# A data-driven analysis of unscheduled flows in the European power system

Maurizio Titz[1,2] and Dirk Witthaut[1,2]

[1] Institute of Climate and Energy Systems: Energy Systems Engineering (ICE-1), Forschungszentrum Jülich, 52428 Jülich, Germany
[2] Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany

**Abstract.** Cross-border electricity trading is becoming increasingly important in the European power system. Unscheduled power flows induce additional costs and may lead to congestion and impair power grid operation. In this contribution we provide a data-centric analysis of unscheduled flows in the Central European power grid. Using methods from explainable machine learning, we identify the main driving factors for unscheduled flows and quantify their impact. Unscheduled flows in the meshed part of the grid can be attributed to transit or loop flows primarily and are well described by a linear model. The performance is substantially worse for unscheduled flows on bridges, with forecast errors being the most important drivers. This performance gap is probably due to data quality issues.

**Keywords:** Unscheduled Flows · Explainable AI · Transit Flows· Forecast Errors.

## 1 Introduction

Unscheduled power flows - physical flows that deviate from commercially scheduled exchanges - are a persistent feature of the European electricity system [2]. These flows result from the mismatch between market-based scheduling and the physical laws governing the flow of alternating current. While they are an expected side-effect of zonal market design, their magnitude and impact can be significant, increasing costs [27] and complicating system operation [38].

Despite ongoing efforts to align market outcomes with physical constraints, such as the reconfiguration of bidding zones [8] and the introduction of flow-based market coupling [34], unscheduled flows remain a problem.

In this article, we provide a data-centric analysis of unscheduled flows in the Central European synchronous power grid based on publicly available data. We apply methods from explainable machine learning (ML) to identify the main driving factors for unscheduled flows and to quantify their dependencies. In the meshed part of the grid, unscheduled flows can be mainly attributed to transit flows and are well predicted from scheduled commercial exchanges. A linear model captures these relationships accurately and enables a causal interpretation. For bridges, the most important features are generally given by forecast errors. However, model performance on bridges is substantially lower.

Our contribution is organized as follows. We review the fundamentals of unscheduled flows and the recent literature in Sec. 2. The data sources and ML methods are described in Sec. 3 and Sec. 4, respectively. The results of our analysis are presented and discussed in Sec. 5 and summarized in Sec. 6. We remark that we focus on AC connections in this contribution, while DC interconnectors have been treated in Ref. [31].

## 2 Background

### 2.1 Cross-border electricity trading in Europe

In liberalized electricity markets, the dispatch of power plants is mainly determined by electricity trading. Trading takes place on various time scales, from long-term contracts that are closed months in advance to intra-day contracts fixed minutes before delivery. Among these options, day-ahead spot markets are particularly relevant [18,41]. Day-ahead trading allows to adapt to the fluctuations of demand and renewable power generation via forecasts [36,16].

In Europe, day-ahead spot markets are auction-based with a trading interval of one hour. Generally, utility companies have to submit their bids and offers until 12:00 (noon) CET of the previous day. Market operators then set the market clearing price so that supply matches demand. [9,29]. All offers below the clearing price and all bids above the clearing price are executed exactly at clearing price, this is commonly referred to as "pay-as-cleared".

The Central European synchronous power system is divided into bidding zones that reflect local market conditions [12]. Most countries correspond to one bidding zone, while some countries, such as Italy, are divided into several regional zones. Some smaller countries such as Luxembourg or Andorra, are integrated into the bidding zones of neighboring country. Each bidding zone constitutes a separate market with its own market clearing price.

The Single Day-Ahead Coupling (SDAC) mechanism allows electricity to be traded across bidding zones [11]. Market coupling aims to increase market efficiency by pooling demand and supply while respecting transmission constraints between bidding zones [28]. The SDAC employs the *EUPHEMIA* algorithm that aims to optimize the dispatch in all bidding zones to maximize the social welfare [3]. EUPHEMIA uses the order books from the 12:00 day-ahead auction and the network constraints and calculates the market price for all bidding zones as well as the commercial electricity exchange between interconnected bidding zones.

Before the introduction of flow-based market coupling (FBMC) [34], available transfer capacities were allocated before market clearing, typically through a bilateral agreement of the involved transmission system operators. In contrast, FBMC uses a grid model to quantify the grid load and allocates transfer capacities during market clearing. FBMC was first introduced for a few countries in Central Western Europe in 2015 and extended to 13 other countries in June 2022 [30].

### 2.2 Unscheduled Flows

The physical power flow between two bidding zones does not necessarily match the commercial exchange scheduled on the day-ahead market. The difference between the physical flow and the scheduled commercial exchange is referred to as unscheduled flow. Transmission system operators aim to mitigate unscheduled flows as they can induce congestion or put the grid at risk [38]. Unscheduled flows can be caused by several different mechanisms (see Fig. 1 for a graphical overview):

– Transit flows: The power flows in an AC network are governed by Kirchhoff's and Ohm's laws and may therefore deviate from schedules. Consider for example a scheduled export of one Gigawatt from Germany to Austria. The physical power flow is not limited to the lines directly connecting Germany and Austria, but can also flow to Austria via Poland or the Czech Republic [35]. Transmission system operators try to reduce transit flows as they can induce congestion [19,38]. We note that the majority of academic articles focus on transit flows when analysing unscheduled flows. Sometimes the two terms are even used interchangeably.
– Forecast errors and intraday-trading: Day-ahead trading is based on forecasts, that are typically not perfect [39]. Forecasts are updated as the time of delivery approaches. A utility company may react via intra-day trading to correct for foreseen changes in demand and renewable power generation [42,6].
– Grid congestion and countertrading: The EUPHEMIA algorithm takes into account limited transmission capacities between bidding zones – but not within bidding zones. Hence, the day-ahead dispatch can induce congestion that must be resolved by the transmission system operators [40]. If congestion occurs close to a border, system operators can apply counter trading, buying electricity in one zone and selling it in another to mitigate congestion [25]. For example, a German grid operator can buy electricity in Germany and sell it in Denmark to reduce congestion-causing imports. [40].
– Frequency control reserves: Some DC interconnectors are partly used for control reserves, where the power flow is adjusted in real time and thus deviates from the day-ahead schedule [31]. Consider for example the BritNed interconnector linking the British and the Continental European synchronous area. If power consumption exceeds power generation in the British synchronous area, the grid frequency drops. To prevent this, the operator can increase imports via the BritNed interconnector. As DC interconnectors have been analyzed in detail in Ref. [31], we will focus on AC connections in this article.

### 2.3 Literature on unscheduled flows

Potential negative effects and hazards of unscheduled flows have been reviewed in Ref. [19]: They can lead to congestion and induce transmission bottlenecks and thus contribute to overloading of transmission lines which may eventually

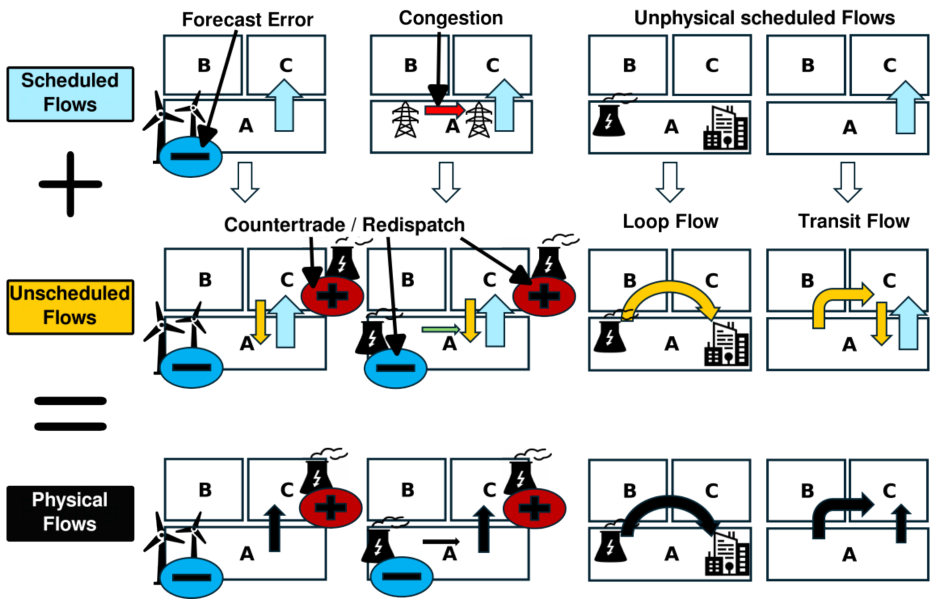


**Fig. 1.** Scheduled flows result from the dispatch determined by the EUPHEMIA algorithm, based on forecasts and market auction outcomes. Unscheduled flows quantify the deviation between scheduled and physical power flows. Forecast errors must be compensated through redispatch or countertrading. If this compensation occurs outside the bidding zone where the error originated, it induces unscheduled flows. Similarly, corrective actions taken by the grid operator to resolve congestion may lead to physical flows deviating from scheduled ones. Unscheduled flows can also occur when scheduled commercial exchanges conflict with the physical behavior of the grid. In the case of loop flows, flow scheduled between two locations within the same zone (e.g., A) may transit through external zones (e.g., B and C). In transit flows, electricity is scheduled from zone A to zone C but physically passes through an intermediate zone B. Since the direct flow between A and C is smaller than the scheduled exchange, this also results in unscheduled flows opposing the scheduled flow.

compromise grid reliability. Notably, unscheduled flows may have contributed to the 2003 North American blackout [7]. A particular example of an unscheduled power flow is the Lake Erie loop flow in the North American grid [5].

Several techniques to manage unscheduled transit flows are reviewed in [38]. Phase-shifting transformers (PSTs) enable the regulation of power flows in AC grids. Their optimized operation can therefore help manage unscheduled flows. [24,21]. Other measures to manage unscheduled flows by topology modifications are discussed in [14]. The estimation of transit flows in power system operation from measurements has been discussed in Refs. [37,32].

Unscheduled flows in the European grid have been analyzed in various articles. An economic analysis focusing on the role of renewable energy and infrastructure policy has been presented in [1]. The impacts of market splitting and network extensions has been investigated using simulation models in Ref. [22] and via statistical analysis of the German-Austrian bidding zone split in Ref. [13]. Studies of the impact of wind power generation can be found in [43,26]. A detailed statistical analysis of unscheduled flows on DC interconnectors and the relation to load-frequency control has been provided in Ref. [31]. An early overview of the integration of European electricity markets and the fundamentals of cross-border trade can be found in [4].

The situation in Central and Eastern Europe has received particular attention. Singh et al. introduced the basic problem and analyzed the impact on the Polish transmission grid using numerical simulations in [35]. Further simulations focusing on the German borders have been presented in [33]. The impact of phase-shifting transformers at the borders of Germany, Poland and the Czech Republic is discussed in [20].

## 3 Data

We use publicly available data from the ENTSO-E Transparency Platform [10]. All data is open access and includes day-ahead generation forecasts and actual generation, day-ahead and actual load, day-ahead scheduled commercial exchanges and physical power flows. For the latter, data is reported for each direction individually. Load data is resampled to hourly resolution using the mean, to match the resolution of the flow data.

We define directional net flows between two zones $i$ and $j$ as

$$\hat{f}_{ij} = f_{ij} - f_{ji},$$

where $f_{ij}$ is the physical flow from zone $i$ to $j$, and $f_{ji}$ the reverse. Net scheduled flows $\hat{s}_{ij}$ are computed analogously. From these we calculate the unscheduled flows as

$$\hat{u}_{ij} = \hat{s}_{ij} - \hat{f}_{ij}.$$

This approach thus captures directional net flow deviations but excludes internal loop flows between two zones.

For the load and the generation we compute the forecast errors as the difference between day-ahead forecasts and actual values

$$l_{i,err} = l_{i,\mathrm{day-ahead}} - l_{i,\mathrm{actual}}.$$

For instance, a load forecast error of $+1$GW means that the day-ahead prediction was 1GW larger than the actual load.

The data on the ENTSO-E Transparency Platform is not complete and requires pre-processing and filtering. For some bidding zones, individual features are not available and are thus discarded in the analysis. Many features miss individual data points, but for most this concerns less than 1% of data. We remove input features if more than 1.3% of the data is missing. We then remove all timestamps for which data is still missing for any input feature.

We excluded some outliers from the data, where flows are implausibly high. Furthermore for some connections we observe many occurrences of zero physical flows on AC links. We consider these values to be implausible under normal operating conditions and we assume this to result from reporting errors or outages. Therefore we excluded these data points from the analysis. While excluding these values has no significant impact on the results is raises questions about the overall data quality, which has been discussed in the past [17].

The European electricity system evolves slowly but continuously, through both physical developments and regulatory changes. Analyzing a long period — including major structural shifts — can obscure consistent patterns. For this reason, we focus on the period from the introduction of flow-based market coupling in June 2022 until the end of 2024.

## 4 Methods

### 4.1 Gradient Boosted Trees

We train gradient boosted decision trees (GBT) to model unscheduled flows. We selected Gradient Boosted Trees due to their strong performance on tabular, unstructured data[15] and because they offer efficient explainability with SHAP, as discussed below. They are also robust to multicollinearity and computationally efficient.

As input features we use scheduled flows, generation forecast errors, and load forecast errors for all European bidding zones, including those outside the Central European synchronous area. We train one separate model for each cross border connection.

We do not use a time-series split, since the model is used for explanation, not for forecasting. Instead we perform a group shuffle split with a block size of one week and a gap size of 24 hours to avoid data leakage from auto-correlated samples. If two consecutive hours have correlated feature and target values, one in the training set and one in the test set, the GBT model can simply learn the correlated data points “by heart”. This is prevented by leaving gaps between groups. Hyper-parameters are optimised using random search.

### 4.2 SHapley Additive exPlanations

SHapley Additive exPlanations (SHAP) [23] is a widely used method for interpreting black-box models by attributing their outputs to individual input features. It uniquely satisfies the properties of local accuracy, missingness, and consistency—criteria that are essential for generating reliable and meaningful feature attributions.

For each prediction, SHAP quantifies the impact of each input feature on the model output. If the model predicts the value $f$ from the feature values $x_1, \dots, x_n$ we have

$$f(x_1, \dots, x_n) = \phi_0(f) + \sum_{j=1}^{n} \phi_j(f; x_1, \dots, x_n), \tag{1}$$

where $\phi_j(f; x_1, \dots, x_n)$ denotes the SHAP values for the $j$th feature.

SHAP values thus give *local* explanations, i.e., explanations of individual predictions. To analyze these explanations across many samples, we use *dependency plots*, which show how the SHAP value of a specific feature $\phi_j(f; x_1, \dots, x_n)$ varies as a function of the feature value $x_j$, see figure 3. If we want to gain a more *global* understanding of the model, and by extension the data it is trained on, we can use SHAP values to quantify the importance of individual features over the whole dataset. The feature importance of the $j$th feature is obtained by aggregating its SHAP values over all samples $s$

$$FI_j = \frac{1}{\mathcal{N}} \sum_{s} |\phi_j(f; x_1^{(s)}, \dots, x_n^{(s)})|. \tag{2}$$

The sum of all feature importances is often normalized to one.

When the number of input features is large, individual feature importances tend to diminish, as relevance is distributed across many potentially correlated variables. To retain interpretability, we aggregate importances for the three groups of features: (i) scheduled flows, (ii) generation forecast errors, and (iii) load forecast errors— and normalize them. This results in three group-level importance values that offer a high-level view of the mechanisms driving unscheduled flows.

### 4.3 Linear Model

We furthermore fit linear regression models to analyze transit flows in the meshed part of the Continental European synchronous area. Focusing on transit flows, we only use the scheduled flows on AC links within the Central European synchronous area as input features. Linear models are inherently transparent and facilitate a causal interpretation. Hence, they complement the previous analysis in terms of GBTs.

When giving the model all scheduled flows as input features, coefficients for lines with high distance can turn out to be high. These effects can only be interpreted as (spurious) correlations, not causal relationships. To restrict the model to more realistic local interactions, we thus constrain the inputs to lines within a topological distance of two from the target line. We fix the intercept to zero. This slightly decreases model performance but forces the model to attribute all unscheduled flows to scheduled flows. This makes the model more intuitive and easier to interpret. The model can be represented as

$$\boldsymbol{u} = \boldsymbol{A}\boldsymbol{s}, \tag{3}$$

where $\boldsymbol{u}$ and $\boldsymbol{s}$ are vectors that summarize the unscheduled flows $\hat{u}_{ij}$ and the scheduled commercial exchanges $\hat{s}_{ij}$ for all border $(i,j)$. As described below, we fix $\boldsymbol{A}_{i,j} = 0$ if the shortest path distance from link $i$ to $j$ is larger than two. In the training and evaluation of the linear model we use a standard random train-test split.

## 5 Results

### 5.1 GBT model

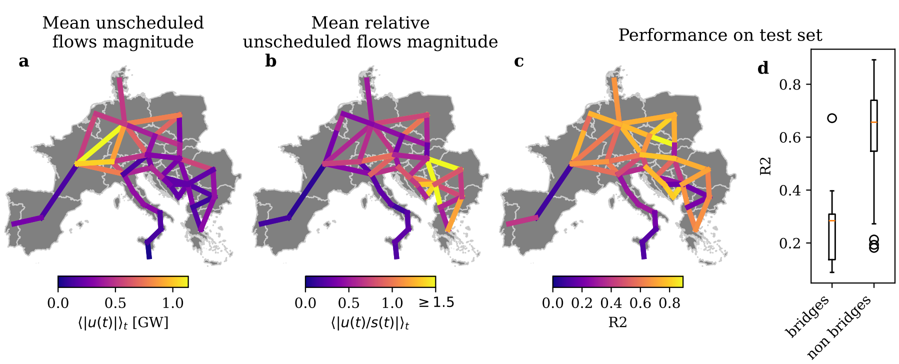


**Fig. 2.** Panel a (b) shows the mean (relative) magnitude of the unscheduled flows, where we divide the values shown in panel a by the mean magnitude of the scheduled flow to get the values in panel b. In panel b, we cropped the color scale at 1.5. The meshed part of western Europe shows the highest magnitudes of unscheduled flows, with DE-FR topping the list. However the relative values show, that this is in part due to the higher scheduled flows. Relatively speaking parts of the Balkan actually have the highest unscheduled flows, with multiple values larger than one. Panel c and d show the performance of the Gradient Boosted Tree (GBT) models. Apart from the DK1-DE link the performance on bridges is much worse than for non-bridges.

Unscheduled flows are largest in the meshed areas of central Europe, see Fig. 2. We compare the absolute value of the unscheduled flows $\langle|u_i(t)|\rangle_t$ and

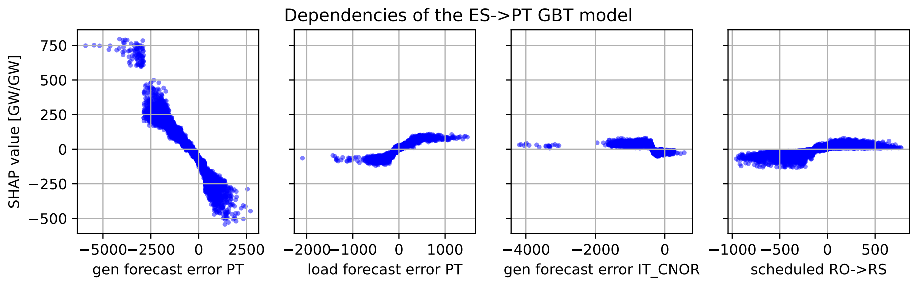


**Fig. 3.** Shap dependencies of the Gradient Boosted Trees model for the unscheduled flows on the line ES-PT, i.e. Spain to Portugal. By far the most important feature is the forecast error for renewable generation in Portugal. When there is a negative error, i.e. more generation than expected, the unscheduled flows are predicted to be higher, i.e. Portugal exports more or imports less than scheduled. Similarly, in the case of lower than expected generation, unscheduled flows to Spain are higher. The next most important feature, the generation forecast error in IT CSUD does already have a questionable relation to the cross border connection, which is most likely purely correlational.

the relative unscheduled flows, i.e. $u_i^{rel} = \langle |u_i(t)|/|s_i(t)| \rangle_t$, where $\langle \cdot \rangle_t$ denotes the average over all timestamps $t$. We observe that relative unscheduled flows are no higher in central Europe than in the eastern Europe and the Balkan area. Indeed, for seven links in the south-east the relative magnitude is larger than one, while it is below 0.6 for DE-FR. On bridges, relative unscheduled flows are generally lower than on non-bridges.

The GBT models perform much better in the meshed region than on most bridge connections.

In the following, we discuss feature importances and dependencies for three representative connections in detail. The connection between Spain and Portugal (ES-PT) is a bridge, i.e. there is no indirect connection which can be subject to transit or loop flows. We find that forecast errors are the most important features in the GBT model. The generation forecast error is the most important factor and its SHAP values are approximately linear. The second most important feature, the load forecast errors in PT is much less important, since they are much smaller than generation forecast errors. The next important features relate to forecast errors or scheduled flows outside the Iberian peninsula. We hypothesize that these features capture only correlations, not causal effects.

The connection between the Czech Republic and Slovakia (CZ-SK) lies in the meshed part of the grid. Here, we find that the scheduled flow on the respective connection is the dominant factor. SHAP values show a near-linear response with saturation at the extremes for the scheduled flow on the line itself. The negative slope indicates that unscheduled flows oppose scheduled flows. On the other hand, scheduled flows from CZ to other zones show a positive slope. This can be explained by transit flows as in the case of CZ-PL. For HU-SK and CZ-

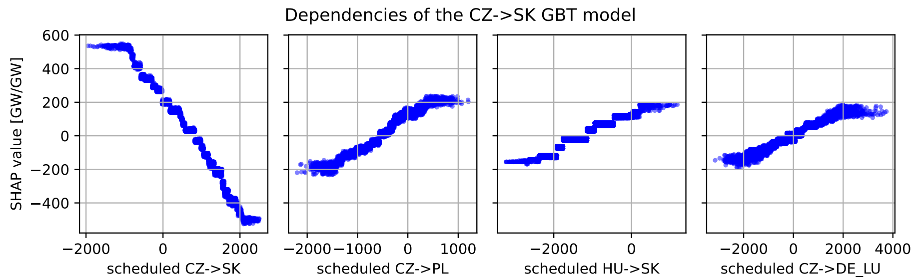


**Fig. 4.** The most important feature for the unscheduled flows on the CZ $\rightarrow$ SK connection is the corresponding scheduled flow. The relation is approximately linear except for extreme values, where the SHAP values saturate. The negative slope indicates that the unscheduled flows tend to oppose the scheduled flows. We also see that exports from CZ in other Zones, increase the unscheduled flows from CZ to SK.

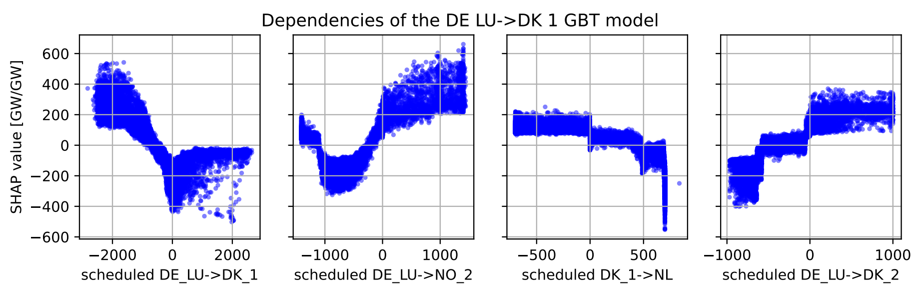


**Fig. 5.** The most important feature for the DE-DK unscheduled flows model is the corresponding scheduled flow. As with the CZ-SK model, we see the expected negative slope, though the dependencies show much stronger nonlinearities. Surprisingly, the next most important features are all scheduled flows on HVDC lines, instead of forecast errors.

DE LU the effect could be that these unscheduled flows are correlated because they are all driven by over-generation in CZ and under-generation in SK.

The bridge with the best performance is the DK-DE link. In contrast to other bridges, it is a bridge only when ignoring HVDC connections. The most important feature for the DK-DE connection is the corresponding scheduled flow. Here the model fits much stronger nonlinear effects than in other models. Surprisingly, the model does not strongly rely on forecast errors as would be expected for bridges. Instead, the next three most important features are all scheduled flows on HVDC lines. Unlike AC connections in the meshed grid, these do not have any direct physical effect on other power flows, since DC power transmission is fully controllable. Accordingly, unscheduled flows can only be affected through market effects but trying to explain these goes beyond the scope of this paper.

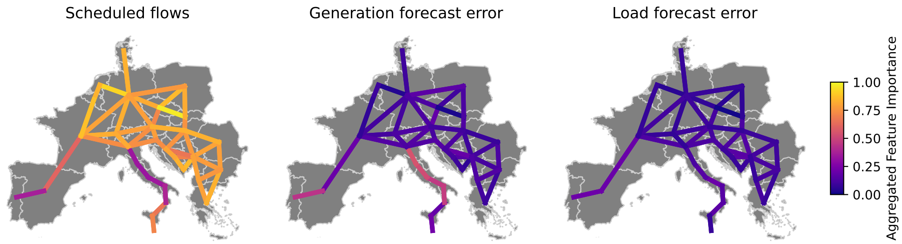


**Fig. 6.** Aggregated feature importances for the three types of input features. Each line corresponds to the GBT model for the unscheduled flows on that line. The feature importances for each model are normalized to sum to one. The importance of all features of each type are then summed to give the aggregated importance. Models for bridges stand out by utilizing forecast errors on a similar level as scheduled flows. For non bridges, they are barely used at all, and the predictions are heavily based on the scheduled flows instead. The DK1-DE LU link is a stark outlier. Load forecast errors are much less important than generation forecast errors.

To conclude, we collect cumulative feature importances results for all connections in the Central European synchronous area. Figure 6 highlights a clear separation between bridges and non-bridges. In meshed areas, predictions rely almost entirely on scheduled flows. On bridges, forecast errors become more important, though scheduled flows still dominate in some cases. The DK1–DE LU connection is an outlier in both importance and performance, as examined above.

Due to the simplicity of the relations for bridges and especially terminal links – unscheduled flows can only be caused from downstream forecast errors – one might expect the best performance here. This hints at problems with the data quality and availability. For instance, missing balancing actions in the generation data would make it impossible to capture these effects.

For all non-bridge and some bridge links the most important features are the scheduled flows. In many cases, the dependencies are close to linear, suggesting that linear models should be able to capture the relations similarly well.

### 5.2 Linear Model

For a detailed analysis of transit flows, we train a linear model to predict unscheduled flows using only scheduled flows as input. The intercept is fixed to zero, such that all unscheduled flows must be derived from scheduled flows. To avoid fitting spurious long-range correlations, we limit inputs to links within a topological distance of two from the target link, see Sec. 4.

The performance is in line with the GBT model (Fig. 7). $R^2$ scores are high for meshed connections and low for bridges. This supports the conclusion that in the meshed grid, unscheduled flows are largely driven by systematic mismatches

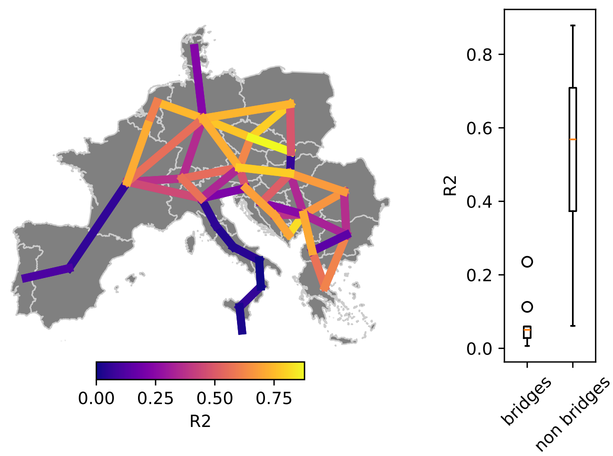


**Fig. 7.** Performance of the linear model. As expected, we see a big difference between bridges and non-bridges. The R2 score for bridges is below 0.1 for most bridges. For non-bridges, the R2 score is almost 0.6 on average. This matches with the GBT performances and feature importances.

between scheduled and physical flows (transit flows), while on bridges they reflect re-balancing effects driven by local forecast errors or other reasons.

We now analyze the effect of individual scheduled flows on the resulting unscheduled flows. This analysis is straightforward using the linear model (3) as the total unscheduled flows pattern decomposes into the sum of the effects from each individual line. These effects are captured in the columns of the learned coefficient matrix $\mathbf{A}$, where the $l$th column represents the impact of scheduled flow on line $l$ across the entire system.

Figure 8 shows the case of scheduled exports from Germany to Austria. The model predicts that a significant portion of this flow —more than 400 MW— will not be directly realized, but instead be rerouted as unscheduled flows, primarily via neighboring countries such as Poland and the Czech Republic. This illustrates how scheduled flows can induce large-scale loop or transit flows even under simple linear assumptions.

Notably, the Czech Republic and Poland have repeatedly experienced congestion due to such transit flows [35]. As a consequence, TSOs have commissioned phase shifting transformers along the borders to Germany to regulate and constrain flows from Germany [20].

## 6 Discussion

To our knowledge, this is the first data-driven study to quantify the drivers of unscheduled flows across all cross-border links in Central Europe using explainable ML techniques. Our analysis demonstrates that unscheduled flows in

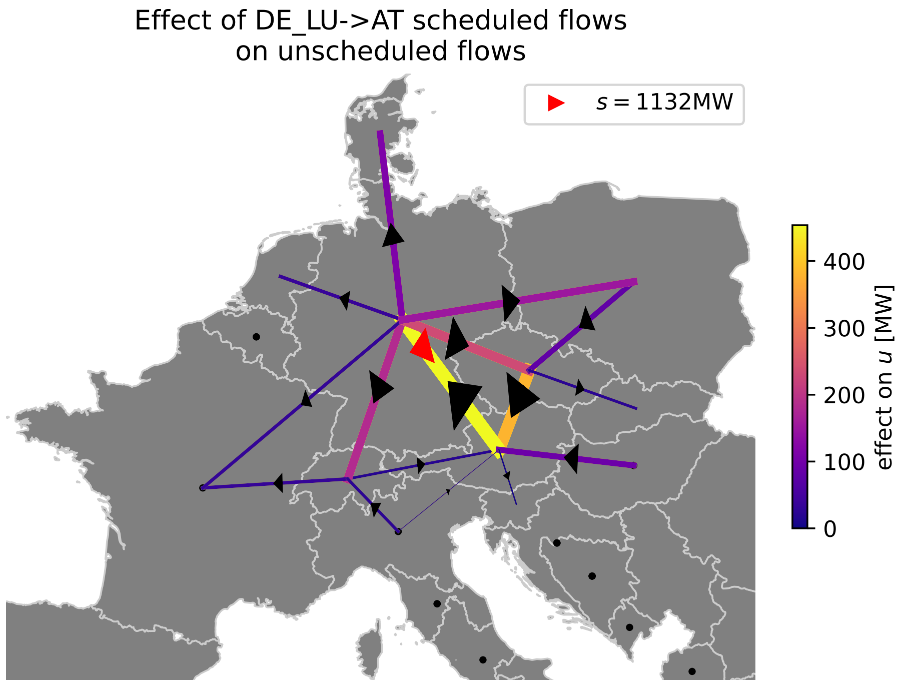


**Fig. 8.** Effect of scheduled flows from Germany to Austria on unscheduled flows as learned by the linear model. The scheduled flow of 1132MW is the mean flow at times when Germany is exporting to Austria. The model learned that unscheduled flows of over 400MW from AT to DE will result from the scheduled flows. In other words, more than one third of the scheduled exports will not take place or take another route. The latter can be seen as transit flows especially via Poland and the Czech Republic.

the meshed part of the Central European power system can be explained to a large degree using only scheduled commercial exchanges. Gradient Boosted Tree (GBT) models perform well on meshed connections, while performance on bridges is consistently worse. SHAP analysis reveals that scheduled flows are by far the most important features across most models, and the relationship between scheduled and unscheduled flows is often close to linear. Forecast errors play a significant role only for some bridge connections in our models. Linear regression models using only scheduled flows as input confirm the dominant role of these features. They perform comparably well in meshed areas, but fail for bridge connections. These findings support the interpretation that unscheduled flows are still primarily caused by a mismatch between market-based scheduling and the physical realities of the power grid. In the meshed core of the grid, systematic transit flows and loop flows emerge due to the simplified network model used in market clearing, particularly in EUPHEMIA. For bridges, generation forecast errors are important drivers of unscheduled flows, while load forecast errors are generally smaller and thus less relevant. Surprisingly, however, even in terminal links, forecast errors do not allow reliable prediction of unscheduled flows at all. This is unexpected, as forecast errors should, in principle, hold all the information about net deviations from schedule. They should thus fully explain unscheduled flows on bridges. The low model performance therefore points to issues with data quality, such as misreported data, or missing balancing actions not reflected in the published data. Forecast errors are likely relevant in meshed areas as well. However, their effects are harder to capture, since the resulting imbalances can be compensated across multiple links. But once again, data quality might also be the problem.

The linear model provides further insight by making the relations indicated by the SHAP analysis explicit. By visualizing the coefficients, we can interpret the model as assigning each scheduled flow a contribution to unscheduled flows elsewhere in the network. In most cases, the resulting patterns correspond well to expected transit flows and confirm systematic deviations from schedules. These effects likely arise from the physical redistribution of power flows that is not fully captured by the simplified market model used in scheduling.

Our analysis is limited by the quality of publicly available data. Beyond missing values we found implausible zero flows. We might very well have been unable to identify further misreported data. Moreover, our models rely entirely on statistical relationships and do not incorporate physical constraints such as Kirchhoff's laws. All interpretations are based on correlation, not physical causation.

Future work could focus on integrating physical modeling or constraints with machine learning to improve interpretability and accuracy. However, the lack of public data on the transmission grid remains a major obstacle. Another direction is to analyze temporal changes in flow patterns in response to regulatory events such as bidding zone redefinitions, flow-based market coupling expansions, or grid extensions. Our preliminary investigations in this direction have shown little measurable effect.

## 7 Appendix

**Table 1.** Performance ($R^2$ score) of the GBT model (first value) and linear model for unscheduled flows. Bridges link are marked by *. In some cases the linear model outperforms the GBT model, because of the different splits used, see method section.

| | | | | | |
|---|---|---|---|---|---|
| CZ-SK | 0.89 \| 0.88 | AT-CZ | 0.67 \| 0.62 | ES-PT* | 0.40 \| 0.11 |
| ME-RS | 0.83 \| 0.82 | CH-FR | 0.67 \| 0.37 | IT NORD-SI | 0.36 \| 0.22 |
| CZ-PL | 0.81 \| 0.79 | DE LU-DK 1* | 0.67 \| 0.24 | HU-SK | 0.31 \| 0.06 |
| BA-ME | 0.80 \| 0.70 | BG-GR | 0.66 \| 0.61 | IT CSUD-IT SUD* | 0.31 \| 0.04 |
| AT-HU | 0.78 \| 0.79 | GR-MK | 0.63 \| 0.57 | IT CNOR-IT CSUD* | 0.30 \| 0.03 |
| BA-HR | 0.77 \| 0.69 | BG-RO | 0.62 \| 0.36 | IT CNOR-IT NORD* | 0.28 \| 0.05 |
| CZ-DE LU | 0.76 \| 0.77 | CH-IT NORD | 0.60 \| 0.56 | HR-RS | 0.27 \| 0.24 |
| AT-DE LU | 0.75 \| 0.72 | DE LU-FR | 0.59 \| 0.56 | BG-MK | 0.21 \| 0.15 |
| HR-SI | 0.75 \| 0.68 | AT-CH | 0.58 \| 0.55 | IT CALA-IT SUD* | 0.20 \| 0.01 |
| DE LU-NL | 0.74 \| 0.74 | FR-IT NORD | 0.56 \| 0.45 | BG-RS | 0.19 \| 0.33 |
| BE-FR | 0.74 \| 0.72 | BE-NL | 0.56 \| 0.60 | BA-RS | 0.18 \| 0.38 |
| DE LU-PL | 0.73 \| 0.74 | HR-HU | 0.55 \| 0.51 | IT CALA-IT SICI* | 0.14 \| 0.06 |
| HU-RO | 0.73 \| 0.67 | CH-DE LU | 0.55 \| 0.36 | IT SICI-MT* | 0.12 \| 0.02 |
| PL-SK | 0.73 \| 0.49 | AT-IT NORD | 0.47 \| 0.34 | ES-FR* | 0.09 \| 0.06 |
| RO-RS | 0.70 \| 0.66 | AT-SI | 0.46 \| 0.48 | | |
| MK-RS | 0.67 \| 0.71 | HU-RS | 0.41 \| 0.35 | | |

## References


1. Abrell, J., Rausch, S.: Cross-country electricity trade, renewable energy and european transmission infrastructure policy. Journal of Environmental Economics and Management **79**, 87–113 (2016)
2. ACER: Methodological paper: Unscheduled flows (2018), `https://www.acer.europa.eu/sites/default/files/documents/en/Electricity/Market%20monitoring/Documents_Public/ACER%20Methodological%20paper%20-%20Unscheduled%20flows.pdf`
3. ALL NEMO COMMITTEE: Euphemia Public Describtion. `https://www.nemo-committee.eu/assets/files/euphemia-public-description.pdf` (2020)
4. Bower, J.: Seeking the single european electricity market: evidence from an empirical analysis of wholesale market prices (2002)
5. Coletta, T., Delabays, R., Adagideli, I., Jacquod, P.: Topologically protected loop flows in high voltage ac power grids. New Journal of Physics **18**(10), 103042 (2016)
6. Cramer, E., Witthaut, D., Mitsos, A., Dahmen, M.: Multivariate probabilistic forecasting of intraday electricity prices using normalizing flows. Applied Energy **346**, 121370 (2023)
7. Dozier, A.Q., Suryanarayanan, S., Liberatore, J.P., Veghte, M.C.: Unscheduled flow in deregulated electricity markets: Bridging the gap between the western electric power industry and academia. In: 2013 IEEE Green Technologies Conference (GreenTech). pp. 451–458. IEEE (2013)
8. of Energy Traders, E.F.: Bidding zones delineation in Europe: Lessons from the past & recommendations for the future (Sep 2019), `https://eepublicdownloads.entsoe.eu/clean-documents/Network%20codes%20documents/Implementation/stakeholder_committees/MESC/2019-09-17/5.2_EFET%20position%20paper_%20BZ%20review_16092019.pdf?Web=1`

9. EPEX SPOT: Basics of the Power Market. https://www.epexspot.com/en/basicspowermarket, accessed on 2024-01-07
10. European Network of Transmission System Operators for Electricity: ENTSO-E Transparency Platform. https://transparency.entsoe.eu/ (2023)
11. European Network of Transmission System Operators for Electricity (ENTSO-E): Single Day-ahead Coupling (SDAC). https://www.entsoe.eu/network_codes/cacm/implementation/sdac/, accessed on 2024-01-07
12. European Union Agency for the Cooperation of Energy Regulators (ACER): Bidding Zone Review. https://www.acer.europa.eu/electricity/market-rules/capacity-allocation-and-congestion-management/bidding-zone-review, accessed on 2024-01-07
13. Graefe, T.: The effect of the austrian-german bidding zone split on unplanned cross-border flows. arXiv preprint arXiv:2303.14182 (2023)
14. Granelli, G., Montagna, M., Zanellini, F., Bresesti, P., Vailati, R.: A genetic algorithm-based procedure to optimize system topology against parallel flows. IEEE Transactions on Power Systems **21**(1), 333–340 (2006)
15. Grinsztajn, L., Oyallon, E., Varoquaux, G.: Why do tree-based models still outperform deep learning on typical tabular data? Advances in neural information processing systems **35**, 507–520 (2022)
16. Han, C., Hilger, H., Mix, E., Böttcher, P.C., Reyers, M., Beck, C., Witthaut, D., Rydin Gorjão, L.: Complexity and persistence of price time series of the european electricity spot market. PRX Energy **1**(1), 013002 (2022)
17. Hirth, L., Mühlenpfordt, J., Bulkeley, M.: The ENTSO-E Transparency Platform – A review of Europe's most ambitious electricity data platform. Applied Energy **225**, 1054–1067 (Sep 2018). https://doi.org/10.1016/j.apenergy.2018.04.048, https://www.sciencedirect.com/science/article/pii/S0306261918306068
18. Huisman, R., Huurman, C., Mahieu, R.: Hourly electricity prices in day-ahead markets. Energy Economics **29**(2), 240–248 (2007)
19. Kavicky, J., Shahidehpour, S.: Parallel path aspects of transmission modeling. IEEE Transactions on Power Systems **11**(3), 1180–1190 (2002)
20. Korab, R., Owczarek, R.: Impact of phase shifting transformers on cross-border power flows in the central and eastern europe region. Bulletin of the Polish Academy of Sciences. Technical Sciences **64**(1), 127–133 (2016)
21. Korab, R., Połomski, M., Owczarek, R.: Application of particle swarm optimization for optimal setting of phase shifting transformers to minimize unscheduled active power flows. Applied Soft Computing **105**, 107243 (2021)
22. Kunz, F.: Quo vadis?(un) scheduled electricity flows under market splitting and network extension in central europe. Energy Policy **116**, 198–209 (2018)
23. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. Advances in neural information processing systems **30** (2017)
24. Marinakis, A., Glavic, M., Van Cutsem, T.: Minimal reduction of unscheduled flows for security restoration: Application to phase shifter control. IEEE Transactions on Power Systems **25**(1), 506–515 (2009)
25. Meeus, L.: The evolution of electricity markets in Europe. Edward Elgar Publishing, Cheltenham (2020)
26. Mohanpurkar, M., Suryanarayanan, S.: A case study on the effects of predicted wind farm power outputs on unscheduled flows in transmission networks. In: 2013 IEEE Green Technologies Conference (GreenTech). pp. 277–284. IEEE (2013)
27. Newbery, D., Strbac, G., Viehoff, I.: The benefits of integrating european electricity markets. Energy Policy **94**, 253–263 (2016)

28. Nominated Electricity Market Operators Committee: Euphemia public description. `https://www.nemo-committee.eu/assets/files/euphemia-public-description.pdf`, accessed on 2024-01-29
29. Nord Pool: Day-ahead market. `https://www.nordpoolgroup.com/en/the-power-market/Day-ahead-market/`, accessed on 2024-01-07
30. Ovaere, M., Kenis, M., Van den Bergh, K., Bruninx, K., Delarue, E.: The effect of flow-based market coupling on cross-border exchange volumes and price convergence in central western european electricity markets. Energy Economics **118**, 106519 (2023)
31. Pütz, S., Schäfer, B., Witthaut, D., Kruse, J.: Revealing interactions between hvdc cross-area flows and frequency stability with explainable ai. Energy Informatics **5**(Suppl 4), 46 (2022)
32. Ronellenfitsch, H., Timme, M., Witthaut, D.: A dual method for computing power transfer distribution factors. IEEE Transactions on Power Systems **32**(2), 1007–1015 (2016)
33. Schneider, M., Barrios, H., Schnettler, A.: Evaluation of unscheduled power flows in the european transmission system. In: 2018 IEEE International Energy Conference (ENERGYCON). pp. 1–6. IEEE (2018)
34. Schönheit, D., Kenis, M., Lorenz, L., Möst, D., Delarue, E., Bruninx, K.: Toward a fundamental understanding of flow-based market coupling for cross-border electricity trading. Advances in Applied Energy **2**, 100027 (2021)
35. Singh, A., Frei, T., Chokani, N., Abhari, R.S.: Impact of unplanned power flows in interconnected transmission systems–case study of central eastern european region. Energy Policy **91**, 287–303 (2016)
36. Staffell, I., Pfenninger, S.: The increasing impact of weather on electricity supply and demand. Energy **145**, 65–78 (2018)
37. Suryanarayanan, S., Farmer, R., Heydt, G., Chakka, S.: Estimation of unscheduled flows and contribution factors based on l/sub p/norms. IEEE Transactions on Power Systems **19**(2), 1245–1246 (2004)
38. Suryanarayanan, S.: Techniques for accommodating unscheduled flows in electricity networks and markets. In: 2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century. pp. 1–6. IEEE (2008)
39. Sweeney, C., Bessa, R.J., Browell, J., Pinson, P.: The future of forecasting for renewable energy. Wiley Interdisciplinary Reviews: Energy and Environment **9**(2), e365 (2020)
40. Titz, M., Pütz, S., Witthaut, D.: Identifying drivers and mitigators for congestion and redispatch in the german electric power system with explainable ai. Applied Energy **356**, 122351 (2024)
41. Wolff, G., Feuerriegel, S.: Short-term dynamics of day-ahead and intraday electricity prices. International Journal of Energy Sector Management **11**, 557–573 (2017)
42. Ziel, F.: Modeling the impact of wind and solar power forecasting errors on intraday electricity prices. In: 2017 14th international conference on the European energy market (EEM). pp. 1–5. IEEE (2017)
43. Zugno, M., Pinson, P., Madsen, H.: Impact of wind power generation on european cross-border power flows. IEEE Transactions on Power Systems **28**(4), 3566–3575 (2013)