Forecasting Electricity Spot Market Merit-Order Curves with Functional Time Series Modeling

Guillaume Koechlin
MOX, Department of Mathematics
Politecnico di Milano
Milan, Italy
guillaume.koechlin@polimi.it

Filippo Bovera
Department of Energy
Politecnico di Milano
Milan, Italy
filippo.bovera@polimi.it

Piercesare Secchi
MOX, Department of Mathematics
Politecnico di Milano
Milan, Italy
piercesare.secchi@polimi.it

***Abstract*—This study introduces a novel approach for forecasting merit-order curves in electricity spot markets by leveraging functional principal component analysis (FPCA) to efficiently represent a pair of supply and demand curves in a vector space and employing multivariate time series models for their prediction. Applied to the Italian day-ahead market during the 2023-2024 period, our approach generates accurate supply and demand curves forecast, and despite not being explicitly optimized for price forecasting, yields price forecasts which outperform state-of-the-art price-based models, highlighting the benefits of a curve-driven methodology.**

***Index Terms*—electricity price forecasting, functional data analysis, merit-order curves**

I. Introduction

The rising integration of non-programmable renewable energy sources and the development of large-scale storage systems, acting both as generation and consumption units, have increased the unpredictability and volatility of electricity spot markets (ESMs) results. This shift calls for the development of more sophisticated forecasting tools. ESMs generally operate under a double uniform-price auction mechanism which implies that their results are entirely defined by the equilibrium price, set as the intersection of the supply and demand merit-order curves, resulting from the aggregation of the individual supply offers and demand bids (Fig. 1). While existing literature on ESM forecasting is turned more towards statistical or machine learning approaches that directly model the resulting price, the approach we follow in this paper instead models the underlying merit-order curves at the origin of price formation. This strategy offers a more comprehensive view of market results, enhances the interpretability of the forecasted price, and may enable more accurate uncertainty quantification of price predictions – though this latter aspect is not explicitly addressed in this study.

Previous work in this area has demonstrated the feasibility of forecasting price through merit-order curves prediction. For instance, [1] introduced the X-model, which estimates supply and demand curves to derive competitive price forecasts. However, their approach relies on a coarse discretization of the curves, introducing approximation errors. [2] advanced this concept by applying functional data analysis to forecast the merit-order curves of the Italian day-ahead market, a methodology similar to that used by [3] for the gas market. The models used in [2], [3], however, do not allow for the incorporation of exogenous variables – such as forecasted demand and renewable generation – which are critical for accurate price prediction. Additionally, previous studies have largely treated supply and demand curves separately, as independent functional time series, overlooking the statistical dependence between them.

In contrast, our study proposes a novel framework that combines functional principal component analysis (FPCA) with multivariate time series modeling to jointly forecast supply and demand curves in any double uniform-price electricity market auction. We focus on short-term forecasting but note that our approach can be extended to longer time horizons by adapting the underlying time series models. Our methodology builds on the X-model but eliminates the need for curve discretization by maintaining the functional form of the merit-order curves. We apply this approach to the Italian day-ahead market (MGP) during the 2023-2024 period and evaluate the accuracy of our curve forecasts in predicting market-clearing prices.

The remainder of this paper is structured as follows: Section II details our forecasting methodology, Section III describes the dataset used, Section IV presents the results, and Section V provides concluding remarks.

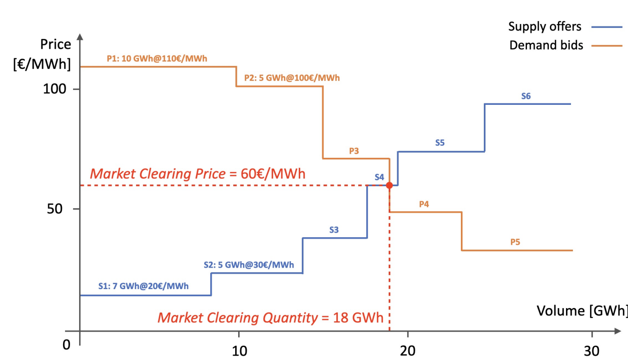


Fig. 1: Example an electricity spot market clearing from the supply (in blue) and demand (in orange) merit-order curves.

II. METHODOLOGY

A. Curves representation

Formally, a merit-order curve represents price as a function of quantity $q \to P(q)$ and is defined from a set of supply offers (demand bids, respectively) characterized by a price and a quantity $(p_1, q_1), (p_2, q_2), \ldots, (p_n, q_n)$ sorted following the merit-order, i.e., increasing order of price $p_1 \leq p_2 \leq \cdots \leq p_n$ for the supply curve, decreasing order of price $p_1 \geq p_2 \geq \cdots \geq p_n$ for the demand curve. Mathematically, a merit-order curve can be defined as follows:

$$P(q) = p_k, \quad \text{if} \sum_{i=0}^{k-1} q_i \leq q < \sum_{i=0}^{k} q_i, \quad k = 1, \ldots, n$$

where we define $q_0 = 0$, for notational convenience. For supply (or demand), $P(q)$ is said to be the marginal price associated to a total offered (resp. demanded) quantity q.

From the offers or bids, we can also define the quantity function, still being a monotone step function:

$$Q^{(s)}(p) = \sum_{i=1}^{n} q_i \mathbb{1}_{\{p_i \leq p\}}, \quad Q^{(d)}(p) = \sum_{i=1}^{n} q_i \mathbb{1}_{\{p_i \geq p\}}$$

That is, $Q(p)$ is the quantity that is offered (or demanded) at a price below (resp. above) p. Note that the definition above does not require the sequence of offers (or bids) to be price-ordered.

The quantity and the price function (or curve) are just two different representations of the same merit-order curve, and we can easily switch from one representation to an other. From now one, when we refer to supply or demand curves without additional precision, we intend the quantity function.

B. A time series of functions

We would like to model and then forecast the time series of curves $Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_T^{(s)}$ and $Q_1^{(d)}, Q_2^{(d)}, \ldots, Q_T^{(d)}$ over a time period composed of T hourly intervals.

In order to forecast functional time series, we follow the guidance of [4] that recommend the use of functional principal component analysis (FPCA) to derive an uncorrelated vector representation of the curves and apply standard multivariate time series forecasting methods to this representation. The forecast is finally translated back in the original curves form using the Karhunen-Loève (KL) expansion [5].

FPCA is a dimensionality reduction technique for functional data, extending classical PCA to infinite-dimensional spaces. For an introduction to functional data analysis, see [6], with a detailed discussion of FPCA in Chapter 8. It decomposes a set of observed functions into an orthonormal basis that captures the dominant modes of variations in the data. This decomposition yields the KL expansion:

$$Q_t(p) \approx \overline{Q}(p) + \sum_{k=1}^{K} \beta_{tk} \xi_k(p)$$

where $\overline{Q}(p)$ is the mean function, $\xi_1(p), \xi_2(p), \ldots, \xi_K(p)$ are the first K functional principal components (FPCs) and $\beta_{t1}, \beta_{t2}, \ldots, \beta_{tK}$ are the corresponding scores for observation t. Intuitively, the FPCs can be seen as specific features common to all curves while the scores are measures of how pronounced are these features in a specific curve.

In our case, we would like to account for possible dependence between the supply and demand curves time series. To do so, we perform FPCA separately for the supply and demand curves and by concatenating the scores for the supply curves with those of the demand curve, we get a K-dimensional representation $Y_1, \ldots, Y_T$ of the paired functional series $(Q_t^{(s)}, Q_t^{(d)})_{t \in \{1,\ldots,T\}}$, where we redefine $K = K_s + K_d$ and

$$Y_t = \left[\beta_{t1}^{(s)}, \beta_{t2}^{(s)}, \ldots, \beta_{tK_s}^{(s)}, \beta_{t1}^{(d)}, \beta_{t2}^{(d)}, \ldots, \beta_{tK_d}^{(d)}\right]^T$$

The advantage of working with Y_t is triple: (i) we turn an infinite-dimension problem into an intrinsically finite-dimension one, (ii) we phase out the intrinsic dependence of functional data, (iii) we jointly model the supply and demand processes.

C. Forecasting the curves' vector representation

We consider here the one-day-ahead forecasting problem. Consistently with existing literature, we regard the hourly vector time series as 24 separate daily vector time series, one for each hour of the day. This is common practice in electricity price forecasting as the 24 hours of the next day are simultaneously settled the day before. In addition, market dynamics vary a lot depending on the hour of the day, justifying to treat them separately. For a detailed motivation of this choice, refer to [7]. We hence change the time indexing of Y_t to consider the value at day d and hour h, $Y_{d,h}$ and we solve 24 one-step-ahead forecasting problems.

We consider four different approaches, all based on the popular parameter-rich ARX model estimated with LASSO used in [8]–[11] in the context of electricity price forecasting. First this framework has showed state-of-the-art performance for price forecasting and second, it was used in the X-model of Ziel and Steinert [1]. In particular, we consider two univariate models treating each component of $Y_{d,h}$ separately as suggested in [12] – the concurrent ARX and the full ARX – and two other multivariate models jointly modeling the vector $Y_{d,h}$ – concurrent VARX and the semi-full VARX. The comparison between the univariate and the multivariate approaches allow to test the added value of modeling the cross-dependence between scores while the comparison between the concurrent and full (or semi-full) approaches allow to test the added value of modeling the cross-dependence between hours. The *AR* part refers to the fact that lagged values are used to predict future values. In our case we include lag 1 and 7, considering therefore information of the day before and 7 days before. Additionally, each model has the suffix *X which means that it makes use of exogenous variables which are available for day $d+1$ on day d (i.e., they are known in anticipation or they are themselves one-day-ahead forecasts). This set of r exogenous variables is represented by the r-dimensional vector $X_{d,h}$. Similarly to [8], for all four models, the L^1-regularization parameter λ is selected as the one that minimizes the Bayesian information criterion (BIC), exploiting the efficient least angle regression (LARS) algorithm [13] for the parameter search.

1) Concurrent ARX (ARX): The ARX models each of the K component $y_{d,h}$ of $Y_{d,h}$ as:

$$y_{d,h} = \phi_{1,h} y_{d-1,h} + \phi_{7,h} y_{d-7,h} + \theta_h X_{d,h} + \epsilon_{d,h}$$

where $\phi_{\cdot,h}$ are the autoregressive coefficients and θ_h is the r-dimensional row vector of exogenous variables coefficients. The *concurrent* denomination relates to the fact that there is no exchange of information between the different hours of the day (nor between the components, for this model).

2) Full ARX (fARX): The fARX models each of the K components $y_{d,h}$ of $Y_{d,h}$ as:

$$y_{d,h} = \sum_{j=1}^{24} \phi_{1,h}^{(j)} y_{d-1,j} + \sum_{j=1}^{24} \phi_{7,h}^{(j)} y_{d-7,j} + \theta_h X_{d,h} + \epsilon_{d,h}$$

where $\phi_{\cdot,h}^{(j)}$ is the linear effect of the lagged hour j on the hour h. The *full* denomination relates to the fact that there is now exchange of information between hour h and any other hour of the day.

3) Concurrent VARX (VARX): The VARX models the full vector $Y_{d,h}$ as:

$$Y_{d,h} = \Phi_{1,h} Y_{d-1,h} + \Phi_{7,h} Y_{d-7,h} + \Theta_h X_{d,h} + \varepsilon_{d,h}$$

Where $\Phi_{1,h}$ and $\Phi_{7,h}$ are the $K \times K$ matrices of autoregressive coefficients and Θ_h is the $K \times r$ matrix of exogenous variables coefficients. Here the model is concurrent but we allow for cross-dependence – described by the off-diagonal coefficients of $\Phi_{1,h}$ and $\Phi_{7,h}$ – between the components of $Y_{d,h}$, contrarily to the two previous models.

4) Semi-full VARX (sfVARX): The sfVARX models the full vector $Y_{d,h}$ as:

$$Y_{d,h} = \sum_{j=1}^{24} \Phi_{1,h}^{(j)} Y_{d-1,j} + \sum_{j=1}^{24} \Phi_{7,h}^{(j)} Y_{d-7,j} + \Theta_h X_{d,h} + \epsilon_{d,h}$$

Where we impose the restriction that the off-diagonal terms of $\Phi_{\cdot,h}^{(j)}$ must be zero whenever $j \neq h$, that is, cross-dependence is allowed between hours, but not between two different components. This restriction is imposed to have a manageable number of parameters. Indeed, without this condition we would have $2 \times [24 \times K] + r = 48K + r$ parameters for each component of $Y_{\cdot,h}$, while with this condition the number of parameters is only $2 \times [23 + K] + r = 2K + 46 + r$. For instance, for $K = r = 10$, the condition implies 76 parameters instead of 490.

D. Curves forecast evaluation

The resulting curves forecasting models are evaluated separately for each hour with the *functional coefficient of determination*

$$R_h^2(p) = 1 - \sum_{d=1}^{n} [\hat{Q}_{d,h}(p) - Q_{d,h}(p)]^2 / \sum_{d=1}^{n} [Q_{d,h}(p) - \overline{Q}_{\cdot,h}(p)]^2$$

which represents the proportion of variance explained by the model, the *average squared correlation*

$$R_h^2 = \frac{1}{p_{\max} - p_{\min}} \int_{p_{\min}}^{p_{\max}} R_h^2(p) dp$$

and as a final summary measure of the model's performance across all hours, the *global average squared correlation*:

$$R^2 = \frac{1}{24} \sum_{h=1}^{24} R_h^2.$$

Additionally, to test for the significance of the outperformance of a model against another, we use the Diebold-Mariano (DM) test [14] where the forecast errors are:

$$L_{d,h} = \int_{p_{\min}}^{p_{\max}} [\hat{Q}_{d,h}(p) - Q_{d,h}(p)]^2 dp$$

which is the squared L^2-norm of the functional residuals. Two one-sided tests are run for each pair of models for each hour.

E. Price forecasting

Though we solve a more general problem than the price forecasting one, we test the ability of the previously described curves forecasting procedure to generate accurate price predictions. The price prediction $\hat{p}_t$ obtained from the curves prediction $(\hat{Q}_t^{(s)}, \hat{Q}_t^{(d)})$ is simply obtained by determining the zero of the function $p \to \hat{Q}_t^{(d)}(p) - \hat{Q}_t^{(s)}(p)$. In order to assess whether the resulting price forecast is competitive with the state-of-art, we compare it with the price-based LEAR model of [8], which is similar to the models for the scores previously described but with a much richer autoregressive structure and dependence for lags of the exogenous variables. The models are compared in terms of Mean Absolute Error (MAE) and with classical DM tests (with squared residuals error).

III. Data

We consider the merit-order curves of the day-ahead electricity market (MGP) of the Italian Power Exchange (IPEX), for the 2023-2024 period. The quantity curves are built according to the procedure described in section II, using the `DomandaOfferta` dataset retrieved from the FTP server of Gestore del Mercato Elettrico (GME) – the entity managing the IPEX – which contains all supply offers and demand bids submitted on the MGP. Additional details are provided in the Appendix A.

The exogenous variables that are used to predict the components of the curves representation are those that are commonly chosen in day-ahead market price forecasting models: day-ahead forecasts for *national load*, *transfer capacity with neighboring zones* and, *wind and solar generation*. We also include a dummy variable indicating the *day type*. More details are provided in the Appendix A-D.

IV. Results

The test period is the whole year of 2024. For each hour h, we therefore have 366 observations in the test set. The curves forecasting models are evaluated with daily recalibration with a sliding window of 358 days, which means that for each day d of the test period, we fit the model to the previous 358 days $d-1,\dots,d-358$ of data and we use it to forecast day d. For the motivation of such a procedure, the reader can refer to [8]. We must precise that we here use a sliding window of 358 days because we only have data from 2023-01-01 and since we need a seven days lag for the vector of components $Y_{d,h}$, the set of training observations starts from 2023-01-08.

To assess the overall models performance, we include in the performance comparison a *Naive* model that simply uses the current day's curve to forecast the curve for the next day: $\hat{Q}_{d+1,h} = Q_{d,h}$.

A. Functional Principal Components Analysis

We found that retaining 5 FPCs for supply and 3 FPCs for demand provides a satisfactory dimensionality reduction while keeping the dimension of Y_t and hence the computational burden low. Fig. 2 shows the curve features that the FPCs capture. We can see that in both cases, the first principal component explains most of the variability. In particular for demand, the FPC1 explains the quasi-totality of the variability and concerns the magnitude of the total demanded quantity. The contribution to the total variance beyond FPC1 is negligible. We decided however to retain the subsequent two components because we believe that they could capture the small scale variability in a domain of the curve where the intersection with the supply curve is likely to occur, and that is therefore of interest. For supply, FPC1 measures the overall supply level at any price, while FPC2 contrasts situations where there is higher supply at low price (below 100€/MWh) compared to high price (above 200€/MWh) and vice versa. FPC3 regards the distribution of the prices above 180€/MWh, and contrasts situations when these prices are higher or lower. FPC4 and FPC5 are more difficult to interpret and regard more complex features of the price/volume distribution.

B. Curves forecasting

The global average squared correlation is presented on Table I. We can see that the Naive model is clearly outperformed by all the other models for both supply and demand curves. Additionally, for the supply curves, VARX, fARX and sfVARX explain all around 80% of the variance while ARX is one step below around 75%. Concerning the demand curves, all models excel at forecasting. This is due to the fact that demand on the MGP has almost zero elasticity which means that the quantity demand curve is more or less constant equal to the forecasted load, which is a predictor to which all models have access to. The analysis of the variables selected by the LASSO (not reported in this paper) confirmed that all models rely almost entirely on the forecasted load for predicting the demand curves' first principal component scores, explaining the quasi-totality of the variance. The R^2 score and the results of the DM test (left in the Appendix B, Fig. 5), suggest that the full approaches significantly outperform the concurrent approaches for the demand curves. This can be due to the fact that though the inelastic demand should depend on the forecasted load for the same hour, it can still be influenced by the load that is forecasted for other hours. The plot of the squared correlation functions is left in Appendix B (Fig. 7).

TABLE I: Global average R^2 scores for the supply and demand curves for each model

	Naive	**ARX**	**VARX**	**fARX**	**sfVARX**
Supply	0.499	0.750	0.798	0.791	**0.809**
Demand	0.326	0.976	0.976	**0.980**	**0.980**

The DM test results for supply confirm the superiority of sfVARX against the other models, as suggested by the R^2 score. The added value of considering lagged cross-dependence between hours as well as FPC scores is confirmed, from the table and the plot. It is also interesting to note the superior performance of the VARX compared to the fARX as it has twice less parameters (in our case, for each component, it has 33 vs 65). This observation may suggest that the cross-dependence is higher between scores than between hours.

C. Price forecast

As indicated in section II, we assess the competitiveness of our curves model for price forecasting by comparing them with the state-of-the-art LEAR model. It is important to note that since the LEAR model has a much richer lag structure than our models for the FPC scores and include lags of exogenous variables, it cannot accept more than three exogenous variables. Therefore we selected the optimal set of exogenous variables by cross-validation, which turned out to be the day-ahead load and solar-wind generation forecasts, without much surprise. The LEAR model was trained using the same rolling window of 358 days as the other models. The computation time of each model, the results of the price forecasts for each hour and the results of the DM tests can be found in the Appendix B in Table III, Table IV and Fig. 6, respectively. The average mean absolute error across hours is shown in Table II, while an example of price forecast from the sfVARX model can be visualized on Fig. 3.

TABLE II: Mean absolute error of the price forecasts obtained from the different models

	Naive	**ARX**	**VARX**	**fARX**	**sfVARX**	**LEAR**
MAE [€/MWh]	11.61	9.91	8.37	9.07	**8.28**	8.62

The best model is sfVARX with a MAE of 8.28€/MWh, followed by the VARX, with a MAE of 8.37€/MWh, both beating the LEAR model that shows 8.63€/MWh. According to the DM test, VARX and sfVARX signficantly outperform LEAR for 7 and 10, respectively, of the 24 hours. Conversely, the outperformance of LEAR over VARX and sfVARX is

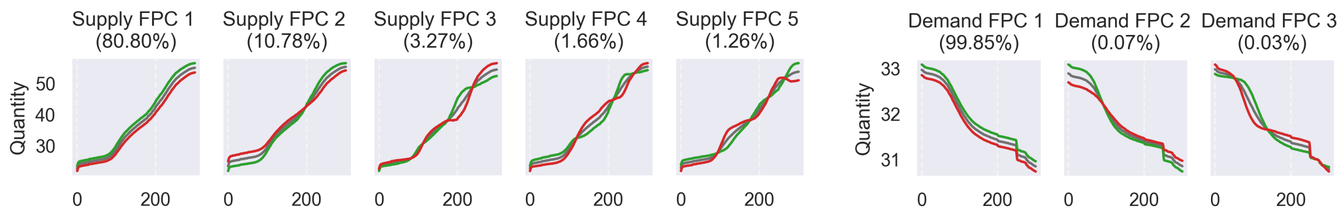


Fig. 2: Visualization of the effects of the functional principal components (FPCs) on the mean curve (in grey). The green curve corresponds to mean curve to which a multiple (25 for supply and 2 for demand) of the FPC is added, while the red curve corresponds to the mean curve to which a multiple of the FPC is subtracted. The number in parenthesis above each plot is the proportion of variance explained by the FPC.

significant for one hour only. An interesting remark is that the naive model is not systematically significantly outperformed by any other model, confirming the relative efficiency of naive forecasting in the context of electricity prices, for some hours of the day at least. A final observation is that the sfVARX significantly outperforms the VARX for only 2 hours of the day and the VARX significantly outperforms the sfVARX for only 1 hour of the day. This very interesting result suggests that the two models are almost equivalent, and show that, accounting for the cross-dependence between hours does not really improve the forecasting performance once the cross-dependence between scores is already modeled. VARX could be preferred here since it is more parsimonious than the sfVARX (33 parameters vs 76).

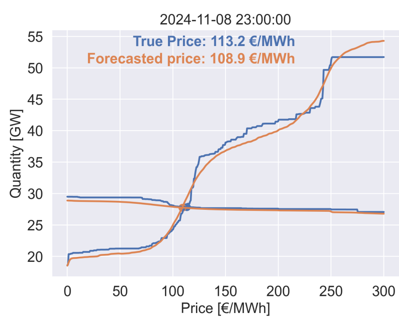


Fig. 3: Example of price forecast with the sfVARX curves forecasting model. The blue lines are the true supply and demand curves while the orange lines are the forecasted curves.

A final point regards a detailed comparison between the LEAR and VARX/sfVARX curves based approach. As can be observed from Table IV, the curves based approach works better for the hours between 10 and 16, which are the hours at which the price occasionally drops due to a low demand and a high renewable generation forecast. We believe that a nice feature of the curves-based model, is that it is very good at predicting these price drops as it precisely models the behavior of the demand and the supply during these moments. It somehow models a non-linear effect on the price from a linear effect on the supply and demand curves. An example of such a situation is plotted on Fig 4. As we can see, one of the three price drops is well anticipated both by the LEAR and the sfVARX model, but the two other drops are predicted only by the sfVARX model.

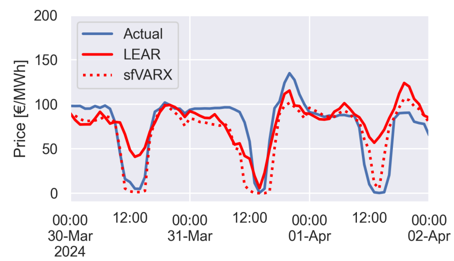


Fig. 4: Comparison of the LEAR and sfVARX price forecasting for three days during which the price dropped at the middle of the day.

V. Conclusion

This study introduced a novel approach for forecasting merit-order curves in electricity spot markets by leveraging functional principal component analysis to efficiently represent a pair of supply and demand curves in a vector space and employing multivariate time series models for their prediction. Our methodology, applied to the Italian day-ahead market, not only demonstrated its effectiveness in forecasting merit-order curves but also produced an interpretable and highly accurate price forecast that outperformed state-of-the-art price-based models.

We tested four variations of our model, each treating the hourly time series as 24 independent daily time series – one for each hour – and differing in whether they accounted for (i) cross-dependence between hours and (ii) cross-dependence between components of the vector representation. Our findings reveal that modeling both types of dependence improves forecasting accuracy, with the cross-component dependence playing a more significant role. In terms of price forecasting, two of our curve-based models outperformed the state-of-the-art LEAR model of [8] while maintaining comparable computational efficiency. Notably, we found that capturing cross-dependence between vector components alone was sufficient to achieve optimal results, eliminating the need to model hourly cross-dependence. This insight is particularly valuable as it allows for a more parsimonious model with fewer parameters, enabling calibration over shorter time windows.

Despite these promising results, further validation is needed through a direct comparison with the X-model of [1], which employs similar vector time series forecasting techniques but a different strategy for curve representation. Additionally, we should benchmark against more complex nonlinear models, such as deep neural networks (DNN) [8] and state-of-the-art architectures like NBEATS-x [15], which were shown to slightly outperform LEAR. Finally, we believe that curves-based price forecasts have certain strengths over traditional price-based forecasts, notably regarding the accurate prediction of price spikes or drops and the enhancement of uncertainty quantification. These aspects will be explored in future research.

REFERENCES

[1] F. Ziel and R. Steinert, “Electricity price forecasting using sale and purchase curves: The X-Model,” *Energy Economics*, vol. 59, pp. 435–454, Sep. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140988316302080

[2] I. Shah and F. Lisi, “Forecasting of electricity price through a functional prediction of sale and purchase curves,” *Journal of Forecasting*, vol. 39, no. 2, pp. 242–259, 2020, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.2624. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2624

[3] A. Canale and S. Vantini, “Constrained functional time series: Applications to the Italian gas market,” *International Journal of Forecasting*, vol. 32, no. 4, pp. 1340–1351, Oct. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207016300620

[4] A. Aue, D. D. Norinho, and S. Hörmann, “On the Prediction of Stationary Functional Time Series,” *Journal of the American Statistical Association*, vol. 110, no. 509, pp. 378–392, Jan. 2015, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2014.909317. [Online]. Available: https://doi.org/10.1080/01621459.2014.909317

[5] L. Horváth and P. Kokoszka, *Inference for Functional Data with Applications*, ser. Springer Series in Statistics. New York, NY: Springer, 2012, vol. 200. [Online]. Available: https://link.springer.com/10.1007/978-1-4614-3655-3

[6] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, ser. Springer Series in Statistics. New York, NY: Springer, 2005. [Online]. Available: http://link.springer.com/10.1007/b98888

[7] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, Oct. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207014001083

[8] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, “Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark,” *Applied Energy*, vol. 293, p. 116983, Jul. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261921004529

[9] F. Ziel, “Forecasting Electricity Spot Prices Using Lasso: On Capturing the Autoregressive Intraday Structure,” *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4977–4987, Nov. 2016, conference Name: IEEE Transactions on Power Systems. [Online]. Available: https://ieeexplore.ieee.org/document/7398175

[10] M. Narajewski and F. Ziel, “Econometric modelling and forecasting of intraday electricity prices,” *Journal of Commodity Markets*, vol. 19, p. 100107, Sep. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405851319300728

[11] B. Uniejewski, J. Nowotarski, and R. Weron, “Automated Variable Selection and Shrinkage for Day-Ahead Electricity Price Forecasting,” *Energies*, vol. 9, no. 8, p. 621, Aug. 2016, number: 8 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/1996-1073/9/8/621

[12] R. J. Hyndman and H. L. Shang, “Forecasting functional time series,” *Journal of the Korean Statistical Society*, vol. 38, no. 3, pp. 199–211, Sep. 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1226319209000398

[13] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, “Least angle regression,” *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, Apr. 2004, publisher: Institute of Mathematical Statistics. [Online]. Available: https://projecteuclid.org/journals/annals-of-statistics/volume-32/issue-2/Least-angle-regression/10.1214/009053604000000067.full

[14] F. X. Diebold and R. S. Mariano, “Comparing Predictive Accuracy,” *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 253–263, Jul. 1995, publisher: ASA Website _eprint: https://www.tandfonline.com/doi/pdf/10.1080/07350015.1995.10524599. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/07350015.1995.10524599

[15] K. G. Olivares, C. Challu, G. Marcjasz, R. Weron, and A. Dubrawski, “Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with NBEATSx,” *International Journal of Forecasting*, vol. 39, no. 2, pp. 884–900, Apr. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207022000413

[16] “ENTSO-E Transparency Platform.” [Online]. Available: https://transparency.entsoe.eu/

APPENDIX A
DATA DETAILS

A. Relaxation of transmission constraints

Around 35% of the time, the country-level economic optimum is not feasible due to transmission capacity constraints, which means that the market cannot be cleared from the country-level supply and demand curves. In this case, GME runs an optimization procedure, known as locational marginal pricing (LMP), which splits the entire country-level pool in a minimum number of "subpools" of market zones such that, in each subpool, the economic optimum is feasible. As a result, there can be several pairs of supply and demand curves – hence several market clearing prices – for a single hour, one for each subpool. To keep the analysis simple, for now, we did as if the country-level economic optimum was always feasible and built the country-level supply and demand curves for any hour, ignoring the problem of transmission congestion. As a result, the curves reflect the true market clearing 65% of the time, the remaining 35% corresponding to a "virtual" market clearing only. The associated market clearing price is what GME calls the *national price without constraints*, and though not always being the true market price, is still a meaningful market indicator.

B. Market coupling

Italy is integrated into the European electricity market through market coupling with several neighboring countries which are France, Austria, Slovenia and Greece. Market coupling allocates cross-border trading capacity in order to harmonize European prices. It therefore imposes the cross-border exchanges between coupled countries. As a result, the purchased quantity in MGP does not necessarily equal the sold quantity: if Italy imports more than it exports, the (domestic) sold quantity will be lower than the purchased quantity because a part of the demand is already satisfied by the imports. Conversely, if Italy exports more than it imports, the sold quantity will be higher than the purchased quantity because part of the supply must satisfy the exports. As a consequence, the market clearing price is not rigorously at the intersection of the supply and demand curves. However, everything happens as if imports were entering in the supply offers at the minimum price – such that the underlying offer is necessarily accepted – causing a right shift of the price *supply* curve, while foreign exports enter in the demand bids at the maximum price – such that the underlying bid is necessarily accepted – causing a right shift of the price *demand* curve. Therefore by incorporating the imports and exports in the supply and demand curves, the market clearing is found exactly at the intersection. In our case, the import and export quantities within the scope of market coupling are retrieved from the `MarketCoupling` dataset and the difference between imports and exports is added to the quantity supply curve, which is equivalent to adding them separately to both curves.

C. Restriction of the curves domain

A last observation regards the domain on which the quantity curves are analyzed. The entire price domain ranges from -500€ to 3000€. Obviously the part of the domain where the intersection point has negligible chances to occur is not of interest, may be subject to arbitrary variability that has no impact of the market outcomes and could just add undesired noise in the model. Therefore, we decided to restrict the quantity curves to a price domain that contains at least all prices in the 2023-2024 period with a generous margin: indeed, we restrict the curves to the domain comprised between 0€/MWh and 300€/MWh.

D. Exogenous predictors

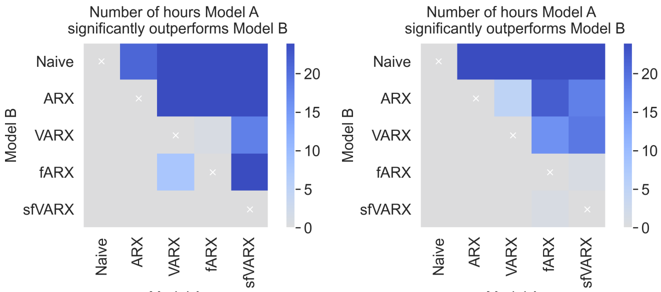
- *Day-ahead load forecast*: National load forecast for Italy, obtained from ENTSOE transparency platform [16].
- *Day-ahead transfer capacities forecast between Italy and neighboring zones*: Obtained from ENTSOE transparency platform. For each neighboring country (excluding Malta and Corsica) we include the import and the export capacity, which makes a total of 12 variables.
- *Day-ahead forecast of wind and solar generation*: Obtained from a private data provider (LSEG Data & Analytics), using forecasts derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) and Global Forecast System (GFS) weather models, country-level. We sum the solar and wind forecasts in one variable.
- *Dummy variable indicating the type of day*: Though this is not really an exogenous variable, we included it in the exogenous variables vector $X_{d,h}$ so we indicate it here. Following previous analyses specific to the electricity market, the days are divided in Mondays, Working days (from Tuesday to Friday), Saturdays and Holidays (Sundays and bank holidays). Thus, three dummy variables are included.

These exogenous predictors are collected in $X_{d,h}$, a 17-dimensional vector. Note, besides, that all these predictors are available one hour before the market clearing at 13:00 and can therefore realistically be used to forecast the MGP results.

APPENDIX B
ADDITIONAL TABLES AND PLOTS FOR THE CURVES AND PRICES FORECASTS

TABLE III: Computational time of the models for price forecasting. A daily iteration corresponds to a model fit on the last 358 days and a prediction for the next day.

	Average time per daily iteration
Naive	0.00s
ARX	1.31s
VARX	2.87s
fARX	6.94s
sfVARX	15.73s
LEAR	2.93s



(a) Supply (b) Demand

Fig. 5: Results of the DM tests for the *curves* forecast

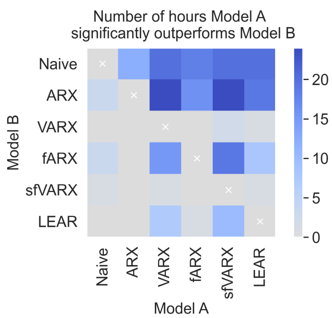


Fig. 6: Result of the DM test for the *price* forecast

TABLE IV: Mean absolute error (MAE) of price forecasts per hour [€/MWh]

h	**Naive**	**ARX**	**VARX**	**fARX**	**sfVARX**	**LEAR**
1	7.76	6.49	5.51	5.32	**5.19**	6.01
2	8.41	7.75	5.95	5.98	**5.67**	6.47
3	9.13	8.24	7.01	**6.76**	6.82	6.77
4	8.72	8.45	7.14	7.08	7.01	**6.97**
5	8.62	8.28	7.17	7.38	**6.85**	6.97
6	7.71	7.30	6.13	6.52	**5.87**	5.90
7	8.87	7.37	5.81	6.15	**5.51**	6.01
8	15.24	10.62	9.55	9.56	**9.02**	9.67
9	17.28	12.35	10.66	11.37	10.48	**10.37**
10	14.85	11.04	8.91	9.74	**8.80**	9.69
11	13.72	10.15	7.76	9.01	**7.70**	8.75
12	15.42	11.70	9.00	10.32	**8.98**	10.17
13	15.00	12.09	**9.79**	10.96	9.96	10.85
14	16.11	11.91	10.41	11.25	**10.15**	11.28
15	17.10	11.69	**9.93**	10.83	9.97	11.48
16	16.13	11.15	**9.74**	10.30	9.93	10.69
17	13.45	11.73	9.77	10.93	9.84	**9.42**
18	11.27	10.24	8.96	9.90	**8.95**	9.16
19	11.18	11.10	9.97	10.87	10.03	**9.50**
20	12.88	13.35	11.90	13.41	12.04	**11.18**
21	10.79	12.48	11.00	12.15	11.10	**10.44**
22	7.91	9.95	8.69	9.66	8.91	**8.01**
23	5.56	6.41	**5.34**	6.29	5.37	5.63
24	5.54	6.02	4.69	5.83	**4.68**	5.50
Avg	11.61	9.91	8.37	9.07	**8.28**	8.62

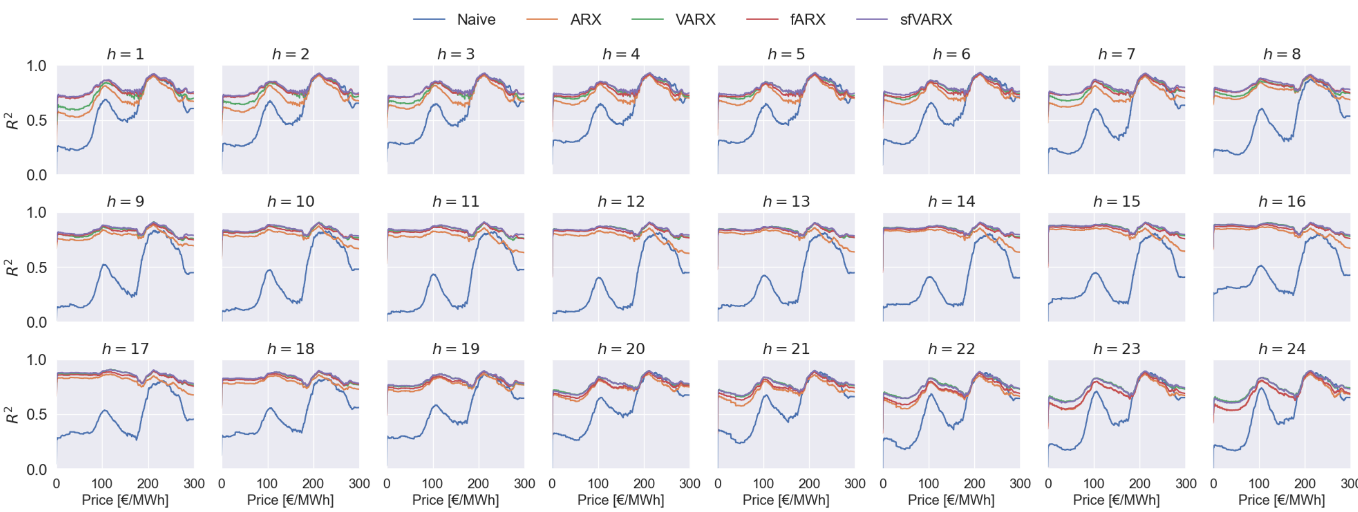


Fig. 7: Functional coefficient of determination of the supply curves predictions