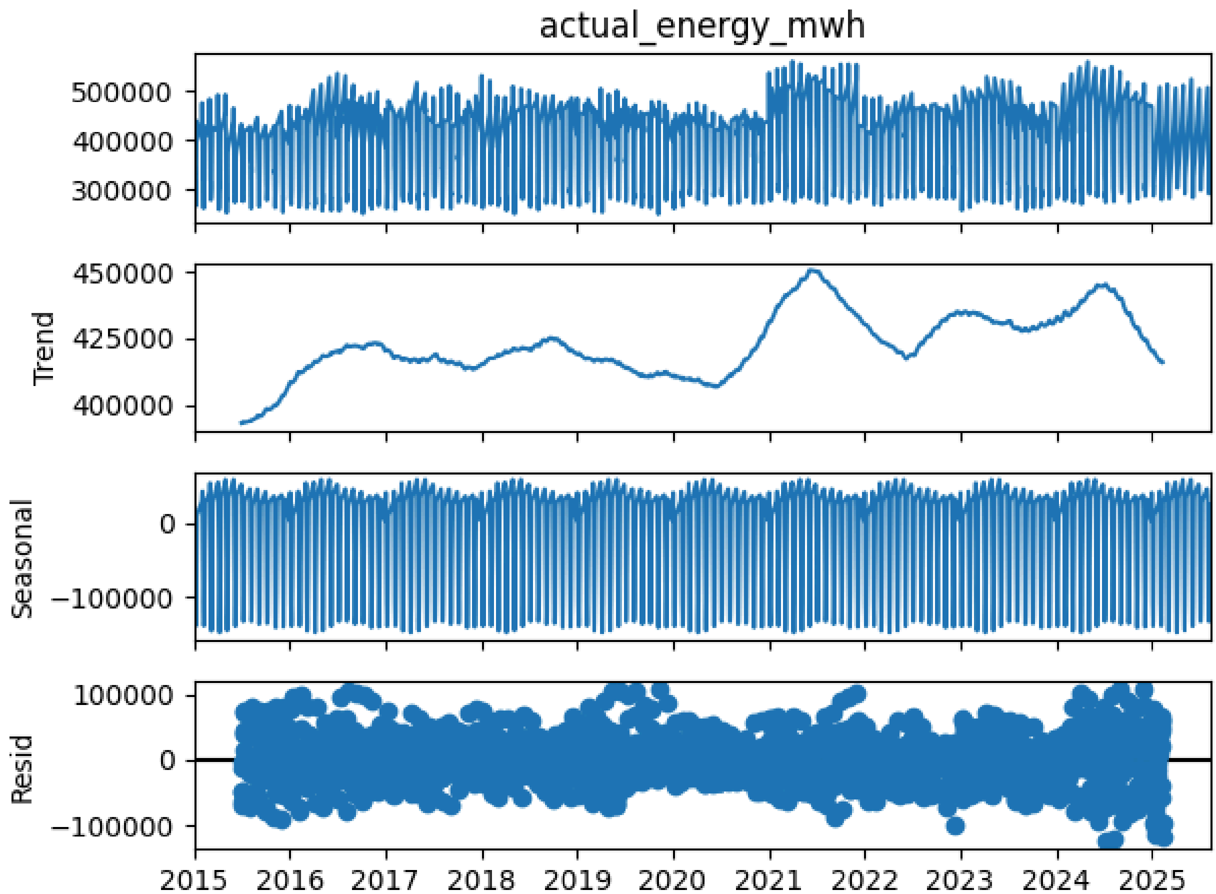

Figure 3. Seasonal decomposition of daily consumption

Correlation analysis between consumption and weather variables (**Figure 4**) confirms that weather plays only a minor role. Temperature variables (mean, min, max) show weak positive correlations with consumption (≈0.06–0.07), while precipitation has almost no effect. In contrast, lagged consumption variables (not shown in the heatmap but evaluated in later modeling) display much stronger predictive power.

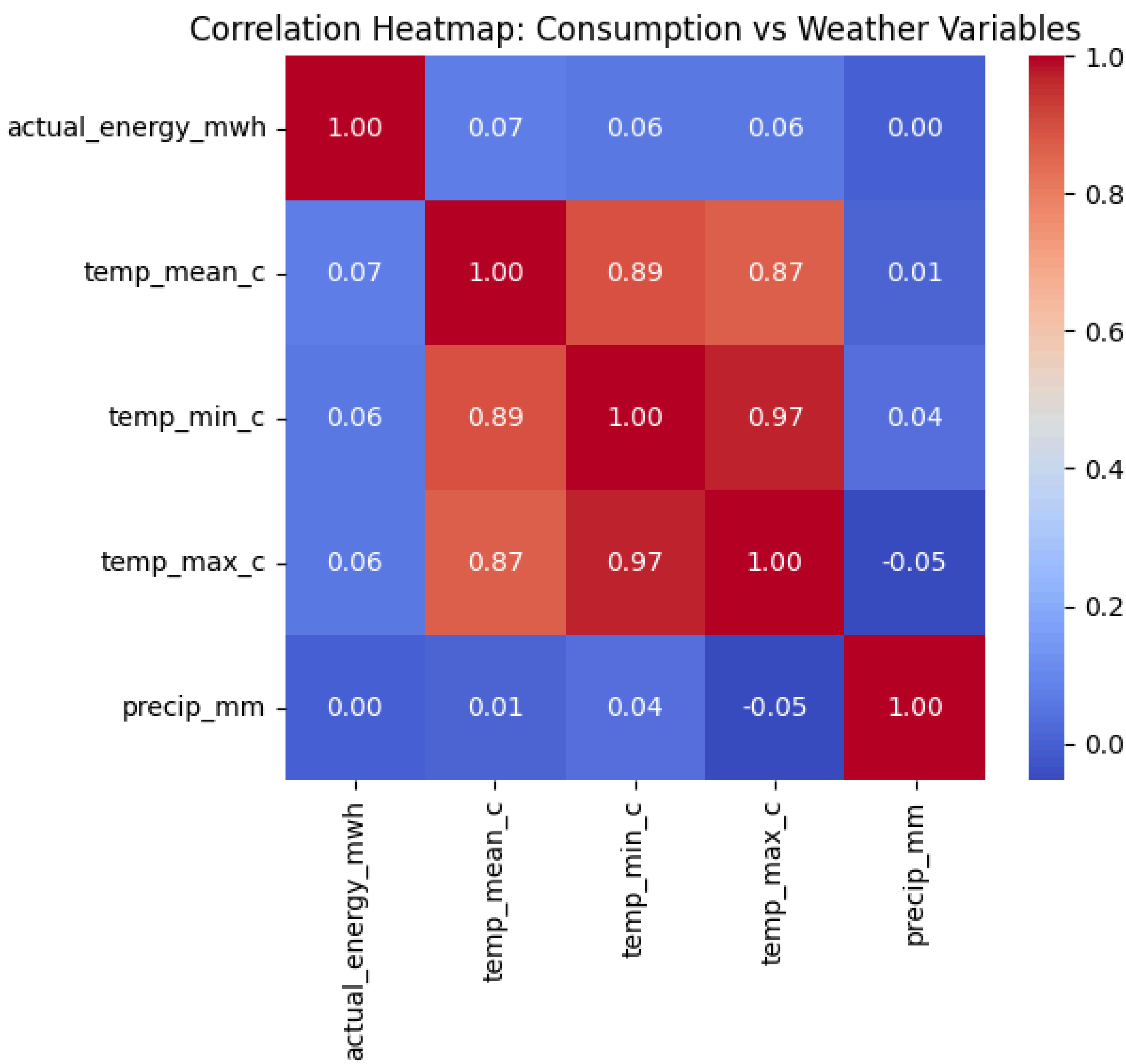

Figure 4. Correlation heatmap: consumption vs. weather variables

5. Modeling and results

5.1 Setup 1: Weather + Seasonality

Linear, Ridge, and Random Forest models were trained on weather and calendar features. All three performed poorly, with negative R^2 scores on the test set. This confirms that weather and seasonality alone cannot explain daily variation in Norwegian electricity demand.

model	RMSE_train	MAE_train	R2_train	RMSE_test	MAE_test	R2_test	n_train	n_test
Linear (WS)	67486.32	52885.27	0.01	69842.25	59957.9	-0.18	3646	224
Ridge (WS)	67605.35	52851.21	0.01	69290.16	59249.43	-0.16	3646	224
RandomForest (WS)	25499.2	20011.19	0.86	69893.85	57989.61	-0.18	3646	224

Table 1. Model performance (Weather + Seasonality)

5.2 Setup 2: Lags only

Using lag features (previous day and previous week consumption) produced strong results. Linear and Ridge regression both achieved test $R^2 \approx 0.92$, while Random Forest overfit heavily (train $R^2 = 0.98$ vs test $R^2 = 0.55$).

model	RMSE_train	MAE_train	R2_train	RMSE_test	MAE_test	R2_test	n_train	n_test
Linear (Lags)	29002.65	18607.93	0.82	18345.44	13488.39	0.92	3646	224
Ridge (Lags)	29002.65	18607.93	0.82	18345.44	13488.39	0.92	3646	224
RandomForest (Lags)	9476.27	5400.95	0.98	42889.95	33758.89	0.55	3646	224

Table 2. Model performance (Lags only)

5.3 Setup 3: Full Model

Combining all features did not significantly improve results. Ridge remained the most robust model, with test $R^2 \approx 0.92$. Ordinary least squares (Linear Regression) exhibited multicollinearity issues, while Random Forest again overfit.

model	RMSE_train	MAE_train	R2_train	RMSE_test	MAE_test	R2_test	n_train	n_test
Linear (Full)	28905.68	18642.35	0.82	18680.35	13728.24	0.92	3646	224
Ridge (Full)	28939.42	18599.53	0.82	18545.32	13559.23	0.92	3646	224
RandomForest (Full)	9261.69	5302.68	0.98	40788.81	31755.13	0.6	3646	224

Table 3. Model performance (Full model)

5.4 Model Comparison Summary

The results are summarized in **Figure 5**, which compares R^2 scores across all setups. Only the lag-based models achieved high predictive accuracy. Weather and seasonality added little to no explanatory power.

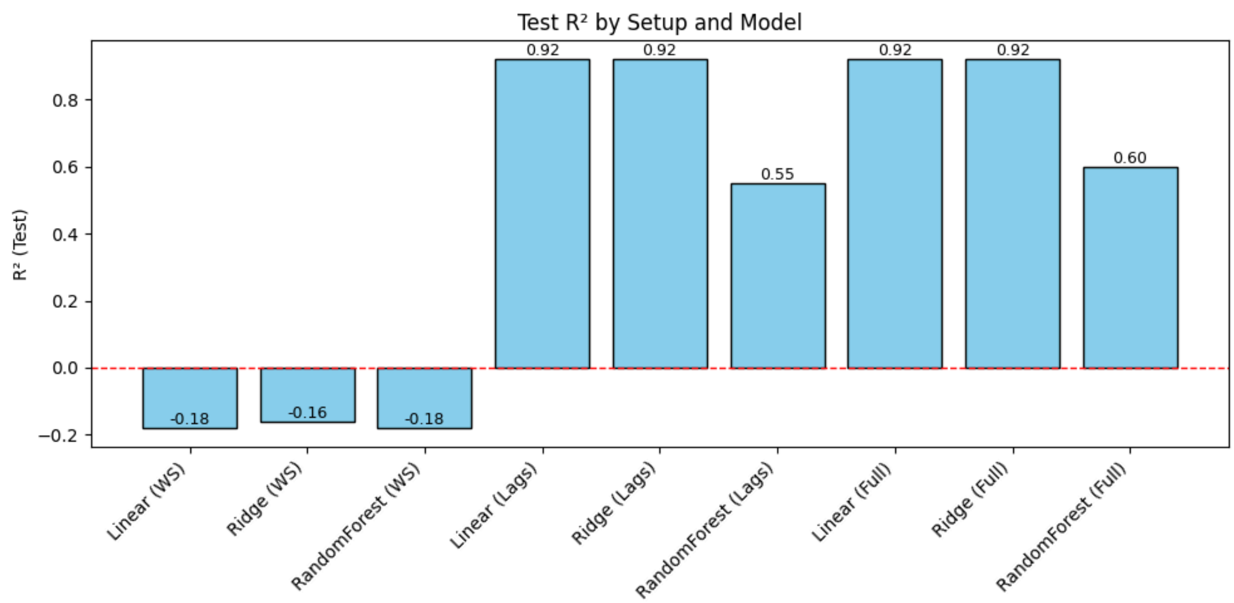

Figure 5. Comparison of R^2 by setup and model

6. Optimization and diagnostics

Ridge regression was optimized through cross-validation, with an α value of 1000 selected. This provided the best balance between accuracy and coefficient stability.

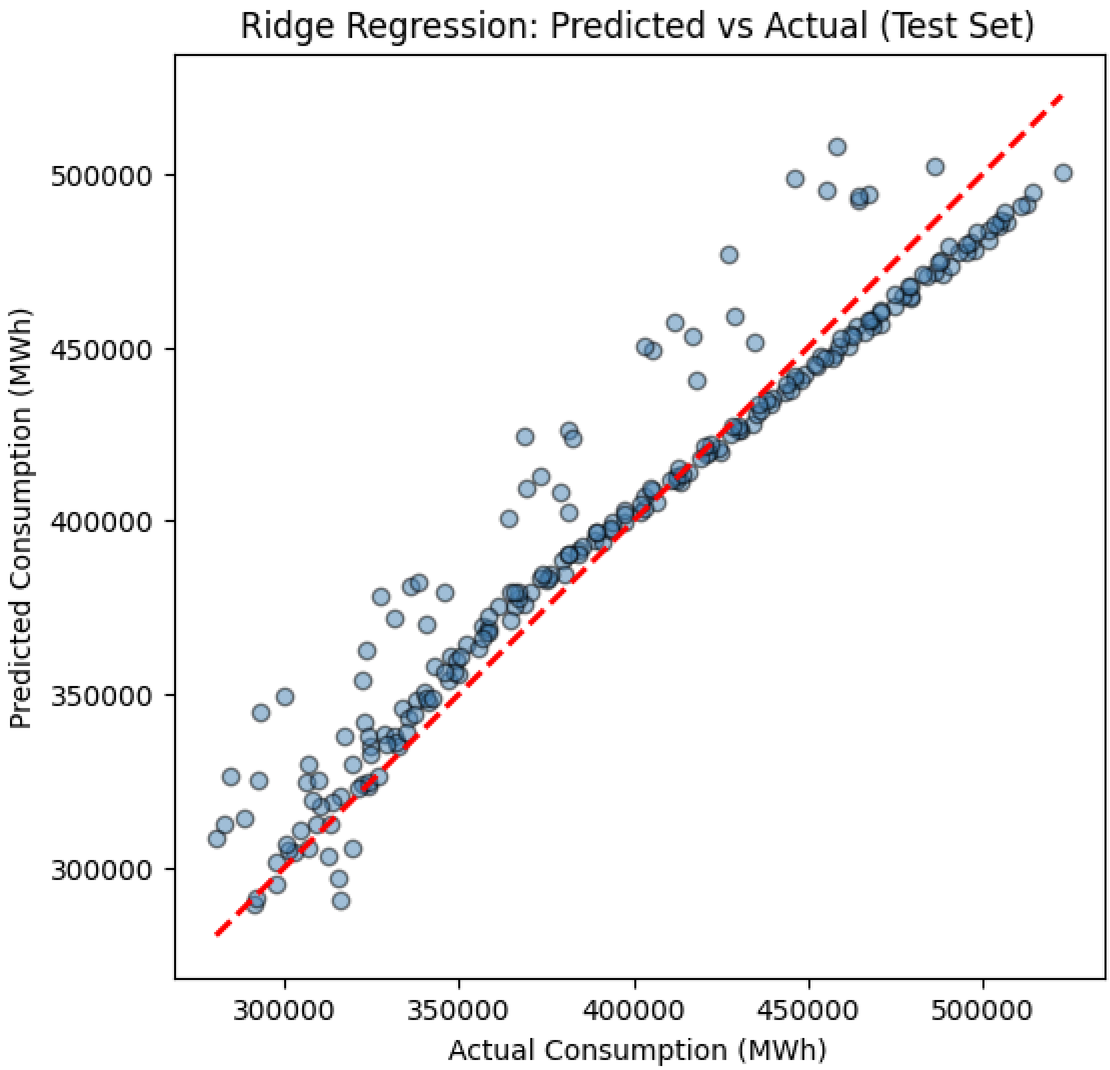

Figure 6. Ridge Regression. Predicted vs. Actual (Test set)

Predicted vs actual consumption (**Figure 6**) shows close alignment, with most points clustering around the diagonal. Residual analysis further validates model performance. Residuals over time (**Figure 7**) fluctuate randomly around zero, with no systematic bias. The distribution of residuals (**Figure 8**) is approximately symmetric, though slightly skewed for extreme values.

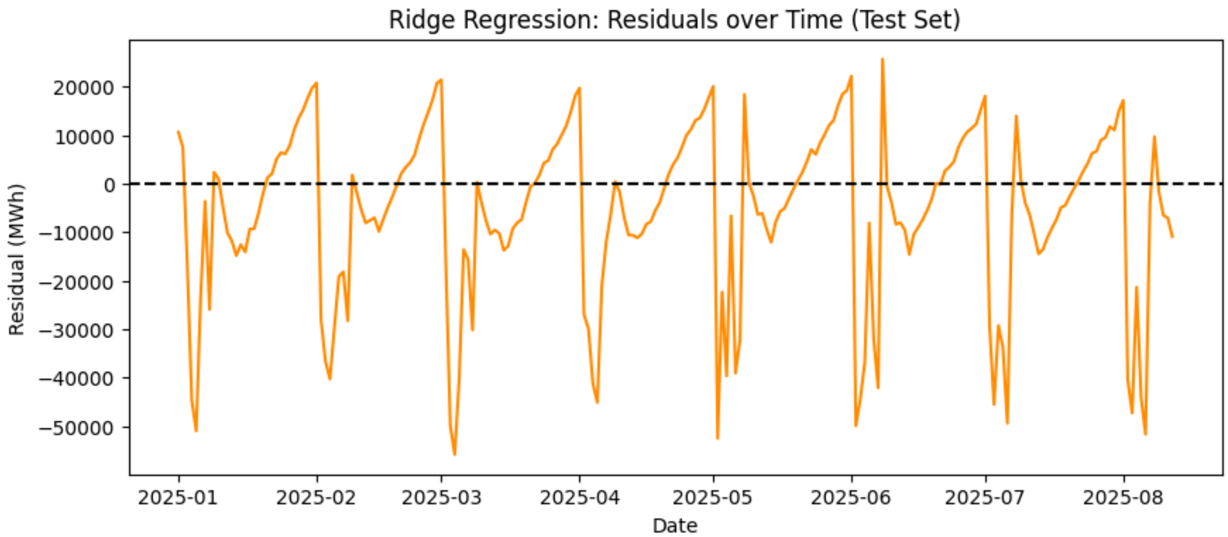

Figure 7. Ridge Regression: Residuals over Time (Test set)

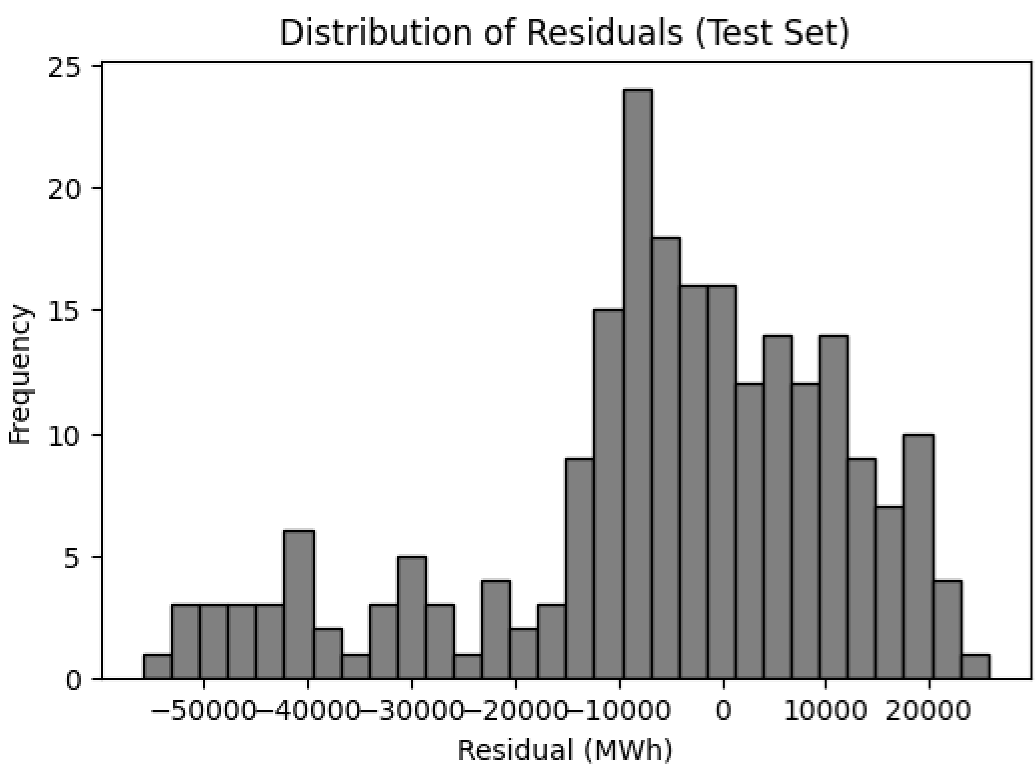

Figure 8. Distribution of residuals (Test set)

These diagnostics confirm that the Ridge model generalizes well and produces unbiased predictions.

7. Ethical and societal considerations

Electricity demand forecasting has clear sustainability benefits, but also raises ethical and societal issues that must be acknowledged:

- Data privacy: This project used only aggregated, public data (ENTSO-E, MET Norway). No individual-level or personal data was included, which minimizes privacy risks.
- Transparency: The modeling workflow is fully reproducible through the open-source notebook and datasets. This transparency reduces risks of “black-box” decision-making.
- Bias and fairness: Lag-based models strongly depend on past consumption patterns. While effective for short-term forecasting, this approach may underrepresent emerging structural changes (e.g., new policies, electrification trends, extreme climate events). This limitation should be communicated clearly to avoid over-reliance.
- Sustainability impact: Accurate forecasts support efficient hydropower reservoir management, reduce waste, and facilitate integration of additional renewable energy. However, poor forecasts could lead to unnecessary imports or inefficient water usage. Responsible deployment therefore requires continuous validation and monitoring.
- Regulatory compliance: The project aligns with European guidelines (e.g., EU AI Act principles), emphasizing human oversight, transparency, and proportionality in AI/ML applications.

By addressing these aspects, the project ensures that its findings are not only technically sound but also socially responsible.

8. Discussion

8.1 Hypothesis vs Findings

The project set out with the hypothesis that temperature and seasonal indicators would be the main drivers of electricity demand in Norway. This was disproved: models using only weather and seasonality performed poorly, with negative R^2 values. In contrast, lagged consumption features explained most of the daily variation, achieving a test R^2 of 0.92.

8.2 Climate and electrification context

Two structural factors explain this outcome. First, recent climate shifts in Norway — colder summers and milder winters — have weakened the traditional seasonal patterns in electricity demand. Second, Norway’s high level of electrification, particularly in transport where nearly

90% of new cars sold in 2024 were electric, has contributed to high but relatively stable consumption. These trends reduce the influence of weather as a predictor while reinforcing the importance of short-term persistence.

8.3 Alignment with forecasting literature

The strong performance of lag-based models is consistent with global forecasting literature, which emphasizes autoregressive approaches for short-term load prediction. Studies confirm that daily consumption tends to follow recent patterns more closely than weather fluctuations, making lagged variables highly predictive.

8.4 Ethical and societal implications

While technically effective, lag-based forecasting also has ethical and societal implications. By relying heavily on historical patterns, such models risk underestimating the impact of emerging factors, including new policies, extreme climate events, or rapid electrification. This underscores the need for transparency in communicating limitations and for human oversight when forecasts are used in decision-making. As highlighted in Section 7, the sustainability impact is significant: good forecasts can optimize hydropower reservoir use, but poor ones could contribute to inefficiency or unnecessary imports.

8.5 Limitations and future work

- Lag features capture short-term demand well, but are less effective for long-term or structural changes.
- Holiday and event effects were not included.
- Future work should evaluate hybrid approaches (e.g., ARIMAX, LSTM) that combine persistence with exogenous factors, while ensuring transparency and compliance with ethical guidelines.

9. Conclusion

This project demonstrated a complete machine learning workflow applied to daily electricity consumption in Norway. While the initial hypothesis was disproved, the analysis delivered a robust model and meaningful insights.

The final model, Ridge Regression with $\alpha=1000$, achieved a test R^2 of 0.92, showing that short-term persistence is the dominant driver of electricity demand. Weather and seasonality contributed little, reflecting climate trends and Norway’s highly electrified society.

The results carry clear sustainability implications. Accurate forecasts support more efficient hydropower reservoir management, reduce reliance on imports, and strengthen the resilience of

Norway's renewable-heavy grid. At the same time, ethical considerations are critical: models must remain transparent, continuously validated, and used with awareness of their limitations.

In conclusion, this capstone not only produced a technically sound forecasting model, but also highlighted how climate change, electrification, and ethical responsibility intersect in shaping the future of energy forecasting in Norway.

9. References

- ENTSO-E. (n.d.). Transparency platform: Total load – Day-ahead values for Norway. Retrieved August 2025, from https://transparency.entsoe.eu
- Norwegian Meteorological Institute. (n.d.). Weather and climate data. Retrieved August 2025, from https://thredds.met.no/thredds/catalog/senorge/seNorge_2018/Archive/catalog.html
- Reuters. (2025, January 2). Norway: Nearly all new cars sold in 2024 were fully electric. Retrieved August 2025, from https://www.reuters.com/business/autos-transportation/norway-nearly-all-new-cars-sold-2024-were-fully-electric-2025-01-02
- Gridserve. (2025, January). Why Norway leads the world in EV adoption. Retrieved August 2025, from https://www.gridserve.com/why-norway-leads-the-world-in-ev-adoption
- Amasyali, K., & El-Gohary, N. M. (2022). Autoregressive methods for short-term power demand forecasting: A systematic review. Energies, 15(18), 6729. https://doi.org/10.3390/en15186729
- Sahoo, B., Chou, J. S., & Huang, H. C. (2017). Short-term load forecasting using ARIMA model. arXiv preprint arXiv:1705.04378. https://arxiv.org/abs/1705.04378