Forecasting Electricity Grid Imbalance Direction: A Machine Learning Approach

Ali ElGuindy

October 2024

Contents

1 Introduction

Accurately forecasting the direction of electricity grid imbalances is essential for effective grid management and stable energy supply. Traditional time series models, such as ARIMA and GARCH, have been commonly used to predict future values by capturing temporal patterns and volatility within the data. However, with the increasing availability of diverse features—such as time-based indicators, lagged metrics, and price-driven signals machine learning models offer the potential to enhance predictive performance by leveraging both temporal and structured data.

In this study, we aim to predict the direction (positive or negative) of electricity grid imbalance at the next time step ($t + 1$). The grid Imbalance represents the net energy imbalance in the Swiss control area, calculated every 15 minutes. This measure helps Swissgrid ensure that electricity supply meets demand within the control area by indicating if there is a surplus (long) or deficit (short) of energy. We begin by constructing time series models (ARIMA + GARCH) to predict the future imbalance values, using the predicted and actual values' signs to compute directional accuracy. Additionally, we introduce various engineered features, including lagged variables, rolling statistics, and interactions, to create a comprehensive dataset that encapsulates short- and long-term patterns, as well as external influences.

To evaluate the efficacy of machine learning in this context, we implement logistic regression as a benchmark and compare its performance with more complex tree-based models, including Random Forest and XGBoost. These tree models are particularly suited to capture intricate feature interactions and nonlinear relationships that traditional time series models may overlook. Hyperparameter tuning is conducted via cross-validation to ensure optimized performance.

Initial results suggest that tree-based models outperform both logistic regression and time series approaches, indicating their potential in capturing the complex dynamics inherent in electricity grid imbalances. Further analysis is provided to assess the specific features contributing most to model accuracy.

2 Data

The data used in this study is provided by Swissgrid, the national transmission system operator (TSO) responsible for managing Switzerland's electricity grid.

The dataset contains 28,796 data points, spanning from January 1, 2024, to October 26, 2024, recorded at 15-minute intervals. The table below outlines the features within the dataset, with the target variable being the total System Imbalance, for which we aim to predict the direction.

Feature	Explanation
Date Time	Timestamp of each observation, recorded at 15-minute intervals. This serves as the index for time series analysis, allowing us to observe changes and patterns over time.
Abgedeckte Bedarf der aFRR+	Activated positive secondary control (Automatic Frequency Restoration Reserve, aFRR+). Represents the demand for upward regulation to increase power generation in the grid, helping to restore frequency when there is an imbalance requiring more energy.
Abgedeckte Bedarf der aFRR-	Activated negative secondary control (Automatic Frequency Restoration Reserve, aFRR-). Represents the demand for downward regulation to decrease power generation, helping to restore frequency when there is excess energy in the grid.
NRV+ (Import)	Net Regulation Volume for imports. Measures the net volume of energy imported into the control area, contributing to balancing needs when there is an energy deficit in the local grid.
NRV- (Export)	Net Regulation Volume for exports. Measures the net volume of energy exported from the control area, helping to relieve surplus energy and avoid overloading the local grid.
Abgedeckte Bedarf der SA mFRR+	Activated Scheduled Manual Frequency Restoration Reserve for positive balancing (mFRR+). Represents the planned manual activation to increase generation capacity as part of balancing requirements over a scheduled period.
Abgedeckte Bedarf der SA mFRR-	Activated Scheduled Manual Frequency Restoration Reserve for negative balancing (mFRR-). Represents the planned manual activation to reduce generation capacity to meet scheduled balancing requirements for anticipated surpluses.
Abgedeckte Bedarf der DA mFRR+	Direct Activation Manual Frequency Restoration Reserve for positive balancing (mFRR+). Indicates the real-time activation of manual reserves to increase generation to meet immediate balancing needs.
Abgedeckte Bedarf der DA mFRR-	Direct Activation Manual Frequency Restoration Reserve for negative balancing (mFRR-). Reflects real-time manual actions to reduce generation when there is an unexpected surplus in the grid.

Abgedeckte Bedarf der RR+	Replacement Reserve for positive balancing (RR+). Represents the amount of reserve capacity available for longer-term upward balancing to address sustained energy shortfalls.
Abgedeckte Bedarf der RR-	Replacement Reserve for negative balancing (RR-). Indicates the amount of reserve capacity available for longer-term downward balancing to manage sustained excess energy.
FRCE+ (Import)	Frequency Restoration Control Error in the positive direction (import). Shows the amount of energy needed to be imported to counteract a deviation in grid frequency and balance the system.
FRCE- (Export)	Frequency Restoration Control Error in the negative direction (export). Reflects the amount of energy that should be exported to counterbalance frequency deviations, maintaining system stability.
Total System Imbalance	Aggregate imbalance within the control area, where a positive (long) value indicates surplus energy and a negative (short) value indicates an energy deficit. This serves as a key measure for balancing requirements.
AE-Preis long	Adjustment Energy Price for surplus energy. The price at which excess energy can be sold back to the grid, providing financial incentives for reducing surplus energy.
AE-Preis short	Adjustment Energy Price for deficit energy. The price at which additional energy can be purchased to cover grid shortfalls, representing the cost of balancing grid deficits.

3 Data Pre-processing

Fortunately, data preprocessing was minimal due to the high quality of the data sourced directly from the official Swissgrid website. However, a notable outlier was identified in the target variable, total System Imbalance, which we aim to forecast. This outlier is readily visible when plotting the time series, as shown in Figure 1.

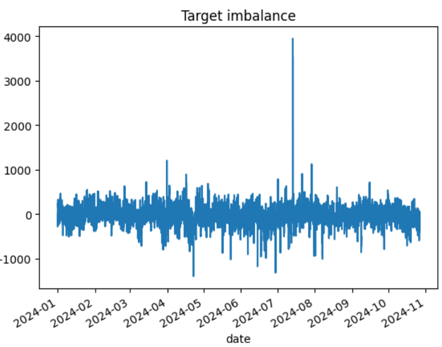

Figure 1: Target imbalance time series plot

To this end, values were capped above the threshold to the 99th percentile value. Hence giving the following plot 2:

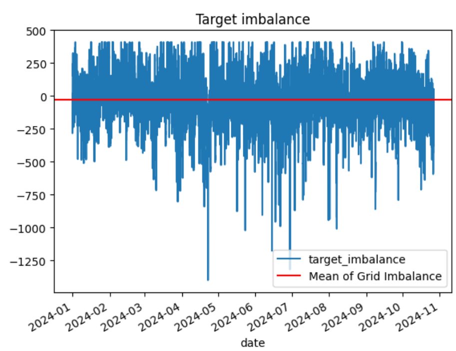

Figure 2: Target imbalance time series plot after outlier removal

The descriptive statistics of the cleaned time series are also presented in the following table 2:

4 Time series modelling

We start by checking the distribution of the target variable . 3. Clearly, the histogram and the QQ- plot show that the imbalance does not follow a normal

Statistic	Count	Mean	Std Dev	Min	25%	50%	75%	Max
Value	28796	-29.47	175.56	-1400.01	-117.08	-24.13	67.67	409.79

Table 2: Descriptive Statistics of Target Imbalance

distribution

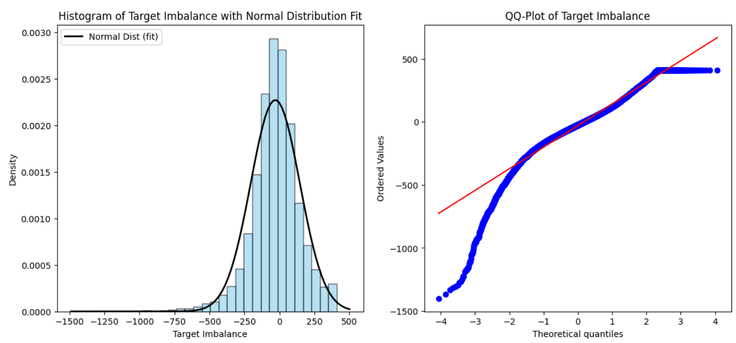

Figure 3: Distribution

The Dickey-Fuller (ADF) and KPSS tests were conducted on the time series to assess stationarity, and both tests indicated non-stationarity in the original series:

- **ADF Statistic**: -16.15, **p-value**: 4.55×10^{-29}
- **KPSS Statistic**: 0.889, **p-value**: 0.01

Both results suggest rejecting the null hypothesis of stationarity, indicating that the series is likely non-stationary.

To address this, we first applied first-order differencing. However, this led to conflicting results between the two tests:

- **ADF Statistic**: -36.52, **p-value**: 0.0
- **KPSS Statistic**: 0.0005

As a next step, we applied seasonal differencing by subtracting hourly and weekly lags. This transformation led to stationarity, with both tests in agreement:

- **ADF Statistic**: -18.62, **p-value**: 2.06×10^{-30}
- **KPSS Statistic**: 0.0051

Since the seasonal differencing transformed the series into a stationary one, we will use this version for modeling and prediction. The following figure 4 shows the autocorrelogram (ACF) and partial autocorrelogram (PACF) for the hourly and weekly lags, helping to identify the nature of the model, whether it's AR, MA, or possibly a combination of both.

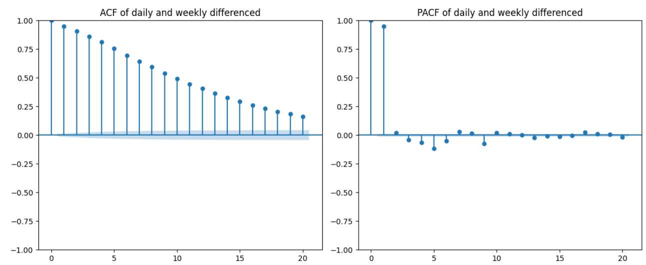

Figure 4: Autocorrlrgam and partial autocorrelogram

The autocorrelation function (ACF) shows a gradual decrease and the partial autocorrelation function (PACF) cuts off (or vanishes) after a the 1st lag, it suggests that the time series may follow an autoregressive (AR) model with p = 1.

Finally, we fit a model using GARCH(1,1) model leading to the following plots 5 and these metrics 3 :

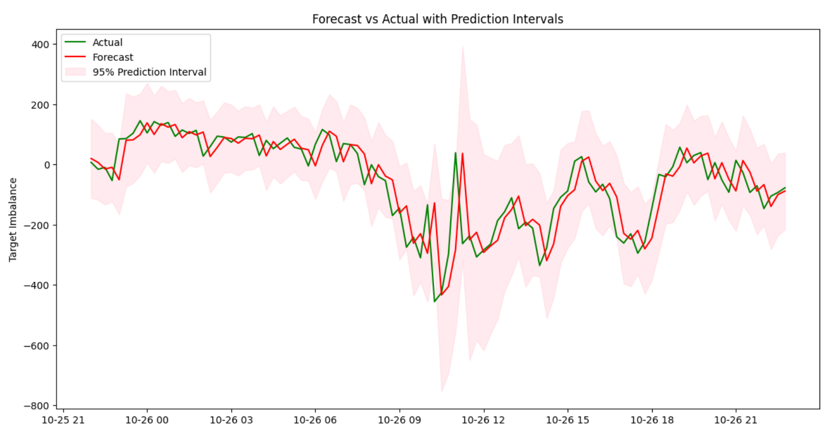

Figure 5: In sample and Out of sample prediction

Finally, using the forecasts values of the grid imbalance at t+1, we evaluate whether the predicted and actual values share the same sign (positive or negative) to compute the model's directional accuracy (i.e., the ability to predict a positive or negative imbalance at the next time step). This approach yielded a directional accuracy of 85% for the time series model

Metric	Value
In-Sample MAE	66.7761
In-Sample RMSE	93.4958
Out-of-Sample MAE	53.7993
Out-of-Sample RMSE	80.1395
In-Sample R-squared	0.9055
Out-of-Sample R-squared	0.6826

Table 3: In-Sample and Out-of-Sample Model Performance Metrics

5 Feature Engineering for ML models

To enhance the predictive power of the model, several feature engineering techniques were applied to capture relevant information about the grid imbalance dynamics and related factors. The following steps were implemented to ensure no forward-looking bias, relying only on historical data up to time t:

5.1 Hourly and Weekly Lags

Hourly and weekly lags were created for each feature, including `target_imbalance`, by shifting values by 1 hour (4 intervals) and 1 week (672 intervals). These lagged values capture the past behavior of each feature up to the current time step, ensuring no forward-looking bias.

5.2 Target Shift for Direction Prediction

The `direction` feature was engineered by shifting the `target_imbalance` variable by -1, allowing the model to predict the direction (positive or negative) of grid imbalance at time $t + 1$ based on available information at time t.

5.3 Rolling Statistics

Rolling statistics capture recent trends and volatility in the imbalance:

- **Hourly Mean**: A 1-hour rolling mean of `target_imbalance` was calculated to capture short-term trend.
- **Daily Standard Deviation**: A 24-hour rolling standard deviation was calculated to quantify daily volatility.
- **Hourly Deviation**: The deviation from the 1-hour rolling mean at each time step captures deviations from recent trends.

Each of these statistics is calculated using data up to time t, preventing any future data leakage.

5.4 Price-Driven Features

Two price-driven features were engineered to capture the effect of energy prices on imbalance:

- **Price Difference**: The difference between `AE-Preis short` and `AE-Preis long`.
- **Price Imbalance Interaction**: The product of `AE-Preis short` and `target_imbalance` to capture their combined impact on the target variable.

5.5 Time-Based Features

The following time-based features were added to capture daily and weekly patterns:

- **Hour of the Day** and **Day of the Week**
- **Weekend Indicator**: A binary feature indicating if the observation falls on a weekend.
- **Business Hour Indicator**: A binary feature indicating whether the time falls within typical business hours (9 AM to 6 PM).

5.6 Lagged Volatility and Extreme Value Indicator

A lagged daily volatility was added, along with an extreme value indicator:

- **Imbalance Volatility (Lag 1)**: Calculated as the lagged daily standard deviation of `target_imbalance`.
- **Extreme Imbalance Indicator**: Flags values above the 95th percentile of `target_imbalance` as extreme.

5.7 Cumulative Sum of Imbalances

The cumulative imbalance up to the current time step was calculated to track the total imbalance trend over time.

5.8 Feature Ratios

The ratio of `FRCE+ (Import)` to `FRCE- (Export)` was calculated to provide a measure of net import versus export at each time step.

5.9 Weekly Sum of Imbalance

A weekly sum of the `target_imbalance` was added by applying a rolling window of one week (672 intervals), capturing the aggregate imbalance over the past week.

All features were calculated using data available only up to time t, thus ensuring no forward-looking bias and maintaining the integrity of the predictive model.

6 ML Models for Predicting Imbalance Direction - Classification Task

In this classification task, we employed machine learning models to predict the direction of grid imbalance. Due to computational constraints, we limited the dataset to 5,000 data points, as running the code on the full dataset was time-intensive.

6.1 Logistic Regression

The logistic regression model achieved an accuracy of 0.7987.

6.2 Random Forest

We performed hyperparameter tuning on the Random Forest model using grid search, resulting in an accuracy of 0.85. The plot below illustrates the feature importance for this model:

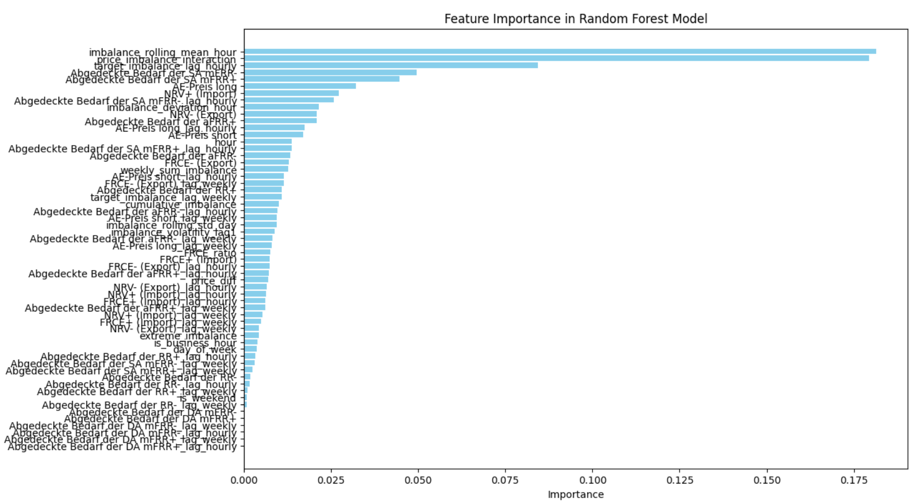

Figure 6: Random Forest - Feature Importance

6.3 XGBoost

The XGBoost model, with hyperparameters optimized via grid search, yielded the highest accuracy at 0.873.

Here is the plot for the feature importance :

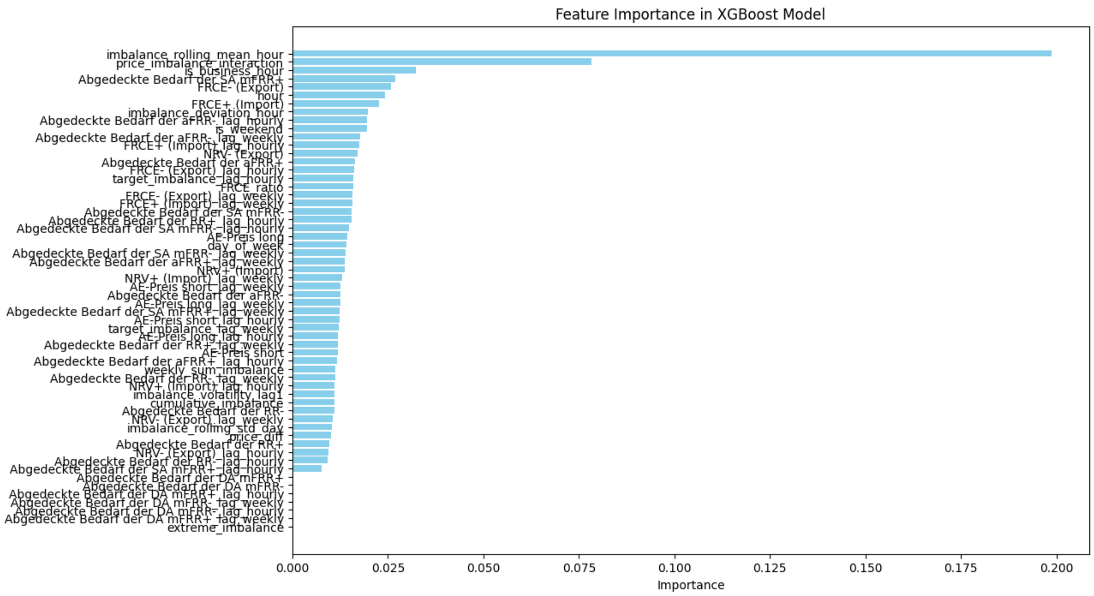

Figure 7: Xgboost- Feature importance

6.4 Xgboost vs Random Forest : Feature importance

The feature importance analysis for both the Random Forest and XGBoost models reveals valuable insights into which features drive predictive power. In both models, the **imbalance rolling mean hour** feature emerges as the most important. This feature likely captures immediate trends in grid imbalance, making it critical for short-term forecasting.

Another highly ranked feature is the **price imbalance interaction**, which measures the interaction between the imbalance and price. This interaction potentially provides valuable context, especially in capturing market responses related to imbalances.

Notably, engineered features such as **is_business_hour** and **is_weekend** also play a significant role, particularly in the XGBoost model. This importance suggests that grid imbalances may follow regular patterns based on the time of day or week, which these temporal indicators help capture effectively. Additionally, lagged features like the **hourly and weekly lags** are prominent in both models' top features, highlighting the relevance of historical data for predicting future states.

In terms of model comparison, Random Forest displays a more even distribution of feature importance across various lagged values, while XGBoost exhibits a sharper focus on top features, with a quick drop-off in importance for others. This characteristic suggests that XGBoost prioritizes capturing specific trends or interactions within the data, whereas Random Forest leverages a broader set of features to improve prediction accuracy.

7 Accuracy: Final results

Here we present the final results of the models tested for predicting grid imbalances at the next timestamp, $t+1$. Accuracy, as shown, indicates how well each model could predict the grid imbalance for the following period.

Table 4: Model Accuracy for Grid Imbalance Prediction at $t+1$

Model	**Accuracy**
Logistic Regression	0.7987
Random Forest	0.8500
XGBoost	0.8730
Time Series Model (TS)	0.8500

The results in table 4 indicate that XGBoost achieved the highest accuracy at 87.3%, suggesting it is the most effective model for predicting grid imbalances at $t+1$, while the Time Series and Random Forest models both performed robustly with accuracies of 85.0%, closely following XGBoost.

8 Conclusion

Predicting the direction of grid imbalance is a challenging task. Nonetheless, using time series models and tree-based models, we were able to achieve satisfactory accuracy, with XGBoost outperforming all other models, reaching approximately 87% accuracy. The intent is that these predictive signals could be valuable for power and gas traders (given the similarities between gas and electricity markets) to take positions or hedge based on the forecasted imbalance direction for $t+1$. Trading strategies can be developed around these signals to optimize profit and loss (P&L) outcomes.

More feature engineering could improve the models adequacy to learn and get better out of sample results.