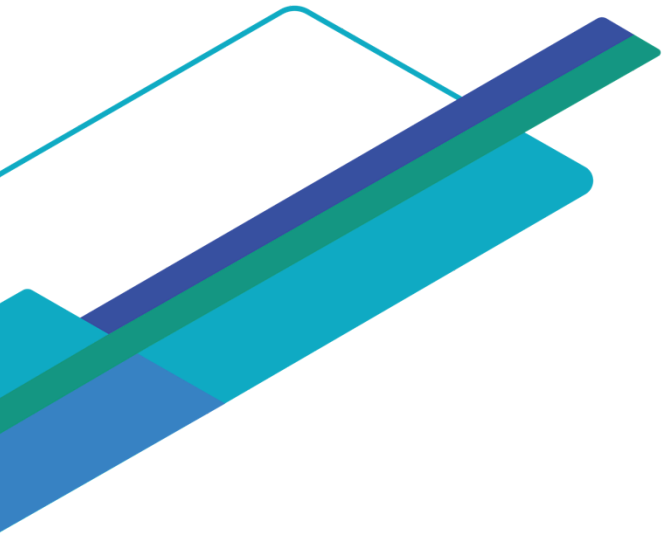
N-SIDE

NGESO - Probabilistic Congestion Management

Phase V

April 2024
Version 1



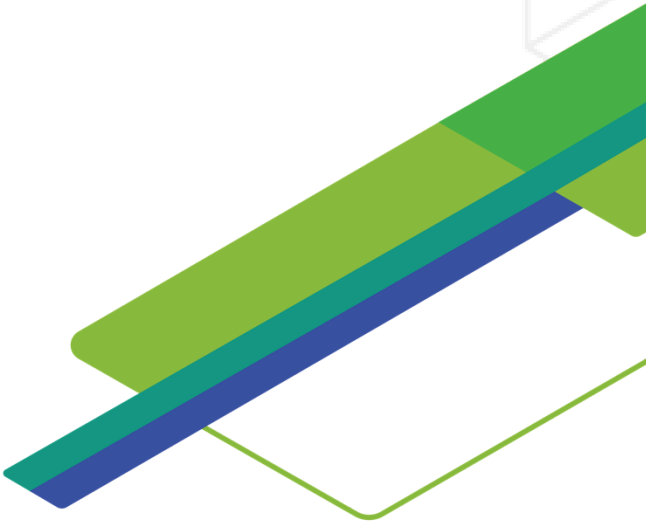


Table of Contents

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

1 Introduction

The GB power grid is becoming more interconnected, and this, together with increased contribution from renewable energy sources, poses some challenges in anticipating energy flows. The volatility inherent to interconnections and renewable energy increases the uncertainty of physical energy flows, complicating the anticipation of internal congestion in the day-ahead market and resulting in more decisions needed within the day by the control room to overcome congestion.

This phase V focuses on the probabilistic forecasting of wind generation after the clearing of the day-ahead market. By accurately estimating next day's wind generation, it becomes possible to assess the impact of uncertainties arising from large wind farms (in addition to those from interconnectors) on network congestion and predict more accurately the probabilistic risk of congestion on specific branches of the power grid.

1.1 High-level Overview of Phase V

This phase of the project involves the development of a supervised learning model to predict the probabilities of day-ahead wind-based electricity generation. This will focus on the three wind-farms located in the South East of England[1]. This phase includes the following steps:

- **Data Mining:** The input features of the model were build through extensive mining of time series data. Once collected, the data was stored in a clean, structured database accessible by the forecasting solution. All time series data align with a standard format using UTC timestamps, ensuring consistency and ease of use.

- **Processing of Input Features:** After executing the necessary manipulations, new data features were developed from existing information, making it usable for the forecasting solution. We construct a selection of time series for use as input features, and we specifically determine the time lags for each time feature.

- **Development of the Wind Generation Forecasting Algorithm:** We select an approach for algorithm development that meets the project use

[1] The selected wind-farms for this project were: Greater Gabbard, London Array and East Anglia One.

case requirements. We anticipated that tree-based regression techniques offer the highest quality probabilistic forecast for this use case.

- **Testing of the Model:** Strict rules were imposed when testing the AI model to ensure reliable and reproducible results. A portion of the dataset is reserved and not used in the testing of the models until the final stage. We further divided the remaining dataset into multiple training and validation datasets to prevent overfitting and ensure accuracy (more details about the forecasting setting are provided in section 3.1).

- **Storage of Predictions, Performance Analysis, and Reporting:** the model generates a point forecast as well as a probability distribution (percentiles) for the wind-based electricity generation associated with each wind farm within scope for each hour of the day. We store these predictions in the database for the period in scope. To evaluate model performance, we focus on the accuracy of the percentile predictions, specifically examining bias and variance.

In the following sections of this document, we detail our approach to the forecasting process. We will outline our data sources, explain our methodology, and provide a thorough analysis and assessment of the performance metrics. Every step will be thoroughly explored to give a complete understanding of our forecasting results.

2 Data Mining

2.1 Data Sources

Data sources about renewable generation, weather and interconnection flows were used to develop the wind generation forecasting algorithm. Some iterations were necessary to test new features and evaluate their impact on performance to find the optimal feature set. This section provides an overview of all the features that were considered during the algorithm development. Although all these features were tested, the final model only relies on the most important features that improved performance. For a detailed examination of the most influential features, please refer to Section 3.6. Below is a list of all the considered features, grouped by country and feature category:

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

Region	Flow features	RES features	Weather Forecast features
GB	DA scheduled flow for all interconnectors Physical flow from the last reference day for all interconnectors	GB solar generation forecast GB wind generation forecast Differences between solar generation forecasts made at different times Differences between wind generation forecasts made at different times	Temperature forecast at different GB locations[2] Wind speed forecast at different GB locations Atmospheric pressure forecast at different GB locations Solar radiation forecast at different GB locations

[2] GB locations considered: London, Manchester, Birmingham, Cardiff, Southampton, Irish sea, North sea, Scottish Highlands.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

Region	Flow features	RES features	Weather Forecast features
BE	DA scheduled flow for all interconnectors Physical flow from the last reference day for all interconnectors	N/A	Temperature forecast at different BE locations[3] Wind speed forecast at different BE locations Atmospheric pressure forecast at different BE locations Solar radiation forecast at different BE locations

Region	Flow features	RES features	Weather Forecast features
FR	N/A	N/A	Temperature forecast at different FR locations[4]

[3] BE locations considered: Offshore (Belgian coast).
[4] FR locations considered: Dunkerque.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

			Wind speed forecast at different FR locations Atmospheric pressure forecast at different FR locations Solar radiation forecast at FR different locations

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

2.2 Data Quality

Before starting the work on the forecasting algorithm, some checks were done on the input sources used in the model. The objective of this initial phase is to identify any issues with the data: accuracy of input features, identification of outliers, missing values, inconsistent definitions, etc.

Missing values

First of all, a complete check on the feature completeness was carried out for the period ranging from May 31st, 2022 until May 31st, 2023. Figure 1 shows the number of missing values for some of the features with the highest number of missing data.

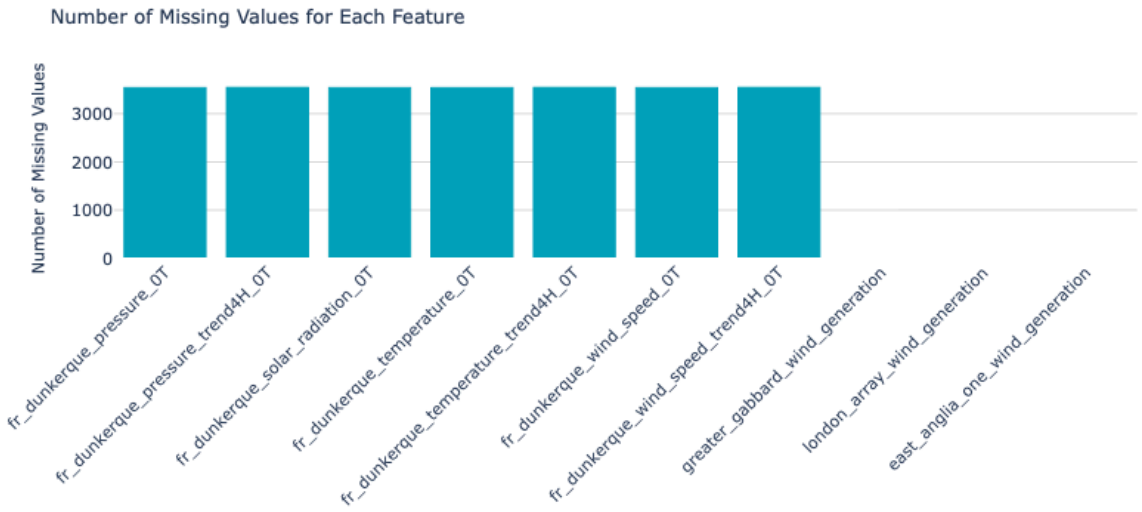


Figure 1. Number of missing values reported for all features characterised by one or more missing values within the training data.

Out of all features considered, 10 of them are characterised by missing values. Amongst those, 7 features actually correspond to weather parameters in Dunkerque, France: all values ranging between May 31st, 2022 and October 25th, 202 are missing due to a late database dump from October 25th for this table.

Regarding the wind farm generation profiles, there's only a single timestamp missing on October 29, 2023, at 11 p.m. This absence can be attributed to the time-shift that occurs on that date in the UK. However, it has a negligible impact on the training process.

Cross-variable relationships

The relationship between some weather variables was also studied to understand how weather parameters correlate with one another.

The first pairwise comparison is between wind speed and atmospheric pressure recorded both at the Cambridge Wind Park. We can expect that higher wind speeds are associated with lower atmospheric pressure as higher wind speeds typically occur in regions with steep pressure gradients, indicating lower atmospheric pressure. This relationship is illustrated in the next Figure 2.

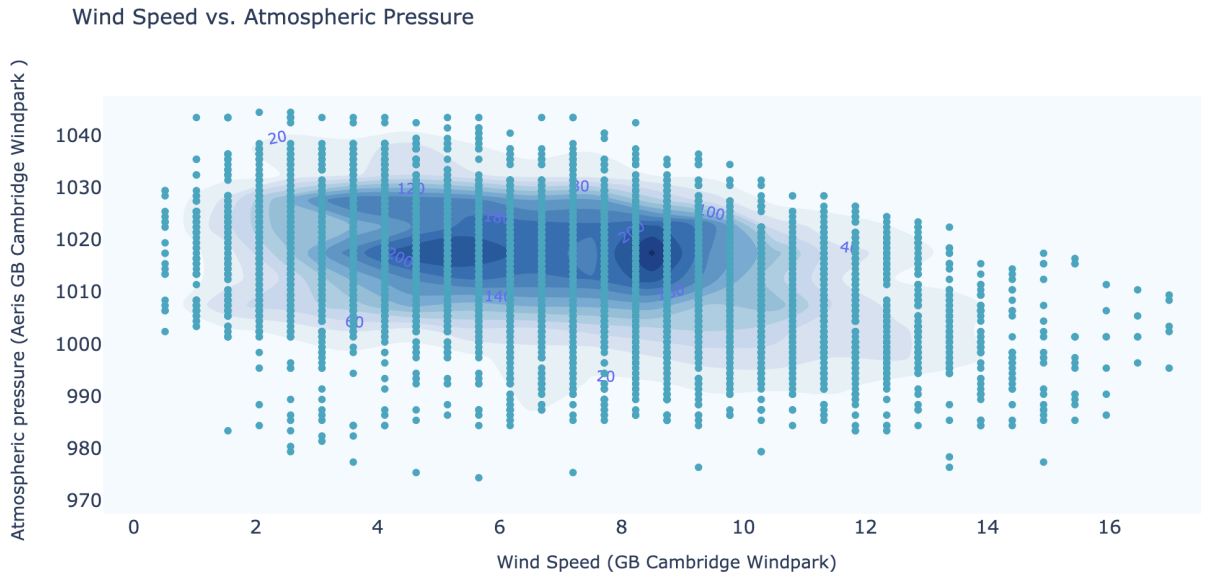


Figure 2: Scatter plot highlighting the relationship between wind speed and atmospheric pressure recorded in the Cambridge Wind Park, GB.

Although some extreme cases are observed, the plot doesn't highlight any irrelevant records. Therefore, all data points from the training data were kept.

The second interesting pairwise comparison is the relationship between wind speed and wind generation.

The following Figure 3 illustrates the relationship between wind speed recorded in the North Sea versus wind generation produced by all three offshore wind farms considered in this phase V.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

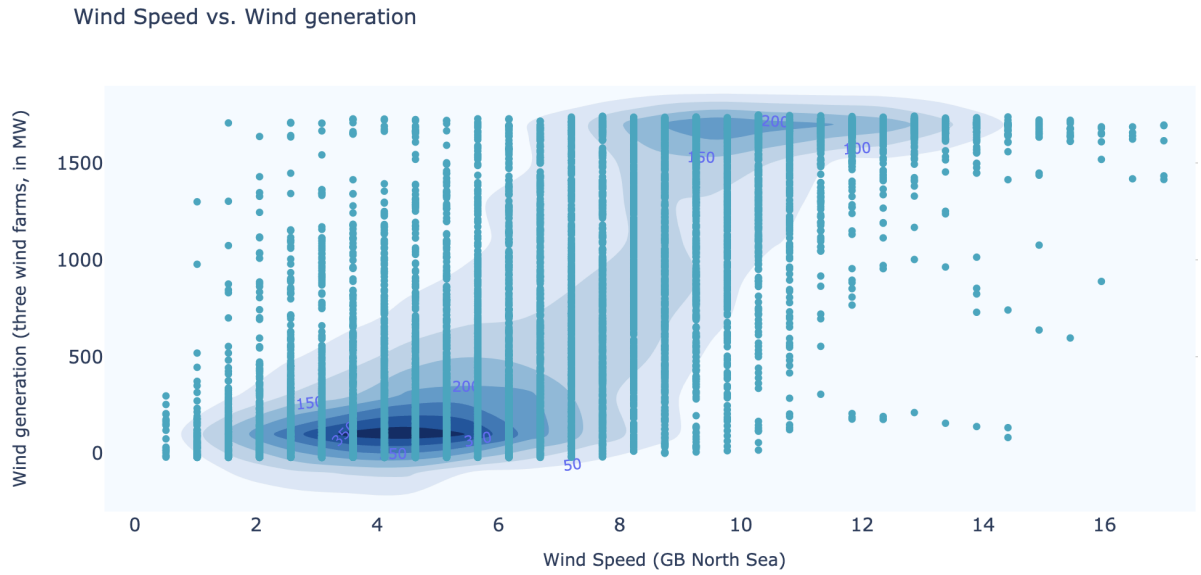


Figure 3: Scatter plot (together with density areas) highlighting the relationship between wind speed and wind generation (produced by three offshore wind farms) in the North Sea.

While a clear correlation is evident, there are instances where exceptionally high wind speeds do not translate into intense electricity production. Notably, no data points were removed from the graph, as the extreme values observed appear consistent with the dynamic nature of wind patterns. This underscores the significance of incorporating wind direction data, an information parameter currently unavailable for algorithm training.

3 Wind Generation Forecasting

3.1 Forecasting Setting

Similar to the setting associated with the interconnector flow forecasting algorithm developed in Phase 1, the forecasting algorithm that would predict 6 p.m. in DA. A visualisation of the timeline is provided in Figure 4.

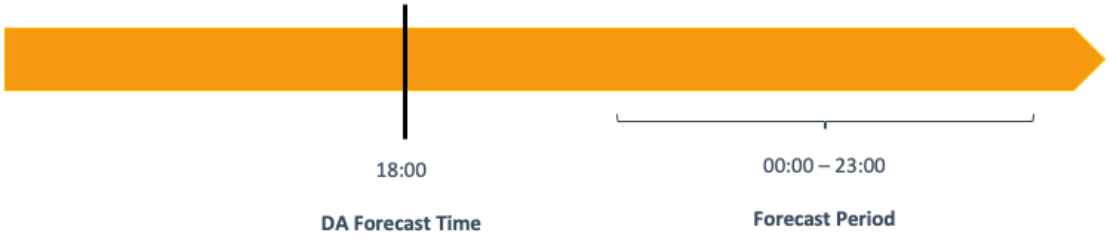


Figure 4: Timeline of the forecaster with the forecast time and forecast period in time.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

During the development of the forecasting algorithm, special care was taken to avoid "data leakage". Data leakage is when the algorithm leverages data that would not be present at the moment of prediction. Moreover, a "rolling window" was implemented to obtain realistic performance estimation. In this setting, the model is trained on historical data (training data) and used to predict unseen data for the following period (prediction data). Re-training happens every month, where the algorithm is trained on shifted training data to make predictions for the following month. A visual representation of this rolling scheme is provided below in Figure 5.

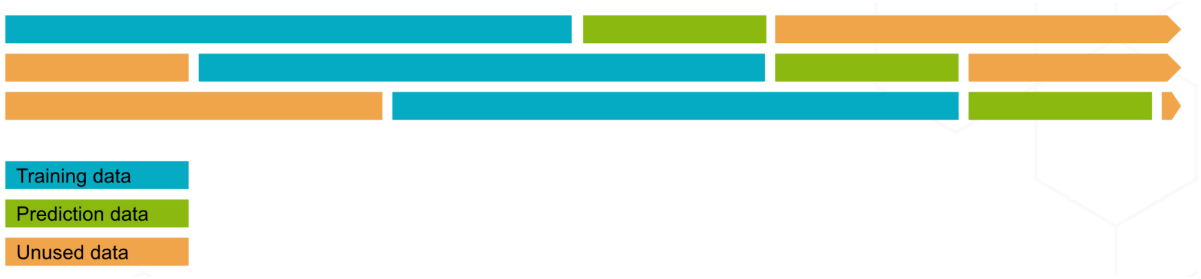


Figure 5: "Rolling window" over time and the re-training strategy used to forecast for the complete prediction period.

In each training/prediction repetition, the training period is 7 months, while the prediction period is 1 month.

3.2 Metrics

Different metrics are used in this project to assess the performance of developed algorithms: Root mean squared error (RMSE), Mean Absolute Error (MAE), bias, and variance.

3.2.1 Root Mean Squared Error (RMSE)

The Root Mean Squared Error is a performance metric to assess to what extent a **point forecast** is close to the actual value:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \hat{x}_i)^2}{N}}$$

with x_i the actual wind generation, $\hat{x}_i$ the predicted wind generation, and N the total number of timestamps for which a prediction is made. The RMSE penalises large errors more compared to smaller errors (due to the squaring of the differences).

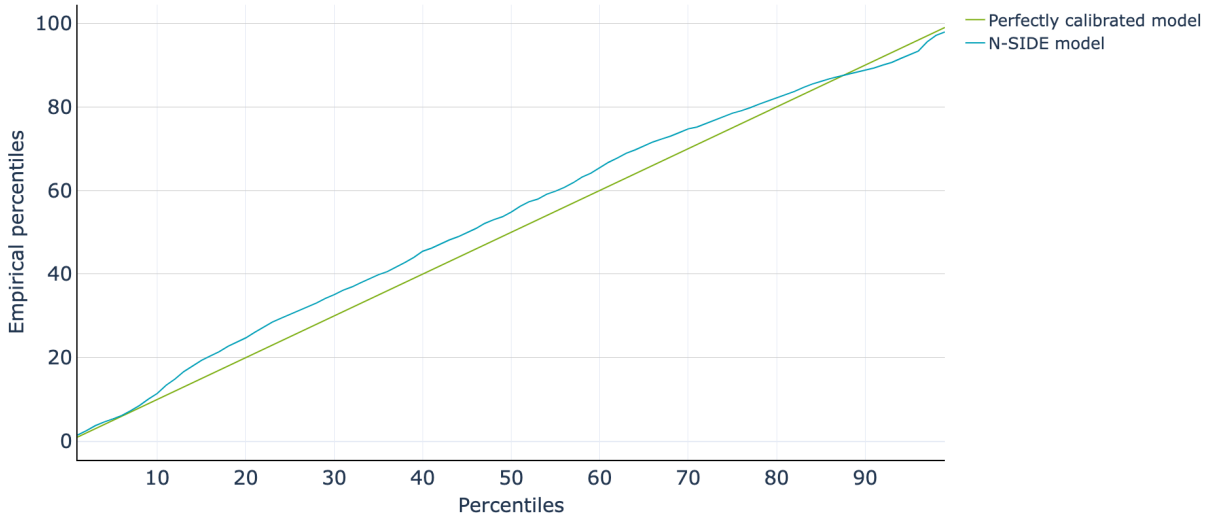
While the RMSE is a good performance metric in point forecasting, RMSE is too limited for probabilistic forecasting.

In a probabilistic setting, one wants not only to have a reasonable point estimate but also a good one for all the considered percentiles of the distribution[5]. Ideally, the probabilistic forecast has (1) a low bias and (2) a low variance. The following sections will elaborate on both metrics to provide the reader with a better understanding of those concepts and the importance of both metrics.

3.2.2 Bias

The bias of the forecasts addresses the question: "**How far off are we?**". In this project, the probabilistic forecast of the wind generation yields 99 values, corresponding to each considered percentile (P1, P2, ..., P99). The value associated with P1 signifies that, when the model is well-calibrated, there is a 1% probability of encountering a value lower than P1.

To assess calibration, the actual wind generation is compared with the corresponding percentile forecasts. A well-calibrated model would have the actual wind production below the P1 forecast 1% of the time, below the P2 forecast 2% of the time, and so on. By plotting the considered percentiles against the empirical percentiles, the bias in predictions can be evaluated. A perfectly calibrated model would result in empirical percentiles matching the considered percentiles, represented by the green curve in Figure 6. The blue curve in the figure depicts empirical percentiles for a certain wind-farm over a specific period.



[5] For this project, 99 percentiles were considered: P1, P2, ..., P99.

Figure 6: Example of the bias plot to assess to what extent the model is well-calibrated.

In this example, the results demonstrate satisfactory performance with a small bias. For most of the cases, it can be observed that the empirical percentiles are slightly higher compared to the predicted percentiles. For example, one observes that approx. 23% of the cases had a higher wind generation compared to the P20 forecast. As it was expected that this would occur in 20% of the cases, there is a slight bias in the predictions.

3.2.3 Variance

The bias is not the only metric to consider when assessing the quality of a probabilistic forecast. Equally important is the variance, which addresses the question: **"How certain are we about our predictions?"** Ideally, the forecaster should have a high level of confidence in the predictions made. Uncertain predicted values correspond to a wide range between the lower and upper percentiles. To illustrate this, please consider the following two figures, which depict the probabilistic forecast under high uncertainty (Figure 7) and relatively lower uncertainty (Figure 8).

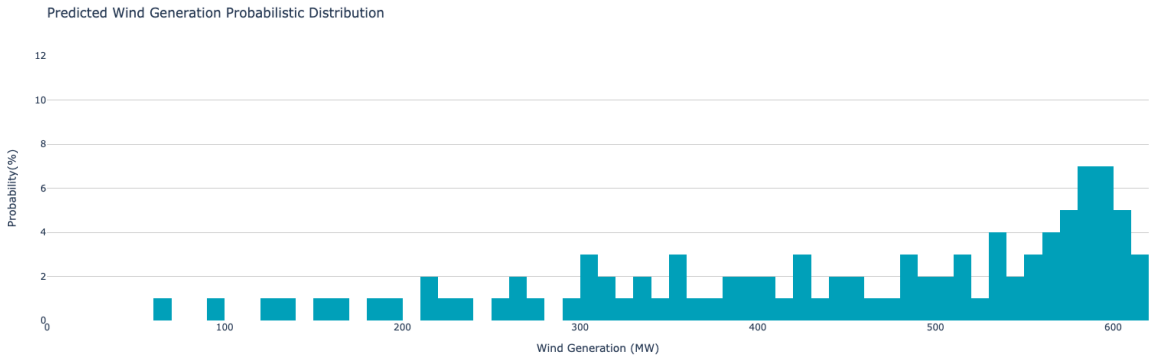


Figure 7: Example of the probabilistic forecast for the wind generation on a certain wind farm for a certain timestamp associated with a high variance.

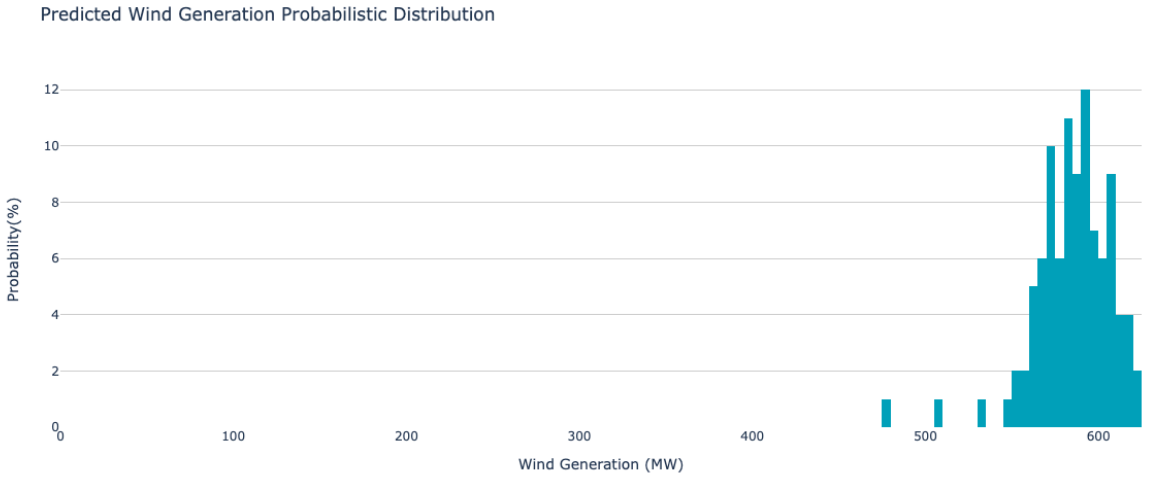


Figure 8: Example of the probabilistic forecast for the wind generation on a certain wind farm for a certain timestamp associated with a low variance.

One way to measure the "uncertainty" in the forecast, is by computing the standard deviation over the predicted percentiles. For each timestamp in the prediction set, one can measure the variance by the standard deviation in the predicted percentiles.

Based on these values, the variance plot can then be constructed by plotting the standard deviation of the predictions (for each timestamp a value) with the standard deviation of actual wind generation (one value of the period of analysis) over the prediction set. Comparing the variance in the predictions with the observed variance would enable us to see periods in which the uncertainty is higher/lower compared to the actuals. Figure 9 shows an example of such a comparison.

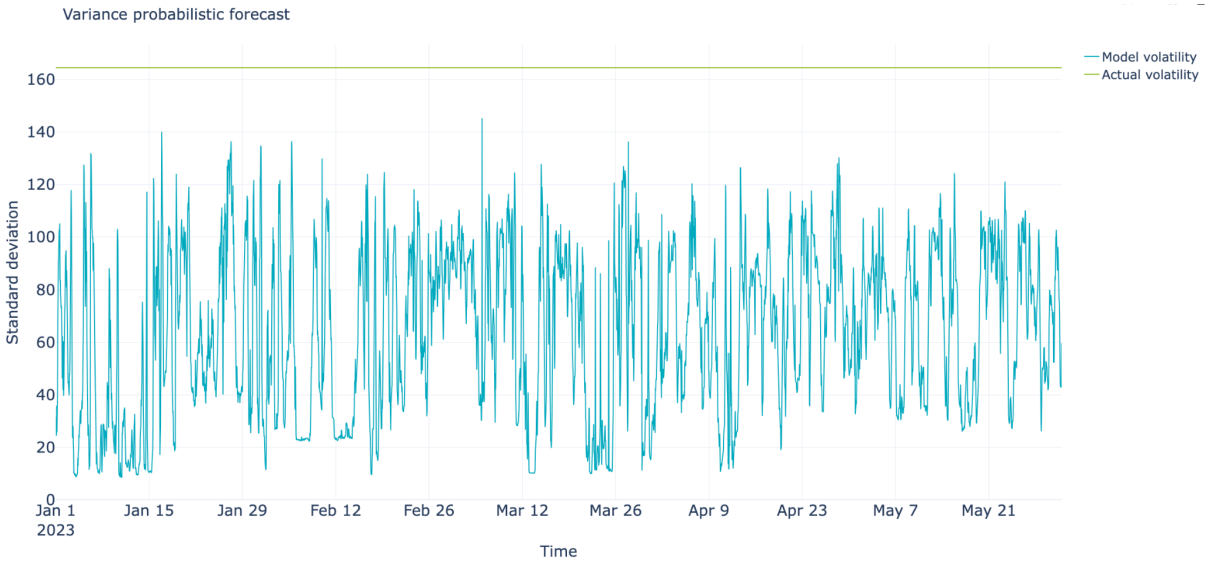


Figure 9: Example of the variance plot. The green horizontal line indicates the actual standard deviation (in MW) in the actual wind generation across the prediction set, whereas the blue curve shows the standard deviation (in MW) for the N-SIDE model for each timestamp ranging from Jan 1st 2023 up to May 31st 2023.

As can be seen, the model is in certain periods characterised by a high variance, whereas the variance is low in other periods. This shows that the model can identify periods of high and low uncertainty when it comes to predicting the wind generation. Some days are more difficult to predict compared to others. Overall, the variance plot also shows that model variance is lower than the actual variance (or volatility).

A low variance in itself is not good enough. For example, imagine a forecaster that always predicts the average of the actual wind generation (as observed in the training set) for all the percentiles. Such a forecaster is characterised by zero variance, but the bias is most likely significant as it does not leverage the situational context to make a forecast. This example shows that the ideal forecaster is characterised by a low bias and variance.

To simplify the comparison of the bias and variance from different models, it was decided to reduce both graphs into two single metrics: the Kolmogorov-Smirnov (KS) statistic and the % reduction in variance.

The **KS statistic** measures the similarity between two distributions. Applied to our specific case, we compute the KS statistic of the perfectly calibrated model and the N-SIDE model in the bias plot. The KS statistic can be measured as the largest difference between both curves. The idea behind this statistical test is that similar distributions should be alike, and thus the largest difference between both curves should be small.The lower the KS statistics the better.

The **% reduction in variance** is computed by comparing the average of the model variance with the actual variance. Ideally, the model can reduce the variance, and thus a higher % reduction in variance is desirable.

3.3 Target Definition

The final objective of the forecasting algorithm is to predict the wind-based electricity generation (in MW) for all the considered wind farms. To delimit the scope of this project, the generation profiles of three large wind farms located in the South East of England were selected for both 2022 and 2023: Greater Gabbard (maximum capacity of 504 MW), London Array (maximum capacity of 630 MW) and East Anglia One (maximum capacity of 714 MW).

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

3.4 Forecasting Model

To model the wind generation, a tree-based algorithm was used. The algorithm aggregates the predictions of (many) individual trees and is by doing so more robust and less prone to overfitting, leading to better generalisation performance.

Intuitively, each tree is rule-based categorisation that divides all the initial cases (training set) into different sets with the idea to group similar cases together and to distinguish distinct cases. Each tree starts with the full training set at the root of the tree. Subsequently, different splits are made which divide the data in two (or more) different categories. A good split would be characterised by having two (or more) groups which have a high heterogeneity between groups and a high homogeneity within each group. In forecasting problems such as this one, typically the RMSE is optimized. A split is thus optimised to create two (or more) groups where the weighted RMSE is reduced.

The following Figure 10 shows visually how a tree is constructed. Intuitively, each box represents a datapoint in the training set. The color of the box would be associated with the wind generation of the wind-farm. A black box could indicate a wind generation value close to the maximum capacity of the wind farm and a white box could indicate a wind generation value close to 0 MW.

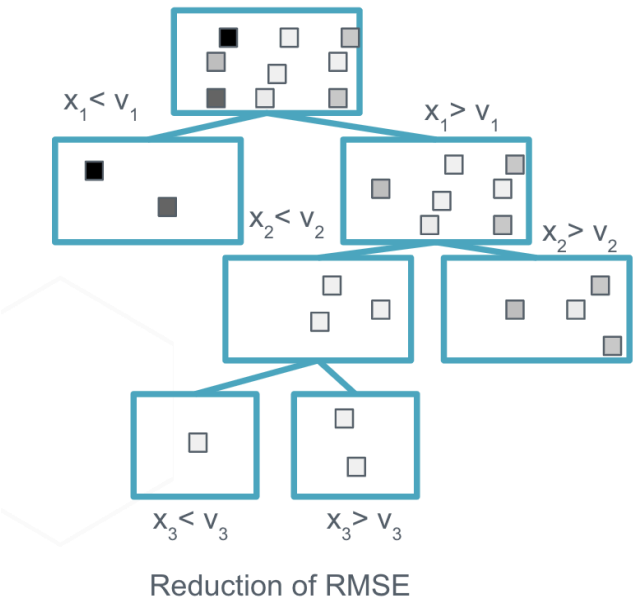


Figure 10: Example tree.

On the top of the figure, the complete training set is contained in the root of the tree. Next, the algorithm looks to optimise the split to minimise the weighted RMSE (across the different child nodes). In this example, the data is split in two categories based on feature X_1 where data points with values for that variable below V_1 would fall in the left group and data points with values above V_1 would fall in the right group. Visually, the dark color boxes (high wind generation) are already to some extent differentiated from the light color boxes (low wind generation). In this case, a good split seems to be found and the model "learns" that X_1 is an important distinguishing feature to determine the wind generation.

In the next step, the model looks for a new splitting rule to further divide each subset of the training data to further reduce the RMSE, i.e. grouping similar generation values together and placing different generation values in different categories.

Continuing this procedure, without any stopping criterion, would lead to a tree in which each group consists only of one box. Such a fully grown tree is perfectly able to predict the correct values on seen data, but will likely fail to achieve good performance on unseen data. Multiple stopping criteria were defined and randomness was introduced to control for the complexity of the individual trees (e.g: maximum depth of the tree, minimum number of training instances in leaf nodes, etc.), leading to better generalisation capabilities.

The final model consists of hundreds of individual trees where the final prediction is defined by the aggregation of the individual predictions (as illustrated in the next Figure 11).

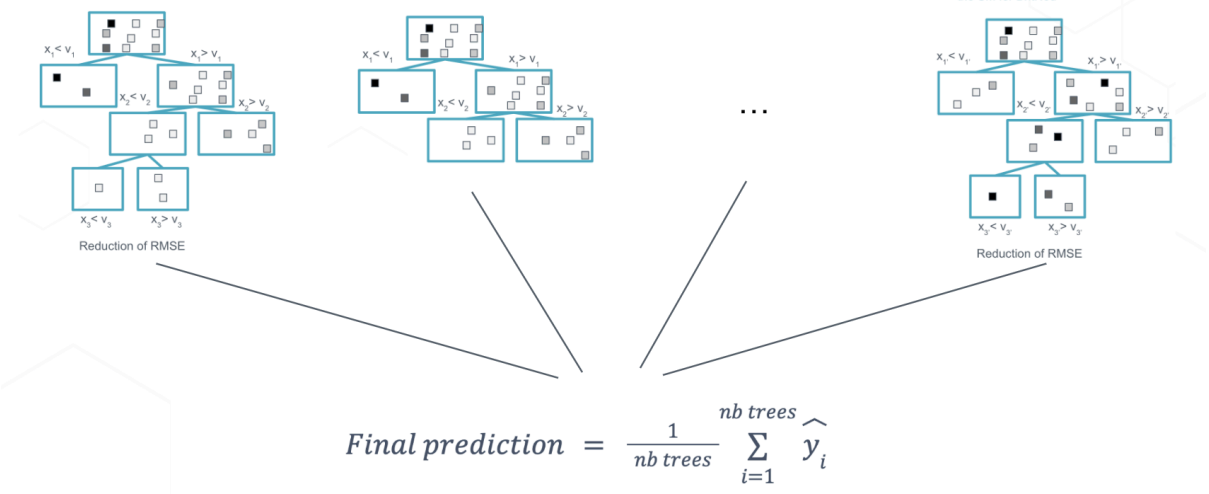


Figure 11: Shows how the individual trees are aggregated in the final "ensemble" model.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

When making a new prediction for a datapoint the splits will be evaluated based on the values of the input features for that datapoint. At each step, it will be determined whether we move left or right. This procedure continues until the datapoint belongs to a group that is not split further, i.e. the leaf node. The prediction of one tree for the datapoint will be the average over the training instances that ended during the training phase in the same leaf node. The idea being: "As these training instances are closely related to the new datapoint in the feature space (as they were evaluated similarly on all the considered splits), it is highly expected that the wind generation will be very similar as well". This procedure is conducted for each tree in the model, leading to a final average across the individual predictions to obtain the final forecasted value.

3.5 Forecasting Performance

3.5.1 RMSE

Looking at the performance over the complete prediction set (01/01/2023 - 31/05/2023), Figure 12 displays the root mean square error (RMSE) for the three wind farms for each month in the prediction set. It can be noticed that the wind farm East Anglia One exhibits the lowest forecast accuracy, while the forecast associated with Greater Gabbard is the most accurate (low RMSE). However, it is important to consider that these wind farms differ in terms of capacity. To be able to do a proper comparison, It is therefore necessary to standardise the metric (RMSE) by expressing it per capacity unit (see next Figure 13).

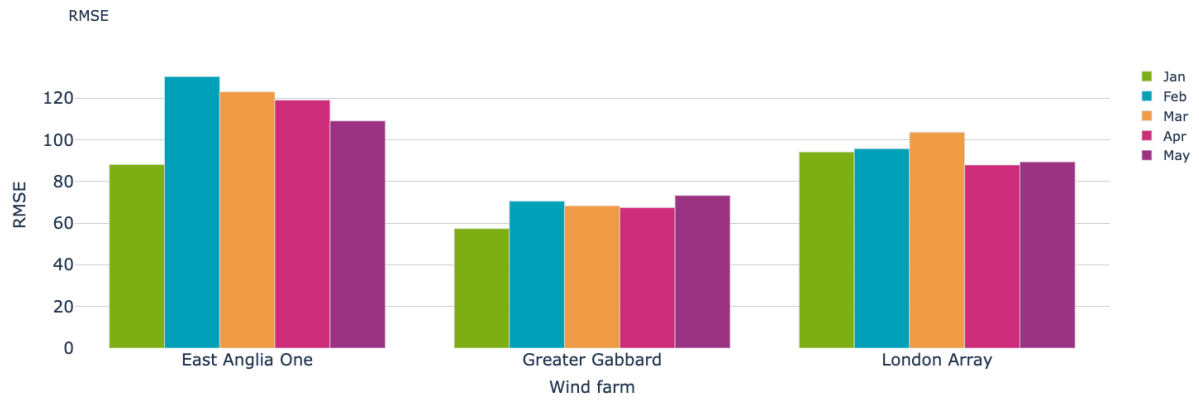


Figure 12: Performance of the N-SIDE model in terms of RMSE for all the windfarms on a monthly basis.

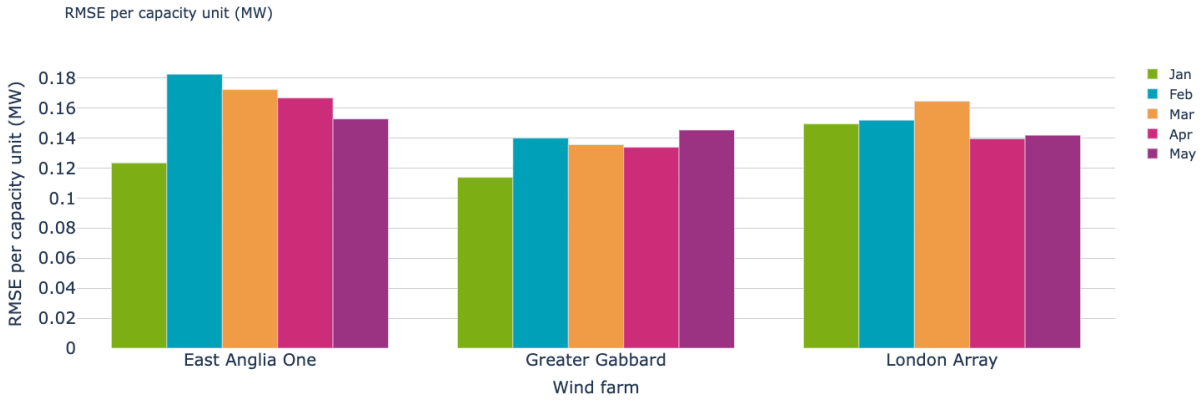


Figure 13: Performance of the N-SIDE model in terms of RMSE standardised per capacity unit for all the windfarms on a monthly basis.

When expressed per capacity unit, it becomes evident that the accuracy associated with the forecasts for Greater Gabbard and London Array is pretty similar. However, for East Anglia One, we still notice a slightly lower accuracy during the months of February, March and April.

From this Figure 13, we interpret that for Greater Gabbard and London Array, the monthly forecast error associated with their generation corresponds to an amount equivalent to between 12% and 15% of their total capacity. For East Anglia One, this error range is a bit wider (between 12% and 18%).

3.5.2 Bias

The bias plot for the East Anglia One wind farm is presented in Figure 14. The bias plot covers the entire prediction period from 01/01/2023 to 01/06/2023. The bias plots for both Greater Gabbard and London Array can be found in Appendix A.

The bias assessment for East Anglia One demonstrates highly satisfactory results. As observed in the figure, the blue curve closely aligns with the green curve, indicating that the model's predictions are well-calibrated. The P-value associated with the KS statistical test confirms that the difference between the two distributions is highly insignificant (p-value: 0.97) even if minor deviations between the model and the ideal outcome can be observed.

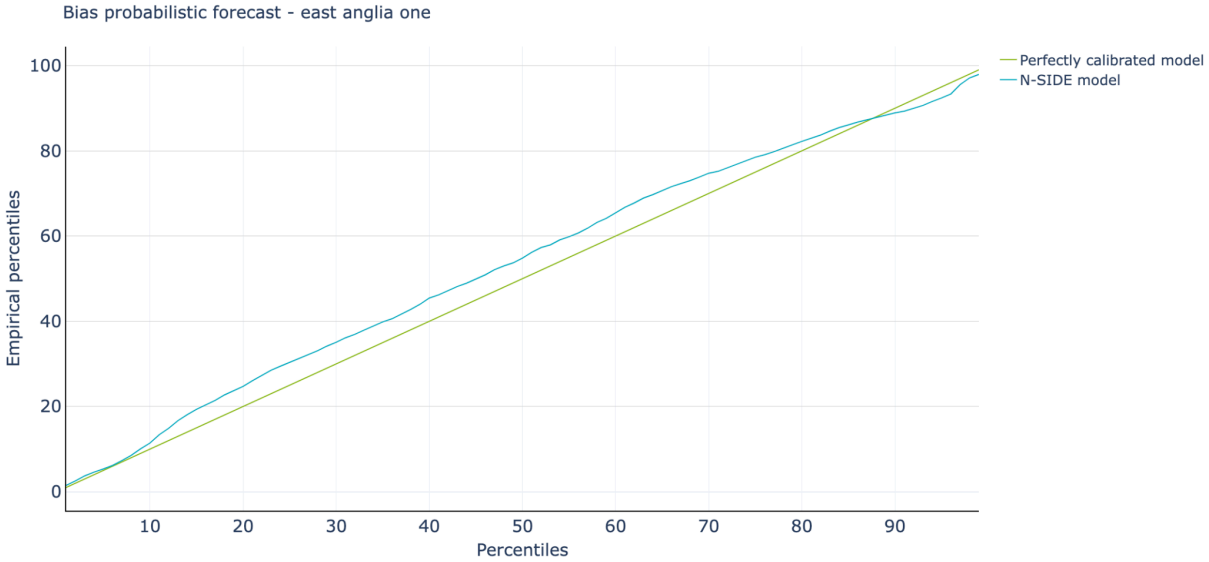


Figure 14: Bias plot for the analysis period (01/01/2023 - 31/05/2023) for the wind farm East Anglia One. The low bias indicates that the predicted distribution is well centred around the actual underlying distribution.

The bias statistical assessment was also carried out for both Greater Gabbard and London Array. This test turned out to be highly insignificant for those two wind farms as well (Greater Gabbard: p_val = 0.36, London Array: p_val = 0.90) .

The bias plot was also created for each month from which the KS statistic was computed. The evolution of the KS statistic can be seen in Figure 15 for all the wind farms.

Figure 15: KS statistic of the N-SIDE model for all the wind farms on a monthly basis.

Comparing the KS statistical values across the five months for each wind farm reveals relevant findings. In May, East Anglia One showed a significant bias (value>=0.2, corresponding to p_value <0.05), while Greater Gabbard demonstrated a similar trend in April. Although the overall bias remains small, these results imply occasional instances of heightened bias in the wind

generation forecasts for these two farms during certain months within the prediction period.

3.5.3 Variance

Figure 16 shows the variance plot for East Anglia One. The results for the other wind farms are presented in Appendix B.

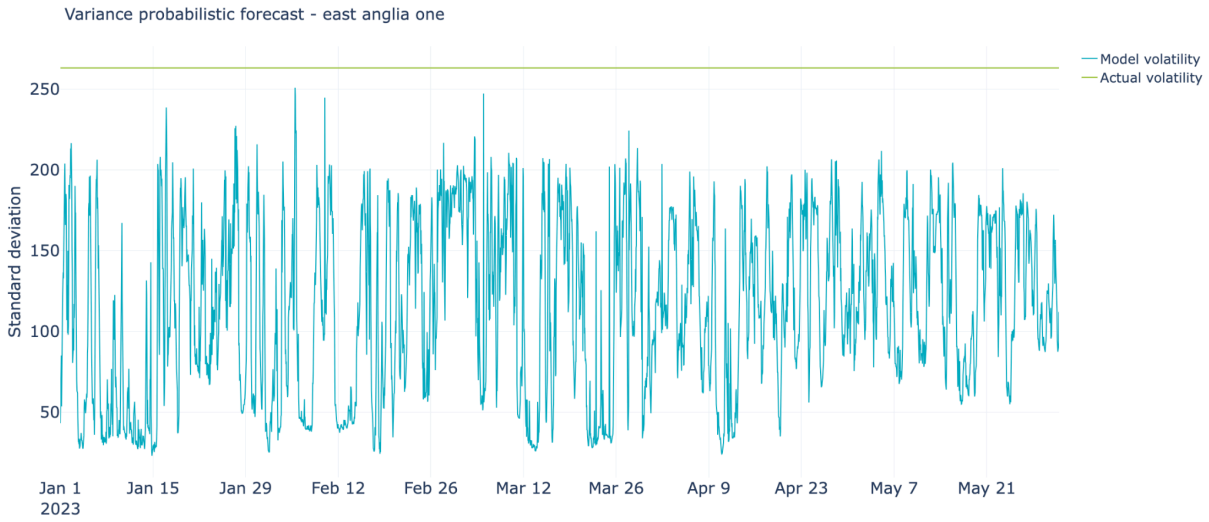


Figure 16: Variance plot for the analysis period (01/01/2023 - 31/05/2023) for wind farm East Anglia One. Model volatility is constantly below the actual volatility, which means the model is able to capture the uncertainty accurately.

The figure reveals that the standard deviation of the actual wind generation over the prediction period is approximately 263 MW. The blue curve represents the standard deviation across different predicted percentiles for a specific period. Periods characterised by high uncertainty, where predicting the wind production is difficult, are reflected by a high standard deviation in the predicted percentiles. Conversely, when the wind generation is more predictable, the standard deviation over the predicted percentiles is lower. The figure demonstrates that the model's standard deviation is always contained below the green curve, indicating its ability to fully capture times of uncertainty in the wind generation. Overall, the standard deviation from the model is constantly below the "actual" standard deviation, indicating that the model successfully reduces the variance in the predictions.

The variance plot was also generated for each month to compute the percentage reduction in variance. Figure 17 provides the evolution of the percentage reduction in variance for all the wind farms, after standardizing this variance reduction as if all three wind farms had the same capacity. First, it is worth noting that the model consistently reduces the variance for all wind farms.

Secondly, the extent of variance reduction showcased by the model is very similar from month to month (mostly between 50-60%). It is important to consider both the variance and bias together to obtain a comprehensive understanding of the model's performance.

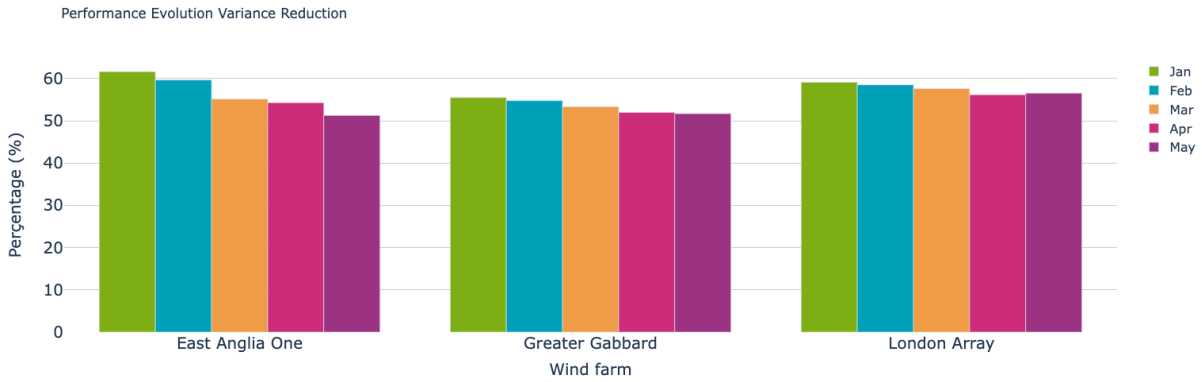


Figure 17: Percentage variance reduction (expressed per capacity unit) of the N-SIDE model compared to the actual variance of the wind generation for all the wind farms on a monthly basis.

3.6 Feature Importance

This section explores the key features that have the most influence on the model. Figure 18 displays the 10 most important input features for the East Anglia One wind farm. Similar plots for the other wind farms can be found in Appendix C.

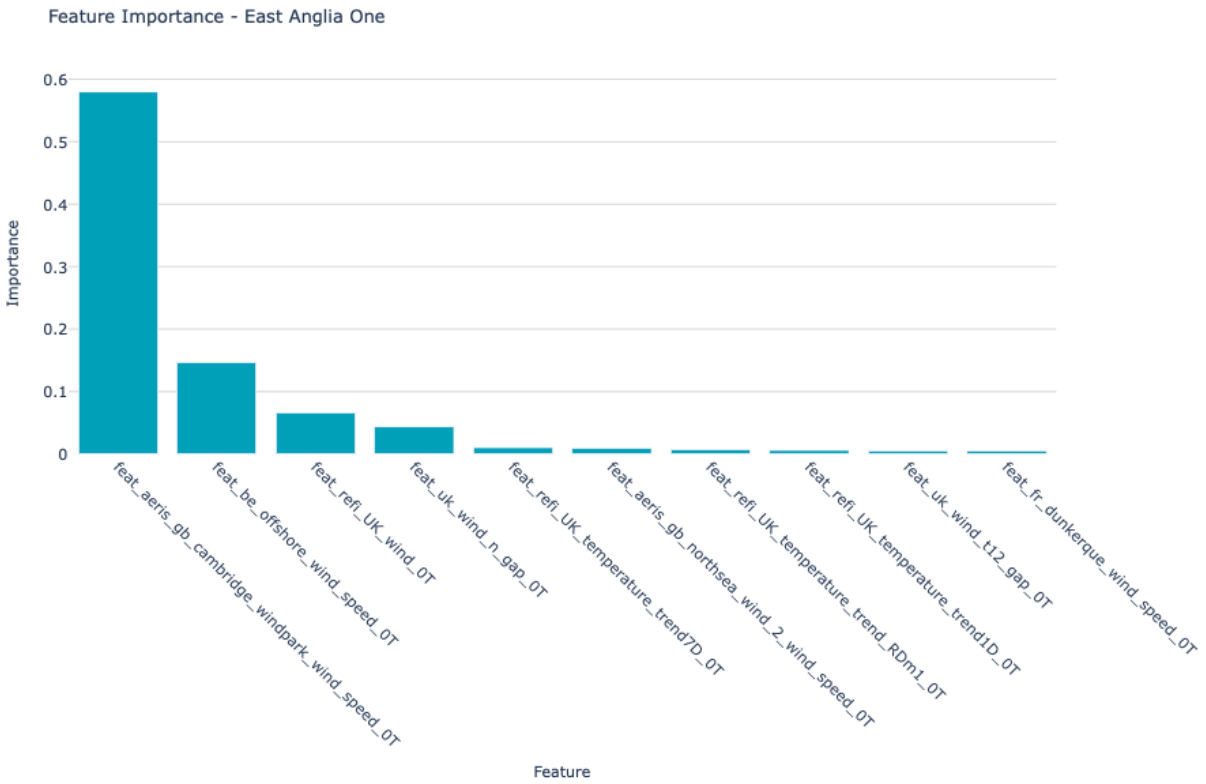


Figure 18: 10 most important features to predict the wind generation by East Anglia One.

The visualisation illustrates that hourly wind speed data from the Cambridge wind park stands out as the most influential feature in the model. Additionally, features related to wind speed from different locations are also deemed highly impactful. This emphasises the substantial influence of wind speed on electricity generation within the East Anglia One wind farm. Furthermore, important features encompass predictions for wind generation in the UK, temperature forecasts, and temperature trends. These insights provide valuable clarity on the key drivers behind the model's predictions for the East Anglia One wind farm.

3.7 Integrating wind farm outage schedules

When building the wind generation forecast, the examination of the wind production profile from Greater Gabbard uncovered instances where it operated below maximum capacity, notably during February and March of 2023 (see next Figure 19). These periods witnessed significant discrepancies between the point forecast and actual production, resulting in diminished accuracy. To address these fluctuations, it was decided to integrate wind farm outage schedules into the forecasting model. As illustrated in Figure 20, integrating this information was able to improve predictions on wind farms undergoing outages.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

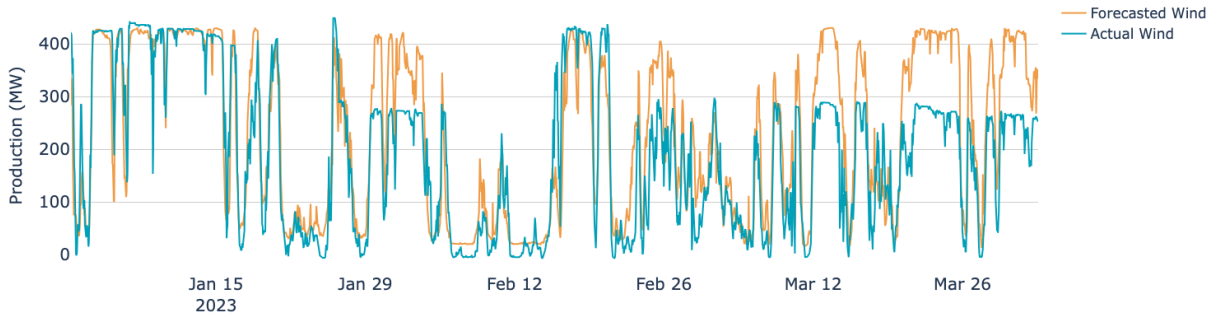


Figure 19. Forecasted versus actuals for Jan-Mar 2023 for Greater Gabbard without considering maintenance schedules. Blue line corresponds to the forecasted generation. Red line corresponds to the actual wind generation.

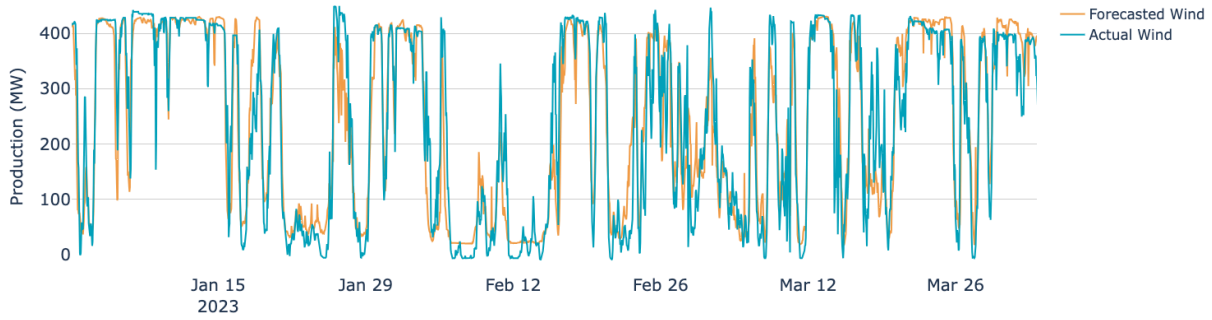


Figure 20. Forecasted versus actuals for Jan-Mar 2023 for Greater Gabbard after integrating maintenance schedules in the forecasting model. Blue line corresponds to the forecasted generation. Red line corresponds to the actual wind generation.

4 Conclusion

Phase V of the project dealt with the development of forecasting models to predict the wind generation on three wind farms. In contrast to point forecasting, the models developed in this project's scope are probabilistic. This enables us in later phases of the project to assess the impact of uncertainties arising from the wind farms on the congestion throughout the GB power grid. Moreover, the developed models also allow for probabilistic inference and are a valuable asset in themselves.

In this report, we introduced the different performance metrics that were used to measure the performance. Apart from the RMSE, the bias and variance were also assessed. Examining the RMSE performance on a testing period ranging from January 2023 to May 2023, East Anglia One demonstrated the lowest forecast accuracy, while Greater Gabbard exhibits the most accurate forecasts. Standardizing RMSE per capacity unit revealed similar accuracy for Greater Gabbard and London Array. However, East Anglia One still showed slightly lower accuracy. Overall, forecast errors correspond to 12-15% of total capacity for

Greater Gabbard and London Array, and slightly wider (12-18%) for East Anglia One.

In terms of bias and variance, all wind farms showed satisfactory performance, with Greater Gabbard displaying the highest bias albeit unsignificant. For the two other wind farms, the bias remained low throughout the testing period. Alongside a significant reduction in variance, we can conclude that the developed forecasting models accurately capture the uncertainties stemming from the wind farms. This outcome highlights the effectiveness of the models in precisely representing the inherent uncertainties.

In terms of influential factors on the predictions, hourly wind speed data from the Cambridge wind park stands out as the most impactful feature in the model, emphasizing its substantial impact on East Anglia One's electricity generation, alongside other wind speed features from different locations, UK wind generation predictions, temperature forecasts, and trends, providing valuable insights into the model's predictive drivers.

The predictions generated by the developed algorithms, together with the interconnector flow probabilistic predictions, will be used to create a diverse range of scenarios for power flow simulations. By sampling various values from the predicted distribution, we can evaluate the effects of uncertainties in the wind farms on congestion within the GB power grid.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

Appendix A: Bias plots for the remaining wind farms

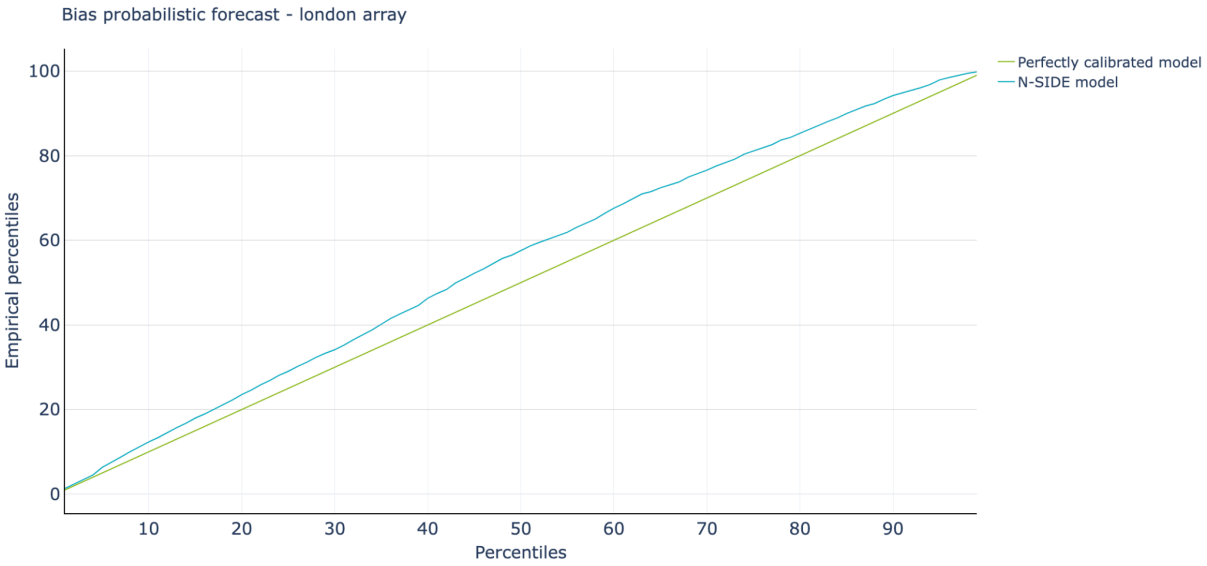


Figure 21: Bias plot for the period of analysis (01/01/2023 - 31/05/2023) for the London Array wind farm. The low bias indicates that the predicted distribution is well centered around the actual underlying distribution.

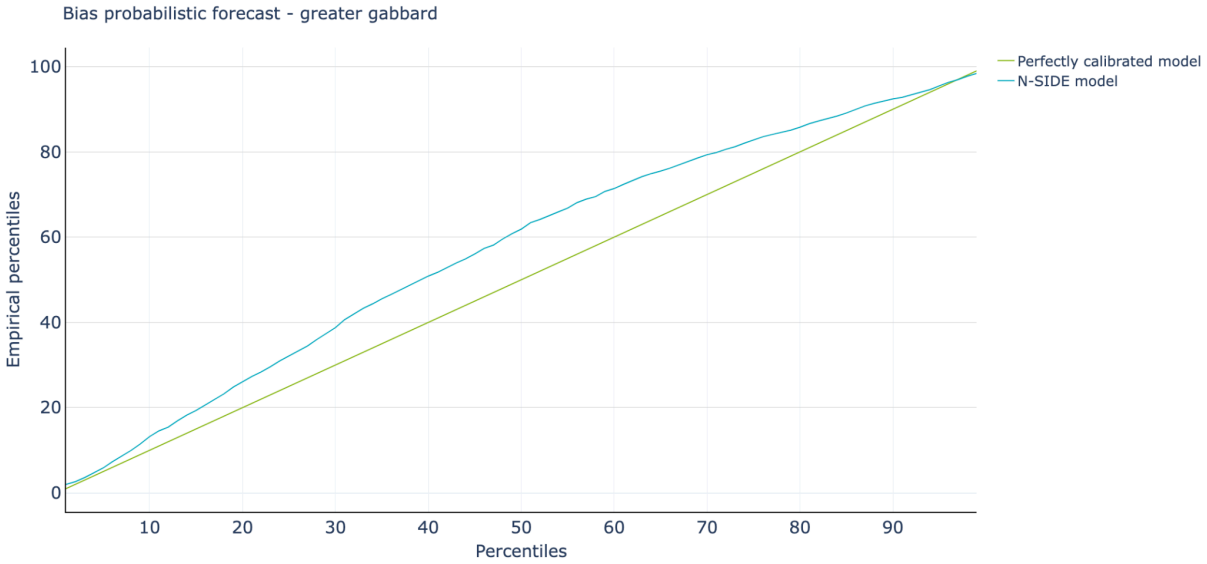


Figure 22: Bias plot for the period of analysis (01/01/2023 - 31/05/2023) for the Greater Gabbard wind farm. A low bias is observed for the lower percentiles, whereas the bias is a bit larger for the higher percentiles. This indicates that the predicted distribution is a bit "wider" on the right hand side compared to the actual underlying distribution.

N-SIDE S.A. Boulevard Baudouin 1er 25, 1348 Louvain-La-Neuve, Belgium
www.n-side.com

Appendix B: Variance plots for the remaining wind farms

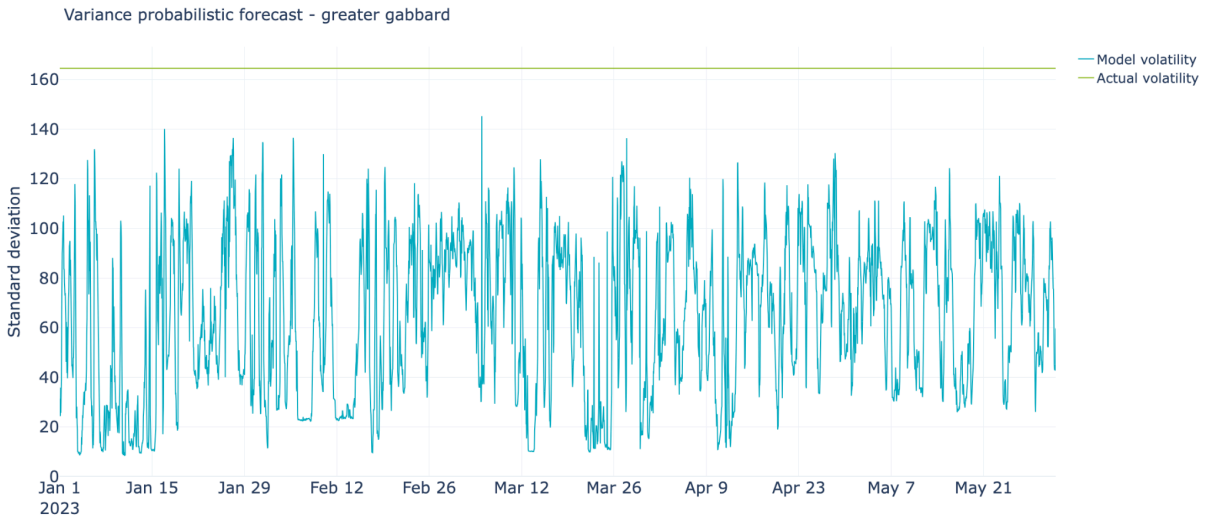


Figure 23: Variance plot for the analysis period (01/01/2023 - 31/05/2023) for Greater Gabbard wind farm. Model volatility is constantly below the actual volatility, which means the model is fully able to capture the uncertainty accurately.

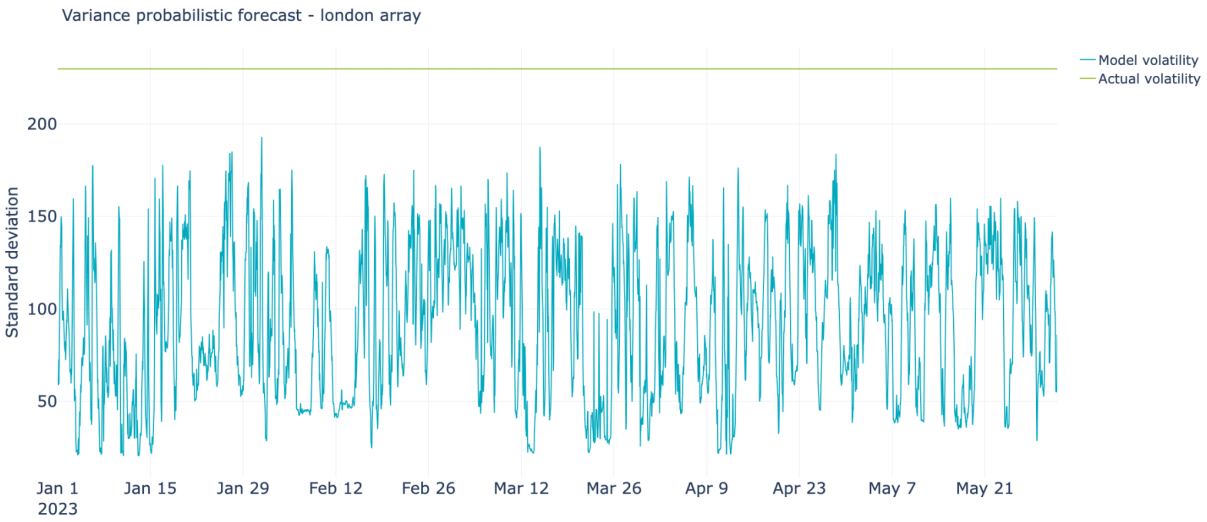


Figure 24: Variance plot for the analysis period (01/01/2023 - 31/05/2023) for London Array wind farm. Model volatility is constantly below the actual volatility, which means the model is fully able to capture the uncertainty accurately.

Appendix C: Feature importance for the remaining wind farms

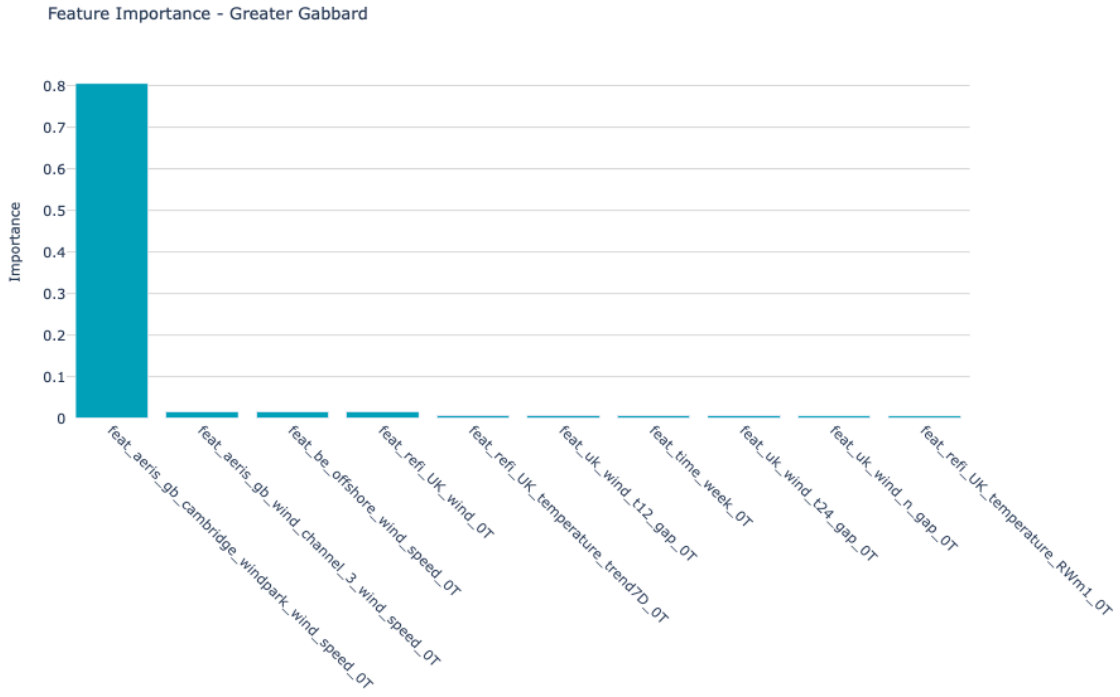


Figure 25: 10 most important features to predict the wind generation on the Greater Gabbard wind farm.

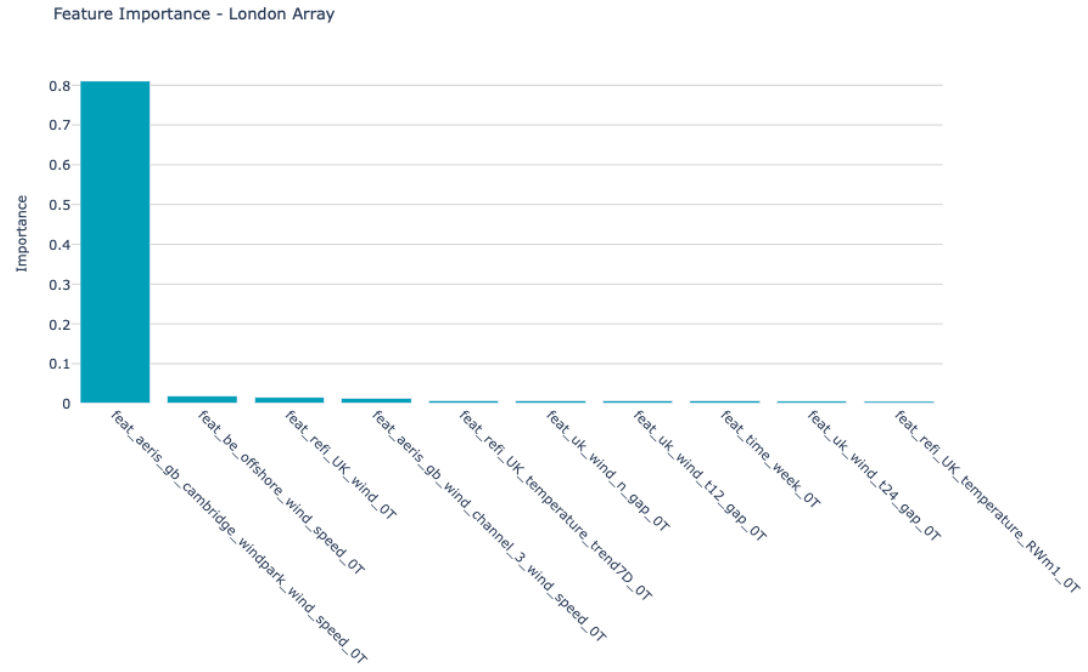


Figure 26: 10 most important features to predict the wind generation on the London Array wind farm.