University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 30/2025 (493)

Is smooth Energiewende possible? Improving the performance of climate policies in Germany by optimizing the risk of electricity delivery

Jakub Bandurski
Eliza Hałatek
Adam Łaziński
Michał Künstler

Warsaw 2025

ISSN 2957-0506

WORKING PAPERS 30/2025 (493)

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

Is smooth Energiewende possible? Improving the performance of climate policies in Germany by optimizing the risk of electricity delivery

Jakub Bandurski[1], Eliza Hałatek[1], Adam Łaziński[1], Micha Künstler[1]*

[1] *Faculty of Economic Sciences, University of Warsaw*
[*] *Corresponding author: m.kunstler2@uw.edu.pl*

Abstract: The Energiewende is a deep-rooted notion in the German economy. The main goal is to achieve climate neutrality by transitioning to renewable energy sources. However, the feasibility of this transition is partially hindered by power grid congestion, which undermines system efficiency and leads to both economic and environmental costs. We address this issue by making a prediction of the likelihood of congestion occurrence within the German TenneT DE electricity network in the years 2020-2023. We propose a twofold approach offering a combination of advanced econometric models and state-of-the-art machine learning methods. We offer separate solutions for up congestion when additional energy needs to be pushed to the network as well as down congestion when energy needs to be pulled away from the network. Analyzing the CatBoost with XAI, we identify factors that play a significant role in driving redispatch events within the German electricity network.

Keywords: energiewende, econometrics, machine learning, climate policy, catboost

JEL codes: Q47, Q48, Q54, C01, C53

Acknowledgements: This paper reflects the views and thoughts of the authors which are solely their own. The research was supported by a database from the World Championship in Econometrics, the Econometric Game, organised by the University of Amsterdam. As such, the authors are also members of the winning team in the 2025 edition. The authors would like to thank the casemakers: Anastasija Tetereva, Onno Kleen, Alla Petukhina for their input and the database distribution.

1. Introduction

Germany's transition towards sustainable development and so-called "green" practices in general has met with great approval in Europe, thus ushering in the concept of the *Energiewende* (Quitzow et al. 2016; Hirschhausen 2014; Byun et al. 2015; Renn and Marshall 2016). Although the transition towards renewables meets with public approval and encourages economic growth (Sievers et al. 2019), it also poses a high risk of congestion events in electricity grid, due to:

1) an irregular production pattern,
2) a long development time for new power lines,
3) an unfavourable weather conditions, which, combined with high production costs and carbon emissions in the event of a blockage, poses an optimisation conundrum for researchers and policymakers (Agency 2016; Sharifzadeh et al. 2017).

"*We want to end the use of nuclear energy and reach the age of renewable energy as fast as possible*" (Merkel 2011)

Although the German government declares to achieve climate neutrality by 2050, very little has been done in terms of predicting failures and overloading of the electricity grid, which then emit carbon dioxide. Specifically, (Davi-Arderius and Schittekatte 2023) points out, there is a more than two-factor relationship of CO2 emissions as a result of redispatched energy to total annual electricity demand met by renewable power grids. Following (Titz et al. 2024), we focus on the hourly data for the electricity grid operated by TenneT DE transmission system for Germany in the period 2020-2023. By doing so, we offer, to the best of our knowledge, the most up-to-date analysis taking into account the post-pandemic period of COVID-19, which has not yet been fully explored (Staudt et al. 2018; Wohland et al. 2018; Rausch et al. 2019; BDEW 2019; Titz et al. 2024). In contrast to (Titz et al. 2024), we focus on the probability of the hourly occurrence of redispatch. Furthermore, we offer distinct analyses for transformer inflow and outflow events, which is a novel approach providing interesting insights for power grids operators. In doing so, we answer the question concerning whether the Energiewende of the German economy, is fully possible. We do this by verifying research hypotheses on the importance of atmospheric factors (wind, solar energy, temperature), natural gas prices, Brent oil, and CO2 emission credits. Our approach expands the existing literature by applying advanced econometric methods tailored to the discreet nature of the dependent variable with state-of-the-art machine learning techniques (Random Forest, XGBoost, CatBoost, which, to the best of our knowledge, have not yet been included). As a result, we obtain results that allow for the creation of a comprehensive climate policy by the government. We argue that our study has wide-ranging applications that can be applied to European economies (Quitzow et al. 2016) that, like Germany, are undergoing the energy transition. The remainder of the article is organized as follows. **In the next section**, a comprehensive review of the literature on electricity congestion is carried out. The description and characterization of the data are presented in **Section 3** together with the autoregressive properties of the target variable. Congestion management measure analyses and prediction of probability of redispatch (*up&down*) along with additional methodology for Logistic Regression are presented in **Section 4**. The results are deferred to **Section 5** and the implications for proposed policies and the conclusion are **Section 6** and **7** respectively.

2. Literature review and hypotheses development

The Energiewende (also known as energy-transition) is a concept introduced by the German government in 2011 as a result of the Fukushima reactor disaster (Bundesregierung 2011). Along the same lines, the renewable and low-carbon vision of the economy has become Germany's focus on climate neutrality (Quitzow et al. 2016; Hirschhausen 2014; Byun et al. 2015; Renn and Marshall 2016), by reducing the use of fossil fuels from 80% to 20% by 2050 (Renn and Marshall 2016). The macroeconomic outlook of the energy transition in Germany points to a favourable relationship between energy efficiency and renewable energy and the development of the economy, with an emphasis on added value in electricity generation (Sievers et al. 2019). The above has driven Germany to become a leader in applied practices (Byun et al. 2015; Curry 2019; Titz et al. 2024).

Although renewable energy sources allow for a smooth transition to a sustainable economy, their production follows a highly irregular pattern and is bound to high levels of uncertainty (Sharifzadeh et al. 2017; Titz et al. 2024), creating a kind of mystery and an optimization dilemma for policymakers. After all, electricity, is one of the major components of energy consumption (Agency 2016; Sharifzadeh et al. 2017), which is why efficient allocation of renewable energy sources is so essential, given their cost (Titz et al. 2024). What's more, in July 2016, the Electricity Market Act (EnWG) came into force in Germany, entailing an orchestrated direction on energy, with a strong emphasis on emergency response and secure supply by transmission system operators (TSOs) and the Federal Network Agency (Comission 2016; Billault-Chaumartin et al. 2020)

"*The electricity grids are essential for our electricity supply. Germany is, however, lagging behind as regards the expansion of the grids. I therefore propose measures that will enable us to finally get going, perceptibly accelerate the grid expansion and upgrade existing grids*." (Altmaier 2018)

Accordingly, the question of network congestion management is drawing more and more scientific attention (Metzger et al. 2021; Titz et al. 2024). Since transmission lines can take up to a decade to be developed (Michiorri et al. 2015; Neukirch 2016; Fernandez et al. 2016; Glaum and Hofmann 2023), and following the Energiewende, the increasing trend towards renewable energy resources, is linked to an increase in power grid congestion (Dorfer et al. 2022). Hence, redispatching leads to an increase in the costs of grid operations, which contradicts the goal of decarbonisation and the whole concept of the Energiewende in Germany (Byun et al. 2015; Quitzow et al. 2016; Renn and Marshall 2016; Dorfer et al. 2022). Given the existing challenges, electricity power systems should be geared towards time and space balancing, as well as the provision of flexible back-up power plants (Elsner et al. 2015; Rodriguez et al. 2014; Schmidt and Staffell 2023).

Furthermore, as the intensity of weather factors, i.e. sun and wind, directly related to renewable energy sources is still a mystery, this creates challenges for power systems as it creates renewable energy droughts (Grochowicz et al. 2023).

"*Increasing numbers of renewable energy units, ranging from offshore wind farms to domestic solar units, are coming on-stream and contributing to a low-emission energy supply. This trend remains a challenge for electricity grid operators. Firstly, electricity generation from wind and solar units varies hugely depending on weather conditions.*" (Energy Systems of the Future (ESYS) 2020)

In addition, challenging the German economy are geographic properties that have the most favourable placement of wind turbines in the North and East (Creutzig et al. 2014; Titz et al. 2024). Jung and Schindler (2020) argue that wind speed significantly improves predictions due to better forecasts of generation from turbines and, consequently, less congestion of lines, which is also in line with recent studies (Emeis 2018; Peper et al. 2024; Titz et al. 2024). Additionally, Grochowicz et al. (2023) also identifies negative weather conditions (low temperatures, low wind) as those associated with difficult periods of investment in energy systems. However, (und Bundeskartellamt. Monitoringberichte 2023; Titz et al. 2024) state that high levels of wind power generation often overload the transmission grid. Consistently, we believe that the consideration of weather factors contributes significantly to the projection of electricity congestion risk.

Hypothesis 1: The low temperatures cause electricity congestion.

Hypothesis 2: The wind speed measurements contribute significantly to the prediction of electricity redispatch.

Decarbonization is an element of the climate transition (Quitzow et al. 2016), embedded by Germany and being practiced across Europe. While the Energiewende creates opportunities for economic growth of the economy (Sievers et al. 2019), Davi-Arderius and Schittekatte (2023) argues that redispatched energy, accounting for up to 4% of total annual electricity demand, contributes up to 11% of the CO2 emissions. Marimoutou and Soury (2015), on the other hand, argue that factors such as prices of Brent oil, natural gas and energy demand are worth considering when examining carbon dioxide emissions. Zhang et al. (2024), on the other hand, shows the existence of electricity-gas price couplings, thus setting the stage for climate policies.

Hypothesis 3: Higher carbon emission credit prices signal a higher risk of redispatches of the electricity.

Hypothesis 4: High Brent oil and gas prices increase congestion risk.

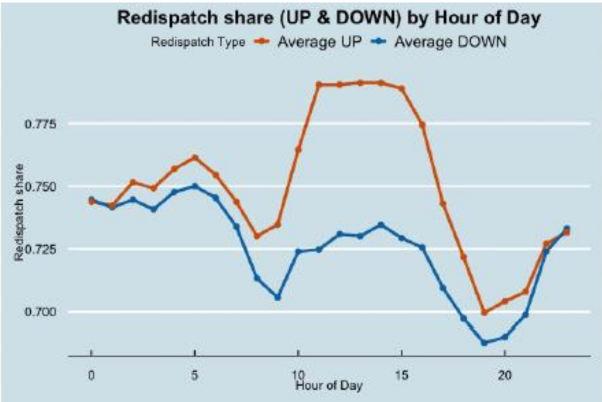
3. Data description

The German electricity grid operated by the transmission system operator TenneT DE spans several regions, including Schleswig-Holstein in the north and Bavaria in the south. Redispatch events represent a significant additional cost to the operator, both financially and environmentally, due to the high carbon emissions often associated with them. Therefore, it is crucial to understand the underlying causes of these redispatches and to accurately predict their occurrence in advance. Some insights into redispatch patterns can be derived through visual exploration. Notably, we focus solely on the occurrence of redispatch events, rather than the number of simultaneous dispatches.

The dataset spans the period from January 2020 to July 2023, a time marked by significant economic and social changes due to the COVID-19 pandemic.

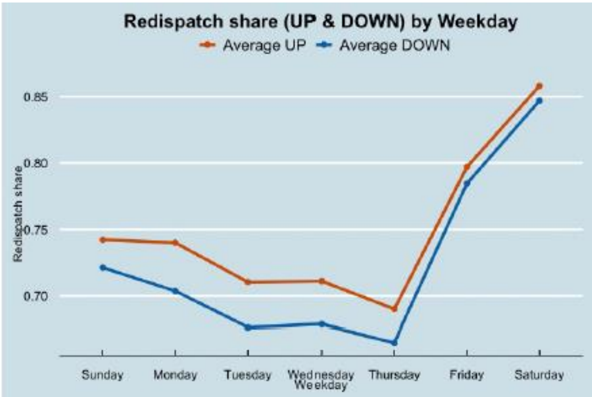
There are two types of redispatch measures—*up* and *down*—implemented by the TSO to manage congestion in the electricity grid. When local generation exceeds what the transmission infrastructure can handle, a downward redispatch instructs certain power providers to reduce output. Conversely, to maintain supply elsewhere in the grid, an upward redispatch is issued, requesting increased generation in less congested areas.

Fig.1a. Seasonal trends in the redispatch share. Share of hourly periods affected by at least one redispatch by hour of day



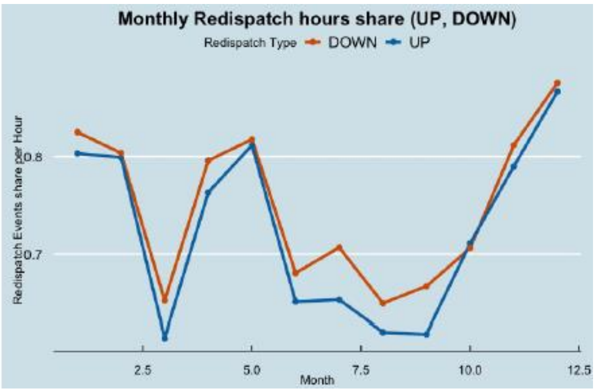
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.1b. Seasonal trends in the redispatch share. Share of hourly periods affected by at least one redispatch by weekday



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.1c. Seasonal trends in the redispatch share. Share of hourly periods affected by at least one redispatch by month



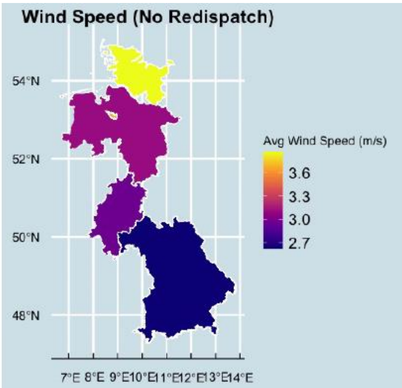
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Figure 1b reveals a clear weekly seasonal pattern in redispatch events. Specifically, the proportion of hours with at least one redispatch is significantly higher during weekends compared to weekdays, indicating a strong weekday effect in the data. This assumption is further supported by data in Figure 1a which highlights hour-of-day seasonality—with redispatch probability decreasing notably during morning and early evening hours, likely due to typical demand patterns. Lastly, Figure 1c illustrates substantial seasonal variation across different times of the year, reinforcing the need for any predictive model to include at least one full year of training data to effectively capture seasonal effects.

Figures 2a and 2b demonstrate that while the spatial distribution of mean wind speeds is similar regardless of redispatch happening or not, the periods with redispatches are characterized by overall higher mean wind speeds, which is in line with (Titz et al. 2024).

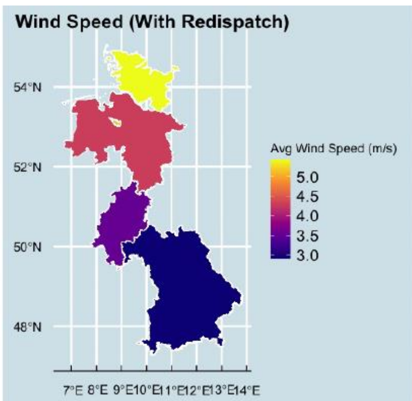
Even though PV produces significantly less energy than wind in Germany, (see Ember et al. 2024), Figures 3a and 3b show that the spatial distribution of sunshine duration seems to influence the redispatch frequencies more than the distribution of wind velocity: redispatches seem to happen more often when there is more sunshine in the south of Germany.

Fig. 2a. Wind speeds in TenneT DE operated regions vs redispatch. Average wind speed in each region in periods not affected by a redispatch



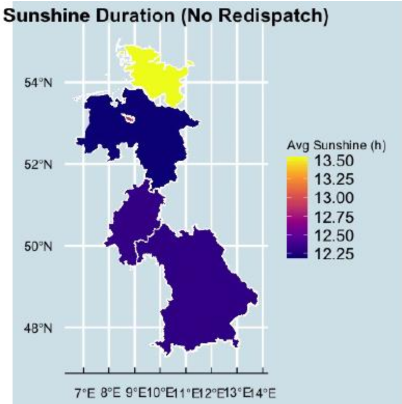
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 2b. Wind speeds in TenneT DE operated regions vs redispatch. Average wind speed in each region in periods affected by a redispatch



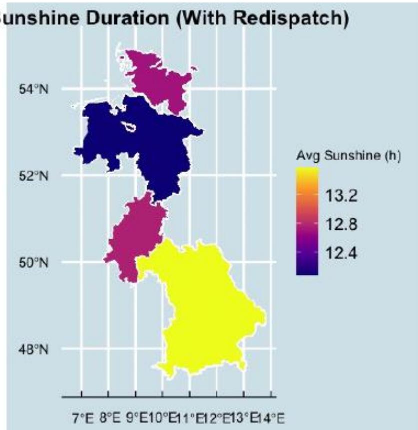
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 3a. Sunshine duration in TenneT DE operated regions vs redispatch. Average sunshine duration in each region in periods not affected by a redispatch



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 3b. Sunshine duration in TenneT DE operated regions vs redispatch. Average sunshine duration in each region in periods affected by a redispatch



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

3.1. Autoregressive properties of the target variable

To assess autocorrelation in binary time series, traditional Pearson-based autocorrelation functions (ACF) can be misleading, especially when the data is imbalanced. Instead, we employed two methods more appropriate for binary outcomes. First, we estimated a logistic regression model using only the previous value of the dependent variable as the independent variable. This allows for a probabilistic interpretation of autocorrelation by quantifying how the odds of a state change over time, controlling for imbalance and preserving the binary nature of the data.

Second, we examined the empirical transition matrix — the estimated probabilities of transitioning between states (e.g., from 0 to 1, or remaining in state 1), similar to analysing a first-order Markov chain. These two approaches are commonly used in binary time series modelling and discrete-state stochastic processes and together provide a robust picture of temporal dependence in binary outcomes. As shown below in tables 1, 2, 3 and 4, both approaches show that both series are heavily auto-regressive, having a very strong dependence on the state at the previous hour.

Moreover, after computing stationary distributions of the transition matrices we arrived at the exact empirical probabilities of the training data as expected (see Norris 1998).

Table 1. Empirical transition probabilities for the *down* series

State at t (*down*)	State at $t-1=0$	State at $t-1=1$
0	93.01%	2.42%
1	6.99%	97.58%

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Table 2. Lagged logistic regression estimates for the *down* series

Variable	Coefficient	Std. Error	p-value
Intercept	-2.5885	0.052	<0.001
down_{t-1}	6.2869	0.073	<0.001

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Table 3. Empirical transition probabilities for the *up* series

State at t (*down*)	State at $t-1=0$	State at $t-1=1$
0	94.13%	2.39%
1	5.87%	97.61%

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Table 4. Lagged logistic regression estimates for the *up* series

Variable	Coefficient	Std. Error	p-value
Intercept	-2.7744	0.053	<0.001
up_{t-1}	6.4833	0.075	<0.001

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

4. Methodology

4.1. Congestion management measure analyses

The achievements to date in congestion management research have undergone evolution from basic econometric methods (BDEW 2019) to machine learning techniques, significantly improving redispatched volume predictions in Germany (Titz et al. 2024). While a wide range of experiments is based on the estimation of aggregated congestion management reduction measures (Koch et al. 2018; Schr¨oder et al. 2018), other approaches, focus strictly on the issue of electric line congestion (Staudt et al. 2018; Wohland et al. 2018; Rausch et al. 2019; BDEW 2019; Titz et al. 2024).

4.2. Prediction of probability of redispatch (up&down)

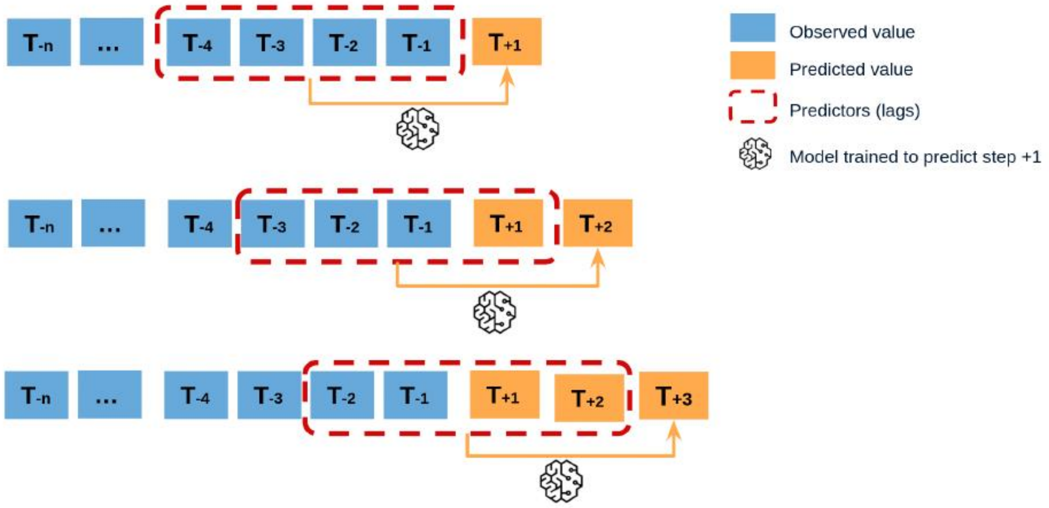
The selection of modelling methods is informed by several notable characteristics of the dataset. First, there is a significant class imbalance, with periods involving redispatch events comprising approximately 76% of the sample. Second, the data exhibits strong autoregressive behaviour, as discussed in the previous section. This results in a substantial portion of the model's explanatory power being captured by a simple AR(1) structure. Third, the dataset displays pronounced seasonality, particularly across weekdays and hours.

The modelling approach is twofold: it combines classical econometric techniques with modern machine learning methods. Each framework offers distinct advantages. Econometric models are statistically rigorous and fully interpretable, while machine learning methods typically offer higher predictive power. However, the former may suffer from lower performance, and the latter often requires eXplainable Artificial Intelligence (XAI) techniques for interpretation. As highlighted in the data description, the AR(1) component is highly influential. Therefore, all models are trained to perform one-stepahead (hourly) forecasts, which are then recursively chained to produce a 24-hour-ahead forecast, as illustrated in Figure 4.

Based on exploratory data analysis and domain intuition, the feature set was reduced to mitigate the risk of overfitting. Only the closing prices of the provided assets were retained, and minimum and maximum sunshine values across all locations were removed. Additional features were engineered from the available data, including the oil price spread, carbon emission futures spread, natural gas price spread, the spread between German and EU CO2 prices, and the spread between CO2 futures and spot prices in Germany, serving as proxies for market volatility.

A weekday indicator was also added to capture the seasonal patterns discussed earlier. The full list of the features used in prediction can be found in the appendix in Table 8.

Fig.4. Recursive forecasting approach. Source: SKforecast library documentation Amat Rodrigo and Escobar Ortiz (2024).



Given findings in the literature, the *up* and *down* redispatch time series are modelled separately. This distinction is important, as the relationship between explanatory variables and each redispatch type may differ—sometimes even in opposite directions. For direction-agnostic forecasting, a bottom-up approach is adopted by aggregating the individual forecasts for *up* and *down* redispatches.

For model training and evaluation, the dataset is divided into a training set (2020–2022) and a test set (January–July 2023). A rolling-window approach is employed, using a one-year window to capture intra-annual seasonality while avoiding training on data affected by systemic grid developments. Evaluation relies on the Area Under the Receiver Operating Characteristic Curve (AUC) (see 8.4), which provides stable performance comparisons under class imbalance. Additionally, the F1 score (see 8.3) is used to calibrate the classification threshold, as it explicitly accounts for the precision–recall trade-off. Robustness is ensured through a 10-fold rolling-window validation scheme. The rolling window approach is chosen over the expanding window in order to limit the risk of systemic error due to training the models based on expired relationships. Hyperparameters are tuned using randomized search on the training set, and final model selection is based on test set performance.

The class imbalance issue is addressed through three strategies. First, the machine learning models employ asymmetric loss functions that assign higher weights to errors involving the minority class. Second, imbalance-robust evaluation metrics are selected. Third, the classification threshold is optimized with respect to the F1 score, rather than defaulting to the 50% cut-off.

As benchmarks, both naive 1-hour-ahead and 24-hour-ahead forecasts are considered. The AutoRegressive Logistic Regression model (Guanche et al. 2013) is used to represent the econometric baseline. The machine learning models applied include XGBoost (Chen and Guestrin 2016), CatBoost (Prokhorenkova et al. 2017), Random Forest (Ho 1995), and L1-regularized AutoRegressive Logistic Regression (following Nardi and Rinaldo 2011). This selection spans a diverse range of architectures and contributes to the novelty of this study by extending the gradient boosting framework used in Titz et al. (2024). The inclusion of both bagging- and boosting-based methods further supports the robustness of the findings.

While neural network architectures—particularly Long Short-Term Memory (LSTM) networks and Temporal Fusion Transformers—have shown promise in related tasks, they require substantially larger training samples and longer training times. This increases the risk of overfitting or capturing systemic errors, and they are therefore not included in the current analysis.

In order to extract policy insights from the ml models, as well as to fully understand them, eXplainable AI (XAI) method is used. XAI is a set of (often) model agnostic methods, which allow to approximate the econometric models' explainability for the more complex blackbox models. The XAI methods allow, for example, to understand for what type of observations the model predictions are inaccurate (waterfall plots), what is the approximate relationship between the target and the individual features (using partial dependence profiles) or show which features are the most significant when it comes to prediction (variable importance and bee swarm). The XAI tools utilized in this experiment are all contained within the SHAP package, which bases on SHapley Additive exPlanations.

4.3. Additional methodology for Logistic Regression

We wanted to implement logistic regression due to its relative simplicity, transparency and interpretability. Since the rest of our approaches focus on machine learning methods which are robust to high number of correlated features, we only had to perform feature selection for logistic regression approach. After preprocessing the data, adding one hot encoded variables and lagged variables due to suspected heavy autoregressive properties of the dependent variable we were left with 93 potential features. Then we implemented Mutual Information (see 8.1) method to isolate the most important features. Since a lot of them exhibit multicollinearity between each other we implemented Variance

Inflation Factor method to get rid of the most correlated features and dropped every feature with VIF (see 8.2) higher than 10.

5. Results

5.1. Logistic Regression

We present the results of logistic regression approach for both target variables below:

5.1.1. Down

The results suggested that only 6 features were significant (P-value less than 0.05) for *down* redispatch prediction show in Table 5.

Table 5. Statistically Significant Variables for *Down* prediction (P-val less than 0.05)

Variable	Coefficient	Std. Error	z-score	P(z)
const	-10.5105	1.495	-7.030	0.000
down_lag_1	6.4663	0.281	23.038	0.000
Brent_oil_spread	0.1521	0.063	2.397	0.017
carbon_emissions_futures_vol	0.0004	0.000	2.376	0.017
max_wind_velocity_schleswig_holstein	0.1264	0.049	2.599	0.009
max_wind_velocity_Niedersachsen	0.1295	0.050	2.607	0.009
min_wind_velocity_Niedersachsen	0.5589	0.190	2.938	0.003

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

When tested on the data from 2023 onwards the ROC-AUC score was 96.4% while the F1-score was 97.5% which suggests that even a relatively simple model such as logistic regression can provide quite satisfactory results.

5.1.2. Up

Similarly, for *up* redispatch prediction there were also just 6 significant features, shown in Table 6. When tested on the data from 2023 onward the ROC-AUC score was 96.6% while the F1-score was 97.7% which points to the same conclusion as in *down* prediction.

Overall, despite the high values of performance measures, the logistic regression is likely to be outperformed by more advanced models. Furthermore, due to subtle, likely non-linear output are found to be insignificant, losing the explanatory power which can be levied from them.

5.2. Model comparison

Given that the logistic regression model struggles to capture more subtle and non-linear relationships — likely due to the dominant explanatory power of the AR(1) component and the imbalanced nature of the data — machine learning approaches were also employed. The objective was to enhance predictive performance while benchmarking the results against the naïve or simpler models to provide a frame of reference.

Table 6. Statistically Significant Variables for *Up* prediction (P-val less than 0.05)

Variable	Coefficient	Std. Error	z-score	P(z)
const	-10.4808	1.403	-7.471	0.000
up_lag_1	6.1410	0.239	25.730	0.000
Natural_gas price_spread	-0.8781	0.381	-2.302	0.021
up_lag_24	0.3981	0.174	2.288	0.022
max_wind_velocity_Niedersachsen	0.1189	0.047	2.527	0.012
max_wind_velocity_schleswig_holstein	0.1018	0.044	2.310	0.021
min_wind_velocity_Niedersachsen	0.5815	0.186	3.126	0.002

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Table 7. Statistically Significant Variables for *Up* prediction (P-val less than 0.05)

Model	*Up* Redispatch		*Down* Redispatch	
	AUC (%)	F1 (%)	AUC (%)	F1 (%)
Naive model t-1	95.09	97.70	94.75	97.51
Naive model t-24	63.29	82.78	62.66	82.28
Logistic regression	96.73	97.69	96.41	97.50
L1-regularized logistic regression	96.69	97.69	96.53	97.51
Random Forest	96.56	97.68	96.57	96.86
XGBoost	97.05	97.70	96.68	97.51
Catboost	97.16	97.70	96.84	97.51

Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

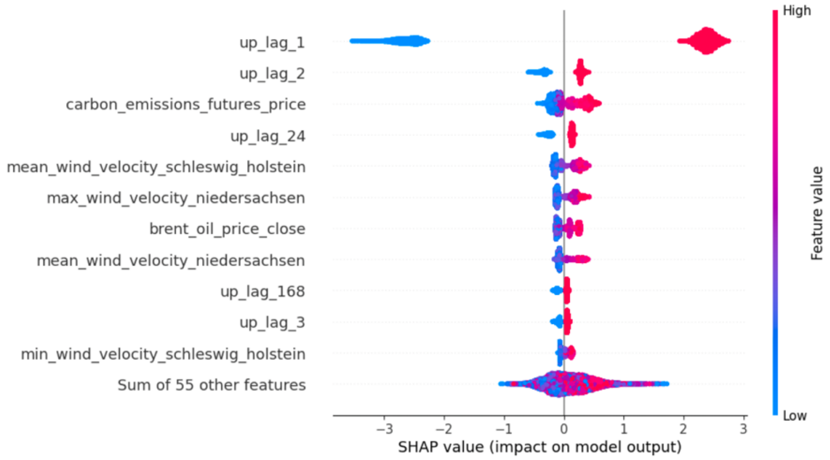
Table 7. presents the performance of each model on the test set. The primary metric used for model comparison is the AUC score, as it captures the quality of the predicted probabilities across all classification thresholds, rather than simply evaluating binary outcomes. CatBoost emerges as the best performing model in both prediction tasks, slightly outperforming XGBoost. Consequently, CatBoost is selected for both the forecasting and explainability (XAI) analyses.

As expected, the boosting algorithms outperform simpler models such as logistic regression, L1- regularized logistic regression, and random forest—a trend consistent with their ability to model complex non-linear relationships. Nevertheless, these simpler models still achieve better AUC scores than the naive baselines, indicating that they can provide valuable insights into the data without the high computation cost required to train and tune complex ML models.

The strong test set performance of the more complex models (CatBoost and XGBoost) suggests that the applied overfitting mitigation strategies have been effective. None of the models was able to outperform the naive AR(1) model in terms of F1 score, indicating that the predicted binary outcomes remained largely consistent across time periods, regardless of model complexity. This outcome is expected, as the majority of the predictive power lies within the AR(1) structure. The additional features primarily contribute to more nuanced shifts in redispatch probabilities, which are better captured by metrics like AUC — where AR(1) performs relatively worse.

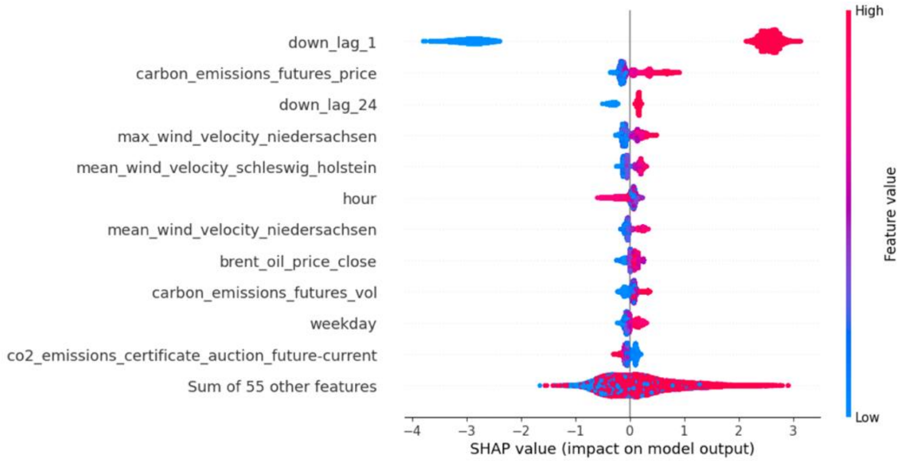
Despite this, CatBoost and XGBoost remain among the top performers in terms of F1 score. Figures 5 and 6 show the most important variables for both CatBoost models. It can be found, that as expected the first lag contributes the most to the prediction in both cases, with other common denominators between the models being wind speed measures, other lags of the target as well as carbon emission credit futures price. The 24th lag is also found to be a good predictor for both models, suggesting that the seasonal trends are noticed by the model for both *up* and *down* redispatches. Interestingly, for *down*-redispatch the hour of day is the 6th best performing feature, with later hours being associated with lower *down*-redispatch risk, the time of day is not found to be among the top predictors for the up-redispatch model.

Fig.5. Shap values beeswarm plot for *up* time series



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.6. Shap values beeswarm plot for *down* time series

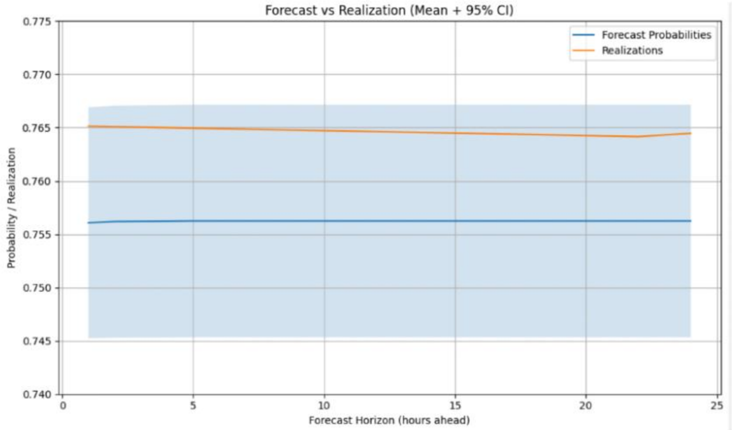


Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

5.3. Forecast results

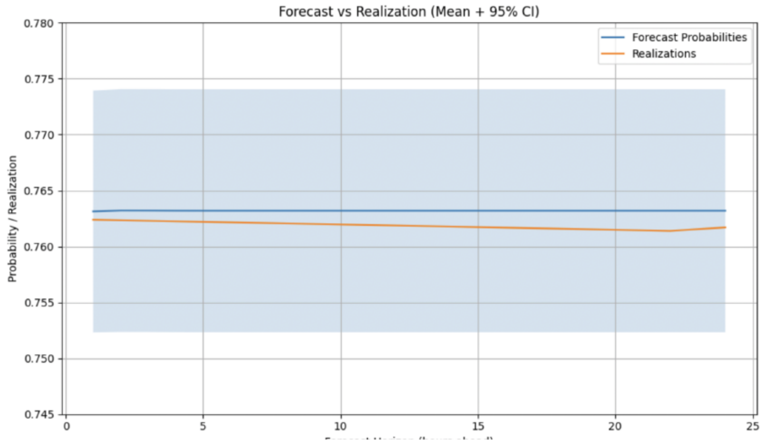
A 24-step recurrent forecast was performed using CatBoost to estimate the probability of both *up* and *down* redispatch events. Due to the recurrent structure of the forecast, prediction errors are expected to accumulate across steps — meaning that the forecast error at the 24th hour is amplified by the preceding 23 predictions. To verify this, the average forecast was compared to the average realization for each period in test sample and for each model. Figures 7a and 7b present these results. The forecasts and realizations are relatively stable over time, with the observed values consistently falling within the 95% confidence bounds of the average predictions. This likely stems from two main factors. First, the models exhibit high predictive accuracy (with F1 scores around 97.5), which leads to minimal error accumulation across the 24 forecast steps. Second, the data's strong autoregressive behaviour makes class transitions within a 24-hour window relatively rare, resulting in stable forecasts. Overall, the forecasting performance appears solid, although there is a slight tendency to underestimate the probability of up-redispatch events.

Fig. 7a. Average CatBoost forecast analysis for the test set – 2023. Average forecast of B *up* model for 2023



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.7b. Average CatBoost forecast analysis for the test set – 2023. Average forecast of CatBoost *down* model for 2023



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

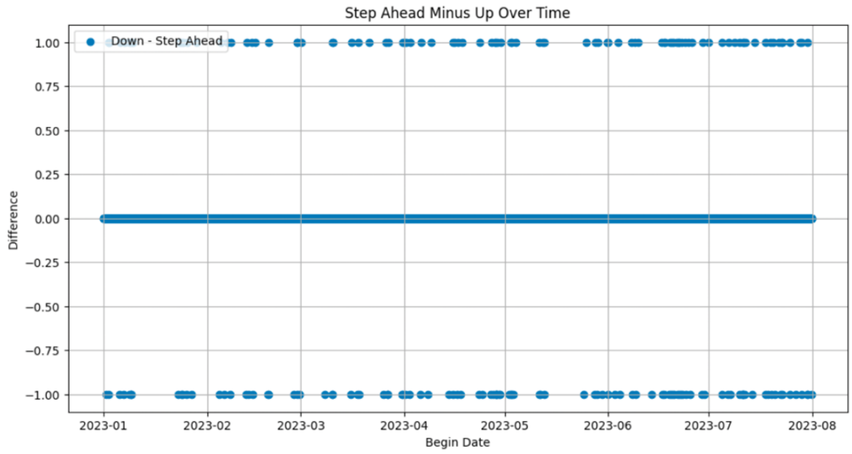
5.4. Forecast misclassification

In this subsection we dive deep to understand why misclassification occurs in our models. The error plot has been used to detect periods of high instability, as in Figure 8. both forecasts notice a drop in quality in the second part of June. Therefore, the 20th of June 2023 has been selected as a candidate as there are two misclassifications in two different directions occurring for both models *up* and *down*.

The False Negative occurred at 10:00 while the False Positive occurred at 19:00. The waterfall of Shapley values for these two specific observations for both *up* and *down* models are shown on Figure 9 and Figure 10. For all cases we can see the overwhelming impact of the AR(1) component which impacted the model prediction.

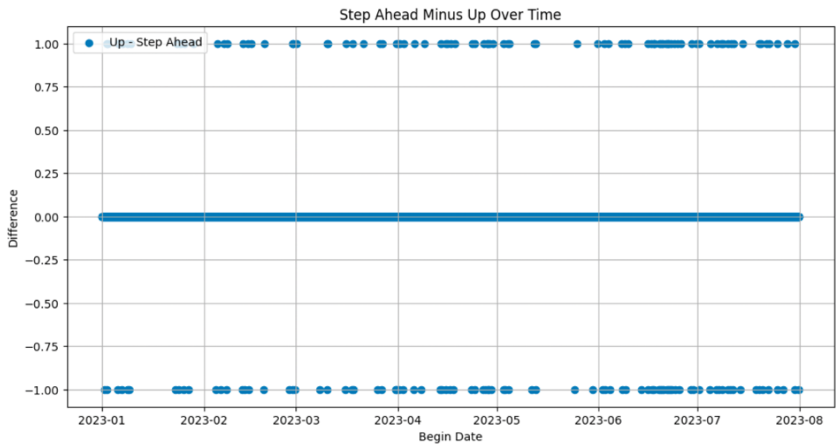
In case of both the False Negative we can notice that if the AR(1) were to be removed the prediction would actually change sign and lead to a correct prediction. At the same time, the forecast for False Positives would still be incorrect, however within much smaller margin.

Fig. 8a. Difference in actual value vs prediction for *up* and *down*. Error in prediction *up* time series



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 8b. Difference in actual value vs prediction for *up* and *down*. Error in prediction *down* time series



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

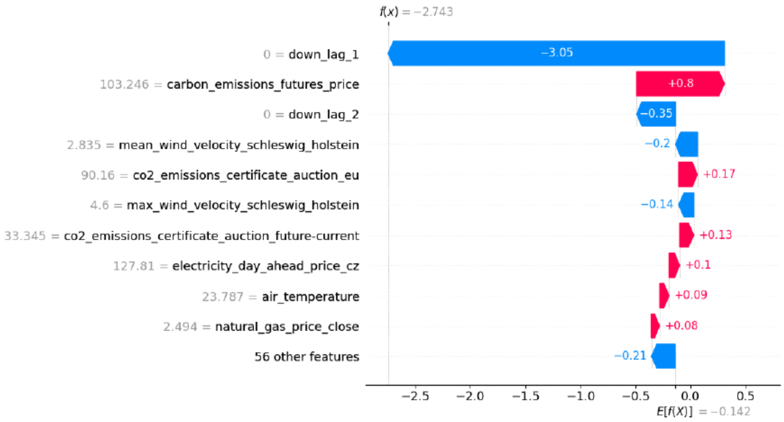
6. Discussion and Policy recommendations

The applied machine learning techniques allowed us to estimate the probability of congestion events within the German electricity network by congestion type *up* - demand for additional power in the local grid vs *down* - need to decrease power production to balance the grid. In order to offer novel, and interpretable results, we propose a dual approach that allows policy proposals depending on the problem the electricity network is facing.

The policies presented below are only a suggestion for the direction of the expected government action. It is important to address that the relationships below are interpretations of the contribution of a given feature to the prediction of electricity redispatch based on the SHAP values obtained. We motivate our approach by the high predictive ability of the model and practices commonly used in the study area (Titz et al. 2024).

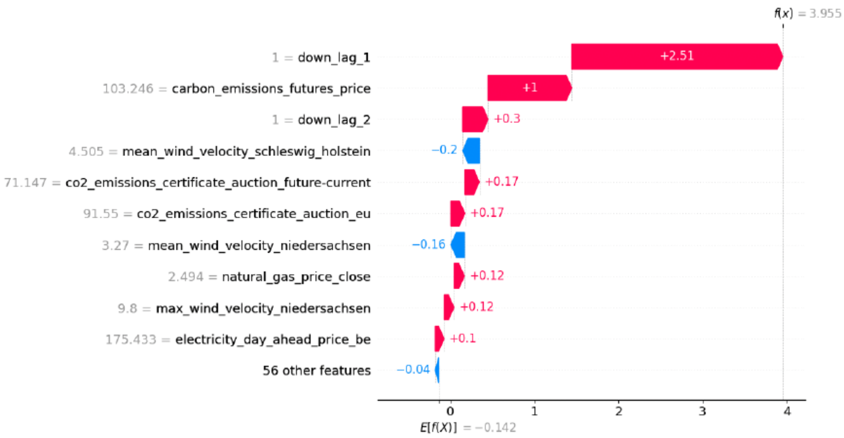
In addition, we note that due to the similar relationship (SHAP values obtained) for the probability of congestion in the '*up*' and '*down*' cases, we will refer to the results together, addressing the discrepancies that exist.

Fig. 9a. Shap waterfall decomposition for misclassification for *down* redispatch. Shap values waterfall composition for 20-06-2023 10:00



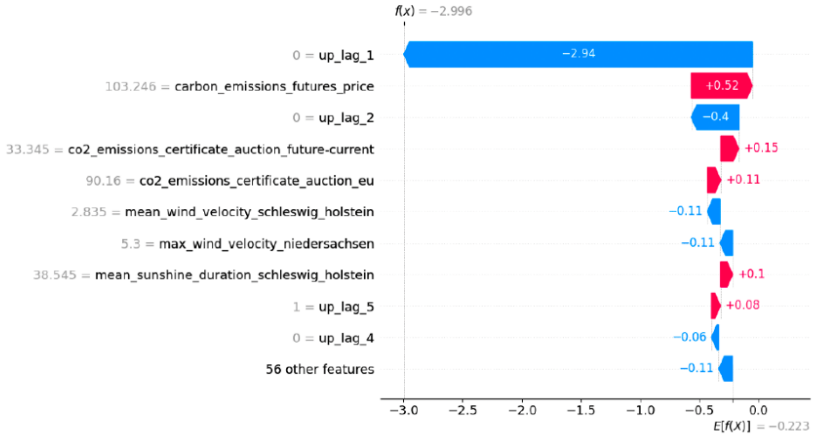
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 9b. Shap waterfall decomposition for misclassification for *down* redispatch. Shap values waterfall composition for 20-06-2023 19:00



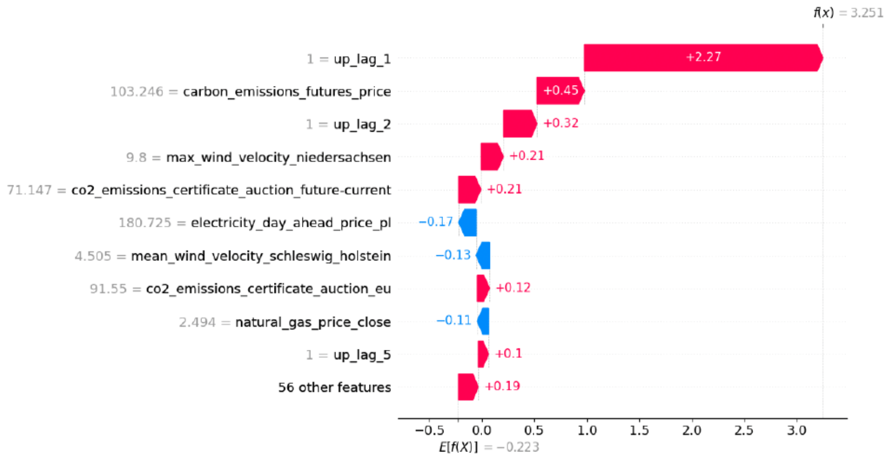
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 10a. Shap waterfall decomposition for misclassification for *up* redispatch. Shap values waterfall composition for 20-06-2023 10:00



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 10b. Shap waterfall decomposition for misclassification for *up* redispatch. Shap values waterfall composition for 20-06-2023 19:00



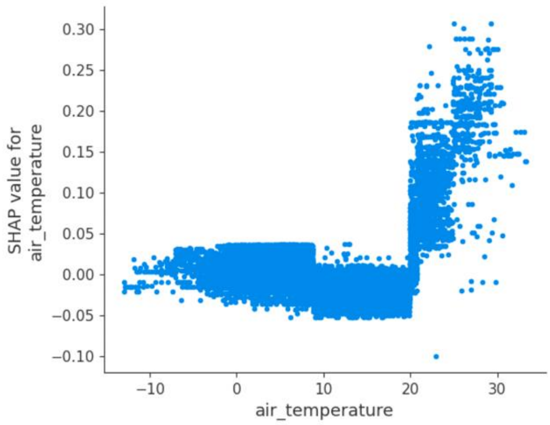
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

6.1. Climate Policies

We argue that a high risk of grid congestion occurs during high wind and heat conditions. This conclusion is drawn from the partial dependence profiles for air temperature and maximum wind Niedersachsen in Figure 11. and Figure 13. The focus on the Niedersachsen region is motivated by high density of wind mill electricity generators, very much dependent on the wind conditions. Notice, that the relationship between the wind proxy feature and impact on the likelihood of redispatch is positive yet not simply linear – underlining the value of machine learning methods. What is more we can see a discrepancy in the influence of air temperature on the likelihood of redispatch. Higher air temperature correlates with higher likelihood for *down* dispatch which most likely is due to higher electricity demand during hotter periods. Conversely, the *up* dispatch likelihood is negatively impacted by higher temperatures possibly because of lower demand. As such, we propose policies that mandate the integration of real-time congestion forecasting tools into energy markets, including day-ahead and intra-day frameworks. Leveraging CatBoost-based grid models informed by meteorological data could enable system operators and market participants to anticipate network constraints—such as those arising from sudden increases in wind generation or reduced cooling system efficiency during heatwaves.

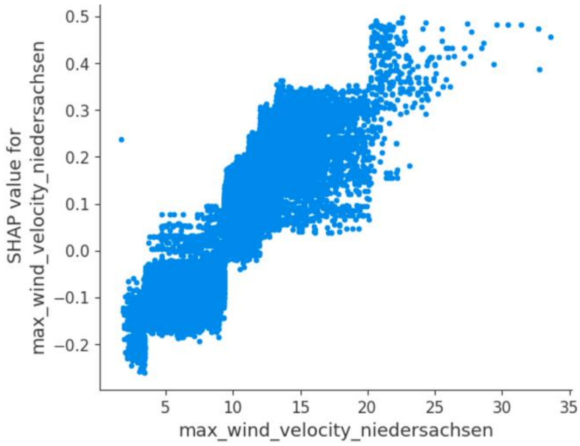
In addition, following (Glaum and Hofmann 2023) we propose Temperature-Aware Transmission Line Ratings (Dynamic Line Rating - DLR). Because static ratings underestimate capacity during cooler wind conditions and overstate during heat waves, the mandate adoption of DLR to better reflect real-time thermal capacity of transmission lines. In addition, given the geographical distribution of wind turbines in Germany—particularly the more favourable wind conditions in the North and East (Creutzig et al. 2014; Titz et al. 2024)—we propose a policy requiring new wind farms to be co-located with battery storage systems or to participate in virtual power plants. Such integration would allow batteries to absorb excess generation during periods of oversupply (mitigating *down*-type congestion) or discharge during curtailment events (addressing *up*-type congestion), aligning with the recommendations of Denholm et al. (2015).

Fig.11a. Partial dependence profiles for *down* time series. Partial dependence profile for air temperature



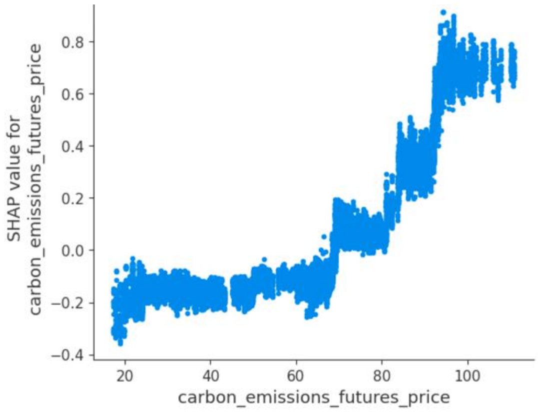
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.11b. Partial dependence profiles for *down* time series. Partial dependence profile for maximum wind speed in the Niedersachsen region



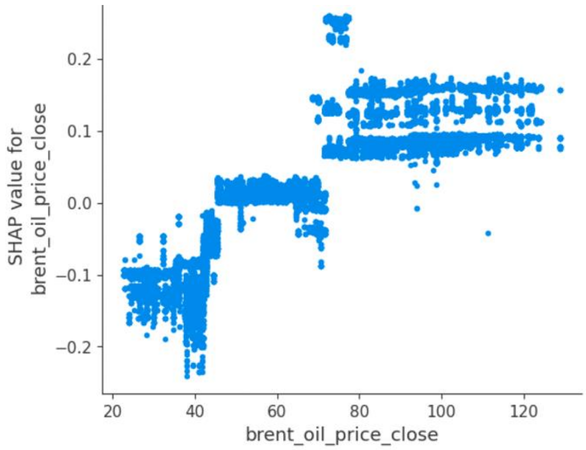
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.12a. Partial dependence profiles for *down* time series. Partial dependence profile for CO2 emissions futures closing prices



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig.12b. Partial dependence profiles for *down* time series. Partial dependence profile for Brent oil closing price

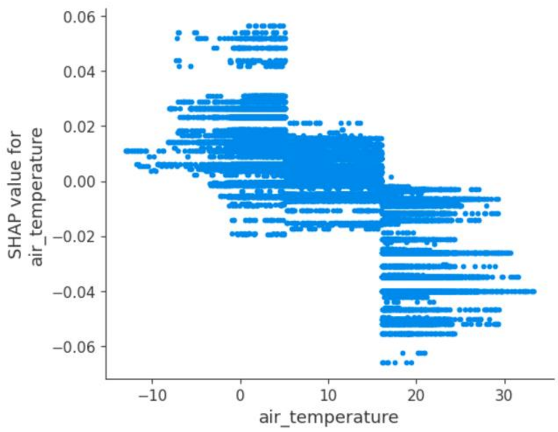


Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

6.2. Policy Recommendations for Market-Driven Congestion Risk

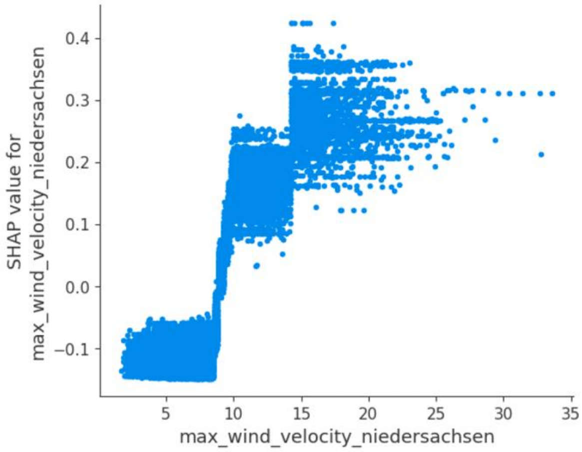
In turn, by analysing the market factors associated with a higher probability of grid congestion, we provide evidence that points to a positive contribution from the Brent oil price, the gas price and carbon dioxide futures (Marimoutou and Soury 2015; Davi-Arderius and Schittekatte 2023; Davi-Arderius and Schittekatte 2023). As shown in Figure 12. and Figure 14. the interplay between CO2 emissions futures prices and Brent oil prices is qualitatively the same for both *up* and *down*. In both cases the relationship is positive meaning that the higher the prices the higher the likelihood of redispatch occurring, regardless of being *up* or *down*. Consistently, we argue that the introduction of a CO2-Oil Risk Indicator for Grid Operation could be a valuable congestion risk metric in system operator dashboards. Consistently, it would act as an early warning when shifts in fossil fuel arbitrage may spike congestion. Since the results indicate a high correlation and volatility between CO2, oil prices and a likelihood of congestion, we propose an adjustment in the regional carbon tax intensity and oil prices dynamics based on grid saturation. In this way, we allow for an opening towards policies that support the Energiewende and the climate transition including the decarbonisation of Germany.

Fig. 13a. Partial dependence profiles for *up* time series. Partial dependence profile for air temperature



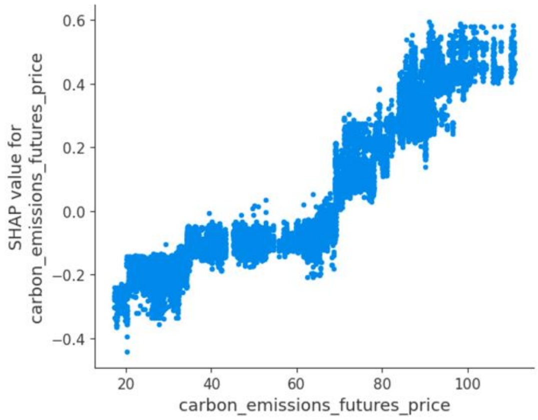
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 13b. Partial dependence profiles for *up* time series. Partial dependence profile for maximum wind speed in the Niedersachsen region



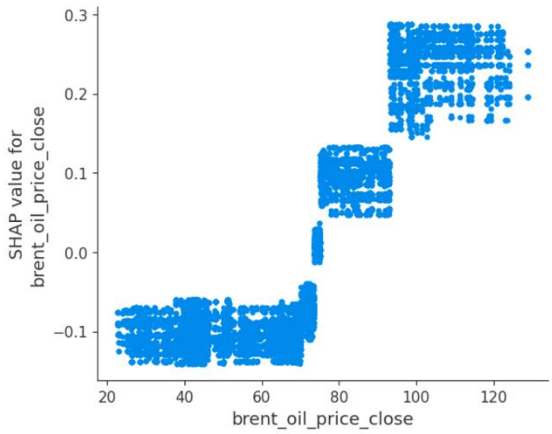
Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 14a. Partial dependence profiles for *up* time series. Partial dependence profile for CO2 emissions futures closing prices



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

Fig. 14b. Partial dependence profiles for *up* time series. Partial dependence profile for Brent oil closing price



Source: Own elaboration based on data provided by Anastasija Tetereva, Onno Kleen and Alla Petukhina during the Econometric Game 2025 organised by the University of Amsterdam.

7. Conclusions

The aim of this study is to minimise the probability of congestion in the German electricity grid and thus propose a climate policy that enables a smooth Energiewende transition. We offer a dual approach, addressing the problem of redispatch as a result of energy surplus and shortage. Using advanced econometric tools as well as a machine learning approach, we find a higher predictive power of the latter and, finding only a few statistically significant regressors of the former, we propose solutions based on machine learning models. The models are trained using a rolling validation approach and interpreted using explainable artificial intelligence (XAI) methods. Due to the key characteristics of the data - namely class imbalance, seasonality and strong autoregressive dynamics - CatBoost proved to be the best performing model on the test set, as measured by the AUC metric.

The analysis indicates that lower temperatures tend to result in an overestimation of electricity demand, thereby increasing the likelihood of *up* redispatch events. However, these same temperature conditions are also associated with a decrease in *down* redispatches. Consequently, we reject the first hypothesis, as the evidence does not clearly support the claim that lower temperatures systematically lead to increased grid congestion.

Wind speed is consistently identified as one of the most important predictors in the final models. This finding supports the second hypothesis, and we therefore fail to reject it. Furthermore, CatBoost models reveal that rising carbon emission prices are associated with an elevated risk of redispatch events, lending support to the third hypothesis. Finally, increases in Brent oil and natural gas prices are also found to be positively associated with redispatch risk, suggesting that broader energy market conditions may serve as valuable signals for congestion forecasting.

Consequently, a dual approach to optimising the risk of electricity congestion in Germany has enabled climate policy proposals and market-driven regulations. Firstly, we argue that the use of CatBoost grid modelling will allow mandating real-time congestion forecasting tools integrated into energy markets. In addition, we argue that the adoption of DLR (Dynamic Line Rating) will allow for more accurate precision of weather conditions, thereby mninimising congestion risk and thus faster supplier response. In turn, considering the market-driven side of the solutions offered, we suggest transformations towards a CO2-Oil Risk Indicator for Grid Operation, which are an important determinant in spike congestion. Additionally, regulating the regional carbon tax together with tracking fuel price dynamics will optimise the probability of redispatch. We argue that the vision of transforming Germany's (Merkel 2011; Altmaier 2018) is possible, but only with continuous monitoring of the atmospheric and market situation. At the same time, we would like to point out that the above remarks are only suggestions setting the direction towards the Energiewende. We argue that a driven-by-data approach, using a state-of-the-art combination of machine learning models with XAI techniques, significantly outperforms existing approaches, thereby filling an existing gap in the literature (Staudt et al. 2018; Wohland et al. 2018; Rausch et al. 2019; BDEW 2019; Titz et al. 2024).

This research, however, is subject to several limitations. First, state-of-the-art neural network architectures were not employed due to the limited length of the available training window. Second, the time period covered by the models includes a high degree of structural instability, particularly due to pandemic-related disruptions (Alasali et al. 2021). As such, extending the analysis to include data from pre-pandemic periods may provide additional insights and improve robustness.

Furthermore, redispatches are modeled as a binary outcome rather than in terms of their associated costs. This may reduce the economic relevance of the findings, particularly if prediction errors are associated with high-cost outliers. Finally, the identification of redispatch events lacks spatial granularity. Since redispatches are not evenly distributed across the grid, the absence of spatial information constitutes a significant limitation, especially given the documented heterogeneity in grid topology and congestion patterns (Creutzig et al. 2014; Titz et al. 2024).

References

Agency, I.E., 2016. Key world energy statistics. Available at: https://www.iea.org/publications/freepublications/publication/KeyWorldStatistics2015.pdf [Accessed: 10.04.2025].

Alasali, F., Nusair, K.N., Alhmoud, L., Zarour, E., 2021. Impact of the covid-19 pandemic on electricity demand and load forecasting. Sustainability. https://doi.org/10.3390/SU13031435.

Altmaier, M., 2018. Minister altmaier takes grid expansion into his own hands. Bundesministerium fur Wirtschaft und Klimaschutz der Bundesrepublik Deutschland.

Amat Rodrigo, J., Escobar Ortiz, J., 2024. skforecast.

BDEW, 2019. Redispatch in deutschland: Auswertung der transparenzdaten april 2013 bis einschließlich dezember 2018. Online.

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

Bundesnetzagentur und Bundeskartellamt, 2023. Monitoringbericht 2024. Avaliable at: https://www.bundesnetzagentur.de/DE/Fachthemen/ElektrizitaetundGas/Monitoringberichte/start.html [Accessed: 10.04.2025].

Bundesregierung, 2011. Die energiewende. Avaliable at: https://www.bundesregierung.de/statisch/jb11-12/Webs/Breg/DE/Themen/Jahresbericht20112012/03Energie/1Energiewende/node.html [Accessed: 10.04.2025].

Byun, S.J., Park, H.S., Yi, S.J., Song, C.H., Choi, Y.D., Lee, S.H., Shin, J.K., 2015. Study on the optimal heat supply control algorithm for district heating distribution network in response to outdoor air temperature. Energy 86: 247–256. https://doi.org/10.1016/j.energy.2015.04.029.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. https://doi.org/10.48550/ ARXIV.1603.02754.

Comission, E., 2016. State aid no. sa.42955 (2016/n-2) – germanynetwork reserve. European Comission, Available at:

https://ec.europa.eu/competition/stateaid/cases/265043/2650431872192912.pdf [Accessed: 10.04.2025].

Creutzig, F., Goldschmidt, J.C., Lehmann, P., Schmid, E., von Blucher, F., Breyer, C., Fernandez, B., Jakob, M., Knopf, B., Lohrey, S., Susca, T., Wiegandt, K., 2014. Catching two european birds with one renewable stone: Mitigating climate change and eurozone crisis by an energy transition. Renewable and Sustainable Energy Reviews 38: 1015–1028. https://doi.org/10.1016/j.rser. 2014.07.028.

Curry, A., 2019. Germany faces its future as a pioneer in sustainability and renewable energy. Nature 567 (7749): S51–S53. https://doi.org/10.1038/d41586-019-00916-1.

Davi-Arderius, D., Schittekatte, T., 2023. Carbon emissions impacts of operational network constraints: The case of spain during the covid-19 crisis. Energy Economics 128: 107164. https://doi.org/10.1016/j.eneco.2023.107164.

Davi-Arderius, D., Schittekatte, T., 2023. Environmental impacts of redispatching in decarbonizing electricity systems: A spanish case study. SSRN Electronic Journal. https://doi.org/10.2139/ssrn. 4339705.

Denholm, P., O'Connell, M., Brinkman, G., Jorgenson, J., 2015. Overgeneration from solar energy in california: A field guide to the duck chart. Technical Report NREL/TP-6A20-65023, National Renewable Energy Laboratory (NREL), Golden, CO, USA.

Dorfer, M., Fuxjager, A.R., Kozak, K., Blies, P.M., Wasserer, M., 2022. Power grid congestion management via topology optimization with alphazero. https://doi.org/10.48550/ARXIV.2211.05612.

Elsner, P., Erlach, B., Fischedick, M., Lunz, B., Sauer, D.U., 2015. Flexibilitatskonzepte fur die Stromversorgung 2050: Technologien - Szenarien - Systemzusammenhange. Germany: acatech.

Ember, Energy Institute, and OurWorld in Data, 2024. Electricity generation from wind power – ember and energy institute. https://ourworldindata.org/grapher/electricity-generation-by-source. Major processing by Our World in Data. Based on Ember, "Yearly Electricity Data"; Energy Institute, "Statistical Review of World Energy".

Emeis, S., 2018. Wind Energy Meteorology: Atmospheric Physics for Wind Power Generation. Cham, Switzerland: Springer.

Energy Systems of the Future (ESYS), 2020. Electricity grid congestion – towards smart distribution grids. https://energiesysteme-zukunft.de/fileadmin/user upload/Publikationen/PDFs/ESYS Position Paper Grid congestion.pdf. Position paper by the Academies' Project "Energy Systems of the Future".

Fernandez, E., Albizu, I., Bedialauneta, M., Mazon, A., Leite, P., 2016. Review of dynamic line rating systems for wind power integration. Renewable and Sustainable Energy Reviews 53: 80–92. https://doi.org/10.1016/j.rser.2015.07.149.

Glaum, P., Hofmann, F., 2023. Leveraging the existing german transmission grid with dynamic line rating. Applied Energy 343: 121199. https://doi.org/10.1016/j.apenergy.2023.121199.

Grochowicz, A., van Greevenbroek, K., Bloomfield, H.C., 2023. Using power system modelling outputs to identify weather-induced extreme events in highly renewable systems. https://doi.org/ 10.48550/ARXIV.2307.13520.

Guanche, Y., Minguez, R., Mendez, F.J., 2013. Autoregressive logistic regression applied to atmospheric circulation patterns. Climate Dynamics 42 (1–2): 537–552. https://doi.org/10.1007/ s00382-013-1690-3.

Hirschhausen, C., 2014. The german "energiewende"—an introduction. Economics of Energy Environmental Policy 3 (2): 1–12.

Ho, T.K., 1995. Random decision forests, in: Proceedings of 3rd International Conference on Document Analysis and Recognition, Volume 1, pp. 278–282 vol.1.

Jung, C., Schindler, D., 2020. Integration of small-scale surface properties in a new high resolution global wind speed model. Energy Conversion and Management 210: 112733. https: //doi.org/10.1016/j.enconman.2020.112733.

Koch, L., Schroder, D., Erdmann, G., 2018. Reasons for curtailment of electricity production in northeastern germany – an empirical analysis. Presented at the 41st IAEE International Conference, Groningen.

Marimoutou, V., Soury, M., 2015. Energy markets and co2 emissions: Analysis by stochastic copula autoregressive model. Energy 88: 417–429. https://doi.org/10.1016/j.energy.2015.05.060.

Merkel, A., 2011. We want to end the use of nuclear energy and reach the age of renewable energy as fast as possible. The Guardian.

Metzger, M., Duckheim, M., Franken, M., Heger, H.J., Huber, M., Knittel, M., Kolster, T., Kueppers, M., Meier, C., Most, D., Paulus, S., Wyrwoll, L., Moser, A., Niessen, S., 2021. Pathways toward a decarbonized future—impact on security of supply and system stability in a sustainable german energy system. Energies 14 (3): 560. https://doi.org/10.3390/en14030560.

Michiorri, A., Nguyen, H.M., Alessandrini, S., Bremnes, J.B., Dierer, S., Ferrero, E., Nygaard, B.E., Pinson, P., Thomaidis, N., Uski, S., 2015. Forecasting for dynamic line rating. Renewable and Sustainable Energy Reviews 52: 1713–1730. https://doi.org/10.1016/j.rser.2015.07.134.

Nardi, Y., Rinaldo, A., 2011. Autoregressive process modeling via the lasso procedure. Journal of Multivariate Analysis 102 (3): 528–549. https://doi.org/10.1016/j.jmva.2010.10.012.

Neukirch, M., 2016. Protests against german electricity grid extension as a new social movement? a journey into the areas of conflict. Energy, Sustainability and Society 6 (1). https: //doi.org/10.1186/s13705-016-0069-9.

Norris, J.R., 1998. Markov Chains. Cambridge University Press.

Peper, J., Kroger, D., Kipp, J., Ziel, F., Rehtanz, C., 2024. Assessing the impact of weather-induced uncertainties in large-scale electricity systems.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A., 2017. Catboost: unbiased boosting with categorical features.

Quitzow, L., Canzler, W., Grundmann, P., Leibenath, M., Moss, T., Rave, T., 2016. The german energiewende – what's happening? introducing the special issue. Utilities Policy 41: 163–171. https://doi.org/10.1016/j.jup.2016.03.002.

Rausch, B., Staudt, P., Weinhardt, C., 2019. Transmission grid congestion data and directions for future research, in: Proceedings of the Tenth ACM International Conference on Future Energy Systems, Phoenix, AZ, USA, pp. 443–446.

Renn, O., Marshall, J.P., 2016. Coal, nuclear and renewable energy policies in germany: From the 1950s to the "energiewende". Energy Policy 99: 224–232. https://doi.org/10.1016/j.enpol. 2016.05.004.

Rodriguez, R.A., Becker, S., Andresen, G.B., Heide, D., Greiner, M., 2014. Transmission needs across a fully renewable european power system. Renewable Energy 63: 467–476. https://doi.org/10.1016/ j.renene.2013.10.005.

Schmidt, O., Staffell, I., 2023. Monetizing Energy Storage: A Toolkit to Assess Future Cost and Value. Oxford University PressOxford.

Schroder, D., Koch, L., Erdmann, G., 2018. Empirical analysis of regional congestion and curtailment clusters in northeastern germany. Presented at the 41st IAEE International Conference, Groningen.

Sharifzadeh, M., Lubiano-Walochik, H., Shah, N., 2017. Integrated renewable electricity generation considering uncertainties: The uk roadmap to 50 Renewable and Sustainable Energy Reviews 72: 385–398. https://doi.org/10.1016/j.rser.2017.01.069.

Sievers, L., Breitschopf, B., Pfaff, M., Schaffer, A., 2019. Macroeconomic impact of the german energy transition and its distribution by sectors and regions. Ecological Economics 160: 191–204. https://doi.org/10.1016/j.ecolecon.2019.02.017.

Staudt, P., Traeris, Y., Rausch, B., Weinhardt, C., 2018. Predicting redispatch in the german electricity market using information systems based on machine learning, in: ICIS 2018 Proceedings, San Francisco, USA, pp. 443–446. Presented at the International Conference on Information Systems (ICIS) 2018.

Titz, M., Putz, 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. https://doi.org/10.1016/j.apenergy.2023.122351.

Wohland, J., Reyers, M., Marker, C., Witthaut, D., 2018. Natural wind variability triggered drop in german redispatch volume and costs from 2015 to 2016. PLOS ONE 13 (1): e0190707. https://doi.org/10.1371/journal.pone.0190707.

Zhang, Q., Li, F., Fang, X., Zhao, J., 2024. Implications of electricity and gas price coupling in us new england region. iScience 27 (1): 108726. https://doi.org/10.1016/j.isci.2023.108726.

Appendixes

1.1. Mutual Information for Feature selection

Mutual Information (MI) is an information-theoretic measure that quantifies the dependence between a feature X and a target variable Y. It is defined as:

$$MI(X;Y) = \sum_{x\in\chi} \sum_{y\in\gamma} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

where $p(x,y)$ is the joint probability distribution of X and Y, and $p(x)$, $p(y)$ are the marginal distributions. MI is zero if and only if X and Y are independent, and it captures both linear and non-linear dependencies.

1.2. Variance Inflation Factor

Variance Inflation Factor (VIF) quantifies the degree of multicollinearity among features in a regression model.

For a given predictor X_j, VIF is defined as:

$$VIF(X_j) = \frac{1}{1 - R_j^2}$$

where R_j^2j is the coefficient of determination from regressing $X_j$on all other predictors. A high VIF indicates that $X_j$is highly linearly correlated with other features, which can inflate the variance of estimated coefficients and lead to unstable or non-interpretable models. In practice, features with VIF higher than 10 should be transformed or excluded from further analysis.

1.3. F1 Score

The F1 score is a performance metric for binary classification that balances precision and recall. It is particularly useful when the class distribution is imbalanced or when false positives and false negatives carry different costs.

It is defined as the harmonic mean of precision and recall:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

where

$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN}$$

with TP = true positives, FP = false positives, and FN = false negatives. The $F1$ score ranges from 0 to 1, with higher values indicating better classification performance.

1.4. AUC score

The Receiver Operating Characteristic - Area Under the Curve (AUC) score is a performance metric for binary classification models, such as the one we are modelling in this analysis. It measures how well the classifer can distinguish between 0 and 1 for each possible cutoff threshold.

Formally, the ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR), and the AUC is the integral under this curve:

$$AUC = \int_0^1 TPR(FPR^{-1}(x))dx$$

An AUC score of 0.5 indicates random performance, the model is no better than a coin toss while a score of 1.0 represents perfect discrimination.

Table A1. Final features list

Variable Name
brent_oil_price_close
carbon_emissions_futures_price
carbon_emissions_futures_vol
electricity_day_ahead_price_ch
electricity_day_ahead_price_cz
electricity_day_ahead_price_be
electricity_day_ahead_price_nl
electricity_day_ahead_price_pl
electricity_day_ahead_price_at
electricity_day_ahead_price_de_lu
natural_gas_price_close
sunshine_duration
air_temperature
wind_speed
day
hour
month
is_workday
is_weekend
is_holiday
consumption_forecast_grid_load_mwh
consumption_forecast_residual_load_mwh
day_ahead_auktion
co2_emissions_certificate_auction_de
co2_emissions_certificate_auction_eu
production_forecast_total_mwh
production_forecast_photovoltaics_and_wid_mwh
production_forecast_wind_offshore_mwh
production_forecast_wind_onshore_mwh
production_forecast_photovoltaics_mwh
production_forecast_other_mwh
mean_sunshine_duration_bayern
mean_sunshine_duration_bremen
mean_sunshine_duration_hessen
mean_sunshine_duration_niedersachsen

Variable Name
mean_sunshine_duration_schleswig_holstein
mean_wind_velocity_bayern
min_wind_velocity_bayern
max_wind_velocity_bayern
mean_wind_velocity_bremen
min_wind_velocity_bremen
max_wind_velocity_bremen
mean_wind_velocity_hessen
min_wind_velocity_hessen
max_wind_velocity_hessen
mean_wind_velocity_niedersachsen
min_wind_velocity_niedersachsen
max_wind_velocity_niedersachsen
mean_wind_velocity_schleswig_holstein
min_wind_velocity_schleswig_holstein
max_wind_velocity_schleswig_holstein
brent_oil_price_spread
carbon_emissions_futures_spread
natural_gas_price_spread
co2_emissions_certificate_auction_spread
co2_emissions_certificate_auction_future-current
weekday
up_lag_1
up_lag_2
up_lag_3
up_lag_4
up_lag_5
up_lag_6
up_lag_24
up_lag_48
up_lag_168

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl
ISSN 2957-0506