Energy Consumption Forecasting Methodology Report

This project is an experiment to investigate the potential of vibe-coding to produce energy forecasts. The project has almost exclusively been generated by several AI agents which have been cross-referenced. The two contributors have very limited experience with and insights on data science and machine learning. The accuracy of the data will give an insight if this method should be employed more often for data heavy problems. Succes would mean little labour and no skills are needed in the future to produce results in this area.

This document outlines the technical approach, modelling techniques, and validation strategy employed for forecasting energy consumption for the Fortum challenge, providing both a 48-hour operational forecast and a 12-month strategic forecast based on three years of historical data.

1. Modelling Techniques and Rationale

The project adopted a dual-strategy approach, selecting distinct models optimized for the specific horizons (short-term vs. long-term) to maximize accuracy and utility for Fortum's operational needs.

48-Hour (Short-Term) Forecast

Technique: SARIMA (Seasonal AutoRegressive Integrated Moving Average)

- **Rationale:** Energy consumption data exhibits strong, clear seasonality (daily and weekly cycles) and auto-correlation. The SARIMA model, which is SARIMAX without exogenous variables, explicitly captures these complex hourly and daily temporal dependencies, making it highly effective for precise, high-frequency, near-term prediction, which is critical for operational scheduling.
- **Constraint:** While initially planned, the incorporation of exogenous variables (like predicted short-term weather) was not pursued due to data unavailability, resulting in a pure time-series SARIMA model.

12-Month (Long-Term) Forecast

Technique: Year-on-Year (YoY) Baseline Projection

- **Rationale:** Given that the initial three-year dataset only yielded **45 monthly data points**, it was determined that this sample size was insufficient to reliably train complex machine learning models or accurately identify long-term seasonality patterns (e.g., strong yearly cycles) for a 12-month projection.
- **Adopted Approach:** The 12-month forecast utilized the previous year's actual consumption data as a simple, highly interpretable baseline. This approach offers stability and, crucially, facilitates **easy introduction of a margin or**

adjustment factor based on market sentiment (e.g., lower or higher energy prices), which is highly valuable for strategic financial planning.

2. Feature Selection & External Data

Data Preprocessing

The initial three-year dataset was preprocessed to ensure stationarity (where needed for SARIMA) and feature richness. Steps included:

1. **Index Conversion:** The timestamp column was converted to a proper DatetimeIndex to facilitate time-series operations.
2. **Missing Data Handling:** Missing values were handled via linear interpolation or forward/backward filling.
3. **Aggregation:** For the 12-month forecast, high-frequency data was aggregated to a monthly total/average to establish the YoY baseline.

Features Used

Impact of External Data Constraint: The inability to secure reliable future weather data meant the short-term model was simplified from SARIMAX to SARIMA, relying entirely on historical temporal patterns rather than dynamic weather inputs.

3. Model Training and Validation

Training and Tuning

All models were trained using a standard time-series split to maintain temporal ordering.

- **SARIMA (48-Hour):** Model orders (p, d, q) and seasonal orders (P, D, Q, s) were determined through a combination of **ACF/PACF plots** and **Grid Search** to minimize the AIC (Akaike Information Criterion) on the training set. Since no exogenous features were available, the focus was purely on optimal temporal parameter selection.
- **YoY Baseline (12-Month):** No traditional training was required. The historical data for the last 12 months in the training set served as the projection.

Validation Strategy

To ensure the models' robustness and ability to generalize to unseen future data, a **Time Series Cross-Validation** approach was used.

- The training data was split into training and validation folds, always maintaining chronological order.
- **Metric:** The primary evaluation metric was the **Mean Absolute Percentage Error (MAPE)**

MAPE is a critical **business-friendly metric** as it provides an intuitive measure of forecast accuracy expressed as a percentage, which is easily understood for budgeting and risk assessment. Nevertheless, it was not achieved to use the MAPE as the actual evaluation metric for model training. Therefore, the absolute error is used to train the data.

4. Business Understanding Alignment

The solution design directly aligns with Fortum’s dual operational and strategic objectives:

- **48-Hour (Operational):** The high-precision SARIMA model supports **real-time risk management and trading operations**. Accurate 48-hour forecasts allow Fortum to efficiently schedule power generation/procurement and minimize imbalance costs associated with deviations from predicted demand.
- **12-Month (Strategic):** The stable **Year-on-Year Baseline** approach supports **long-term financial and capacity planning**. Its simplicity is a feature, enabling direct manipulation for risk assessment—for example, by using the baseline consumption as a proxy for next year's consumption (or price) and then easily introducing a fixed margin based on forward-looking market sentiment.

5. Results Summary

The **SARIMA model** for the 48-hour horizon successfully ensures that the immediate, high-frequency dynamics—the hourly peaks and troughs driven by human activity—are modeled with high precision, leading to a near-term forecast that is operationally actionable despite the lack of exogenous weather data.

The **Year-on-Year Baseline** provides a transparent and robust strategic forecast. While more complex models may yield marginally better statistical metrics, the YoY baseline offers unparalleled **interpretability and manipulability** for strategic financial teams, which is often prioritized for budgeting and risk-hedging over a complex, black-box model.