Congestion Ahead: Learning to Forecast Redispatch Load in Renewable-Driven Grids

Sciences Po

April 2025

1 Introduction

In this report, we build upon the analysis conducted in Case A, which investigated the key drivers and determinants of redispatch events in the German electricity network, with a particular focus on the TenneT DE region. Case A combined econometric analysis and machine learning techniques to model redispatch occurrences, emphasizing direction-specific congestion. The study identified several critical predictors—such as lagged congestion, wind speed, day-ahead electricity prices, and regional weather patterns—as central to understanding the dynamics of redispatch.

Case B extends this analytical framework to address a more comprehensive and continuous forecasting challenge: predicting the *total redispatch load* across three horizons—day-ahead, week-ahead, and month-ahead—on an hourly basis. This task is of significant economic importance, especially in the context of increasing reliance on intermittent renewable energy sources and the associated risk of congestion, which is projected to rise sharply in the coming years.

The main objective of Case B is to develop robust forecasting models capable of accurately predicting the hourly total redispatch load using a broad set of explanatory variables. Leveraging the insights gained from Case A, we strategically focus on the most influential features, which allows us to improve model performance while remaining mindful of computational constraints. The resulting forecasts not only enhance operational understanding but also offer valuable insights for policymaking in energy systems management.

The structure of this report follows the standard empirical workflow: Section 2 provides a comprehensive description of the dataset along with initial exploratory analyses. Section 3 outlines the forecasting methodology and validation strategies. Section 4 presents empirical results and model performance across different forecast horizons. Section 5 discusses the implications of the findings in an economic and policy context. Finally, Section 6 concludes with key takeaways and potential directions for future research.

2 Data Description

Describe the dataset used, including sources, sample size, and time period. Explain key variables, transformations, and any data cleaning steps. Provide summary statistics and visualizations where relevant.

Data and forecasting model input choice

In order to understand the main trends at play, and to have a rapid and clear overview of the main determinants of total redispatch load, we proceed to an analysis of key summary statistics and use simple econometric tools to select the most important variables. First, we display the main summary statistics of the dataset. Because we want to understand how they interact and potentially drive the total redispatch load, we decided to present summary statistics separately for each quartile of total load. These are displayed in Table 1 below.

Variable	**Q1**	**Q2**	**Q3**	**Q4**
Air Temperature	10.25	9.77	10.27	8.66
Wind Speed	3.02	3.51	3.51	4.37
Wind Production (Offshore)	2334.13	2377.55	2383.43	2399.68
Wind Production (Onshore)	3808.27	3890.82	3898.00	3948.11
Sunshine Duration (Bayern)	12.40	14.03	14.35	12.28
Wind Velocity (Bayern)	2.63	2.81	2.70	3.25
Wind Velocity (Bremen)	3.67	4.55	4.73	6.16
Wind Velocity (Niedersachsen)	3.10	3.85	3.95	5.12
Wind Velocity (Schleswig-Holstein)	3.88	4.81	5.03	6.49
Brent Oil Price (Open)	64.12	72.10	76.79	76.16
Carbon Emissions (Futures price)	661.05	879.60	927.36	1062.80
Day-ahead Electricity Price (CZ)	121.49	135.18	143.09	129.34
Day-ahead Electricity Price (NL)	119.76	133.61	144.43	131.32
Day-ahead Electricity Price (PL)	102.93	110.44	118.37	111.89
Lag Total Load (1hr)	16.93	228.58	606.33	1593.89
Lag Total Load (2hr)	37.55	250.27	610.20	1547.77
Lag Total Load (3hr)	60.59	272.06	616.63	1496.58
Lag Total Load (12hr)	233.07	407.22	639.32	1166.83
Lag Total Load (1week)	420.08	516.20	640.91	872.59
Workday	0.80	0.71	0.63	0.71
Holiday	0.02	0.04	0.05	0.04

Table 1: Summary statistics for each quartile of total load.

We can observe some similarities with what we observed on redispatch dummy originally, in Case A. Interestingly, this uncovers the difference in a set of variables, among quartile of total load. This enlightens the role of electricity supply-related variables (such as wind velocity and day-ahead electricity price), as well as that of consumption-related seasonal patterns (workday for instance) and the persistence as well as seasonality of total load.

Cross-correlation analysis Then, in order to choose the most important variables in the most robust way, let us perform correlation analysis and Lasso regression selection. Lasso regression allows us to identify the most important variables by shrinking coefficients of less relevant predictors to zero. This is what we do, for different specifications, allowing us to isolate the most important variables, driving total load the most. Then, to enlighten that aspect, we present the correlation between total load and the key variables identified through Lasso regression. It is displayed in the Table 2, below.

Table 2: Correlation with Total Load

Variable	**Corr.**
Wind speed	0.38
Air temperature	-0.10
Wind production (offshore)	0.04
Wind production (onshore)	0.04
Sunshine duration (Bayern)	-0.02
Wind velocity (Bayern)	0.22
Wind velocity (Bremen)	0.41
Wind velocity (Niedersachsen)	0.42
Wind velocity (Schleswig Holstein)	0.45
Brent oil price (opening)	0.20
Carbon emission futures volume	0.15
lag_total_load_1hr	0.96
lag_total_load_2hr	0.91
lag_total_load_3hr	0.85
lag_total_load_1w	0.22
Workday	0.02
Holiday	-0.01

Auto-correlation and seasonality analysis Finally, let us observe our data and its patterns and seasonality. Indeed, redispatch displays strong seasonality, at different horizons, that we need to understand and take into account in our forecasting model. Figure 1 displays the auto-correlation of total load, over a month. If it appears clearly, as for redispatch overall, that the first hours lags are very important, we can also observe seasonality at the weekly level. Moreover, there appears to be clear signs of daily seasonality, with auto-correlation peaks around every 24 hours.

To have a clearer view of that last trend, we display the auto-correlation over a week. This allows to observe clear seasonality, with daily auto-correlation, although less pronounced than weekly seasonality.

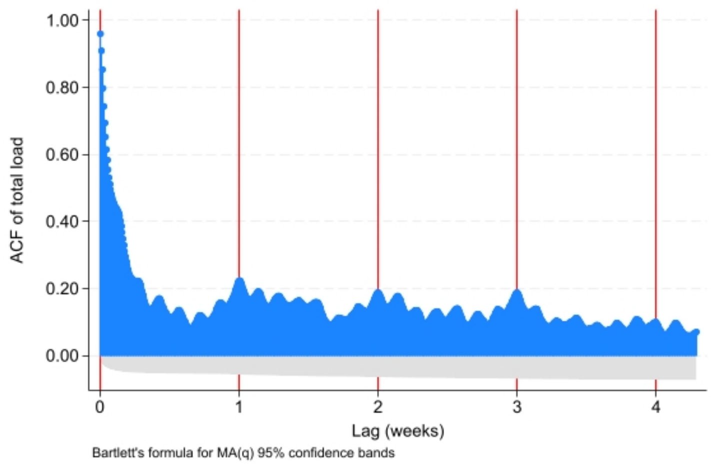

Figure 1: Auto-correlation function of total load (weekly)

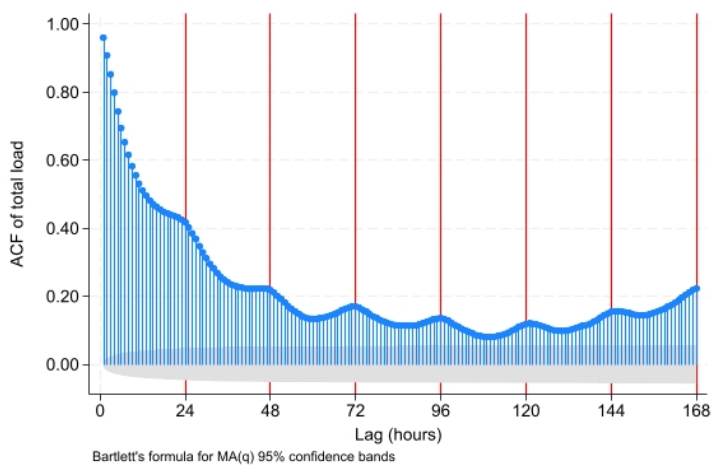

Figure 2: Auto-correlation function of total load (daily)

3 Methodology

The primary machine learning model employed for forecasting is the *LightGBM Regressor*, selected for its efficiency and predictive accuracy on structured time series data. For the sake of clarity, we present the methodology in the context of the 24-hour ahead forecast; the same procedure is applied for the 1-week and 1-month horizons.

As outlined in the previous section, the most informative predictors for hourly total load include lagged values of the load itself and wind speed, among other relevant features. However, when forecasting 24 hours ahead, we do not have access to observed lagged values for wind and load during the forecast window. To overcome this limitation, we employ a recursive forecasting strategy supported by a two-model architecture. Ideally, we would have more computing power and would have included one prediction model for every major features.

Two-Model Architecture:

- A **wind speed model** trained to predict future wind speeds using historical wind speed lags.
- A **load model** trained to forecast total redispatch load based on lagged load values, either lagged or predicted wind speed features, and other relevant variables that we identified above as being important determinants of total load.

Recursive Forecasting Strategy:

To obtain the forecast for load at $t + 24$, the system recursively simulates both wind speed and load from $t + 1$ through $t + 24$. For each forecast step $h \in \{1, \dots, 24\}$:

1. Wind speed at $t + h$ is predicted using the wind model, relying on previously observed or predicted wind lags.
2. The predicted wind speed is included in the feature set for the load model.
3. The total load at $t + h$ is predicted using the updated features.
4. The lag structures are updated to include the newly predicted values for both wind and load.

Moreover, the dataset includes periods characterized by exceptional external shocks—most notably the COVID-19 pandemic and the war in Ukraine. These episodes introduced atypical patterns in energy consumption, price volatility, and redispatch behavior. Consequently, our forecasting framework must account for the classical *bias-variance trade-off*. While training on such data may reduce bias by capturing complex temporal dynamics, it increases variance by overfitting to non-repeating anomalies. As a result, such models may perform suboptimally in more stable future conditions.

To partially mitigate this issue and adapt to evolving data patterns, we implement a rolling training window. For each forecast timestamp t, models are trained using only the two most recent months of data: $\text{train\_start} = t - 2\,\text{months}$, and $\text{train\_end} = t$. This rolling window ensures that each model is trained on the recent trends.

4 Empirical Results

Day-Ahead Predictions

Figure 3 illustrates the performance of the day-ahead ($t+24$) total load forecasting model across 2023. The blue curve represents the actual observed load, while the orange curve corresponds to the predicted values produced by the recursive forecasting framework.

Overall, the model demonstrates strong performance in capturing the general temporal structure and seasonality of the total load. Daily and weekly patterns are well reproduced, and the predicted series tracks the direction and magnitude of load fluctuations in most typical scenarios. This indicates that the rolling window training procedure and recursive prediction pipeline are effective in modeling baseline dynamics.

However, the model exhibits visible limitations in predicting extreme load values, particularly sharp upward spikes. To mitigate this issue, we introduced an additional feature — `load_delta` — designed to capture the recent dynamics of the load signal. Specifically, it measures the difference between short-term lagged values of the load (i.e., the rate of change), which can indicate rapid upward trends preceding a peak. The rationale was that this temporal gradient could provide the model with early warning signals of impending extremes.

While the inclusion of `load_delta` improved the model's ability to follow rising trends and moderately increased its responsiveness to load ramps, it was not sufficient to resolve completely the underestimation of the most pronounced peaks. This suggests that predicting such extremes may require more specialized modeling approaches tailored to rare event detection.

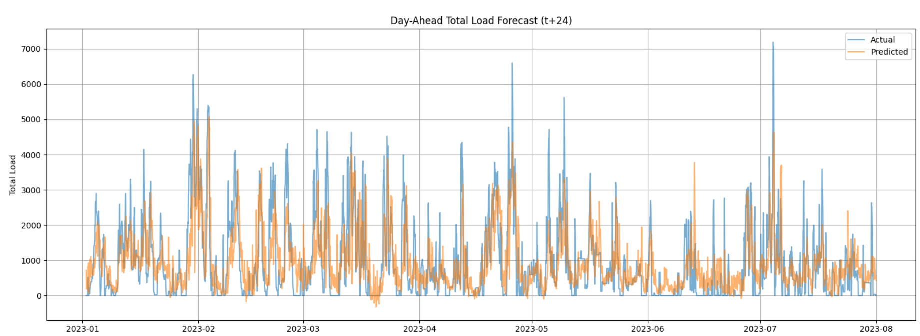

Figure 3: Day-Ahead Total Load Forecast (t+24)

Week-Ahead Predictions and the Effect of Recursive Error Accumulation

When performing long-horizon forecasting using recursive models, a well-known issue arises: *error compounding*. At each prediction step, the model uses its own previous predictions as input features, which introduces and accumulates inaccuracies over time. As the forecast horizon extends, this accumulation can lead to a significant degradation in predictive performance.

This phenomenon becomes particularly evident when forecasting 168 hours (one week) into the future. As shown in Figure 4, the model's accuracy notably deteriorates compared to the day-ahead setting.

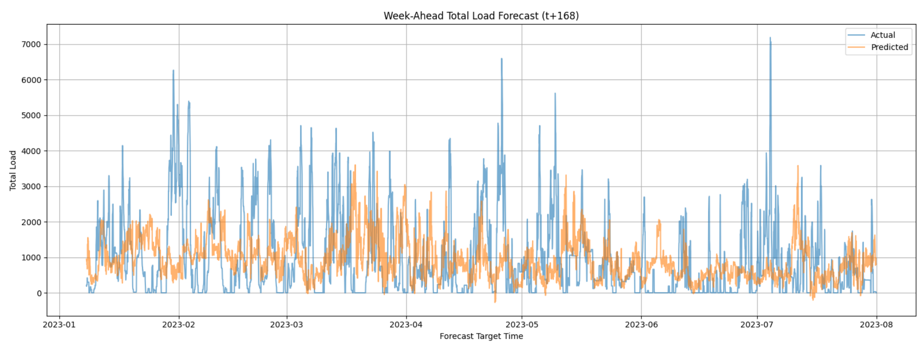

Figure 4: Week-Ahead Forecast (t+168): recursive model with predicted wind input

Our hypothesis was that this degradation was primarily due to the compounding of wind forecast errors over the week-long prediction horizon. While the wind model performs adequately in the day-ahead scenario, its recursive predictions over 168 steps become increasingly unreliable.

Month-Ahead Predictions

Due to computational constraints and limited time, we were not able to fully execute the recursive forecasting pipeline for the 1-month horizon. However, the noticeable degradation in predictive performance beyond the 1-week horizon already points to a key conclusion: without access to reliable long-term weather forecasts, accurate prediction of total load at a 1-month horizon is not feasible. This underscores the importance of high-quality meteorological inputs for extended forecasting windows.

5 Discussion and Economic Interpretation

Our findings have direct real-world implications, as they allow one to forecast and anticipate the total redispatch load at different horizons. It has direct policy implications, and may be useful both for policy-makers and market operators. Furthermore, in line with our results from case A, they propose a reliable and rich overview of the drivers of total redispatch load. As for our previous findings, they go in line with literature insisting on the role of wind power generation in leading to redispatch measures, as well as the link with electricity prices.

Then, they entail similar measures and come as a confirmation of the policy recommendations that we proposed in Case A, after analysing redispatch probability. They included the need for better coordination between consumers and suppliers, including incentives to smooth consumption in order to better match supply and avoid congestion, as well as the importance of densifying the transmission grid system, especially in frequently congested and renewable-intensive production areas.

We also called for a more granular data analysis, at the geographical scale, to better understand geographical imbalances and the way to circumvent them. This should, logically, lead to a better integration of geo-spatial inequalities and features, including in public commands and new industrial projects. In particular, and due to the importance of some geographical-specific variables in explaining total redispatch load (e.g. wind velocity in Bayern), one may realise the need to collect additional data and take more area-specific measures to that aim.

Large-scale policy recommendations

At a broader scale, our findings entail important policy changes and choices, and underline the need to enlarge the policy-making horizon. Indeed, climate change is going to affect the electricity market and redispatch load in multiple ways. First, electricity consumption is going to increase drastically in the coming years, particularly due to the electrification of public uses, and of cars in the first place. This is particularly true in Germany, which has taken strong measures in support of the electrification of vehicles, including professional ones (incentive programs, bonuses, etc.).

On the supply side, too, the coming years bring their share of uncertainty. Indeed, the increasing use of renewable energy, among which wind power generation occupies an important place, will lead to more production intermittence. Above all, and consequently, as we saw that strong and rapid winds are associated with more important total dispatch load, the congestion risk is going to increase, and associated redispatch expenses to aggravate. This is a direct consequence of the anticipated higher occurence of extreme wind events in the future, that is pointed to, along with its economic and welfare costs, by Bilal and Kanzig (2024).

Finally, there are broader implications, touching to the German as well as European industrial policy, and even, at last, on welfare and economic growth. Indeed, understanding the intensity and cost of load redispatch, policy-makers are faced with deeper decisions to take. In line with some of our earlier recommendations, it appears that the overall structure of both the electricity market and transmission grid, and the industrial German network are affected and should be designed adequately. Rethinking, and potentially redrawing, the map of either the electricity generation or the industrial production network would require wide policy implications and a vast, coordinated, effort from the German government. Indeed, it would imply changing both the rules and norms framing their installation and location decisions. It would also affect the way in which public markets are attributed, and large-scale plants are built and located. All of that cannot be designed and thought of without a deep dedication of the policy-makers.

Eventually, this would have broader spatially-distributed welfare implications, that need to be thoroughly analysed beforehand. Indeed, local labor markets would be affected by such decisions, as well as local economies and consumers. These implications cannot be measured here but additional data would be necessary to asses the side effects of the proposed policy changes.

6 Conclusion

We have managed to forecast accurately total redispatch load in the German TenneT DE network, at different horizons. Due to prediction variability and noise, our further ahead in time predictions, are less accurate than our day-ahead forecast. In line with our findings from the analysis of redispatch probability (case A), we found out that redispatch total load is very persistent over a few hours, and display seasonality at the daily as well as weekly basis. We also pointed to wind velocity as a main driver of load redispatch, in line with the related literature.
This led us to confirm our earlier policy recommendations and to propose larger-scale insights into needed action and discussion at the policy-making level.
Further research should definitely levy more precise and granular data, especially at the geo-spatial level. It would also be interesting to gain advantage of more recent data, as recent years have seen extreme events taking place, such as Covid, Russian invasion of Ukraine and subsequent inflation, especially on the energy market.
Finally, as discussed in the empirical section, it should also seize the opportunity to turn to real weather forecast model and data, in order to enrich predictive power and forecast accuracy, especially at longer horizons.

7 References

Bilal, Adrien and Känzig, Diego R. "The Macroeconomic Impact of Climate Change: Global vs. Local Temperature", *NBER Working Paper*, 2024.

Council of European Energy Regulators, *Redispatching Arrangements in Europe against the Background of the Clean Energy Package Requirements*, 2021.

European Commission and Joint Research Centre (Thomassen, G. and Fuhrmanek, A. and Cadenovic, R. and Pozo Camara, D. and Vitiello, S.), "Redispatch and congestion management – Future-proofing the European power market". *Publications Office of the European Union*, 2024.

Maurizio Titz and Sebastian Pütz and Dirk Witthau. "Identifying drivers and mitigators for congestion and redispatch in the German electric power system with explainable AI", *Applied Energy*, 2024.

K. Van den Bergh and D. Couckuyt and E. Delarue and W. D'haeseleer. "Redispatching in an interconnected electricity system with high renewables penetration", *Electric Power Systems Research*, 2015.