Econometric Game 2025

XGBoost for Redispatch Forecasting in Germany: A Counterfactual-Driven Approach

Team 11

Amsterdam, The Netherlands - April 2025

Abstract

The integration of renewable energy into Germany's already established electricity infrastructure has led to unforeseen operational challenges, especially in managing grid congestion. This study develops a high-frequency forecasting framework for redispatch events within the TenneT DE transmission zone, we transform the task into a binary classification problem. We use a sparse dataset with weather metrics, electricity prices from neighboring countries, market signals to train classification models that distinguish between positive and negative congestion. The results show elevated prediction accuracy, particularly with T-ROSE -rebalanced models, that account for disproportionate class representation within our data, resulting in important enhancements in sensitivity and F1-scores. This study also provides actionable insights through counterfactual analysis, identifying key covariates that influence congestion occurrences with the aim of proposing relevant policy measures.

1 Introduction

The ongoing energy transition in Europe has subjected electricity grid operators to an unparalleled level of operational complexity. In Germany, the rapid integration of renewable energy sources—especially wind and solar—has fundamentally transformed the conventional supply and demand dynamics, resulting in considerable volatility within the system. Although this transition is essential for attaining carbon neutrality, it presents a novel array of technical and economic challenges. Chief among these is the management of congestion within the high-voltage electricity grid, particularly in the TenneT DE region, where redispatching power flows has become a commonplace yet expensive strategy to maintain stability.

Redispatching denotes the urgent modification of electricity generation or consumption to avert transmission line overloads. That is a mitigation approach to address issues of excess supply or demand, often through regional redirection of energy. These interventions, while effective, are expensive. National expenditures on congestion management have soared past €1.4 billion annually—costs that are ultimately borne by consumers. While according to Clara Büttner (2025) efficient redispatching can lead to an approximate cost reduction of 70% over the alternative. The increasing frequency and scale of redispatch events underscore a systemic inefficiency that urgently calls for a predictive and preemptive approach. Rather than responding reactively to bottlenecks, system operators must anticipate them. This pivot from reactive to predictive management hinges on our ability to accurately forecast congestion events using a range of exogenous signals—from weather conditions and electricity demand to fuel prices and regulatory events such as holidays.

The core objective of this study is to develop a high-frequency, day-ahead forecasting model for redispatch events in the German electricity grid. Specifically, we seek to predict the likelihood of congestion within the TenneT DE transmission zone at an hourly resolution. This task is not trivial. Redispatch events are inherently rare, making the target variable highly imbalanced. Moreover, electricity systems are driven by a tangled interplay of nonlinear relationships across weather patterns, market dynamics, and grid constraints. Effective prediction, therefore, demands more than just traditional statistical tools—it requires a blend of modern machine learning, robust validation schemes, and interpretability-aware design manifested through sensitivity analysis.

The study uses empirical modeling to analyze data structure and interdependencies in power system dynamics. A forecasting methodology is developed using binary classification models, such as extreme gradient boosting Wang & Guo (2020) trained on data from 2020 to mid-2022 and validated in the second half of 2022, and its performance evaluated on 2023 data. Importantly, in this study we use all the knowledge obtained from data visualization, machine learning forecasting and data driven insights to propose a framework for the estimation of feature marginal effects leading to invaluable information towards policy calibration. We aim to disentangle and understand the most important drivers of congestion in order to propose actions that can be taken from the Federal Government in order to address within region infrastructure deficiencies, account for cross country international trade, as well as provide clear economic intuition on the construction of a strategy focused on incentivizing citizens to demand energy when it is most convenient for the grid.

2 Literature Review

In this section we present the literature on congestion management and redispatch in power systems of Germany. In order to provide valuable insights into methodologies, findings, and policy implications.

Zhou et al. (2011) develop a structural forecasting model using convex hull algorithms to predict congestion in wholesale markets. Their method maps congestion status and marginal generator flags to load-space convex sets, enabling precise forecasts with a 1% lower error rate compared to GARCH models in NYISO data during extreme load conditions. Additionally, this approach reduces data requirements by 40% compared to traditional OPF-based methods, enhancing scalability. Billault-Chaumartin et al. (2020) investigate redispatch patterns, emphasizing the interplay between reserve power plants and energy market dynamics. Their analysis reveals significant regional disparities: northern Germany, characterized by abundant wind energy, frequently requires downward redispatch, while southern regions with higher demand depend on upward adjustments. This imbalance leads to cost inefficiencies, as reserve power plants often operate suboptimally due to grid bottlenecks. The study underscores the importance of aligning renewable energy deployment with grid expansion to reduce reliance on costly remedial measures. Probabilistic Forecasting for Congestion Mitigation Gürses-Tran et al. (2020) proposes a probabilistic load forecasting framework aimed at mitigating day-ahead congestion. Their approach integrates uncertainty quantification into load forecasts, reducing grid congestion risks by 15–30% in case studies. The model demonstrates superior performance compared to deterministic benchmarks, particularly during peak demand periods. This methodology facilitates proactive grid management, thereby minimizing the need for real-time redispatch. Titz et al. (2024) employ explainable AI techniques to identify key drivers and mitigators of redispatch. They identify wind power generation as the primary driver of congestion, accounting for approximately 60%, alongside cross-border electricity trading. Conversely, factors such as hydropower flexibility and optimized intraday trading reduce redispatch needs by up to 25%. The authors advocate for aligning renewable incentives with grid topology and implementing nodal pricing to lower annual redispatch costs by €1.2–2.1 billion.

3 Data

The provided redispatch dataset includes measurements of potential features for modeling grid congestion in Germany. The data is structured in a time series format on an hourly basis and the variables are categorized into three main groups:

1. **Weather Data** that includes hourly indices for solar radiation, air temperature, and wind speed, which represent the influence of weather conditions on electricity generation and consumption.

2. **Electricity Data** with measurements of electricity prices of neighbor countries, providing insights into energy demand and supply dynamics per hour.

3. **Market Data** that includes prices of oil and gas, carbon emissions, as well as indicators for working days and holidays per day.

This comprehensive dataset, comprising more than 70 time series features, enables the development of robust and highly accurate forecasting models for grid congestion.

3.1 Data Preprocessing

Given the hourly nature of the time series data, we incorporated time lags of 24 hourly intervals for features available on an hourly basis, such as weather data and electricity data. To manage computational complexity, additional lags were not considered. Additionally, all features were aggregated over the preceding week to capture broader temporal patterns. The significance of those two time frames are highlighted in the study by Billault-Chaumartin et al. (2020) through the application of a Fast Fourier Transform (FFT).

For the market data, which includes daily open, close, high, and low prices for oil and gas as well as carbon emissions, we considered only the open price at time t. This decision aligns with the real-world application of forecasting grid congestion on an hourly basis, only the open price of the day would be known and available for use.

For the weather data, which includes hourly measurements such as sunshine duration and wind speed, along with their minimum, maximum, and mean values for each city and hour, we opted to use only the mean values. This approach was selected due to the high temporal resolution of the dataset, ensuring a more streamlined and computationally efficient analysis.

Considering the nature of the response variable—specifically, the count of concurrent redispatches—it was transformed into a binary variable. This modelling decision serves our purpose of identifying under which circumstanced congestion arises, rather than how many times it occurs on an hourly level. However, this resulted in an imbalanced class representation within the response, which was addressed using the methodology outlined in Section 4.3.

3.2 Visualization

The study explores redispatch dynamics in real-world grid operations, focusing on the strong empirical interrelation between positive and negative events. Positive redispatches happen when there is need for more energy in order to meet high demand while negative redispatches occur in the directly opposite scenario. Both types of actions tend to co-occur during periods of elevated wind velocity, aligning with the compensatory nature of redispatching (Figure 1). The data-driven approach treats redispatch events as jointly determined phenomena, better supporting holistic policy analysis and forecasting Titz et al. (2024). At the beginning of 2021, the macroeconomic context is linked to carbon emissions futures, with the implementation of Phase 4 of the EU Emissions Trading System causing a significant rise in CO2 futures prices (Figure 2). This policy shift has unintentionally intensified congestion by amplifying generation in regions not yet supported by adequate transmission infrastructure. The study also juxtaposes redispatch activity with fossil fuel prices, demonstrating moderate co-movement with redispatch episodes (Figure 3). A significant institutional shift in German grid operations occurred with the introduction of the

Netzausbaubeschleunigungsgesetz—the so-called “grid extension acceleration law”—which came into force in 2019. The legislative reform aimed to expand redispatch responsibilities by integrating smaller generating units into congestion management. The Redispatch 2.0 mechanism, implemented in June 2022, has led to a significant increase in renewable energy installations Titz et al. (2024), a shift from traditional centralized power plants (Figure 4 bottom). This granular approach to congestion mitigation introduces new complexity into forecasting and policy design. Figure 5 illustrates all key predictors into a correlation matrix, revealing the strongest positive correlation with redispatch frequency, while renewable generation and weather indicators are negatively associated. The complex interplay among predictors calls for machine learning frameworks capable of handling high-dimensional, non-linear structures.

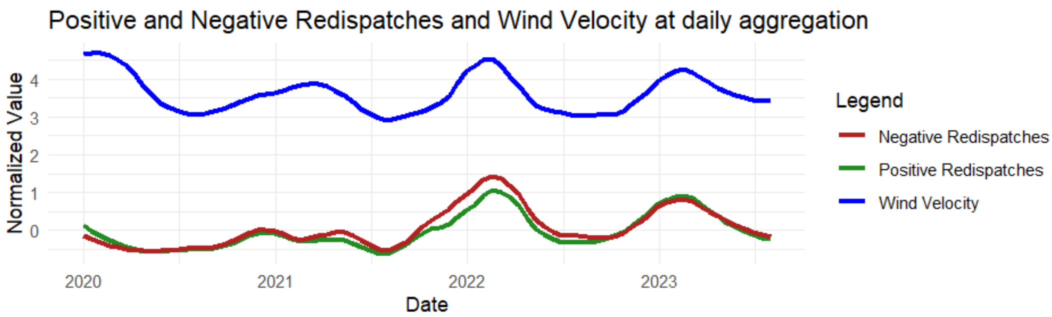

Figure 1: Positive and Negative Redispatches and Wind Velocity at Daily Aggregation

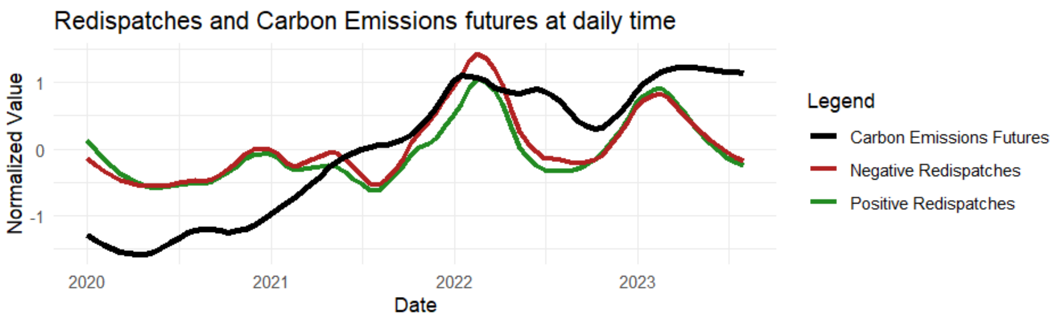

Figure 2: Redispatches and Carbon Emissions Futures at Daily Time

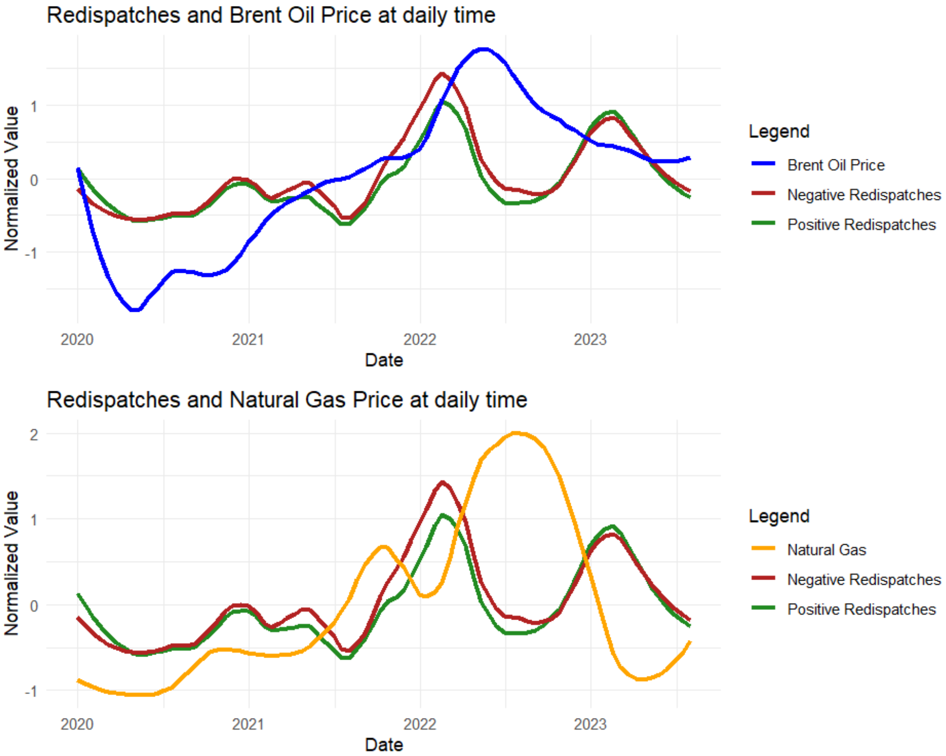

Figure 3: Redispatches and Brent Oil Price (top) and Natural Gas Price (bottom) at Daily Time

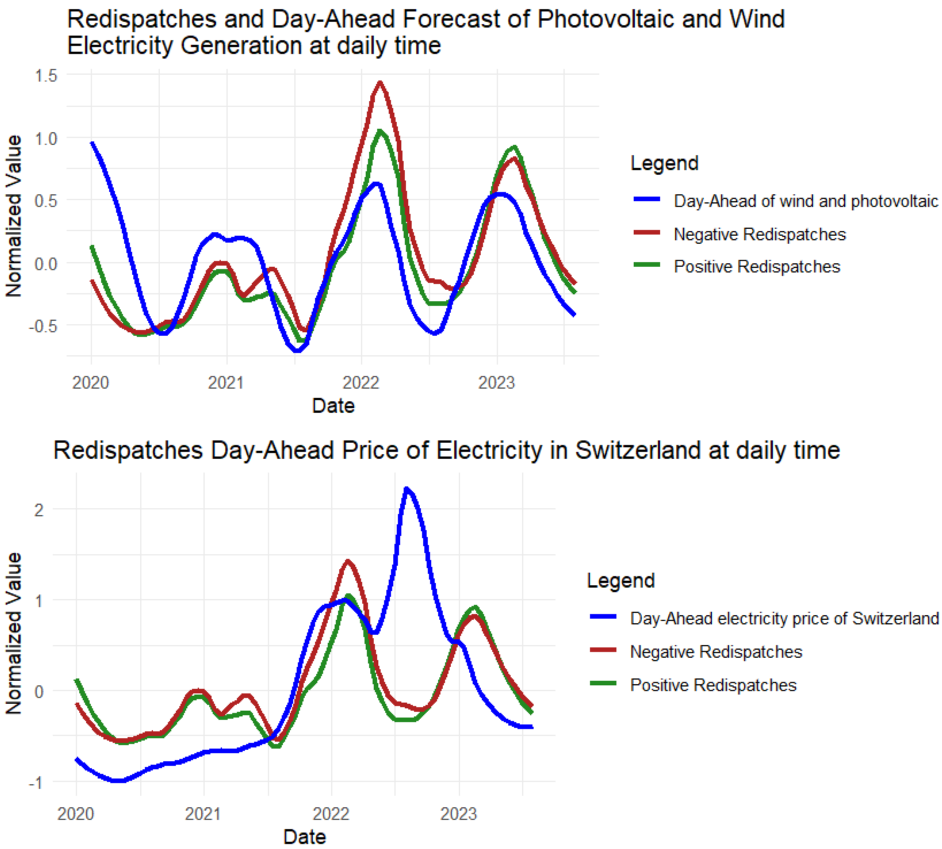

Figure 4: Redispatches and Day-Ahead of wind and photovoltaic electricity generation (top) and Day-Ahead Price of Electricity in Switzerland (bottom) at daily time

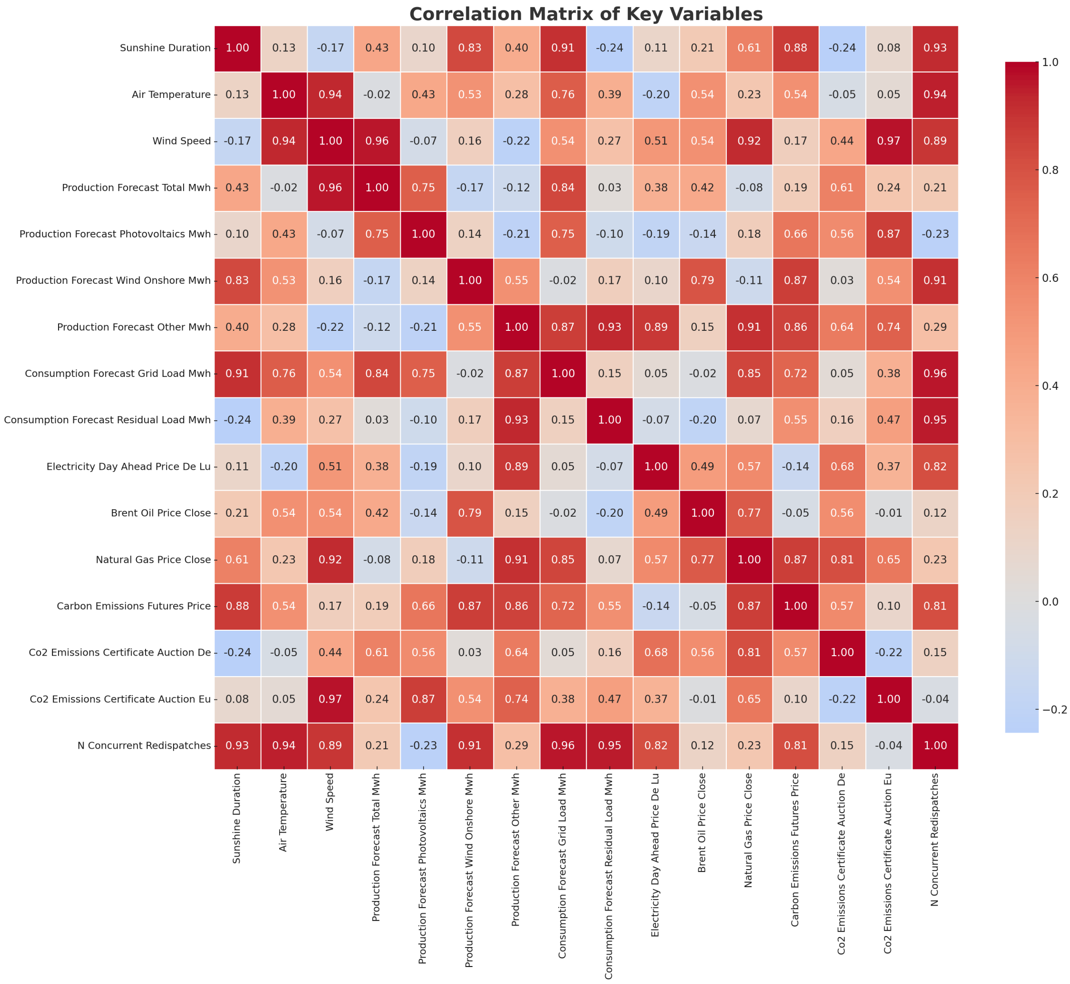

Figure 5: Correlation Matrix of Key Variables

4 Methodology

This section delineates the econometric and machine learning techniques employed to forecast day-ahead redispatch occurrences within the German power grid, specifically concentrating on the TenneT DE network. Given the binary nature of the target variable, `n_concurrent_redispatches`, which we define as 1 if the number of redispatch events exceeds 1 and 0 otherwise, we adopt a probabilistic classification framework. We also explicitly address challenges such as class imbalance, time series dependence, and high-dimensionality of features. Our methodology combines economic insight, domain-specific feature design, and state-of-the-art classification techniques to predict redispatch events accurately. The approach is scalable, interpretable, and policy-relevant, offering both short-term forecasting capability and long-term planning value.

4.1 Problem Formulation

Let $y_t \in \{0, 1\}$ denote the binary target variable indicating the occurrence of a redispatch event at hour t. Let $\mathbf{x}_t \in \mathbb{R}^p$ denote the p-dimensional feature vector at time t. Our goal is to model the conditional probability:

$$P(y_t = 1 \mid \mathbf{x}_t) = f(\mathbf{x}_t; \boldsymbol{\theta}), \tag{1}$$

where $f(\cdot)$ is a parameterized classifier and $\boldsymbol{\theta}$ denotes the parameters to be estimated.

4.2 Modeling Framework

XGBoost Classifier: XGBoost is a gradient boosting decision tree model designed to optimize binary logistic loss with integrated L1 (Lasso) and L2 (Ridge) regularization techniques. It is particularly adept at managing sparse and heterogeneous datasets.

The model produces a probability estimate, denoted as $\hat{p}_t = \hat{P}(y = 1)$. Predictions are classified based on a threshold $\tau = 0.5$, where instances with $\hat{p}_t \geq \tau$ are assigned to the positive class. To evaluate the model's performance, we use metrics such as accuracy,sensitivity,specificity, F1-Score, and log loss.

4.3 Handling Class Imbalance

A standard way to approach disproportionate representation in the sample, is to use Random Oversampling Examples (ROSE), proposed by Menardi & Torelli (2012), in order to create synthesized observations from the minority class. Although established, the technique does not account for the temporal structure of the data and thus isn't suited for time-series analysis. This is why we construct a variation of ROSE with a temporal resampling component that limits the window of possible near-neighbors. This Temporal Random Oversampling is accomplished via bootstrapping. While we run the risk of making too many similar minority class observations, thus potentially making the model memorize instead of generalize, we use a very diverse and rich set of features that create the necessary variability within the covariates to address the issue. The novel T-ROSE oversampling technique (that is actually in part inspired by T-SMOTE proposed by Zhao et al. (2022)is a time-aware, new strategy specifically designed to tackle class imbalance in time-series data—in our case, to counteract the imbalanced prevalence of 1s (minority initially at 33%) and push it towards a more balanced distribution (up to 77%). This approach generates artificial minority instances (those with binary value 0 in our response variable) from the mere usage of available past observations in the 24-hour lookback window and thus is obedient to the causal order and the inherent temporal structure of the data rigidly.

The approach identifies candidate neighbor observations to each minority observation (the "anchor") from the past 24 hours. Subsequently, a neighbor is drawn at random from this set, and synthetic predictor variable values are generated through domain-specific perturbations: binary variables are generated by drawing one of the two anchor or neighbor values randomly (maintaining them strictly 0 or 1), and strictly positive variables are perturbed through a truncated normal between 0 and the maximum value of the variable, shielding from any out-of-range negative outcomes. In order to model consistency, we do so to ensure that we do not recreate the structure of our features but rather render the synthetic observation barely different from the genuine ones. The synthetic observation adopts the timestamp of its anchor with a constant offset (e.g., 30 minutes) added to it to avoid duplicate time points while enforcing temporal coherency. Also, to allow for the possibility of more than one value at the same exact time, we add half an hour to the anchor's start date, in this way no two values are going to be in the same time interval and synthetic observations are easily detectable.

This new fusion of 24-hour temporal translation with personalized feature perturbation not only maintains the oversampled data within the realm of realism and in accordance with the operational dynamics of the grid but, more importantly, significantly improves the balance of the dataset in terms of minority representation to improving the downstream predictive model performance and reliability and enabling robust policy analysis.

4.4 Model Validation and Forecasting Strategy

We split the data as follows:

- Training: January 2020 to June 2022
- Validation: July 2022 to December 2022
- Test (Evaluation): January 2023 to December 2023

4.5 Modeling Redispatch Asymmetry

We also experiment with modeling **positive and negative redispatch separately**, motivated by the hypothesis that their drivers differ (e.g., supply surplus vs. deficit). Two separate classifiers are trained:

$$P(y_t^+ = 1 \mid \mathbf{x}_t) = f_+(\mathbf{x}_t), \tag{2}$$

$$P(y_t^- = 1 \mid \mathbf{x}_t) = f_-(\mathbf{x}_t), \tag{3}$$

where y_t^+ and y_t^- are indicators of positive and negative redispatch events.

4.6 Variable Selection and Explainability with XGBoost

To ensure predictive accuracy and transparency in modeling redispatch events, we use **XGBoost-based Recursive Feature Elimination (RFE)**. This methodology allows us to (1) filter out uninformative predictors, and (2) provide robust interpretability of the final model.

4.6.1 Recursive Feature Elimination with XGBoost

We implement a recursive feature elimination (RFE) strategy using the XGBoost algorithm. The key idea is to iteratively remove features with low importance until a subset of the most predictive variables remains. The process is based on feature importance measures extracted from XGBoost's tree ensemble.

Let $\mathcal{F} = \{x_1, x_2, \ldots, x_p\}$ be the set of all predictors.

XGBoost Feature Importance Metrics

After training the model, the importance of each feature $x_j \in \mathcal{F}$ can be computed using:

- **Weight:** Number of times x_j is used to split across all trees
- **Gain:** Average improvement in the objective function due to x_j
- **Cover:** Number of observations affected by those splits

We define the gain importance as:

$$\text{Gain}(x_j) = \frac{1}{T_j} \sum_{t=1}^{T_j} \Delta L_t(x_j)$$

where:

- T_j is the number of times feature x_j is used
- $\Delta L_t(x_j)$ is the gain in the loss function from splitting on x_j at node t

Threshold-Based Elimination

We retain only those features for which:

$$\text{Gain}(x_j) \geq \tau \cdot \max_{x_k \in \mathcal{F}} \text{Gain}(x_k)$$

where $\tau \in [0, 1]$ is a user-defined threshold (e.g., $\tau = 0.10$) that filters out all variables contributing less than 10% of the maximum gain. This process is repeated recursively until a desired feature count or performance is reached.

4.6.2 XGBoost with RFE and Sensitivity Analysis

Given the intricate nature of our response variable's directionality, we present all classification accuracy metrics—separately for positive and negative redispatch, as well as combined. For policy evaluation, however, we disregard the direction of the effect and focus solely on the total number of grid congestion occurrences. This choice is grounded both in the literature and our in-sample insights.

To ensure a parsimonious and interpretable model, we use the logistic specification of the XGBoost model, allowing us to obtain probabilities for grid congestion occurrences. Our primary approach for estimating and identifying policy-relevant insights is sensitivity analysis Christopher Frey & Patil (2002) based on the trained model. Specifically, we create and evaluate counterfactual experiments by artificially tweaking input features, enabling us to assess shifts in congestion probabilities.

Our goal is to identify key covariates through recursive feature elimination, using predictive performance to guide model selection. The result is a concise, well-defined, and not overly complex model.

This ensures that our model is both **parsimonious** and **transparent**, two essential qualities in high-stakes forecasting for energy systems and policy-making.

5 Empirical Results

5.1 Predictions

In order to assess model performance we follow the standard procedure of spliting the dataset into a training and test set. This split happens around the 30th of June 2022 in order to predict until 31st of December 2022. We evaluate performance through logloss, a quantification of the difference between actual class and predicted probabilities (the smaller the better), sensitivity, which measures how well the model detects positive instances, specificity, which reflects the ratio at which the model identifies true negatives, accuracy, which is calculated on the basis of total correct classification divided by the totality of cases and finally F1-Score which is the harmnonic mean of precision and recall see Table 1 for hour per hour predictions and Table 2 for 24 hours ahead predictions. It's important to note that when including the lagged hourly values of the response variable, the model seems prone to overfitting bias and that's why we choose not to work with this specification in the following section where we propose policies that can possibly mitigate grid cognestion across Germany. All metrics show very good fit, although there has been intensive hyperparemeter tuning, overfitting cannot be totally rejected. The resampled dataset achieves greater accuracy in predicting minority class true negatives and has overall greater predictive ability, indicating efficiency gains through T-ROSE resampling. In terms of predictions, we evaluate recursively, that is include the forecasted values in the forecasts. While we experience some loss in almost all categories, the result is miniscule suggesting great out-of-sample forecasting abilities.

	Raw Data			Resampling Data		
	Up	Down	Both	Up	Down	Both
Logloss	0.14	0.15	0.15	0.12	0.12	0.12
Sensitivity	0.92	0.92	0.91	0.96	0.96	0.96
Specificity	0.98	0.98	0.98	0.98	0.98	0.98
Accuracy	0.96	0.96	0.96	0.97	0.97	0.97
F1-Score	0.98	0.98	0.98	0.98	0.98	0.98

Table 1: Classification Evaluation metrics of 2023 for 'hour per hour' prediction

	Raw Data			Resampling Data		
	Up	Down	Both	Up	Down	Both
Logloss	0.15	0.15	0.15	0.12	0.13	0.12
Sensitivity	0.92	0.91	0.91	0.96	0.96	0.96
Specificity	0.97	0.97	0.97	0.97	0.98	0.98
Accuracy	0.96	0.96	0.96	0.97	0.97	0.97
F1-Score	0.97	0.97	0.97	0.98	0.98	0.98

Table 2: Classification Evaluation metrics of 2023 for '24 hours ahead' prediction

5.2 Discussion and Policy Recommendation

In short, the metrics show improved performance with the Temporal ROSE resampling technique as the oversampling of the minority class effectively balances the binary outcomes resulting in higher sensitivity, as well as specificity and accuracy see Table 3.

	Raw Data			Resampling Data		
	Up	Down	Both	Up	Down	Both
Sensitivity	0.09	0.09	0.11	0.49	0.53	0.55
Specificity	0.94	0.94	0.95	0.99	0.96	0.95
Accuracy	0.74	0.74	0.76	0.80	0.80	0.80
F1-Score	0.84	0.84	0.85	0.86	0.85	0.86

Table 3: Classification Evaluation metrics of 2023 for 'hour per hour' prediction without the time lags of the response

Feature	Orig. Prob	Mod. Prob	Change	% Increase	Counterfactual
Wind Forecast (Offshore)	0.5118	0.5182	0.00639	0.12	+20%
Workday Indicator	0.5118	0.5146	0.00275	0.09	+20%
Grid Load Forecast	0.5118	0.5126	0.00078	0.12	+20%
Carbon Futures (Open)	0.5118	0.5114	-0.00038	0.07	+20%
Day-Ahead Price Change (CZ)	0.5118	0.4993	-0.01251	0.27	+20%

Table 4: Effect of Feature Modification on Predicted Probabilities via Sensitivity Analysis

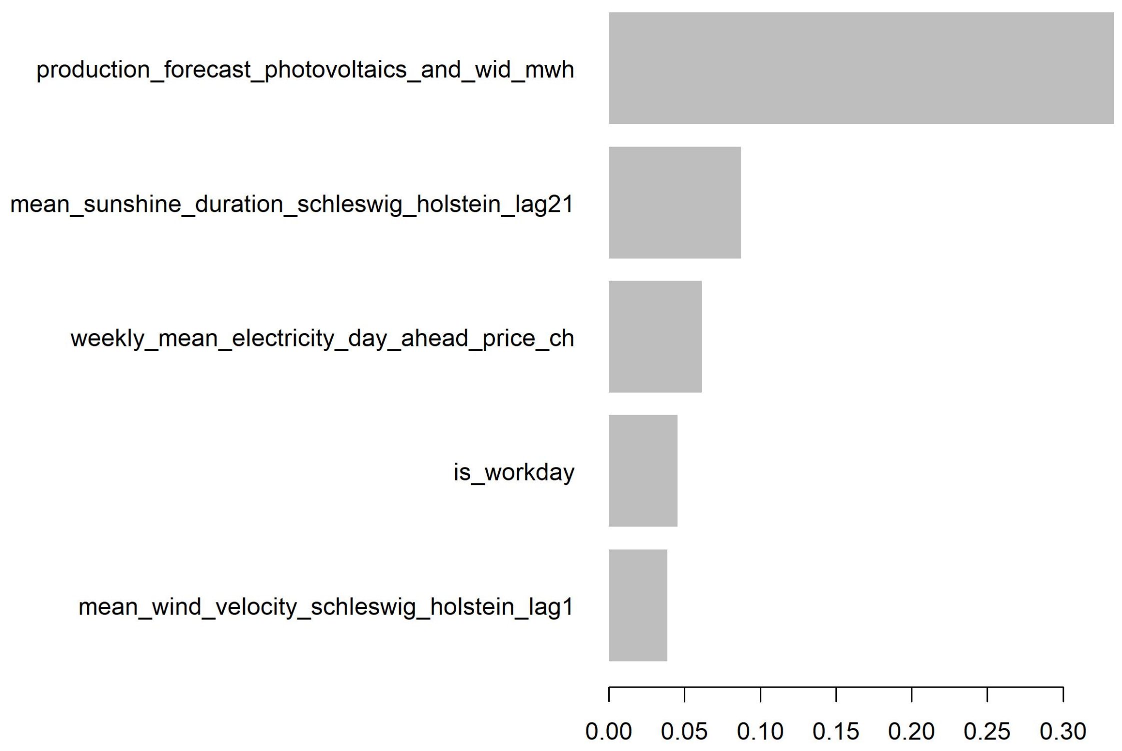

Figure 6: Important features of resampled data

Following Titz et al. (2024), we use RFE on the resampled data to narrow down our features to the absolute most important ones (Figure 6). However, we use a less arbitrary approach, as we do not choose a threshold but instead calculate F1-scores for all iterations and identify through Figure 7 the maximum value of the metric. We keep this model and proceed to identify, through sensitivity analysis the effects that each one of the most important 15 variables has on the probability of grid congestion occurring. Sensitivity analysis is detrimental to this cause because it allows us to use the specification of the model that we obtained through intensive CV and hyperparameter tuning and assess how a change in the features would directly affect the logistic probability of congestion.

We focus on five specific features (see Table 4) that are not only, important but can also be used when considering policy implications. More specifically, we see that congestion is more likely to happen on workdays. This is well expected as the vast majority of the population is employed , so the backbone of the economy operates in full flow. Providing incentives such as cyclical prices, evening or night discounts could prove detrimental in not overloading the grid during operating hours. Moreover, we understand that wind speeds (and thus wind-generated electric energy) exhibits a counter-intuitive effect. Although the very point of renewable energy sources is to harvest as much as possible when available, the higher the offshore value of the wind in MWh, the higher the probability of congestion. This result is significant in understanding the advancements required in infrastructure, efficiency, intra-region energy transfers, and cross country transfers since the current grid network tends to underperform when dealing with excess supply of energy from wind turbines. Moreover, investing in batteries to store energy could prove very helpful during winter spells with limited sun exposure and low wind levels. Having access to such tools would allow Germany to be flexible and redistribute energy when it is most needed, that is during high demand and low supply. Finally, consistent with the international trade flow approach of Titz et al. (2024) we identify the day-ahead price of Switzerland as a feature of importance. To be precise, we want to generalize this finding and recommend that Germany takes advantage of arbitrage opportunities among its closest neighbors. The solution to grid congestion can even come from outside the country, as international trade agreements or partnerships could prove beneficial. Transferring energy from the southern part to the northern part of Germany can be extremely inefficient when neighboring countries such as the Netherlands, Czech Republic, Luxembourg and Poland can offer a more timely and cheaper solution to electricity redistribution. To further solidify our findings and provide better interpretability , we use Partial Dependence Plots (PDPs) to visualize the upward trend between a selected feature on the estimated logistic probability obtained from XGBoost. Figure 8 shows the marginal effect of offshore wind production on the predicted probability while holding all other features constant, or rather to be more precise, while averaging across all other observations. The marginal effect is highest around the 200–400 MWh range — that is when small increases in offshore wind production lead to large increases in probability of congestion occurring. From this point onwards, we see a threshold effect where further increases in wind speed do not affect the response, this is likely due to the grid being already in a state of excess supply. Optimizing the grid to handle energy generated by wind of this magnitude would be a step towards greater network stability and energy redistribution.

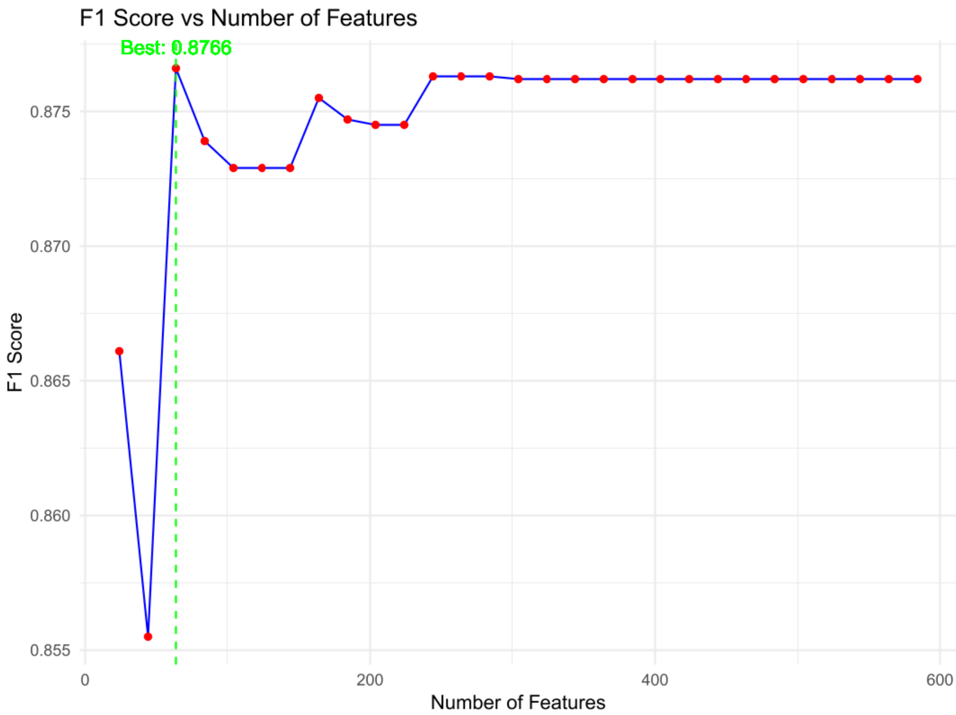

Figure 7: F1 Scores vs Number of Features

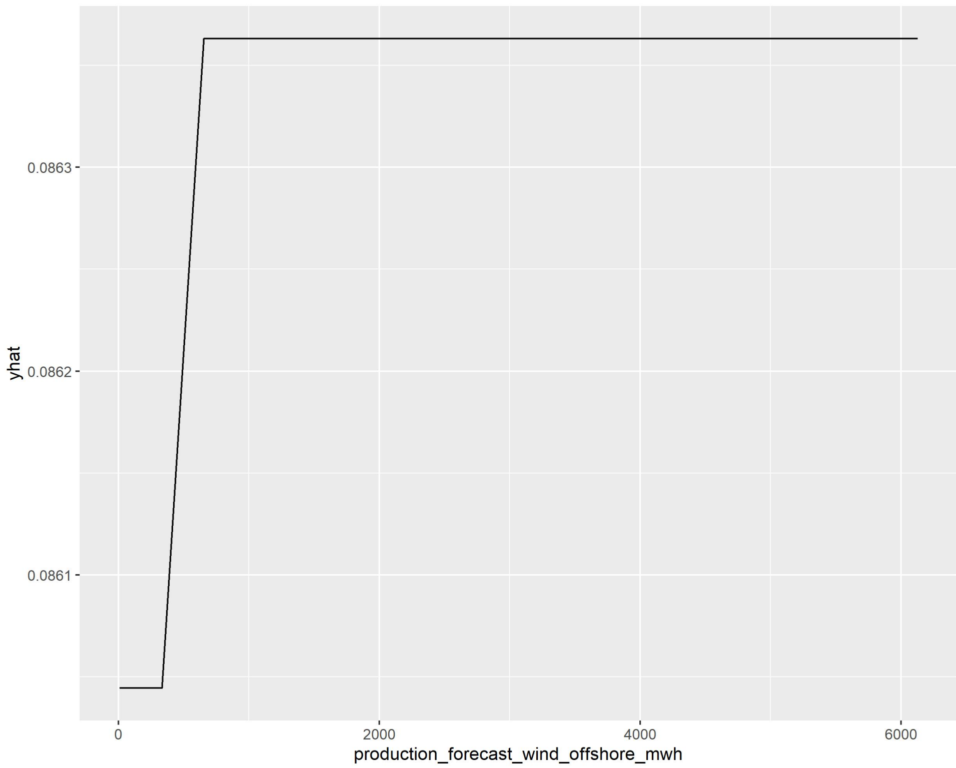

Figure 8: Partial Dependence Plot (PDP) of Wind Forecast (Offshore)

6 Conclusion

The current energy transition in Germany has revealed fundamental weaknesses in the national grid, especially regarding congestion management. Our study offers a data-driven response to this challenge, proposing a robust and interpretable forecasting framework that predicts day-ahead redispatch events with remarkable accuracy. By combining advanced machine learning techniques with domain-specific knowledge and a rich set of exogenous features, we demonstrate that congestion is not an unforeseeable anomaly, but a manageable phenomenon—provided the right tools are in place. One interesting modeling approach that we take is the adjustment of the standard ROSE algorithm to correctly respect and handle the temporal persistence of our dataframe. This strategy not only improved model sensitivity and generalization, but it also increased the realism of synthetic data, guaranteeing that predictions are operationally feasible. We isolate the key drivers of redispatch using XGBoost-based recursive feature removal which include offshore wind predictions, grid load expectations, and market calendar effects. Most importantly, we uncover the underlying mechanisms that drive congestion occurrences, thus rendering them of probabilistic nature rather than a statistical anomaly. Our counterfactual sensitivity analysis demonstrates that even minor changes in high-impact characteristics can drastically alter congestion probability, providing the way for preventive policy interventions. This closes the gap between predictive analytics and real-world utility. More than just predicting redispatch, this study aims to reimagine grid management as a forward-thinking organization that blends forecasting precision with strategic flexibility. As renewable penetration increases and volatility rises, such models can aid in both operational decision-making and long-term infrastructure planning, ensuring that Germany's energy transition is not just sustainable, but also stable, efficient, and equitable.

Although this study offers a comprehensive approach to predicting grid congestion events for Germany, there is still much to be explored regarding other research opportunities. To begin with, extending the forecasting period beyond the day-ahead scope could be beneficial. Exploring week-ahead and month-ahead models could enhance the value of long-term planning and investment decisions, especially in planning infrastructure expansions and cross-border energy alliances. Given enough time, even more complex and dynamic forecasting structures could be explored like Recurrent Neural Networks.

Secondly, not accounting for feedback loops does make the model dynamic but only subjectively through interrelational dependencies of the predictors and congestion risk.

Changes in the power market's structure, policies adopted, or technological advancements require more consideration which can be effectively addressed by time-varying parameter models or regime switch approaches. Furthermore, data on national and international tradeflows could help quantify the nature of the response and the true relationship between positive and negative congestions.

Another obvious improvement is the addition of geospatial data. As of now, the approach works on a national level, but is inherently spatial in nature, considering congestion is a regionally bound phenomenon. Focusing on regional transmission limits, weather conditions specific to a region, and nodal pricing of electricity might greatly help in achieving better accuracy improving the targeted measures to be undertaken. Methodologically, future research might compare the performance of deep learning models such temporal convolutional networks or transformers against tree-based methods, notably their capacity to simulate long-term dependencies. However, model complexity must be balanced with interpretability, particularly in policy-relevant applications. Finally, it would be useful to translate model results into cost-effective policy recommendations. Estimating the economic worth of congestion mitigation under various scenarios—such as additional storage capacity, updated pricing systems, or international energy exchange—would assist policymakers in prioritizing measures based on both technical feasibility and financial impact. Given a broader working horizon, we could work around more complex or more finely tuned counterfactual exercises.

References

Billault-Chaumartin, F., Eising, M. & Motte-Cortés, A. (2020), Redispatch patterns of energy market and reserve power plants in germany, *in* '2020 17th International Conference on the European Energy Market (EEM)', IEEE, pp. 1–6.

Christopher Frey, H. & Patil, S. R. (2002), 'Identification and review of sensitivity analysis methods', *Risk Analysis* **22**(3), 553–578.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/0272-4332.00039*

Clara Büttner, U. M. (2025), 'The impact of redispatch on grid and storage expansion planning in the german energy system', *Energy Informatics Review* .

Gürses-Tran, G., Flamme, H. & Monti, A. (2020), Probabilistic load forecasting for day-ahead congestion mitigation, *in* '2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)', IEEE, pp. 1–6.

Menardi, G. & Torelli, N. (2012), 'Training and assessing classification rules with unbalanced data', *Data Mining and Knowledge Discovery* .

Titz, M., Pütz, S. & Witthaut, D. (2024), 'Identifying drivers and mitigators for congestion and redispatch in the german electric power system with explainable ai', *Applied Energy* **356**, 122351.

Wang, Y. & Guo, Y. (2020), 'Forecasting method of stock market volatility in time series data based on mixed model of arima and xgboost', *China Communications* **17**(3), 205–221.

Zhao, P., Luo, C., Qiao, B., Wang, L., Rajmohan, S., Lin, Q. & Zhang, D. (2022), T-smote: Temporal-oriented synthetic minority oversampling technique for imbalanced time series classification, pp. 2381–2387.

Zhou, Q., Tesfatsion, L. & Liu, C.-C. (2011), 'Short-term congestion forecasting in wholesale power markets', *IEEE Transactions on Power Systems* **26**(4), 2185–2196.