Synthesis of artificial bid sets for day-ahead power exchange models

1st Dávid Csercsik
Institute of Economics, Centre for Economic and Regional Studies
Budapest, Hungary and
Pázmány Péter Catholic University
Budapest, Hungary
csercsik@itk.ppke.hu

***Abstract*—We consider the scenario, when detailed bid data of a day-ahead power exchange is not accessible due to privacy issues, but the statistical extract of the data is available. More precisely, we assume that in the case of standard bids, the approximate joint distribution of bid quantity and price is available, while in the case of block bids we know the average and standard deviation of the bid parameters. Based on these statistical features we re-generate the bid set, and analyze, how much the market outcome differs from the original result, depending on the detail of the statistics used. In addition, we analyze how much the description details affect the performance of the regenerated bid set.**

***Index Terms*—Day-ahead electricity markets, simulation**

I. Introduction

Day-ahead electricity auctions or power exchanges (DAPXs) accept supply and demand bids from participants in various formats, and determine the set of accepted (incl. partially accepted) and rejected bids. Moreover the auction outcome defines the payoff for every accepted bid through the market clearing prices (MCPs) [1]. Since the introduction of this trading format [2], DAPXs have been the subjects of various scientific studies. On the one hand, the computationally efficient clearing of such markets is a challenge because of the fundamentally non-convex nature of the problem [3]. In addition to exact methods like [3], there are also heuristic approaches, which aim to utilize the characteristic properties of the problem in order to reduce computational time [4]–[6]. On the other hand, from the perspective of market participants, the question of strategic bidding arises naturally (for reviews on strategic/optimal bidding in DAPXs see e.g. [7], [8]).

If one proposes a novel efficient clearing method [9] or a potential approach for strategic bidding [10], it is a straightforward question, how the proposed algorithm performs in the case of realistic problems. This is the point when some issues may arise. Although DAPXs usually publish the resulting trading quantities and clearing prices (the latter is a necessary signal for potential investors), there is no publicly available information about the details of the submitted bids. This is not surprising, since the nature of the auction is 'sealed bid', if a participant is aware of the parameters describing the bid of its competitor, it may gain advance. Based on this factor DAPXs usually do not publish bid data, even years after the auction (when it may be already obsolete for research purposes in some aspects). In this paper we argue that it would be very desirable from the research point of view to publish statistical extracts of bidding data, which ensure anonymity and privacy but in the same time provide valuable insights into the nature of the market in question. Such meta-data publishing would have other benefits as well (considering e.g. portfolio analysis or deeper market analysis), however in the current paper we restrain ourselves only to a few aspects.

In the case of market clearing algorithms, testing requires a data set, which fits the actual market problems *in size*. In other words, the resulting MCPs published by the DAPXs may be the result of 100 or 10000 submitted bids as well, while the underlying optimization problem sets a very different challenge. In addition, these numbers may strongly vary with individual power markets (depending on the actual number, and bidding strategy of participants).

In the case of strategic bidding methods however, knowing the number of submitted bids is not enough to have a 'realistic' scenario. While the resulting MCPs are published, if a generating unit has high enough capacity, it may exert market power, i.e. its bidding behavior may affect the resulting MCPs. If there is no data available about the bids of the market at all, it is impossible to estimate the effect of specific bidding strategies on the MCP.

In the current paper we introduce two slightly different statistical descriptions of the DAPX bid sets, based on which a 'realistic' artificial (or 'in silico') bid set may be generated, and analyze how much the description details affect the resulting market outcomes in the case of the reproduced data.

II. Materials and Methods

In this study we consider a multiperiod market with stepwise bids only, but including also block bids, which define interdependencies between individual periods. There are two

This work has been supported by the ÙNKP-21-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, and Funds FK 137608 and K 131545 of the Hungarian National Research, Development and Innovation Office. Dávid Csercsik is the grantee of the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

types of bids in our market model: (1) Standard bids, which are submitted for single periods and have 2 parameters, namely quantity and price. The bid price describes the minimal price (per unit) upon which the bid may be accepted in the case of supply bids, and the maximal price to be paid in the case of demand bids. These bids may be partially accepted, and they must be accepted if the respective MCP is appropriate (if not, they must be rejected – partial acceptance is allowed if the bid price equals the MCP). (2) Block bids, which are submitted by producers (suppliers) potentially for multiple periods have 4 parameters: index of the start period, index of the end period (may be equal to the start period as well), quantity and price. Any block bid *may* be accepted, if the clearing price is appropriate in each of the included periods (block bids are not necessary accepted if the prices are appropriate - see the phenomena of paradoxically rejected block orders e.g. in [11]). Block bids have the fill or kill property, e.g. they may be entirely accepted or rejected (no partial acceptance is allowed). The market clearing algorithm may be implemented as on optimization problem. Bid acceptance constraints may be described by big-M conditions using auxiliary variables (see e.g. [12]) or via the use of the primal-dual framework [1].

A. Reference bid set

We assume the existence of a reference bid set, from which the auctioneer extracts the statistical data. In the case of this study, this bid set covers 10 time periods, includes 250 standard supply and 250 standard demand bids, and 20 block bids. The supply bids include price-taker bids (assume e.g. renewable, non-controllable sources) as well. Standard supply bids range from 80 MWh to 257 MWh regarding quantity and from 23 to 367 EUR/MWh regarding price. Standard demand bids range from 78 MWh to 259 MWh regarding quantity and from 92 to 377 EUR/MWh regarding price. The quantity of block bids is between 51 and 90 MWh, while their price is between 110 and 356 EUR/MWh. The average length of a block bid is 3.85 periods.

B. Statistics used in the method

Regarding standard bids, we assume that the approximate joint distribution of bid quantities and prices is accessible. The levels of detail, denoted by d_q^s and d_q^s in the case of quantity and price respectively, defines the number of bins in the joint distribution. The width of a bin (w_q, w_p) may be calculated as

$$w_q = \frac{q_{max} - q_{min}}{d_q^s}, \quad w_p = \frac{p_{max} - p_{min}}{d_p^s} \tag{1}$$

where q_{max}, q_{min}, p_{max} and p_{min} denote the maximal and minimal price/quantity value among bids. These extreme values may be calculated either for the whole bid set (in this case 24 periods) or for each period distinctly. According to this, we will distinguish two methods: In the case of the first method (*UB*), the extreme values, and thus the bins are defined universally over the periods (universal bins), while in the case of the *PB* method, the extreme values, bin limits and bin widths are calculated for each period individually (periodic bins).

After the bins along the two dimensions (q and p) have been defined, the next step is to count the bids relevant for each bin to obtain an empirical probability density function of the joint distribution. Strictly speaking, we assume that not only the distribution (which is normalized by nature), but also the exact number of bids corresponding to each bin in each period is available. Let us define Ψ^S and Ψ^D as 3-dimensional arrays. $\Psi^S(t,i,j)$ ($\Psi^D(t,i,j)$) denotes the number of standard supply (demand) bids corresponding to period t, which fall into the i-th bin regarding the quantity and into the j-th bin regarding the bid price.

In the case of block bids, we assume that the average value and standard deviation of the four bid parameters (start period, end period, quantity, price) is available in addition to the number of block bids.

C. Bid generation, based on the statistics

In the case of the standard bids, the approach for data generation based on the statistics is very simple. Considering Ψ^S and Ψ^D, and the bin limits, for each nonzero entry the respective number of bids are generated randomly, assuming uniform distribution in the quantity and price range defined by the respective bin limit values.

In the case of block bids, we generate the appropriate number of block bids (the number of block bids is assumed to be known), with the following parameters. The start time of the bid (t_e^{BB}) is determined as described in eq. 2

$$t_s^{BB} = max\left(min\left(RI([\bar{t}_s - 2\sigma(t_s)], [\bar{t}_s + 2\sigma(t_s)]), T\right), 1\right) \quad , \tag{2}$$

where $RI(k,m)$ stands for the random integer function regarding the interval $[k,m]$ (k and $m \geq k$ are integers), $[x]$ denotes the integer closest ro $x \in \mathcal{R}$, $\bar{t}_s$ is the average value of the start time of block bids, and $\sigma(t_s)$ is the standard deviation of the start time of block bids. T is the number of periods considered in the market.

Similarly, the end time of the bid t_e^{BB} is determined as described in eq. 3

$$t_e^{BB} = max\left(min\left(RI([\bar{t}_e - 2\sigma(t_e)], [\bar{t}_e + 2\sigma(t_e)]), T\right), 1\right) \quad , \tag{3}$$

where $\bar{t}_e$ is the average value of the end time of block bids, and $\sigma(t_e)$ is its standard deviation.

The quantity of a block bid (q^{BB}) is determined as a realization of a random variable with uniform distribution in the interval

$$[\overline{q^{BB}} - \sigma(q^{BB}) \ , \ \overline{q^{BB}} + \sigma(q^{BB}) \quad ,$$

where $\overline{q^{BB}}$ and $\sigma(q^{BB})$ denote the average value and standard deviation of block bids. The price of block bids is derived using the same approach.

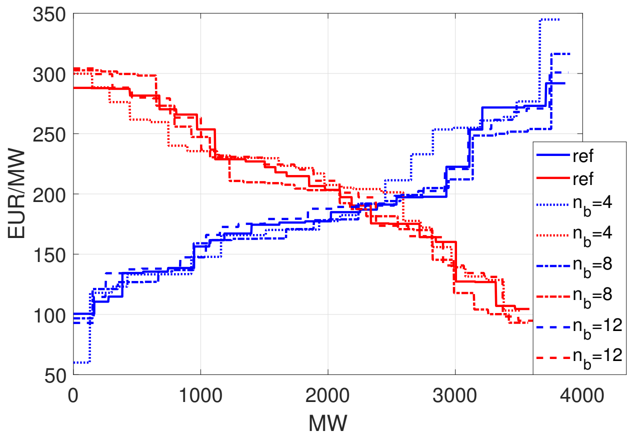

Fig. 1. Reference and reproduced bid curves (UB method) of period 1, omitting block bids.

D. Comparision

After the reproduced bid sets are generated based on the previously introduced statistics, we perform the market clearing using the reproduced bid sets, and compare the outcome of the clearing with the outcome of the reference bid set. We summarize the results in the next section.

III. Results

We analyzed the performance of method UB and PB assuming 3 different details of the statistics. In the first case, 4 bins have been used along each dimension ($n_b = 4$), in the second case 8, while in the third case 12.

To give an impression about the relation and original bid curves, one reproduction of the resulting bidding curves of standard bids is depicted in Fig. 1. It is not surprising that the higher number of bins results in a more accurate reproduction of the original bid curve.

As the reproduction is random (bids are generated randomly inside their respective bins), we use multiple instances of the reproduced bid sets (in this case 10), and perform the market clearing for the original and all of the reproduced bid sets.

Figures 2 and 3 depicts the resulting MCPs in the case of the UB and PB method respectively, compared to the original, reference bid set.

If one aims to numerically characterise the performance of the methods in the context of the resulting MCP, a very simple norm may be introduced.

$$AD_{MCP} = \frac{\sum_i \sum_t ||MCP_i(t) - MCP_{ref}||}{n_r} \quad (4)$$

The averaged deviation of the MCP (AD_{MCP}) defined by eq. 4 calculates the 2-norm of the MCP deviation for each period and sums these values over the periods for each reproduced bid set. Following this, the resulting values are averaged. The numerical results of the simulations are summarized in table I.

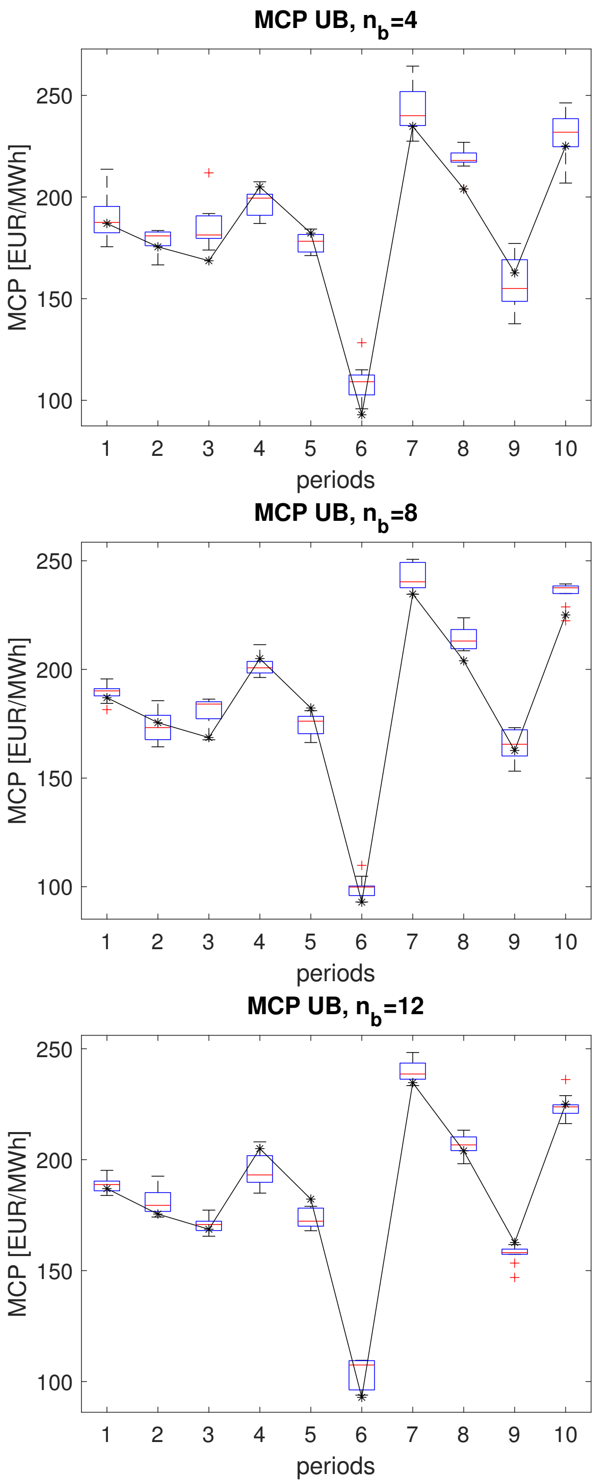

Fig. 2. MCP values resulting from the original bid set (continuous curve), and from statistics-based reproduced bid sets (box plots) in the case of the UB method.

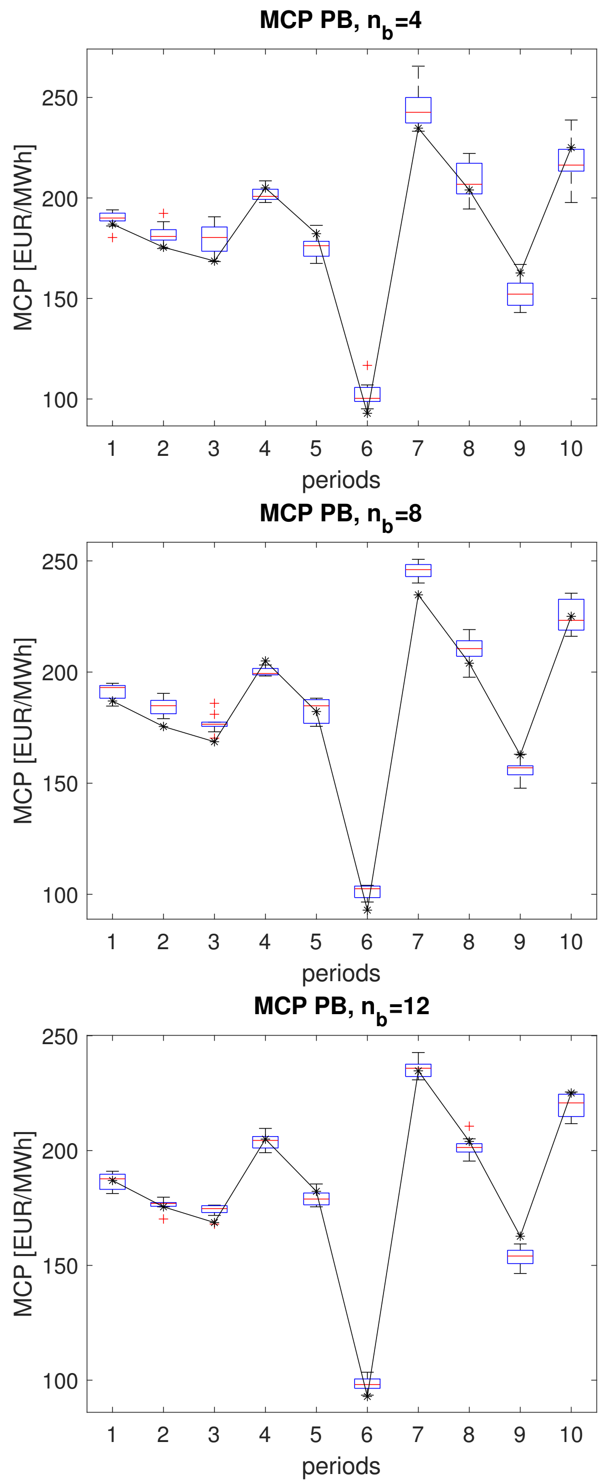

Fig. 3. MCP values resulting from the original bid set (continuous curve), and from statistics-based reproduced bid sets (box plots) in the case of the PB method.

	UB	PB
$n_b = 4$	40.80	32.15
$n_b = 8$	27.95	25.55
$n_b = 12$	24.94	17.37

TABLE I
VALUES OF AD_{MCP} IN THE CASE OF THE UB AND PB METHOD, ASSUMING DIFFERENT DETAIL LEVELS (n_b)

ID	q	p	ID	q	p
D1	11	11	S1	11	11
D2	8.9	8.9	S2	7.1	7.1
D3	6.9	6.9	S3	5.1	5.1
D4	4.9	4.9	S4	3.1	3.1
D5	1	1	S5	1	1

TABLE II
EXAMPLE BID SET

IV. DISCUSSION

A. UB vs PB

In table I, on the one hand we may see that the MCP-deviance norm introduced in equation 4, decreases as the number of bins is increased, i.e. more detailed statistics provide more accurate results compared to the original MCP values. On the other hand, table I also shows, that the nature of the representation also significantly matters. If the bins are defined to each period individually (PB), a 33% lower level of detail ($n_b = 8$ vs 12) implies almost the same resulting AD value compared to the UB case (25.55 vs 24.94). This highlights the importance of the methodology details, namely that if appropriate data is available, it is advisable to adjust bin limits to maximum and minimum bid prices of individual periods. These maximum and minimum bid prices may significantly vary with e.g. the appearance/disappearance of price-taker bids of renewable-based generating units over different periods.

B. Bias of the reproduced MCPs

Figures 2 and 3 show that in some of the periods, the MCPs resulting from the reproduced bid sets may be different from the original MCP values (see e.g. period 3 or 10 in the UB case when $n_b = 8$). At first sight this phenomena is not straightforward. The explanation lies in the fact that the relations of the original bid values and the bin limits strongly determine the properties of the reproduced bid sets. To highlight the underlying origins of the bias, let us consider a simple 1-period example and take e.g. the reference bid set summarized in Table II (no block bids are present this time).

If we suppose 5 bins along each dimension, the minimum and maximum values define the following bin limits for both q and p: $[1, 3, 5, 7, 9, 11]$. The Ψ^S and Ψ^D matrices will be diagonal, and since in the reproduction process we assume uniform distribution of bid parameters in each bin, the expected value of the parameters of the reproduced bid set will be in the middle of the bins (i.e. equal to 2, 4, 6, 8 and 10 respectively).

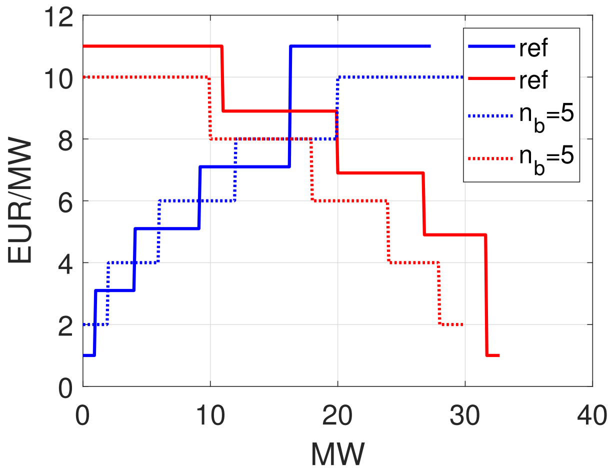

Fig. 4. Reference and reproduced bid curves, based on table II.

Figure 4 depicts the resulting bid curves of the original bid set and the reproduced one (assuming that the reproduced bid set is the realisation corresponding to the expected values).

We can see, the intersections of the curves do not coincide (regarding neither quantity or price), thus the expected value of the MCP resulting from the reproduction is different compared to the original MCP. This example shows the limitations of this approach. Unless the detail of the statistics is very high, there is no guarantee for the precise reproduction of the original MCPs in the case of using a statistic-based artificial bid set.

V. Conclusions and future work

In this paper we have considered a scenario when statistical details of DAPX bid data is available, and analyzed the

A. Future work

The joint distribution-type representation may be also used in the case of block bids. In this case however, the respective Ψ arrays will be 4 dimensional (since a block bid has 4 parameters). If such data is not available, heuristics may be used to generate a set of block bids, which result in the same statistical features as block bids of the original bid set (the method described in this article for block bids does not ensure the exact match of the known statistical values, but gives an acceptable approximation).

performance of the reproduced bid set, regarding the market outcomes (more precisely MCPs). We found that the detail of the statistics naturally affects the precision of the reproduction, but also found that the methodical details of the statistics (i.e. the global or period-wise determination of bin limits) has also significant effect. The study shows that assuming any detail of the statistics, if on wishes to create an artificial data set matching the original MCPs as good as possible, it is advisable to use period-dependent definition of bins, and use this data in the reproduction process.

References

[1] M. Madani, "Revisiting european day-ahead electricity market auctions: Mip models and algorithms," Ph.D. dissertation, Université catholique de Louvain, 2017.

[2] F. P. Sioshansi, *Competitive electricity markets: design, implementation, performance*. Elsevier, 2011.

[3] M. Madani and M. Van Vyve, "Computationally efficient mip formulation and algorithms for european day-ahead electricity market auctions," *European Journal of Operational Research*, vol. 242, no. 2, pp. 580–593, 2015.

[4] D. I. Chatzigiannis, G. A. Dourbois, P. N. Biskas, and A. G. Bakirtzis, "European day-ahead electricity market clearing model," *Electric Power Systems Research*, vol. 140, pp. 225–239, 2016.

[5] L. H. Lam, V. Ilea, and C. Bovo, "European day-ahead electricity market coupling: Discussion, modeling, and case study," *Electric Power Systems Research*, vol. 155, pp. 80–92, 2018.

[6] K. Derinkuyu, "On the determination of european day ahead electricity prices: The turkish case," *European Journal of Operational Research*, vol. 244, no. 3, pp. 980–989, 2015.

[7] A. K. David and F. Wen, "Strategic bidding in competitive electricity markets: a literature survey," in *2000 Power Engineering Society Summer Meeting (Cat. No. 00CH37134)*, vol. 4. IEEE, 2000, pp. 2168–2173.

[8] G. Li, J. Shi, and X. Qu, "Modeling methods for genco bidding strategy optimization in the liberalized electricity spot market–a state-of-the-art review," *Energy*, vol. 36, no. 8, pp. 4686–4700, 2011.

[9] D. Shah and S. Chatterjee, "An efficient common cumulative power method for market clearing in day ahead electricity market," in *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*. IEEE, 2019, pp. 1–6.

[10] Y. Zheng, H. Yu, Z. Shao, and L. Jian, "Day-ahead bidding strategy for electric vehicle aggregator enabling multiple agent modes in uncertain electricity markets," *Applied Energy*, vol. 280, p. 115977, 2020.

[11] M. Madani and M. Van Vyve, "Minimizing opportunity costs of paradoxically rejected block orders in european day-ahead electricity markets," in *11th International Conference on the European Energy Market (EEM14)*. IEEE, 2014, pp. 1–6.

[12] D. Csercsik, Á. Sleisz, and P. M. Sőrés, "The uncertain bidder pays principle and its implementation in a simple integrated portfolio-bidding energy-reserve market model," *Energies*, vol. 12, no. 15, 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/15/2957