Probabilistic DAM price forecasting using a combined Quantile Regression Deep Neural Network with less-crossing quantiles

van der Heijden, Ties; Palensky, Peter; Abraham, Edo

DOI
10.1109/IECON48115.2021.9589097

Publication date
2021

Document Version
Final published version

Published in
IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society

Citation (APA)
van der Heijden, T., Palensky, P., & Abraham, E. (2021). Probabilistic DAM price forecasting using a combined Quantile Regression Deep Neural Network with less-crossing quantiles. In *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society: Proceedings* Article 9589097 (IECON 2021 - 47TH ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY). IEEE. https://doi.org/10.1109/IECON48115.2021.9589097

Probabilistic DAM price forecasting using a combined Quantile Regression Deep Neural Network with less-crossing quantiles

1st Ties van der Heijden
Electrical Sustainable Energy
Water Resources Engineering
Delft University of Technology
Delft, the Netherlands
T.J.T.vanderHeijden-2@tudelft.nl

2nd Peter Palensky
Electrical Sustainable Energy
Delft University of Technology
Delft, the Netherlands
0000-0003-3183-4705

3rd Edo Abraham
Water Resources Engineering
Delft University of Technology
Delft, the Netherlands
0000-0003-0989-5456

IECON 2021 - 47th Annual Conference of the IEEE Industrial Electronics Society | 978-1-6654-3554-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/IECON48115.2021.9589097

***Abstract*—In this paper we propose a Quantile Regression Deep Neural Network capable of forecasting multiple quantiles in one model using a combined quantile loss function, and apply it to probabilistically forecast the prices of 8 European Day Ahead Markets. We show that the proposed loss function significantly reduces the quantile crossing problem to (near) 0% in all markets considered, while in some cases simultaneously increasing forecasting performance based on classical point forecast metrics applied to the expected value of the probabilistic forecast. The models are optimized using an automated approach with an elaborate feature- and hyperparameter search space, leading to good model performance in all considered markets.**

***Index Terms*—Quantile Regression, Electricity Price Forecasting, Deep Neural Network, Day Ahead Market, Crossing Quantile Problem**

I. Introduction

Due to the increasing penetration of renewables in energy systems around the world, electricity generation is becoming more volatile. Consequently, electricity prices can become more volatile [1], and harder to forecast [2]–[5]. Forecasting prices accurately benefits market-based demand response, resulting in a shift in demand due to price differences. In Europe, the main market for short-term trading is the Day Ahead Market (DAM). On the DAM energy is traded in hourly blocks, with hourly prices. Market participants make a bid before 12:00 AM at d-1, after which the market closes and the Market Clearing Price (MCP) is decided. The actual price is unknown when making a bid, motivating research in Electricity Price Forecasting (EPF) in the context of the DAM.

Many different modelling approaches have been applied to EPF in DAMs. Machine Learning (ML) methods have been proven to be effective in EPF [6]–[8]. The MultiLayer Perceptron was previously successfully applied to the Spanish and Pennsylvania-New Jersey-Maryland (PJM) electricity markets [9], [10]. The Deep Neural Network (DNN) has shown success in forecasting Belgian [11], [12], Dutch [13], Nordpool, German, France and PJM markets [12]. While the DNN has been shown to be successful in price forecasting, the Lasso Estimated Auto Regressive (LEAR) [14] was shown to be a competitive non-ML method, even leading to the highest forecasting accuracy in a benchmark study on the Dutch market. However, as prices are becoming more volatile, DNNs could start outperforming regularized linear regression models like the LEAR [13].

Large forecasting errors can lead to sub-optimal dispatching and a loss in both system efficiency and profits for the user/producer. Since electricity prices are becoming harder to forecast, probabilistic forecasting can be added value since it gives a prediction interval, which is an indication of the forecast uncertainty. It allows for risk management and stochastic bidding/optimisation of assets [15]. Probabilistic forecasting gained track in the energy sector after GEFCOM2014, where the probabilistic forecasts outperformed point forecasting methods [16].

One way to perform a probabilistic forecast is through Quantile Regression [17] (QR). With QR, a model is trained using a Quantile Loss or Pinball Loss function, where over- and under-predictions are penalised differently. Using QR, a conditional estimate (i.e. conditioned by a set of explanatory variables or features) of the probability distribution can be constructed that is likely to contain the real value [18]. QR has been around for many years, but has recently been applied to Neural Networks [19]–[21].

When quantiles are estimated independently using multiple DNNs, it is possible that quantiles do not monotonically increase in value. This is known as the crossing quantile problem, and it is regarded as a serious modelling problem, possibly leading to invalid response distributions [22]. Multiple approaches have been applied to prevent crossing quantiles. A 2-stage model was made to estimate the quantiles after a point forecast is made [21] but quantiles are estimated simultaneously. Another approach is to develop a specific NN where the model is constrained using training [19]. Also, smoothing the loss function has been applied with a penatly and weight constraints during training [20].

This work was supported by: TKI Watertechnologie.

In this paper, we propose a combined quantile DNN (CQR-DNN) trained with a mean quantile loss function. Through this loss function, several quantile forecasts are combined into a single DNN. The method is applied it to 8 European Day Ahead Markets, using an increasing amount of quantiles. The respective forecast performance is evaluated in both point forecast metrics as probabilistic forecasting metrics. In order to validate the added value of the proposed methodology, a benchmark is performed using separate QR-DNNs that are trained to forecast a single quantile. The models are optimized by constructing an elaborate feature- and hyperparameter search space and using the Tree Parzen Estimator algorithm to optimize both features and hyperparameters simultaneously. The automated approach leads to good forecasting performance for all markets considered. Using the CQR-DNNs, quantile crossing is substantially reduced compared to the separate QR-DNNs. We evaluate the forecasts on their Probability Interval Coverage Percentage (PICP), the Probability Interval Normalized Averaged Width (PINAW), the Winkler score [23], and on the Mean Absolute Error (MAE) and relative MAE (rMAE) [24] of the expected value of the forecasts. We show that in some cases the expected value of the forecast improves over a single model trained with the mean absolute error, possibly indicating an improved generalisability of the models.

II. Methodology

In this section we will discuss the classical QR-DNN, the proposed CQR-DNN, the Tree Parzen Estimator for feature- and hyperparameter optimisation, and several evaluation criteria.

A. Quantile Regression Deep Neural Network

A DNN is a feed forward Neural Network, trained using back-propagation. During training, a loss function is minimized to optimise the weights and biases of the activation function in the neurons. In our case, the ReLU activation function is applied. The Adam [25] optimizer was used to optimise the weights and biases. The loss function indicates the goal of the optimisation, in EPF it is common to minimise the Mean Absolute Error (MAE) for point forecasting. For quantile forecasts, the pinball loss function [17]

$$L_\tau = \max(\tau e, (1-\tau)e), \tag{1}$$

can be applied. In the equation, L_τ is the pinball loss of the quantile τ and error e. In the pinball loss, over- and under-forecasting is penalised differently. This leads to a quantile forecast where $\tau\%$ of the observations would be lower than the forecast value. When the loss function is applied to train neural networks to forecast separate quantiles (e.g. the QR-DNN), the crossing quantile problem occurs. The crossing quantile problem occurs when the quantile forecasts are not monotonically increasing over the probability, which is contradictory to the definition of quantiles. The phenomenon can be partially explained due to differences in the stochastic optimisation during training, however the black-box and non-linear properties of DNNs complicate investigating the cause. Neural Networks have been trained with non-crossing constraints [19], however this was found to be quite computationally expensive. We propose a combined quantile loss function, where the mean pinball loss of all forecast quantiles is minimised. The combined quantile loss function penalizes each output node of the Neural Network differently, according to a different quantile loss. The average pinball loss over all quantiles

$$L_{CQ} = \frac{1}{N}\sum_{n=1}^{N} L_{\tau_n}, \tag{2}$$

is minimised when training the DNN for N quantiles. In Figure1, the CQR-DNN is shown for 5 quantiles (or percentiles). Forecasting all quantiles with the same model makes sure the stochastics in training can't affect the quantiles differently, since they are trained with exactly the same routine. Besides that, combining the quantiles allows the model to learn relationships between quantiles, like non-crossing properties.

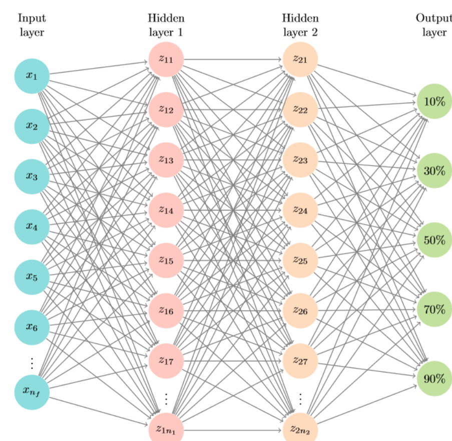

Fig. 1. A Combined Quantile Regression Deep Neural Network with n_f features, n_1 nodes in hidden layer 1, n_2 nodes in hidden layer 2 and 5 quantile output nodes of the same random variable.

B. Tree Parzen Estimator

The Tree Parzen Estimator [26] (TPE) algorithm was applied to optimise features and hyperparameters simultaneously. The TPE is an efficient Sequential Models Based Optimisation approach [27], [28], where a surrogate model is built using Bayes rules and a defined search space. The surrogate model describes the probability of the loss being higher ($h(x)$) and lower ($l(x)$) than a certain threshold value (y^*), as a function of the search space instantiation

$$p(y|x) = \frac{p(x|y) * p(y)}{p(x)}, \tag{3}$$

where y is model performance and x is a search space instantiation. Model performance is therefore estimated as a function of the features and hyperparameters, where $p(x|y)$ is defined as

$$p(x|y) = \begin{cases} l(x) & \text{if } y < y^* \\ h(x) & \text{if } y \geq y^*. \end{cases} \tag{4}$$

In the TPE algorithm, samples are taken from both $l(x)$ and $h(x)$, after which the fraction $\frac{l(x)}{h(x)}$ is evaluated for all samples. The next suggested candidate is the candidate with the highest expected improvement, i.e. the candidate with the largest ratio between low and high probability in $l(x)$ and $h(x)$ respectively.

C. Evaluation criteria

To evaluate a model's forecasting performance, several metrics are used. Both point-forecast metrics and interval metrics are used. When a point forecasting metric is applied to a stochastic forecast, the expected value of the forecast is used to calculate the metric. As custom in EPF using point forecasts, the Mean Absolute Error (MAE) and the relative-MAE [24] (rMAE) are calculated for all models. The rMAE shows the MAE relative to the MAE of a naive forecast. Besides that, the Prediction Interval Coverage Percentage (PICP) and Prediction Interval Normalised Average Width (PINAW) are used. The PICP shows the percentage of variables that are covered in the 80% prediction interval (PI) (quantile 0.1 - 0.9). The PINAW represents the average width of the PI as a percentage of the maximum observed price range. Besides that, we show the percentage of observed crossing quantiles to demonstrate the added value of the combined quantile loss function of Equation (2). Finally, the Winkler Score (WS) [23] is applied to score the quantile forecasts for certain prediction intervals. It is calculated using

$$W_{\alpha,t} = \begin{cases} (u_{\alpha,t} - l_{\alpha,t}) + \frac{2}{\alpha}(l_{\alpha,t} - y_t) & \text{if } y_t < l_{\alpha,t} \\ (u_{\alpha,t} - l_{\alpha,t}) & \text{if } l_{\alpha,t} \le y_t \le u_{\alpha,t} \\ (u_{\alpha,t} - l_{\alpha,t}) + \frac{2}{\alpha}(y_t - u_{\alpha,t}) & \text{if } y_t > u_{\alpha,t}, \end{cases} \tag{5}$$

where the Winkler score $W_{\alpha,t}$ is calculated for the α% (PI), given by $[l_{\alpha,t},\ u_{\alpha,t}]$. The Winkler score calculated the PI width, and penalises any observation y_t that falls outside of the PI.

III. Data

The data and features included in the model are open-source data from the ENTSO-E transparency platform [29] exclusively. The data that was used are the historic DAM prices, historic load and the day-ahead load forecast and the day-ahead renewable generation forecasts. Due to the lack of good actual generation data by energy source on ENTSO-E, the day-ahead generation forecasts are used. The day-ahead forecasts might even contain more information on DAM prices than actual generation, due to the day-ahead market closure of the DAM. For the analysis performed in this paper, data from 2015-2019 was used. Data from 2019 was used to test the models. To be able to include the most recent data (2018) for training, data from 2017 was used as a validation set for feature selection, hyperparameter selection and early stopping during training. Information leakage is limited to a minimum through lagged variables. No data leakage occurs in the test set. Figure 2 shows the price, load and load forecast of France with the train, validation and test sets clearly indicated by color.

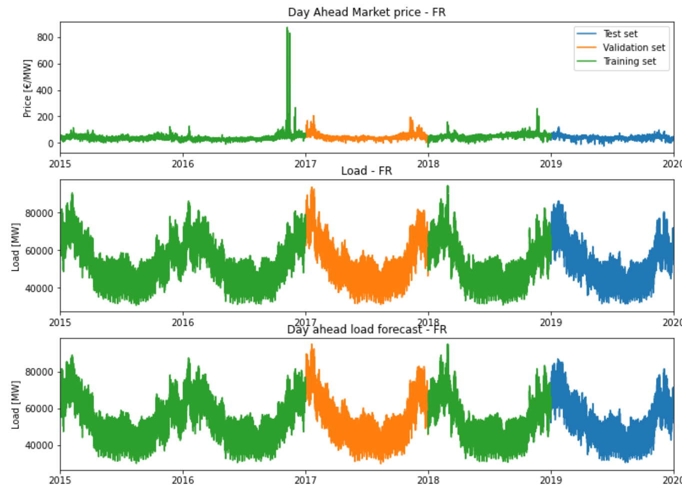

Fig. 2. French historic price, load and load-forecast with train, validation and test splits indicated for model training.

IV. Feature and hyperparameter optimisation

Features and hyperparameters are optimised simultaneously using the TPE algorithm described in Section II-B. The hyperparameters considered in the optimisation of the DNN and the corresponding search range in the optimisation are shown in Table I.

TABLE I
HYPERPARAMETER SEARCH SPACE FOR THE DNN MODEL OPTIMISATION USING THE TPE ALGORITHM.

Hyperparameter	Variable type	Search space
Number of layers	Integer	[1,2]
Nodes per layer	Integer	Layer 1: 50 - 450 Layer 2: 50 - 250
Dropout rate	Continuous	0 - 0.5
Batch size	Integer	7^1 - 7^4
Regularisation (l_2-norm)	Continuous	$1e^{-5}$ - $5e^{-2}$
Batch normalisation	Binary	[False, True]
Random seed	Integer	1 - 300

The feature search space is quite extensive to account for large difference between markets. First, the length of the training data is considered in the search. Because markets aren't stationary, the optimal data training length can differ between the different bidding zones. The considered training data lengths are 1 (2018), 2 (2018, 2016) and 3 (2018, 2016, 2015) years. Different price series can have different auto-correlation, therefore the optimal amount of lagged prices that are used in the forecast can differ as well. The prices of d-1, d-2, d-3 and d-7 are always considered, however the lagged prices from d-4 to d-6 are included in the search space. The load of d-1 is always considered, where the load of d-2 to d-7 are included in the search space. The load of d-7 is added to the search space as a separate binary choice option, so it can be included as a feature independently from the other lagged variables. Similarly, the load forecast for the next day is included in the search space as a binary choice option. Renewable energy generation forecasts is included in the search space for features as binary choice option per renewable generation source (solar, wind onshore, wind offshore). The generation forecasts are included in the same way as the load is included in the model features, meaning that the generation forecasts are included on the days at which either the actual load or load forecast are included as features as well. Finally, including European market integration features in a price forecasting model has been shown to improve forecast accuracy in both the Netherlands [13] and Belgium [30]. For all considered markets, the external market with highest feature importance [13] was picked as a candidate feature. For simplicity, only load features are considered as they were shown to be an important feature in the Belgian market [30]. Similarly to the renewable energy forecasts, the market integration features 'follow' the native load features. The features included in the search space, and their corresponding search range are shown in Table II.

TABLE II
FEATURE SEARCH SPACE FOR THE DNN MODEL OPTIMISATION USING THE TPE ALGORITHM.

Feature	Variable type	Search space
N of years in training data	Integer	1 - 3 years
N of lagged price days	Integer	3 - 6 days
N of lagged load days	Integer	1 - 6 days
Load of d-7	Binary	[False, True]
Load forecast	Binary	[False, True]
Solar energy generation	Binary	[False, True]
Onshore wind generation	Binary	[False, True]
Offshore wind generation	Binary	[False, True]
EU market integration	Binary	[False, True]

V. Results and discussion

In this section we will discuss some of the features and hyperparameters resulting from the TPE search, the forecasting performance of the point forecasts, and the quantile regression forecasts.

A. Feature and hyperparameter optimisation

The feature optimisation results can be seen in Table III. In all considered markets, the maximum amount of training data resulted in the lowest validation losses. The amount of lagged price and load days do vary over the markets. The load forecast seems to be a good feature in all markets, and the load of d-7 in most. For a considerable amount of markets, external market features are included in the best performing model. In all markets except the French, which relies heavily on nuclear energy generation, renewable energy generation was included as a feature. For the French case, their (inflexible) nuclear power generation capacity could explain why no market integration features were selected. Similarly, the flexible hydro-dominated Norwegian generation could explain why Norwegian features are good for other markets, but the relationship goes one way [13].

B. Point forecast performance

Figure 3 shows the rMAE of the models, showing the relative performance of the CQR-DNN compared to the naive forecast. The feature- and hyperparameter optimisation procedure resulted in competitive forecasting performance. For the Dutch case, the model beats the best performing DNN in an earlier benchmark [13]. A benchmark involving multiple markets reports slightly worse performance in general, but these were tested on a different year. Table III also shows the absolute MAE in [€/MWh]. Figure3 also shows that in

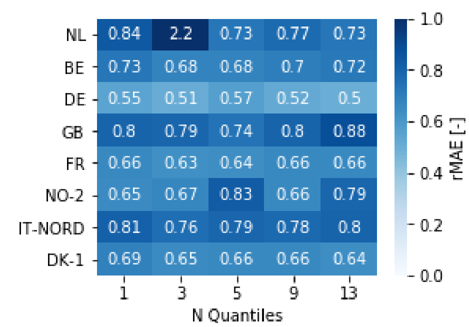

Fig. 3. rMAE of the expected value of the QR-DNN's over the amount of quantiles that are forecast.

general, the point forecast loss is decreased when multiple quantiles are included in the model. This could be due to an increase generalisation ability of the model. Although for the Dutch case, an outlier is present at the CQR-DNN with three included quantiles in the forecast. Possibly this is due to overfitting on the validation data, leading to a high test loss and low validation loss.

C. Quantile Regression

Figure 4 shows the day-ahead probabilistic forecast for France and the Netherlands using the CQR-DNN forecasting 13 simultaneous quantiles. We evaluate the quantile forecasting models on four main criteria: the percentage of crossing quantiles, the PICP, the PINAW and the Winkler score. The PICP, PINAW and Winkler score are calculated for the 80% prediction interval. Figure 5 shows the percentage of crossing quantiles for both the CQR-DNN and the regular QR-DNN. It shows that with our proposed loss function described in Equation 2, quantile crossing occurs significantly less. Especially when a large amount of quantiles are included in the model.

Figures 6 and 7 show the PICP and the PINAW for the CQR-DNN and the QR-DNN. The PICP and PINAW don't differ over the amount of quantiles for the separate quantile forecasts, since the quantiles are only trained once and the prediction interval stays the same (80%). The method resulting in the highest PICP differs between the markets, but the PINAW is generally lower for the CQR-DNN. A narrow PI with similar coverage probability would be preferable for a probabilistic forecast.

Figure 8 shows the Winkler score at the 80% PI for the CQR-DNN and the separate quantile forecasts. A low Winkler

TABLE III

FEATURES RESULTING FROM THE TPE OPTIMISATION, WHERE Θ_1 IS THE LENGTH OF THE TRAINING DATA IN YEARS, Θ_2 THE AMOUNT OF LAGGED PRICE DAYS, Θ_3 THE AMOUNT OF LAGGED LOAD DAYS, Θ_4 WHETHER THE LOAD FORECAST IS USED, Θ_5 WHETHER THE LOAD OF D-7 IS USED, Θ_6 SHOWS WHETHER EU MARKET INTEGRATION FEATURES ARE CONSIDERED WITH THE EXTERNAL MARKET IN BRACKETS, Θ_7, Θ_8 AND Θ_9 INDICATE WHETHER SOLAR, ONSHORE WIND AND OFFSHORE WIND GENERATION DATA IS USED. THE MODELS WITH LOWEST MAE IN [€/MWH] IS PICKED OVER ALL QR-DNNS WITH VARYING AMOUNT OF QUANTILES.

	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	Θ_6	Θ_7	Θ_8	Θ_9	N Quantiles	MAE
BE	3	6	1	True	False	True (FR)	False	False	True	3	5.31
DE	3	3	1	True	True	False (GB)	True	True	True	13	4.74
DK-1	3	4	1	True	True	True (NO-2)	False	True	True	13	5.18
FR	3	5	3	True	True	False (NO-2)	False	False	False	3	4.14
GB	3	3	2	True	True	True (NL)	True	True	True	5	4.04
IT-NORD	3	3	1	True	False	True (GB)	False	True	False	3	4.27
NL	3	6	3	True	True	True (NO-2)	False	True	False	5	4.14
NO-2	3	3	1	True	True	False (IT-NORD)	False	False	True	1	1.68

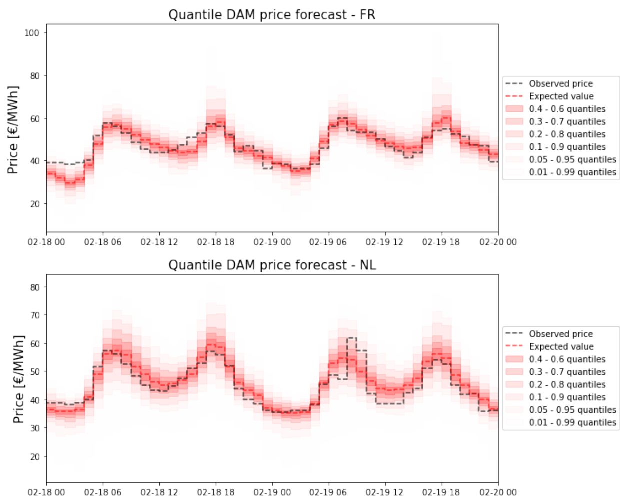

Fig. 4. Day-ahead quantile forecast for French and Dutch DAM price.

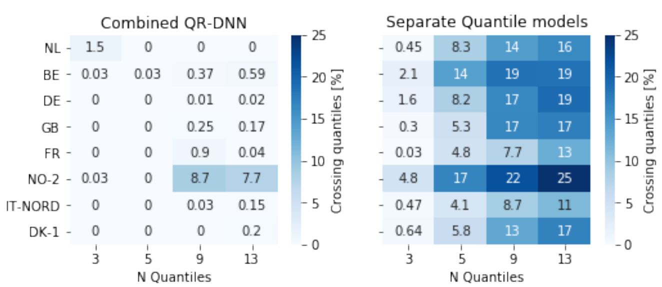

Fig. 5. Percentage of crossing quantiles per amount of quantiles included in the forecast, for both the QR-DNN and the separate quantile forecasts.

score indicates a higher performance. The score is calculated using the PI width, while penalising observations that fall outside the PI. For all considered markets, the CQR-DNN leads to a quantile forecast with lower Winkler score than the separate QR-DNN's.

VI. CONCLUSIONS

In this paper, we propose a loss function that can be applied to forecast multiple quantiles using a single DNN. The resulting CQR-DNN is applied to forecast the DAM prices of 8 European bidding zones, and compared with regular QR-DNN's trained to forecast a single quantile per model. The models were optimised for both features and hyperparameters, using an elaborate search space and the TPE algorithm. The automated approach lead to good results for all markets considered. We show that by using the CQR-DNN, quantile crossing occurs significantly less, with most CQR-DNNs not showing any crossing quantiles. This can be explained by the elimination of differences in training between quantiles due to the stochastic nature of Machine Learning methods. Also, by having all quantiles present in the output layer allows the model to find relationships between quantiles, like non-crossing properties. The proposed CQR-DNN also outperforms the standard DNN on point-forecast metrics applied to the expected value. This can be explained by an increased generalisation ability of the model, due to the

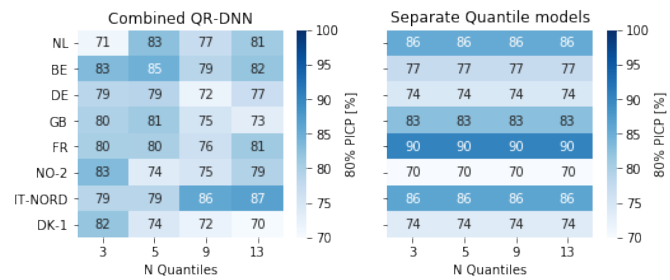

Fig. 6. The 80% prediction interval coverage probability for the CQR-DNN and the QR-DNN's.

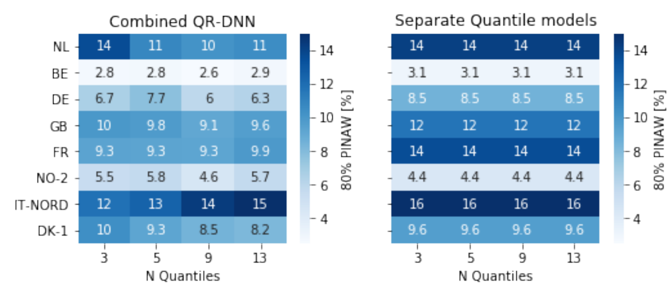

Fig. 7. The 80% prediction interval normalised average width for the CQR-DNN and the QR-DNN's.

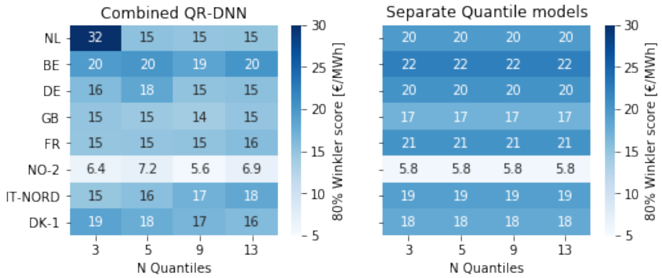

Fig. 8. The 80% PI Winkler score for the QR-DNN's and the separate quantile forecasts.

inclusion of multiple quantiles. Also, the loss function could be smoothed due to the combination of all quantile losses, leading to more efficient training. Finally, it could be that the knowledge of surrounding quantiles helps with estimating the expected value. Based on the PICP, the two approaches are equally good. However, the PINAW of the combined QR-DNN is lower for most markets. A narrow PINAW, combined with a high PICP is preferable for a stochastic forecast. The Winkler score does favor the CQR-DNN, with lower scores for all considered markets compared to the separate quantile forecasts.

References

[1] S. Mosquera-López and A. Nursimulu, "Drivers of electricity price dynamics: Comparative analysis of spot and futures markets," *Energy Policy*, vol. 126, no. October 2018, pp. 76–87, 2019. [Online]. Available: https://doi.org/10.1016/j.enpol.2018.11.020

[2] K. Maciejowska, "Assessing the impact of renewable energy sources on the electricity price level and variability – A quantile regression approach," *Energy Economics*, vol. 85, p. 104532, 2020. [Online]. Available: https://doi.org/10.1016/j.eneco.2019.104532

[3] L. Grossi and F. Nan, "Robust forecasting of electricity prices: Simulations, models and the impact of renewable sources," *Technological Forecasting and Social Change*, vol. 141, no. May 2018, pp. 305–318, 2019. [Online]. Available: https://doi.org/10.1016/j.techfore.2019.01.006

[4] A. Gianfreda, L. Parisio, and M. Pelagatti, "The impact of RES in the Italian day-Ahead and balancing markets," *Energy Journal*, vol. 37, no. Special Issue 2, pp. 161–184, 2016.

[5] C. Brancucci Martinez-Anido, G. Brinkman, and B. M. Hodge, "The impact of wind power on electricity prices," *Renewable Energy*, vol. 94, pp. 474–487, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.renene.2016.03.053

[6] 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, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.ijforecast.2014.08.008

[7] M. Narajewski and F. Ziel, "Ensemble Forecasting for Intraday Electricity Prices: Simulating Trajectories," *Applied Energy*, vol. 279, no. April, p. 115801, 2020. [Online]. Available: http://arxiv.org/abs/2005.01365

[8] C. Monteiro, I. J. Ramirez-Rosado, L. A. Fernandez-Jimenez, and P. Conde, "Short-term price forecasting models based on artificial neural networks for intraday sessions in the Iberian electricity market," *Energies*, vol. 9, no. 9, 2016.

[9] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied Energy*, vol. 162, pp. 218–230, 2016.

[10] P. M. Bento, J. A. Pombo, M. R. Calado, and S. J. Mariano, "A bat optimized neural network and wavelet transform approach for short-term price forecasting," *Applied Energy*, vol. 210, no. October 2017, pp. 88–97, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2017.10.058

[11] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, no. January, pp. 386–405, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2018.02.069

[12] J. Lago, G. Marcjasz, B. D. Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Renewable and Sustainable Energy Reviews*, 2020 (Under review).

[13] T. van der Heijden, J. Lago, P. Palensky, and E. Abraham, "Electricity price forecasting in European Day Ahead Markets: a greedy consideration of market integration," *Under review at IEEE Access*, 2021.

[14] B. Uniejewski, J. Nowotarski, and R. Weron, "Automated variable selection and shrinkage for day-ahead electricity price forecasting," *Energies*, vol. 9, no. 8, 2016.

[15] C. Pape, A. Vogler, O. Woll, and C. Weber, "Forecasting the Distributions of Hourly Electricity Spot Prices - Accounting for Serial Correlation Patterns and Non-Normality of Price Distributions," *SSRN Electronic Journal*, 2018.

[16] R. Juban, H. Ohlsson, M. Maasoumy, L. Poirier, and J. Z. Kolter, "A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1094–1102, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ijforecast.2015.12.002

[17] R. Koenker and G. Bassett, "Regression Quantiles," *Econometrica*, vol. 46, no. 1, p. 33, 1978.

[18] J. R. Andrade, J. Filipe, M. Reis, and R. J. Bessa, "Probabilistic price forecasting for day-ahead and intraday markets: Beyond the statistical model," *Sustainability (Switzerland)*, vol. 9, no. 11, 2017.

[19] A. J. Cannon, "Non-crossing nonlinear regression quantiles by monotone composite quantile regression neural network, with application to rainfall extremes," *Stochastic Environmental Research and Risk Assessment*, vol. 32, no. 11, pp. 3207–3225, 2018. [Online]. Available: https://doi.org/10.1007/s00477-018-1573-6

[20] K. Hatalis, A. J. Lamadrid, K. Scheinberg, and S. Kishore, "A novel smoothed loss and penalty function for noncrossing composite quantile estimation via deep neural networks," *CoRR*, vol. abs/1909.12122, 2019. [Online]. Available: http://arxiv.org/abs/1909.12122

[21] Q. Xu, K. Deng, C. Jiang, F. Sun, and X. Huang, "Composite quantile regression neural network with applications," *Expert Systems with Applications*, vol. 76, pp. 129–139, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2017.01.054

[22] D. Ouali, F. Chebana, and T. B. Ouarda, "Quantile regression in regional frequency analysis: A better exploitation of the available information," *Journal of Hydrometeorology*, vol. 17, no. 6, pp. 1869–1883, 2016.

[23] R. L. Winkler, "A decision-theoretic approach to interval estimation," *Journal of the American Statistical Association*, vol. 67, no. 337, pp. 187–191, 1972.

[24] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[25] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 12 2014.

[26] J. Bergstra, D. Yamins, and D. Cox, "Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures," *Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, 2013*, 2013.

[27] F. Madrigal, C. Maurice, and F. Lerasle, "Hyper-parameter optimization tools comparison for multiple object tracking applications," *Machine Vision and Applications*, vol. 30, no. 2, pp. 269–289, 2019.

[28] K. Eggensperger, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Efficient benchmarking of hyperparameter optimizers via surrogates," *Proceedings of the National Conference on Artificial Intelligence*, vol. 2, pp. 1114–1120, 2015.

[29] ENTSO-E, "Entso-e transparancy platform," 2018. [Online]. Available: https://transparency.entsoe.eu/

[30] J. Lago, F. De Ridder, P. Vrancx, and B. De Schutter, "Forecasting day-ahead electricity prices in Europe: The importance of considering market integration," *Applied Energy*, vol. 211, no. July 2017, pp. 890–903, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2017.11.098