Investigating Anomalies in Bidding Curves on the Electricity Spot Market: A Machine Learning Approach.

[Jilles De Blauwe, Technical University of Denmark (DTU), +32 471 80 13 21, jcdbl@dtu.dk]
[Dogan Keles, Technical University of Denmark (DTU), +45 93 51 09 06, dogke@dtu.dk]

Overview

Since the liberalization of the electricity sector, electricity has become a commodity freely traded on several markets. A large share of the short-term trading takes place in the Day-Ahead Market (DAM). This DAM, also called the spot market, is typically the reference point for the intraday and balancing power market prices. On the DAM, bids to supply or consume electricity are placed for each time interval, usually one hour, of the next day. It is the intersection of the aggregated supply and demand curve for that hour that determines the Market Clearing Prices (MCP) of the specific hour.

Generally, the supply curve is considered rather elastic, resulting from variable Renewable Energy Sources (vRES) in the market, whereas the demand is mostly inelastic. Therefore, the placement of the supply bids is most significant for the MCP. Now, we have traditionally assumed the generators place supply bids according to the marginal cost (MC) of their production units. However, generators might deviate from bidding at MC and instead explore alternative bidding strategies (S. Li & Park, 2018) if these can potentially increase revenues, which has changed spot market dynamics (European Commission: Directorate-General for Energy et al., 2019). Such cases where the bidding curve deviates strongly from the expected bidding can be termed as "*bidding anomalies.*" However, these cases of bidding anomalies should of course be monitored by the regulator agencies, as we should verify that such bidding anomalies did not result from illegal actions, such as market manipulation or the abuse of market power. However, monitoring anomalies has become increasingly complex. First, the expected bidding behaviour has changed, following the high level of vRES penetration in our system, which has influenced spot market dynamics (Paschen, 2016). Second, the manual methods of monitoring spot market behaviour might not be up to date with the high complexity and numerous actors on today's spot markets (Tiwari et al., 2022), which calls for the development of more advanced methods to study spot market bidding. Fortunately, automated methods, such as Machine Learning (ML) have become available, which are effective at analyzing large quantities of raw data.

Thus, this study has developed new approaches to investigate and understand the bidding behaviour on the spot market using ML. Further, building from these techniques, we have also developed new methods to detect and analyze anomalies in spot market bidding data. We then apply our methods to several years of bidding data from the Belgian, German, and Danish spot markets, performing an in-depth case study to determine the expected versus anomalous bidding behaviour. Furthermore, the anomalies we here detect in the bidding patterns are discussed, as we explore how a variety of factors, such as strategic bidding, might have caused these anomalies.

Methods

For this analysis, we first need to determine what the expected bidding behaviour is. Therefore, we investigate the relation between the hourly bids placed in the market and several key market features that drive bidding, such as the expected vRES Generation, Expected Load, and Temperature. Therefore, we apply clustering to the bidding curves, based on the method of (Z. Li et al., 2024), and relate the formed clusters to the exogenous data. This clustering is based on the inter-curve distance between hourly supply curves. We then apply a hierarchical clustering algorithm on generated pairwise distances, creating a number of distinct clusters with curves of a similar shape. Next, we determine the mean values for the market drivers per cluster, which indicates the relation between these drivers and the expected bidding curves.

In the second part of our study, we use the generated clusters to determine anomalies where the bidding deviated from the expected bidding. We have developed four different anomaly detection methods:

1. First, we identify hours with bidding curves that deviated significantly from any other bidding curves in our data, termed the *"shape"* anomalies.
2. Second, we determine hours with exogenous driver data that significantly deviated from the mean values for other hours in the same cluster, termed "*exogenous driver"* anomalies.
3. Last, we fit two separate prediction models to our set of driver data, and the generated clusters. On one hand, we use a Multi-Nomial logit (MNL) prediction model, and second, a Random-Forest classification model, which are then both used to predict probabilities of the hours belonging to each cluster. Now, the hours with

high residual values for the predicted cluster as compared to the actual cluster, thus the hours that were hardest to predict, form the “*MNL”* and “*RF”* anomalies for both models, respectively.

Then, after the anomalies are generated, we follow a set approach to examine under which circumstances the anomalies occur, and what causes these might have. First, we explore for differences in driver data for the anomalies as compared to the data for other hours. Then, we explore whether additional features, such as Generation Unavailability, Transmission Capacity, and Cross-Border drivers, might be used to understand some of the anomalies. Last, for the remaining anomalies, we explore where factors such as strategic bidding, or market power, might have been the cause.

Results

Our methods were then applied to bidding data from years 2020-2023, for the Belgian, German, and Danish bidding zones. However, in explaining the anomalies, we focus specifically on the German Market, for the year 2022. The German market is the largest out of those considered, and both highly competitive and liquid. However, the market is quite tight in hours with low production from vRES, meaning the available peak generation is only limited compared to the total demand, making it suitable for strategic bidding being applied.

Our case study resulted in the following preliminary findings:

- Generally, clustering into three clusters seems to result in clearly defined clusters. These correspond to distinct market states: High vRES, Generation Surplus, Low MCP – Low vRES, No Generation Surplus, and High MCP, and an intermediary case, with moderate prices. This is displayed in the Figure 1, with corresponding driver data means in Table 1.

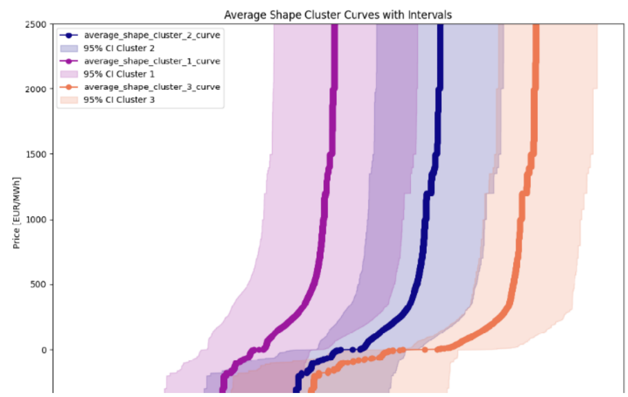

Figure 1: Average curves for three clusters of bidding curves based on shape, for the German bidding zone, 2022

	Temperature [°C]	Load [MW]	Solar [MW]	Wind [MW]	Generation [MW]	Price [EUR/ MWh]
Cluster_1	8.857	54991.7	3095.677	12099.55	53541.23	264.512
Cluster_2	14.391	60204.72	14758.51	17713.3	63808.03	181.385
Cluster_3	13.092	56811.93	14322.06	29259.92	66953.89	44.29

Table 1: Mean driver values for clustering the German bidding data into three clusters based on curve shape, 2022

- As zones have increasingly high shares of vRES capacity installed, the placement of supply bids are also fully decided by the RES forecast.
- The *shape anomalies,* or extreme outliers in the bidding curve shape, result mostly from a number of bidding hours where very high vRES generation leads to an especially flat curve, with many near zero price bids.
- We find a number of *MNL* and *RF anomalies* can be explained by high exports for the zone, where the excess of vRES generation within the zone is fully exported, as generators are attracted by cross-border price differences.
- For detecting the use of bidding strategies from a single market is highly complex, and detection thereof might thus require the using data from other markets (balancing or intraday), or identification of the placed bids and generation plants.

Conclusions & Policy Recommendations

Based on our preliminary results, we observe that in today’s markets with high renewable penetration, vRES are the main driver of spot market bidding, where the strength of the relation increases with more installed renewable capacity. Second, the potential to earn cross-border price differences drives the export of generation excesses. Thus, the resulting MCP at different times of equal excess renewable generation still varies, depending these exports. Last, automated methods can prove highly valuable as a first step in detecting anomalous bidding behaviour on the spot market. However, relating anomalies to specific bidding strategies or manipulative behaviour being applied, still requires additional analysis, likely requiring more explicit data about the actors making the bids.

References

European Commission: Directorate-General for Energy, Moser, A., Bracht, N., & Maaz, A. (2019). *Simulating electricity market bidding and price caps in the European power markets – S18 report*. Publications Office. https://doi.org/10.2833/252345

Li, S., & Park, C. S. (2018). Wind power bidding strategy in the short-term electricity market. *Energy Economics*, *75*, 336–344. https://doi.org/10.1016/j.eneco.2018.08.029

Li, Z., Alonso, A. M., Elías, A., & Morales, J. M. (2024). Clustering and forecasting of day-ahead electricity supply curves using a market-based distance. *International Journal of Electrical Power & Energy Systems*, *158*, 109977. https://doi.org/10.1016/j.ijepes.2024.109977

Paschen, M. (2016). Dynamic analysis of the German day-ahead electricity spot market. *Energy Economics*, *59*, 118–128. https://doi.org/10.1016/j.eneco.2016.07.019

Tiwari, S., Bell, G., Langseth, H., & Ramampiaro, H. (2022). Detection of Potential Manipulations in Electricity Market using Machine Learning Approaches: *Proceedings of the 14th International Conference on Agents and Artificial Intelligence*, 975–983. https://doi.org/10.5220/0010991800003116