Warsaw, 31 October 2023

Prof. dr hab. Michał Rubaszek
Head of Financial Markets Modelling Unit
Warsaw School of Economics
michal.rubaszek@sgh.waw.pl

DOCTORAL DISSERTATION REVIEW

Title: **Deep learning in point, probabilistic and ensemble forecasting of electricity prices**

Author: **mgr Grzegorz Marcjasz**

Supervisor: **prof. dr hab. Rafał Weron**

The review is based on a letter from prof. Rafał Weron, the Chairman of the Scientific Council of the Management and Quality Studies Discipline at Wrocław University of Science and Technology.

Research objective

The main objective of the thesis is to explore if deep neural networks (DNN) deliver superior performance in point, probabilistic and ensample electricity price forecasting (EPF) compared to best traditional benchmarks (e.g. LASSO). There are four detailed research objectives:

O1: Develop a set of best practices in machine learning EPF.

O2: Develop an interpretable and well performing DNN model for point EPF.

O3: Develop a well performing DNN model for density EPF.

O4: Develop profitable trading strategies based on DNN forecasts.

Structure and content of the dissertation

The thesis is divided into two parts. The first one consists of Introduction, the description of electricity markets, a primer on neural networks and summary. The second part wraps five articles, which constitute the main achievement to be evaluated in this report.

A1. Marcjasz G. (2021), Forecasting Electricity Prices Using Deep Neural Networks: A Robust Hyper-Parameter Selection Scheme, Energies, 13(18), 4605.

Article **A1** proposes an original and interesting method of DNN hyper-parameter calibration by using data from the other market, rather than past observations from the analyzed one. The Author claims that this delivers better EPF performance.[1] Second, the article shows that ensemble DNN performs slightly better than the LASSO benchmark, even though this is hardly seen for individual DNN vs LASSO competitions.

A2. Lago J., G. Marcjasz, B. De Schutter, R. Weron (2021) Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark, Applied Energy, 293, 116983

In my subjective view, article **A2** delivers the most valuable contribution to the EPF literature among the five reviewed papers. Predominantly, it proposes robust guidelines on how to design EPF competition. They are related to the size of the evaluation period, selecting tough benchmark, replicability or information on computing time. I like and agree with these guidelines, but was somehow missing a broader discussion on how they relate to general forecasting principles, e.g. those applied in forecasting competitions.[2] The article, which has already gathered 212 citations on google scholar (20 October 2023), is also accompanied by Python EPFtoolbox, which can be freely downloaded from GitHub. Finally, it provides comprehensive evidence that DNN performs well compared to LASSO in point EPF.

The contribution of Grzegorz Marcjasz is evaluated at 30% and concerned designing the study, developing the methods, concluding analyses and drafting the paper. It should be

---

[1] I write "the Author claims" as the results are hardly readable given the lack of the legend in Figures 3 and 4.

[2] E.g. Castle, J.L.; Doornik, J.A.; Hendry, D.F. (2021). Forecasting Principles from Experience with Forecasting Competitions. Forecasting 3, 138-165

mentioned that the journal is not classified in the Ministerial list to the "Management and Quality Studies" discipline.

A3: Olivares K., C. Challu, G. Marcjasz, R. Weron, A. Dubrawski (2022), Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with NBEATSx, International Journal of Forecasting, 39(2), 884-900.

Article **A3** applies guidelines as well as data from article **A2** to evaluate a new EPF method – NBEATSx. This method, which wraps up three components (trend, seasonal and exogenous) into point forecast, turned out to be significantly better compared to benchmark models (LASSO/DNN). Consequently, the contribution relies on introducing a new competitive method into point EPF literature.

The share of Grzegorz Marcjasz is evaluated at 25% and concerned designing the study, taking part in method development, validating the results and drafting the paper.

A4. Marcjasz G., M. Narajewski, R. Weron, F. Ziel (2023), Distributional neural networks for electricity price forecasting, Energy Economics, 125, 106843.

Article **A4** proposes a distributional DNN (DDNN), in which the output layer consists of distribution parameters (normal or Johnson distribution), and applies it in density EPF. DDNN forecasts are compared to LASSO/DNN combined with quantile regression averaging. The new method outperforms the benchmarks, both in terms of MAE and CRPS. I very like that the article presents the value of various strategies to load and unload a battery of 2MW total capacity with efficiency factor of 0.81. It is showed that this kind of battery can deliver the profit of about 12.4k EUR within 554 days. Assuming the price of batteries at around 200k EUR per MWh[3] and given the current level of interest rates, this result illustrates why in the environment of no subsidies storage based on batteries is hardly profitable, even if based on very complex trading strategies.

The contribution of Grzegorz Marcjasz is evaluated as 30% and concerned designing the study, developing and testing the models, analyzing the results and drafting the paper.

---

[3] See e.g. https://atb.nrel.gov/electricity/2023/commercial_battery_storage

A5. Marcjasz G., T. Serafin, R. Weron, (2023). Trading on short-term path forecasts of intraday electricity markets with distributional neural networks

Article **A5** extends the analysis on trading strategies based on DDNN presented in **A4**, but applying it to intraday rather than day-ahead electricity market. This change to continuous intraday market required to account for temporal dependency between the consecutive subperiods, which is done by using copula functions. In brief, the proposed DDNNC (DDNN-Copula) model outperformed the benchmark LQC (LASSO-Quantile-Copula) in terms of both the strategy profits and the accuracy of probabilistic EPF. I would notice that these results are based on 200-day testing sample, which is not entirely consistent with the guidelines formulated in article **A2**.

The contribution of Grzegorz Marcjasz is evaluated as 40% and concerned designing the study, preparing and testing the DDNN model, analyzing the results and drafting the paper.

Overall assessment

In my opinion, there are multiple reasons to consider the dissertation as valuable and even outstanding contribution to the discipline of Management and Quality Sciences. The most important ones are as follows:

1. It addresses an important and up-to-date topic on the use of machine learning methods in forecasting.
2. It delivers new results on the usefulness of DNN in point and density EPF for day-ahead and intraday markets.
3. It proposes comprehensive guidelines in designing forecasting competitions with the use of ML methods.
4. It delivers software codes available at GitHub (e.g. Python EPFtoolbox).
5. It shows that DNN forecasts are not only valuable in statistical sense, but also deliver economic value in terms of trading strategy.

Comments

I very like and highly evaluate the dissertation and, in general, I don't have any major comments. I would only raise three polemical issues that are related to the thesis and which I would be happy to discuss during thesis defense.

C1. In the first part of the thesis I was missing the opinion of Grzegorz Marcjasz why complex models perform relatively well in EPF compared to simple ones. For instance, which of the "golden rules" formulated by Armstrong et al.[4] are relevant in the context of the results presented in the thesis? Can we formulate new "golden rules" or eliminate the old ones in the case of EPF?

C2. My second comment, somewhat related to the first one, is that I was missing a broader interpretation of the results presented in part 2. In other words, after reading the thesis I know that DNN delivered competitive forecasts. However, I still don't know what are the sources of this superior performance. Can we establish which features of DNN make them so good in EPF? Is it related to a better extraction of deterministic trends from the data or rather more accurate measure of the relationship with the exogenous variables?

C3. I wonder to what extent comparing DNN to ARMAX or LASSO is a truly fair competition. In particular, in the case of ARMAX we have one specification of the model with limited options to fine-tune the hyperparameters. In the case of DNN, we can bult almost infinite number of DNN models, which differ depending on our choice related to net structure, optimization algorithms, activation functions, number of epochs, etc. If "all models are wrong but some are useful" (George Box), and assuming that all models are equally bad, theoretically 50% of DNNs should perform better and 50% worse than ARMAX/LASSO. Would it be possible to present the distribution of DNN performance with numerous hyperparameter settings compared to ARMAX/LASSO and show that significantly more than half of DNN forecasts is more accurate than those from the benchmark? In other words, to what extent can we generalize the results from the dissertation and say that the specific DNN is the best model in all instances, rather that to say that closely undefined DNN (or the ensemble of few DNN) performed well in the past?

---

[4] Armstrong, Green, and Graefe, 2015. Golden rule of forecasting: Be conservative, Journal of Business Research 68(8): 1717-1731.

I would also mention about three minor issues:

- I don't understand the purpose of sections 2 and 3 in the first part of the thesis. The basics on the functioning of electricity markets as well as a short description of basic DNN is in my opinion of little use to understand the very sophisticated results from articles **A1-A5**.
- The information about the contribution of Grzegorz Marcjasz was not transparent enough to fully evaluate the contribution to coauthored articles. For instance, to me the most important contribution of article **A2** relies on formulating forecasting guidelines. To what extent Grzegorz Marcjasz participated in formulating these guidelines?
- The citation of the article I coauthored on page 5 it is not needed, especially as it is incorrect. The sentence *Interestingly, the joint prediction of all 24 hourly prices, e.g., using vector autoregressive models, generally underperforms (Ziel and Weron, 2018), in contrast to the natural gas and crude oil markets (Rubaszek et al., 2020).* wrongly suggests that my analysis was based on hourly data.

Additional information

Grzegorz Marcjasz has authored or co-authored numerous scientific publications, which is not usual at this stage of his scientific career. The Google Scholar profile shows 31 items (most published in top field forecasting or energy journals) that have accumulated 918 citations and h-index of 12 (date 20 October 2023). This kind of statistsics are hardly seen even in applications for the next scientific degree (habilitation).

Conclusion

The research carried out by Grzegorz Marcjasz shows that he has profound understanding of machine learning methods, the functioning of electricity markets and forecasting literature. Moreover, he shows exceptional skills in terms of processing large and sophisticated databases. Without doubts, the presented dissertation constitutes valuable contribution to the EPF literature. Consequently, it meets (and even surpasses by a wide margin) the requirements for doctoral theses as described in the Act of 20 July 2018 on Scientific Degrees and Academic Title. I recommend to admit the thesis to the next stages of the doctoral dissertation proceedings. Additionally, I recommend awarding the thesis in accordance with the regulations of the University.

Prof. dr hab. Michał Rubaszek

Zakład Modelowania Rynków Finansowych

prof. dr hab. Michał Rubaszek