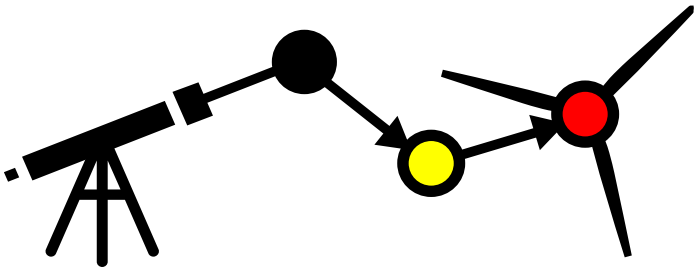
BeFORECAST



Deliverable: Performance assessment of wind-farm power now- and forecasting

Deliverable No.: D5.2

economie
FOD Economie, K.M.O., Middenstand en Energie

This project has received funding from the Energy Transition Fund of the Belgian Federal Public Service for Economy, SMEs, and Energy (FOD Economie, KMO, Middenstand en Energie)

DOCUMENT CONTROL SHEET	
OCUMENT INFORMATION	
Document title	Performance assessment of wind-farm now- and forecasting
Author(s), (organization)	Simone Gremmo (VKI) Wim Munters (VKI) Stijn Ally (VUB) Geert Smet (RMI) Dieter Van den Bleeken (RMI) Joris Van den Bergh (RMI)
Deliverable No.	D5.2
Work Package No.	WP5
Lead beneficiary	VKI
Dissemination level	Public
Date of issue	30/11/2025

DOCUMENT SUMMARY

This deliverable (D5.2) evaluates the performance of wind farm power forecasting methods, combining physics-based and data-driven approaches to improve accuracy for offshore wind farms in the Belgian North Sea. The work spans intraday/day-ahead forecasting and nowcasting, addressing wake effects and uncertainty quantification.
Three main developments are presented:

- Wake-aware NWP forecasting: Implementation of wind farm parameterizations (WFP) in mesoscale models (ALARO and WRF) and post-processing corrections using PyWake engineering wake models.
- Neural Power Net (NPN): A machine-learning model trained on historical wind speed, direction, and power data, which significantly reduces bias and mean absolute error compared to traditional power curve methods.
- Data-driven nowcasting: A modular deep-learning framework introduced by VUB (Ally et al., 2025), integrating LSTM-based architectures and Gaussian-process layers for uncertainty-aware predictions. This approach improves RMSE by up to 25% and provides probabilistic forecasts, enabling risk-aware operational decisions.

Comparing the baseline reference ALO4_CP with the most performant implementation ALO4_WFP_NPN, which features a wind farm parameterization and a neural network postprocessing layer (see Figures 5 and 6), we find that the bias can be reduced from 5% to a virtually unbiased power forecast, whereas mean absolute error improves by approximately 15% to 20%, depending on lead time and wind direction.

Table of Contents

1. Introduction

Accurate power prediction of offshore wind farms is critical for operational planning, grid stability, and overall integration of offshore wind into the electricity system. Specifically for the Belgian case, with 2.2 GW installed in close proximity, weather systems impact the entire offshore fleet as a whole leading to potentially significant power excursions, affecting the national electricity system and further stressing the importance of accurate forecasting. Additionally, traditional forecasting models often do not account for wind farm wake effects. With the Belgian offshore wind farms having a high power density, incorporating these effects in power forecasts is essential.

This deliverable reports on the performance assessment of wind-farm power prediction algorithms at conventional intraday / day-ahead time scales (forecasting) as well as very short time scales (nowcasting). The work reported here was performed in the context of tasks T5.2 (Wake models for power postprocessing from NWP output), T5.3 (Intraday and day-ahead wind and power forecasting), and T5.6 (Observation-based nowcasting of wind and power potential). The improvements in forecasting accuracy enabled by the research activities in the project are quantified here. A twofold approach is employed. On the one hand, RMI and VKI algorithms for intraday and day-ahead forecasting based on numerical weather prediction (augmented with wind farm effects and machine learning postprocessing layers) are benchmarked using the 3E performance assessment framework developed within the project (see Deliverable D5.1). On the other hand, a fully data-driven modular methodology developed by VUB is introduced and discussed.

More specifically, the work presented here builds on previous developments in the BeFORECAST project, introducing 4 complementary strategies: (1) post-processing corrections using PyWake, an open-source tool that applies engineering wake models to account for turbine interactions; (2) the implementation of wind farm parameterizations (WFP) in mesoscale models such as ALARO and WRF, enabling dynamic representation of wake effects within the atmospheric flow; (3) the development of a Neural Power Net (NPN) postprocessing layer, a machine learning model trained to convert wind forecasts into power predictions while implicitly capturing wake-related patterns; and (4) a modular deep-learning based fully data-driven methodology converting publicly available weather forecasts into power predictions. In addition, a preliminary evaluation of the Ma & Archer WFP in WRF is included, offering insights into its potential advantages over the widely used Fitch scheme for wind power modeling.

The deliverable combines long-term verification against power data from the European Network for Transmission System Operators for Electricity (ENTSO-E) and Belgian TSO Elia with case studies of high-impact events, such as February 2022 Storm Eunice, to quantify improvements in forecast accuracy. By comparing bias, mean absolute error (MAE), and other performance metrics across different methods, this report provides a comprehensive evaluation of how wake-aware modeling and data-driven techniques can significantly reduce forecast errors and improve reliability for offshore wind farms.

The remainder of this deliverable is structured as follows. Section 2 elaborates on the NWP-driven forecasting methodologies (i.e., based on NWP models run by the consortium). The conversion from wind speeds into wind farm power is elaborated, and an intercomparison in performance of different tools is presented. Next, Section 3 discussed the data-driven modular methodology introduced by VUB. Section 4 highlights the main conclusions from the technical work presented in the preceding sections. Finally, Section 5 provides an overview of the dissemination activities performed as part of this deliverable.

2. NWP-driven intraday and day-ahead wind power forecasting

2.1. Power forecasting with ALARO

We describe three methods for deriving power forecasts from NWP wind speed forecasts. The currently used operational method, with power curves, is briefly explained in Section 2.2. A method to correct these power forecasts, using PyWake is described in Section 2.3. Finally, Section 2.4 presents a comparison of a neural network with the previous methods. All methods are applied to both the operational ALARO forecasts, and the experimental forecasts from an ALARO model with wind farm parameterization (Van den Bleeken et al., 2025) as described in BeFORECAST deliverable *D4.2 Intercomparison of mesoscale models including wind-farm parameterization implementations*.

2.2. Power curves

In the power curve method, wind forecasts V = (u,v) from NWP models at turbine height (or 100m) are converted to wind power P as

$$P = \frac{1}{2}\rho\, C_P(|V|)A_{rotor}|V|^3 \quad (1)$$

where ρ is the air density, A_{rotor} the area spanned by the rotor blades and, C_P the power coefficients of the turbine.

Operationally, when taking forecasts from the ALARO 4 km model are used as input for the power forecasting and are then further corrected using quantile regression (and historical power production data), for some anonymous wind farms in the Belgian offshore zone (BOZ). We will use these operational forecasts as a benchmark, and refer to them as **ALO4_CP**.

To calculate power forecasts for the ALARO 4km model with wind farm parameterization (WFP), we also apply equation (1), but then without further quantile regression, and refer to the results as **ALO4_WFP_CP**.

2.3. PyWake

PyWake is an open-source, Python-based wind farm simulation tool developed at DTU capable of computing flow fields, power production of individual turbines (Pedersen et al., 2023). The PyWake software package, freely available through an MIT license, can be used to calculate the wake interaction in a wind farm in a computationally inexpensive way for a range of steady state conditions. In this way, it can compute the power production of a wind farm while considering the wake losses depending on the wind farm layout configuration. An important feature of PyWake is the possibility to use different engineering wake models for the simulation, such as the NOJ and Bastankhah wake deficit models.

For computation of wake interactions, PyWake needs turbine layout configuration, and details on the turbine type (rotor diameter, Cp and Ct coefficients, ..). The turbine layout for the BOZ windfarms was extracted from publicly available shapefiles provided by the Management of the marine environment (MUMM, 2025) Scientific Service. Turbine specifications were taken from (confidential) RMI data, obtained in the context of development of the Elia Storm Forecast Tool (Smet et al., 2019).

We implemented three engineering wake models, with following wake deficit models:

- BastankhahGaussian deficit (Bastankhah & Porté-Agel, 2014)
- Niels Otto Jensen (NOJ) deficit (Jensen, 1983)
- Fuga deficit + Fuga blockage (Ott & Nielsen, 2014)

All are implemented in PyWake, where Fuga also includes the Fuga blockage model (Pedersen et al., 2023). Using these models, we created lookup tables for each windfarm, for combinations of wind speeds and wind directions, using 16 wind direction bins and wind speed bins from 0 to 30 m/s in 0.5 m/s intervals. These tables give the total wind power output for each wind farm, depending on the free flow input wind speed and wind direction. A graphic example of such a table is shown in Figure 1 for BastankhahGaussian deficit at an anonymous wind farm in the BOZ.

The BOZ wind farms turbine layout (turbines in blue, offshore high-voltage station in red) is shown in Figure 2 BOZ wind farms turbine layout (turbines in blue, offshore high-voltage station in red).Figure 2. The three wake models implemented in PyWake show markedly different levels of skill when applied to the ALARO 4 km forecasts. The Fuga deficit and blockage configuration (**pyw_FB1**) reduces the performance relative to the operational, wake-free baseline (**ALO4_CP**), while the NOJ model (**pyw_NOJ1**) provides only marginal improvement. In contrast, the BastankhahGaussian model (**pyw_BG**) yields a clear and systematic gain in forecast skill. This is evident both in the reduction of wind power bias (Figure 3 left) and of mean absolute error (Figure 3 right). For this reason, only the BastankhahGaussian-based PyWake correction is retained in subsequent comparisons with other forecasting methods.

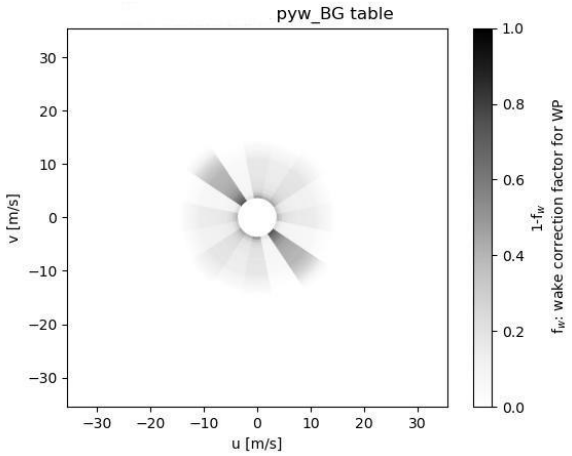


Figure 1 Example of PyWake correction for BastankhahGaussian deficit at an anonymous wind farm in the BOZ.

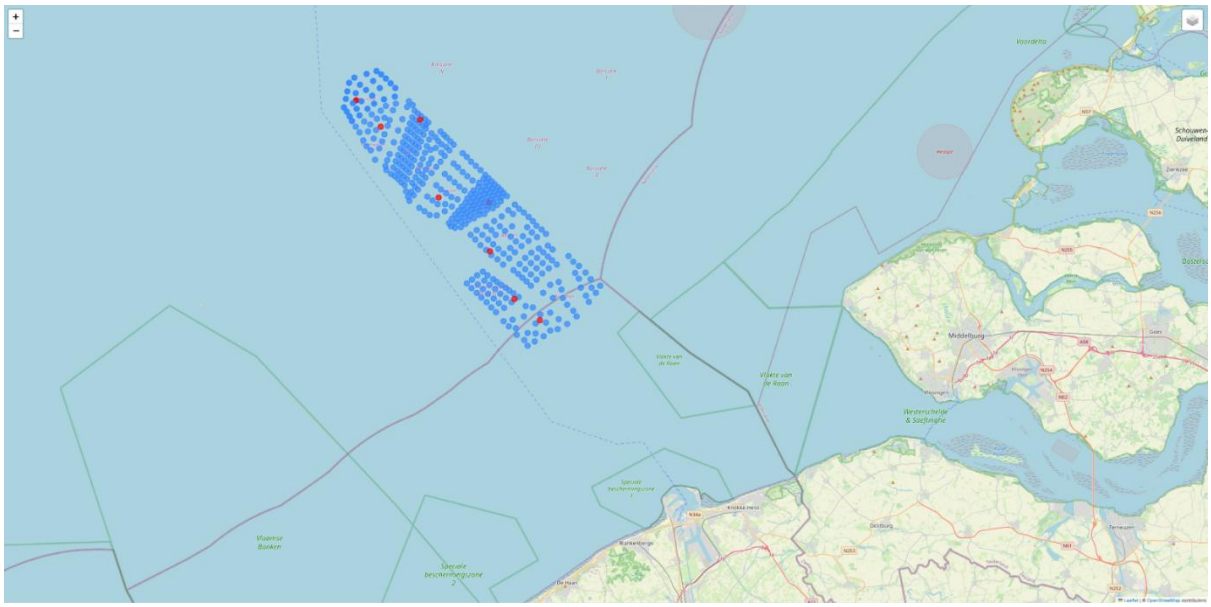


Figure 2 BOZ wind farms turbine layout (turbines in blue, offshore high-voltage station in red).

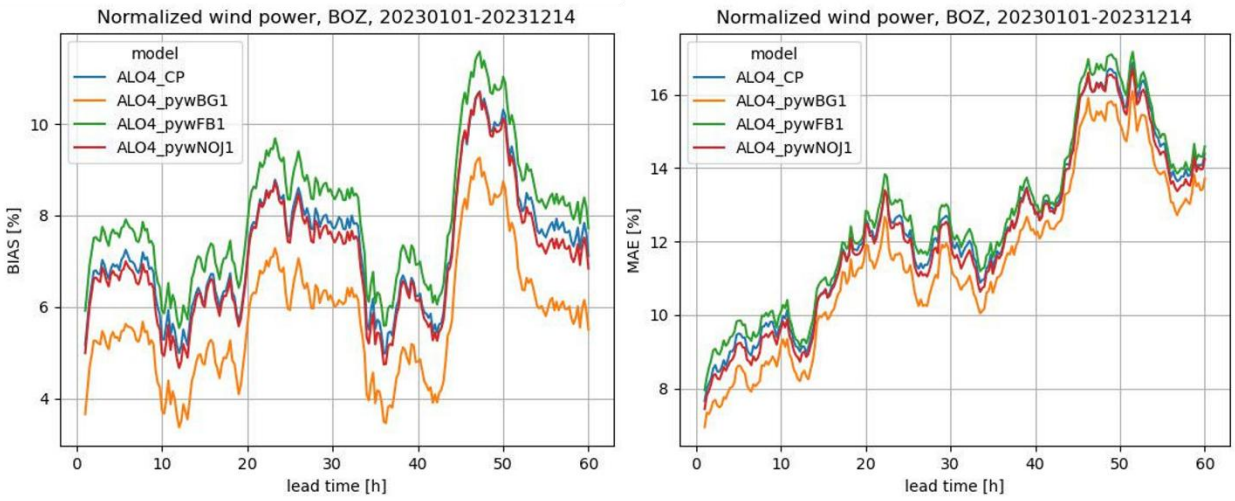


Figure 3 BIAS (left) and Mean Absolute Error (MAE, right) of operational ALO4_CP and three PyWake corrected ALO4 at BOZ.

2.4. Neural Power Net (NPN)

2.4.1. Neural network architecture

A neural network (Van den Bleeken et al., 2025) was designed and trained to predict wind power from three NWP forecast variables: wind speed, wind direction, and lead time. Including wind direction enables the model to account for wake effects (Barthelmie et al., 2009; Platis et al., 2018), while lead time captures diurnal and forecast-horizon-dependent biases.

The model is a multi-layer perceptron (Almeida, 2020) with a three-node input layer and a single-node output layer. After testing several configurations, the optimal setup comprised three hidden layers of 64 nodes each. The network was trained using mean absolute error (MAE) as the loss function, the Adam

optimizer with stochastic gradient descent, and early stopping. This configuration, sketched in Figure 4, is referred to as the Neural Power Net (NPN).

The NPN was trained individually for each wind farm in the BOZ using one year of data (2022), combining 15-minute ALARO forecasts of wind speed and direction with Elia wind power production records. Validation was performed on 2023 data. The 2022 dataset was randomly divided into 85% for training and 15% for hyperparameter tuning.

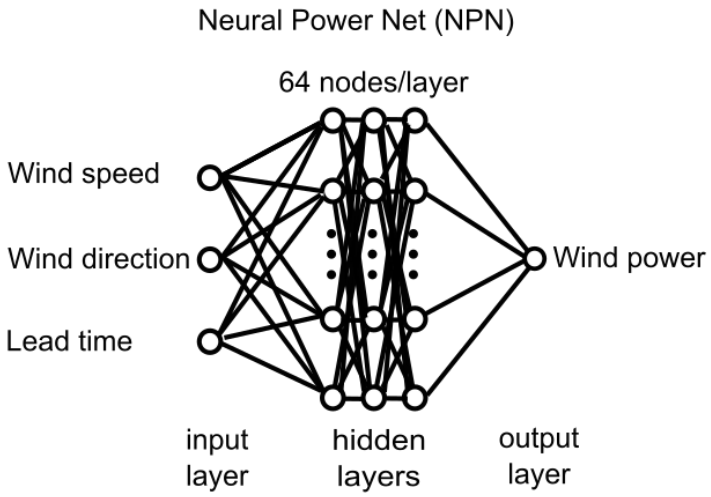


Figure 4: Neural Power Network (NPN) architecture

2.4.2. Comparison with power curve method *and PyWake corrections*

Figure 5 and Figure 6 compare the performance of the NPN-based forecasts with the operational power-curve approach and with the PyWake-corrected forecasts. Figure 5 shows that the operational **ALO4_CP** forecasts considerably overestimate wind power production, which is expected since wake losses are not accounted for. Applying PyWake corrections slightly reduces this positive bias but still underestimates the total wake effect, most likely because inter-farm interactions are not included. Introducing the wind farm parameterization (**ALO4_WFP_CP**) lowers the predicted power output substantially, but the resulting negative bias indicates that the wake effect is overestimated, consistent with the negative wind speed bias reported in D4.2.

The best performance is obtained when using the data-driven neural power conversion. Both NPN-based configurations (**ALO4_NPN** and **ALO4_WFP_NPN**) reduce the bias to values close to zero, and the combination of wake-parameterized wind forecasts with NPN (**ALO4_WFP_NPN**) provides the most accurate results overall.

A similar ranking is visible in the mean absolute error (Figure 5 right). Again, the NPN-based forecasts outperform both the operational and PyWake-corrected methods. The negative bias of **ALO4_WFP_CP**

affects its MAE less than the large positive bias of **ALO4_CP** and the PyWake-based forecast, but its absolute performance remains inferior to the neural approaches.

Figure 6 further illustrates these results as a function of wind direction. The improvement of NPN is most pronounced for wind coming from the north-west and south-east, directions where wake effects are strongest. PyWake performs reasonably well for north-westerly winds but underestimates wake losses for south-easterly flows, whereas both NPN configurations reproduce the expected behaviour more consistently. The added skill of **ALO4_WFP_NPN** relative to **ALO4_NPN** demonstrates that NPN benefits from wake-corrected wind input and not only from statistical learning.

Overall, all three new approaches have improved upon the current operational forecasts. Applying NPN to convert wind to power is more effective than solely improving the wind forecasts with a wake parameterization, but their combination leads to the highest skill.

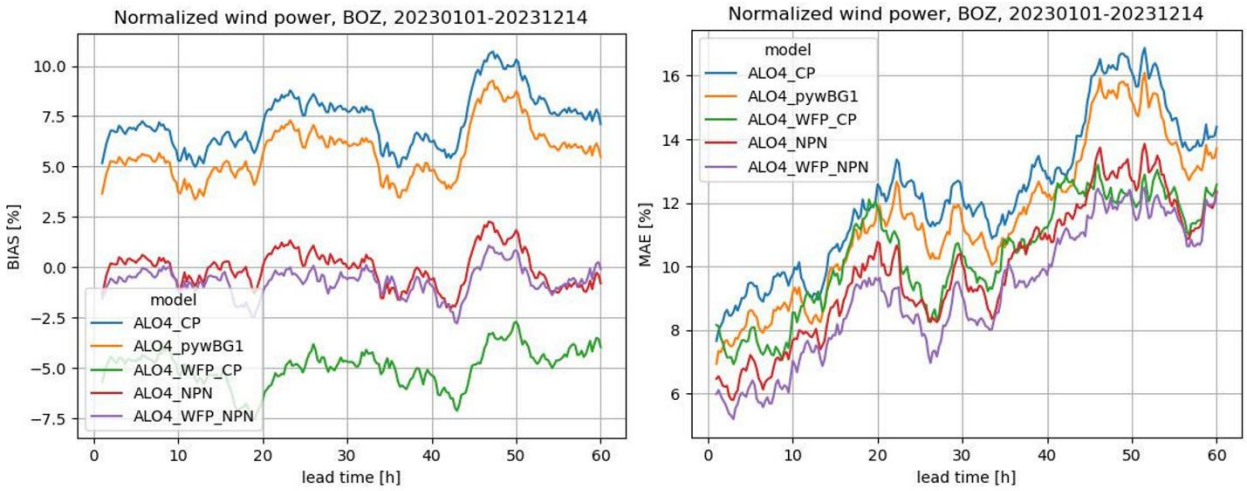


Figure 5 BIAS (left) and MAE (right) of operational ALO4_CP, ALO4_WFP_CP, and their corresponding NPN versions, together with ALO4 corrected using PyWake BastankhahGaussian deficit, averaged over the period 1 January–14 December 2023, for total power production at BOZ.

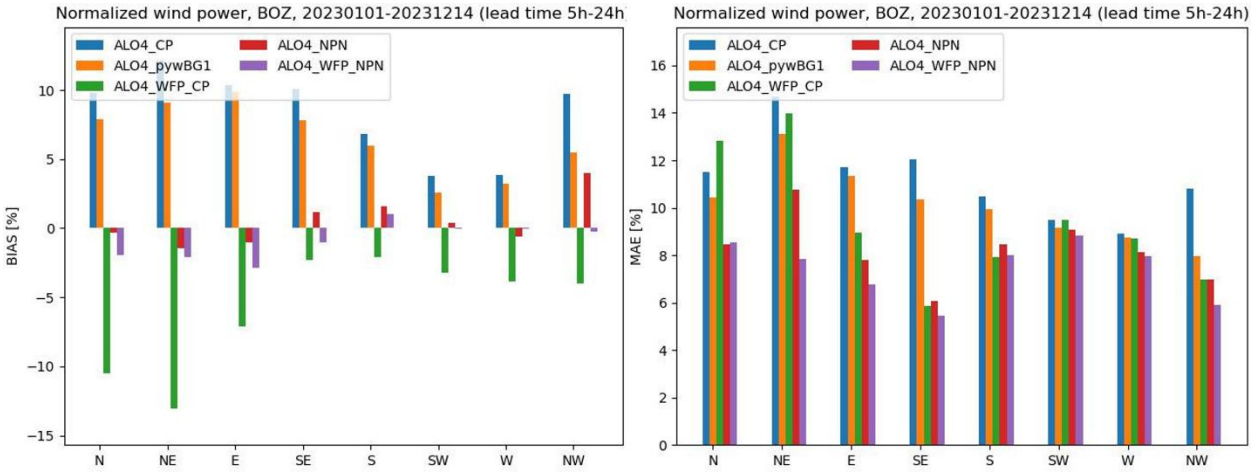


Figure 6 BIAS (left) and MAE (right) of operational ALO4_CP, ALO4_WFP_CP, and their corresponding NPN versions, together with ALO4 corrected using PyWake BastankhahGaussian deficit, averaged over the period 1 January–14 December 2023, for total power production at BOZ, split per wind direction.

2.4.3. *Combining* NSN *with NPN*

The neural speed network (NSN) introduced and validated in BeFORECAST deliverable D4.3 can be combined with the neural power network (NPN) described above. We take the NSN wind speed forecasts of 2023, which were trained on 2022 data, and use the first half of 2023 as training input of NPN, and validate forecasts on the second half of 2023. We will refer to these forecasts as NSN_NPN, or more accurately, depending on the forecast model that NSN was trained on, as **ALO4_NSN_NPN**, **ALO4_WFP_NPN** and **ALO4_WFP_NR_NPN**. A comparison of NPN and NSN_NPN forecasts is shown on Figure 7 left (BIAS) and Figure 7 right (MAE). Consistent with results from other authors, e.g. Phipps et al. (2022), post-processing to power directly is often just as good (or even better) than first post-processing wind speed. Note that in Figure 7, we compare NSN_NPN (trained on speed for 2022 and power for the first half of 2023), with NPN (trained on power for 2022). However, the conclusions remain qualitatively the same (not shown), when comparing with a version of NPN trained only on data in the first half of 2023.

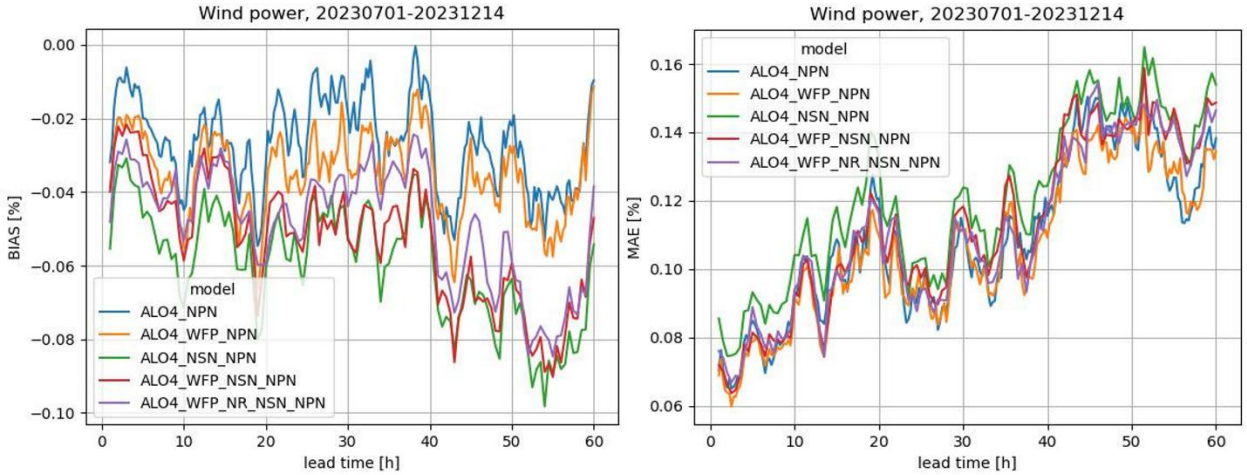


Figure 7 BIAS(left) and MAE (right) of NSN_NPN and NPN for ALO4 and ALO4_WFP at an anonymous wind farm.

2.5. Power forecasting with WRF

An overview of the Weather and Research Forecast (WRF) model (Skamarock et al. 2021) is provided in deliverable D4.2, together with the numerical setup. Here, the focus is computing wind power from WRF.

In WRF, wind turbine power output is calculated using turbine-specific performance data and local wind conditions. This requires providing power curves, expressed through power and thrust coefficients as functions of wind speed, along with geometric properties such as hub height and rotor diameter, and the exact turbine locations in a tabulated format. Based on these inputs, WRF computes the generated power for each grid cell in the computational domain.

By default, WRF does not report power at the individual turbine level. For simulations using the Fitch wind farm parameterization, this information can be obtained through a simple post-processing step that derives turbine power from the velocity field output by WRF. In contrast, for simulations employing the Ma & Archer WFP, the model was modified to directly save turbine-level power along with all other relevant outputs, as post-processing is not feasible in this case. The total power for each wind farm is then calculated by summing the contributions from all turbines belonging to the same wind farm.

Geometric characteristics and turbine locations were sourced from publicly available information. Similarly, power and thrust curves were generated using the PyWake utility, which produces realistic curves based on turbine specifications such as rotor diameter, rated power, and hub height. PyWake generates power and thrust curves by using analytical models to approximate realistic performance behavior. The process starts with turbine parameters such as rotor diameter, hub height, and rated power. From these, PyWake applies aerodynamic principles and standard turbine performance assumptions to compute the power coefficient (C_P) and thrust coefficient (CT) as functions of wind speed.

The resulting curves typically follow the characteristic shape of real turbine performance:

- At low wind speeds, power output is zero until the cut-in speed is reached.
- Between cut-in and rated speed, power increases according to CP and the available kinetic energy in the wind.

Beyond the rated speed, power is capped at the turbine rated power until the cut-out speed, where it drops to zero for safety.

For C_T, the curve reflects how thrust varies with wind speed, peaking near rated conditions and decreasing at higher speeds. These curves are then tabulated and used by WRF or other models to calculate turbine-level power and wake effects.

2.6. Ma & Archer wind farm parametrization

To assess the sensitivity of forecasting performance to specific WFP implementations, additional model runs with the Ma & Archer (Ma et al., 2022), which includes an ensemble of engineering models to account for sub-grid scale interactions between wind turbines. These forecasts cover the period of February 2022, which also includes storm Eunice.

The Fitch wind farm parameterization is the most commonly used scheme in the Weather Research and Forecasting (WRF) model for representing the effects of wind turbines on atmospheric flow and for estimating power production. Its popularity stems from its simplicity and integration within the model framework. However, the Fitch scheme applies the same wind speed to all turbines within a grid cell and treats wake interactions in a simplified manner. These limitations can lead to inaccuracies in power estimation, especially for large wind farms with complex layouts or high turbine density.

To address these limitations, we tested the Ma & Archer WFP, which introduces several improvements over the Fitch scheme. Ma & Archer incorporates analytical wake models that represent wake expansion and velocity deficits at a sub-grid scale. It also accounts for the uncertainties of engineering models by creating an ensemble of wake models and superposition methods. This enables a more detailed representation of flow variability within a wind farm. These enhancements are expected to improve the accuracy of power predictions and reduce biases in scenarios where wake losses are significant.

We compared Fitch and Ma & Archer to evaluate whether the new scheme improvements justify its adoption for mesoscale simulations. While the WRF forecast dataset is available with the Fitch scheme for the full year, we tested Ma & Archer over one month as a preliminary assessment (i.e February 2022). This is a preliminary comparison, and further analysis is needed to determine the conditions under which Ma & Archer provides measurable benefits in power estimation, particularly for offshore wind farms and high-density layouts.

Figure 8 presents the comparison of power forecasts for the entire month of February 2022 at Mermaid wind farm. The one-month time series was constructed by chaining the 24-hour lead time forecasts from the WRF dataset. Overall, the power predictions from both the Fitch and Ma & Archer schemes show good agreement, with only a few specific events where Fitch predicts higher power output than Ma & Archer. When compared to the actual power production (black line), both schemes perform similarly.

Figure 9 provides a closer look at Storm Eunice on February 18. Here again, the two WFPs produce very similar power estimates, with only minor differences during the transient phases of the storm. The small discrepancies are likely due to Ma & Archer's more detailed wake representation, which slightly reduces power during periods of strong directional changes.

Finally, Figure 10 displays a scatter plot comparing the power predictions from Fitch and Ma & Archer. A clear trend emerges: most points lie below the red dashed line, indicating that Fitch generally predicts higher power output than Ma & Archer. This systematic bias reflects Fitch's simplified treatment of wake effects, which does not fully capture the cumulative impact of multiple wakes within large wind farms.

An interesting avenue for further research beyond the scope of BeFORECAST is to assess the grid sensitivity of the different WFPs and isolate performance for specific high-impact wake events.

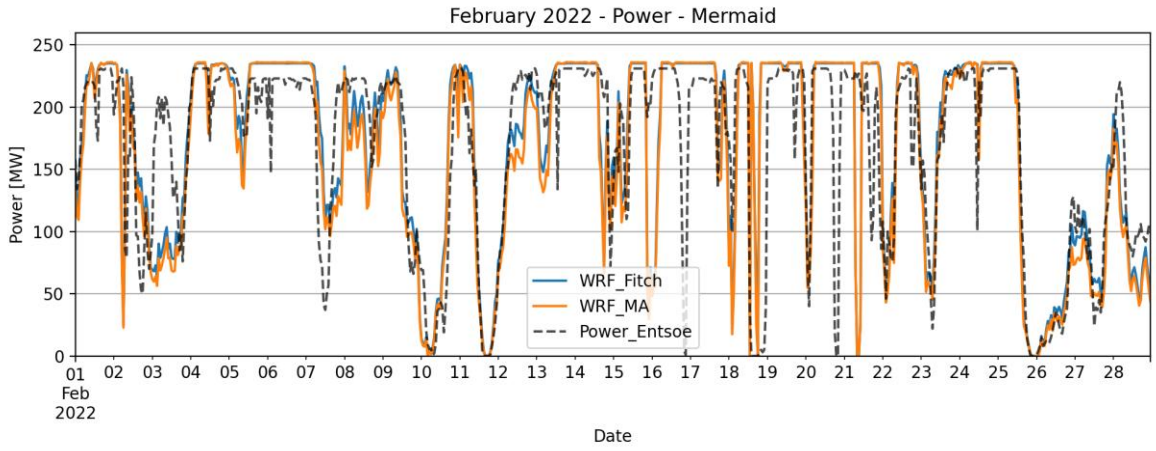


Figure 8 February 2022 - Comparison of Power forecasts for WRF Fitch (blue) and WRF Ma&Archer (orange). Real power production added for comparison (Entsoe - black dashed). Forecasts for the two WFPs is very similar, except in specific events.

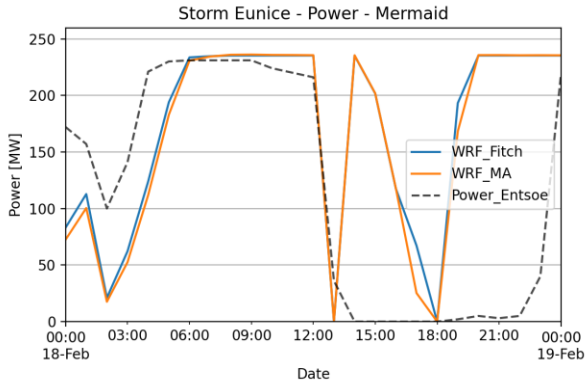


Figure 9 Storm Eunice – 18 February 2022 - Comparison of Power forecasts for WRF Fitch (blue) and WRF Ma&Archer (orange). Real power production added for comparison (Entsoe - black dashed).

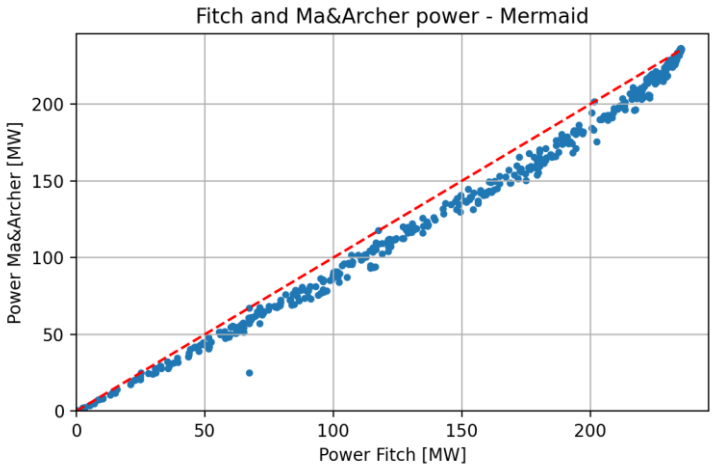


Figure 10 February 2022 - Scatter plot of Fitch and Ma&Archer. Fitch tends to overestimate the power production compared to Ma&Archer

2.7. Intercomparison of NWP-based forecasts

In deliverable D4.2, the comparison between mesoscale models focused on wind speed and wind speed deficit. In this section, the analysis shifts to power prediction. The objective is to evaluate how ALARO and WRF perform in estimating power output, considering both long-term verification scores for the period January to June 2022 and a specific storm event, i.e., Storm Eunice on 18 February 2022. All power calculations are obtained from the simplest power estimation method, i.e., by interpolating wind speed onto the turbine power curve, as described in Section 2.2.

To ensure consistency, wind speed was sampled at specific reference locations for each wind farm. Two approaches were adopted: (1) the average position of all turbines within the wind farm, and (2) operational locations used in ALARO, referred to as RMI locations, to maintain alignment with previous analyses conducted at RMI.

Figure 11 shows the bias and MAE for ALARO and WRF without wind farm parameterization (WFP), calculated against ENTSO-E observations. A positive bias indicates that the models forecast higher power production than observed, which is expected since turbine interactions are not accounted for in these configurations.

When WFP is activated (Figure 12), the bias becomes negative, meaning the models now underpredict power production. Activating WFP improves accuracy overall, as both bias and MAE decrease compared to simulations without WFP. Specifically, the bias is reduced from 0.10 to -0.04, while MAE drops from 0.15 to 0.09. It is also noteworthy that ALARO and WRF perform similarly when WFP is enabled, whereas ALARO shows slightly better performance without WFP.

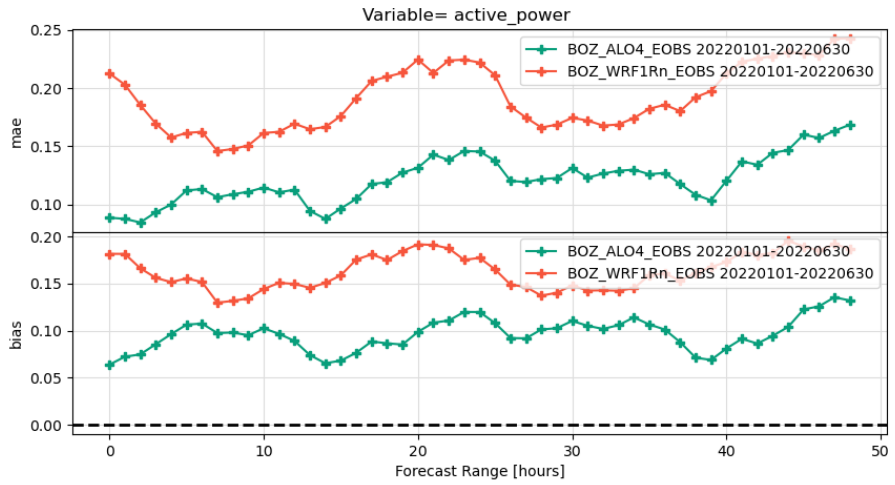


Figure 11 BIAS and MAE from verif3E tool for ALO4 and WRF (with WFP switched off) at shared RMI locations (see explanation in D4.3), against ENTSO-e power observations.

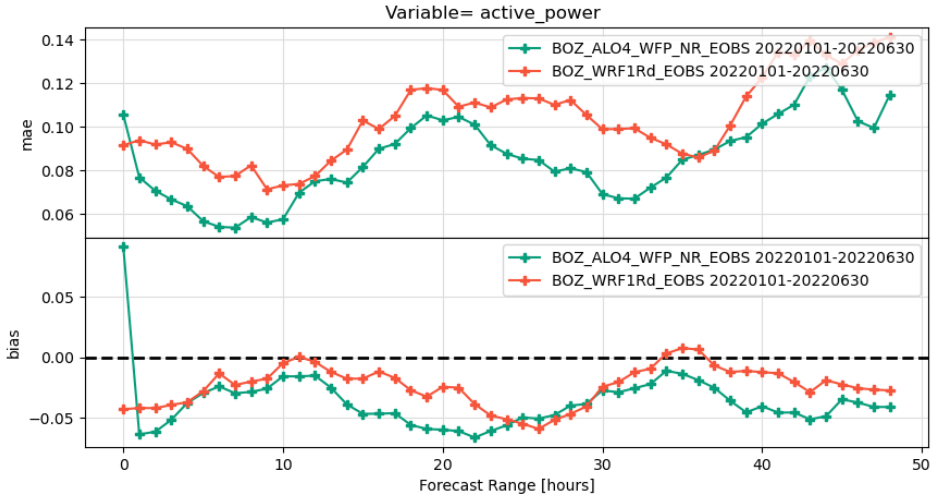


Figure 12 BIAS (bottom) and MAE (top) from verif3E tool for ALO4_WFP and WRF at shared RMI locations (see explanation in D4.3), against ENTSO-e power observations.

Figure 13 compares bias and MAE against Elia observations. ALARO without WFP exhibits the largest positive bias. When WFP is activated, both models show similar bias, but ALARO achieves a smaller MAE. **ALO4_WFP** relies on the "operational" locations used by ALO4, which are confidential, whereas **ALO4_WFP_NR** employs newly tested locations to assess their suitability. These "experimental" locations

are also utilized in WRF (e.g., WRF1Rd). This difference affects bias but has little influence on MAE.

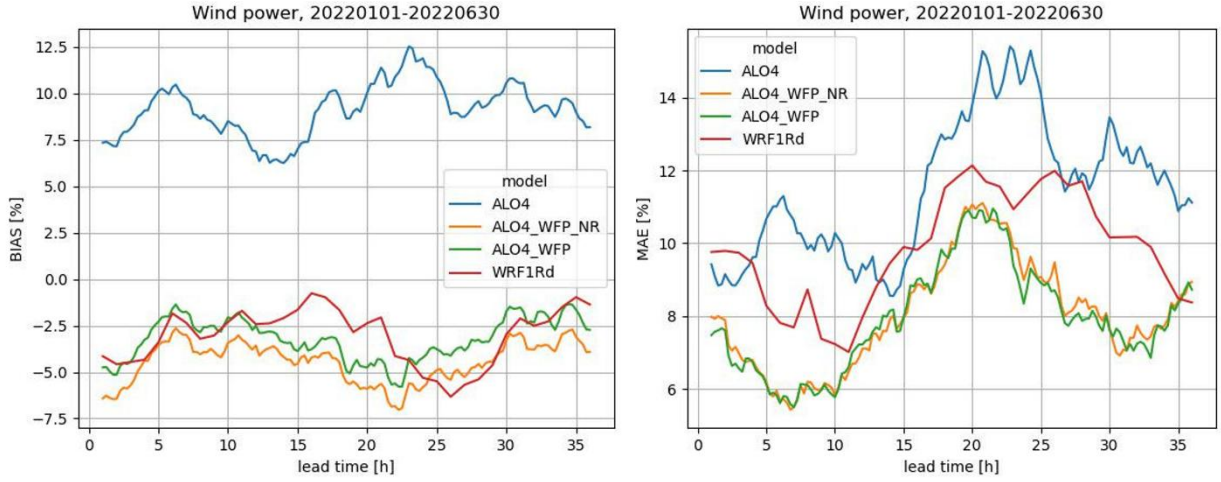


Figure 13 BIAS (left) and MAE (right) of ALO4, ALO4_WFP, ALO4_WFP_NR and WRF against Elia power observations.

Figure 14 breaks down the results by wind direction. For ALARO without WFP, bias remains positive across all directions, while activating WFP in both models generally results in negative bias. Exceptions occur for SE and NW directions, where both ALARO and WRF show a slight positive bias.

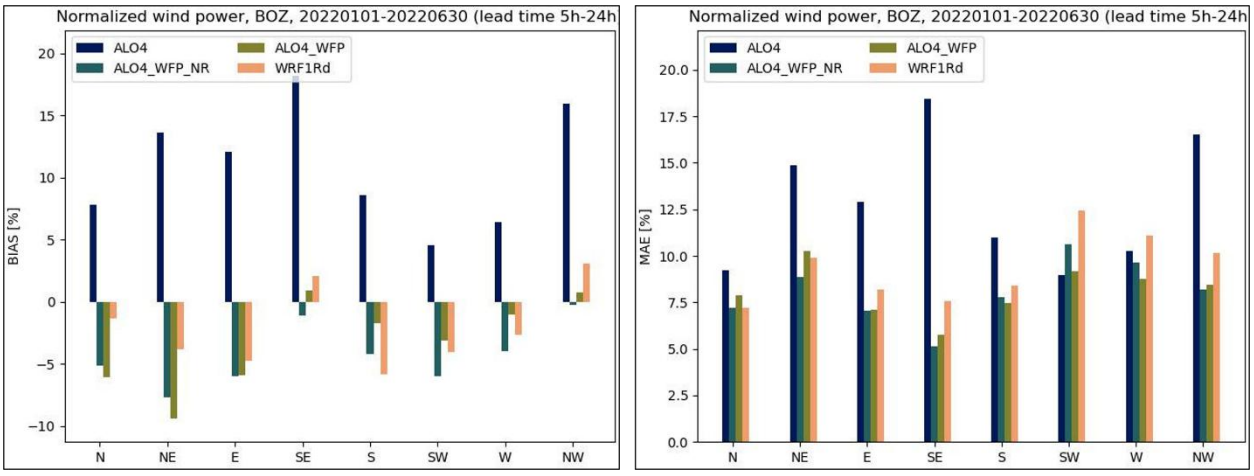


Figure 14 BIAS (left) and MAE (right) of ALO4, ALO4_WFP, ALO4_WFP_NR and WRF against Elia power observations, split per wind direction.

Figure 15 illustrates normalized wind power during Storm Eunice (18 February 2022), comparing ALARO and WRF predictions to Elia observations. All models capture the ramp-down caused by high wind speeds. All models capture the ramp-down caused by high wind speeds, but there is a notable difference between the ALO4 models and WRF. While the former predicts zero power, consistent with observations, WRF estimates about 80% power. This discrepancy arises because WRF, for this specific event, fails to accurately capture the wind speed peak (as noted in D4.2). Moreover, the power ramp-up (after the storm peak) from the observations is not consistent with the wind speed observations and might be related to

a direct control from the wind farm operators. The spread in power predictions during the ramp-up phase is significant because, in region 2 of the power curve, i.e. above cut-in but below rated wind speeds, small differences in wind speed translate into large differences in power.

Finally, Figure 16 shows wind power deficit for ALARO at two resolutions (1 km in blue and 4 km in orange) and for WRF. WRF tends to overestimate losses compared to ALARO, while increasing ALARO's resolution from 4 km to 1 km does not lead to a substantial difference.

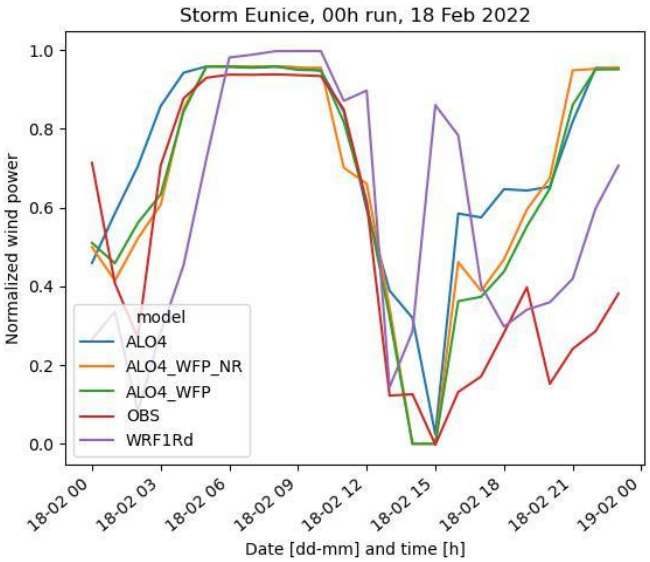


Figure 15 Time series of ALO4, ALO4_WFP, ALO4_WFP_NR and WRF against Elia power observations, for storm Eunice.

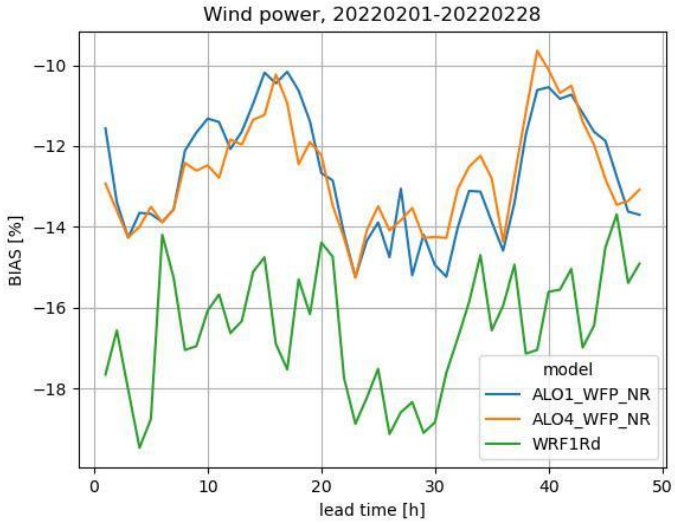


Figure 16 Wind farm effects (wind power deficit) for ALO1, ALO4 and WRF, in the BOZ, as a percentage of installed capacity.

3. Data-driven wind power now- and forecasting

In addition to the methodologies described in the previous section based on NWP model runs performed within the project, an alternative is to use publicly available reanalysis and forecast datasets to train and validate data-driven algorithms without having to perform a dedicated NWP model run. Such models can be used for wind power prediction and nowcasting timescales (i.e. minutes up to a few hours ahead) and forecasting timescales (i.e., several hours up to day-ahead). Note that the impact of including recent weather and SCADA data on prediction performance at nowcasting timescales were also included in **Deliverable D3.3**. In the remainder of this section, we focus on the efforts by VUB on developing and testing a fully data-driven Weather Forecast Correction (WFC) to recover power forecasts from publicly available data. A full description of the methodology and results can be found in Ally et al. (2025). Here we describe the main concepts and conclusions of the study.

Figure 17 depicts the architecture of the proposed ensemble deep-learning framework for probabilistic wind power nowcasting. The design integrates multiple components: an input layer that ingests meteorological forecasts and SCADA observations, feature preprocessing modules for normalization and temporal encoding, and a sequence-to-one prediction block based on Long Short-Term Memory (LSTM) units to capture temporal dependencies. On top of this, a Gaussian-process layer estimates predictive uncertainty, enabling the generation of probabilistic forecasts rather than deterministic outputs. The figure highlights how these elements interact to produce both point predictions and confidence intervals, ensuring robust performance under varying atmospheric conditions and supporting risk-aware operational decisions. Note that the methodology differs in complexity from the relatively simple deep learning architecture for power prediction described in Section 2.4.

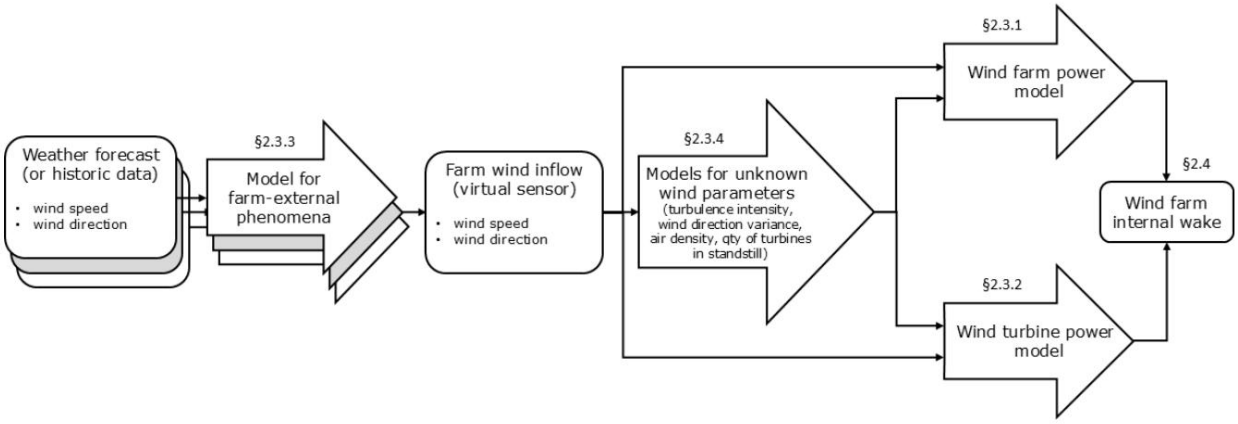


Figure 17: Modular data-driven wind power forecasting architecture. Reproduced from Ally et al. (2025).

The approach combines:

1. **Feature engineering**: Incorporating atmospheric stability indicators, turbine-level wind speed and direction, and temporal encodings.
2. **Hybrid architectures**: Ensemble deep-learning models integrating LSTM layers for temporal dependencies and Gaussian-process components for uncertainty quantification.
3. **Uncertainty-aware predictions**: Providing probabilistic power forecasts rather than deterministic outputs, which is critical for grid operators managing variability in offshore wind production.

Figure 18 compares the performance of the proposed WFC methodology to power forecasts directly converted from a commercial weather forecast for a range of time scales from nowcasting to 3 days ahead for an anonymous Belgian offshore wind farm. It can be observed that the WFC significantly reduces error metrics and improves the correlation with SCADA data.

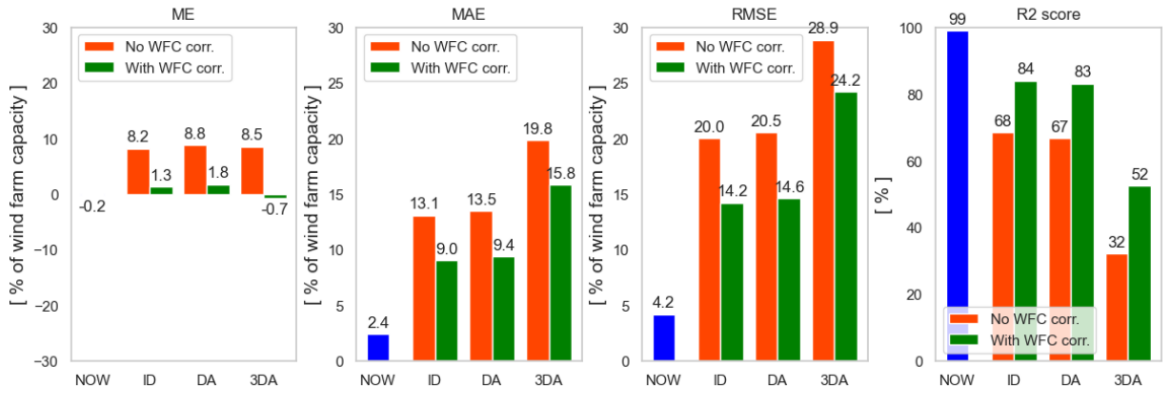


Figure 18: Wind power forecast error metrics on nowcasting (NOW), intra-day (ID), Day-ahead (DA), and 3-Day-ahead (3DA) timescales obtained directly from a commercial weather forecast (No WFC corr.) and by applying the WFC of Ally et al. (2025). Error metrics include Mean Error (ME, lower is better), Mean Absolute Error (MAE, lower is better), Root-Mean-Square Error (RMSe, lower is better), and Correlation (R2, higher is better). Reproduced from Ally et al. (2025).

Key findings from the study include:

1. **Performance gains**: Compared to IEC-based power curve binning, the proposed models reduce RMSE by up to 25% and improve reliability during high-impact events.
2. **Robustness under varying inflow conditions**: The models adapt well to changes in atmospheric stability and wake interactions, outperforming traditional regression-based methods.
3. **Operational relevance**: The uncertainty estimates enable risk-aware decision-making for reserve allocation and congestion management. Furthermore, the methodology is computationally efficient, allowing power forecasts to be produced from weather forecasts in the order of milliseconds.

4. Conclusion

The integration of physical and data-driven approaches in BeFORECAST delivers substantial improvements in offshore wind power forecasting. Physics-based enhancements, such as activating wind farm parameterizations in mesoscale models, reduce systematic overprediction and improve wake representation, cutting bias by 60% and MAE by 40%. Neural Power Net (NPN) further boosts accuracy, achieving near-zero bias and outperforming both operational and PyWake-corrected methods, especially when combined with wake-aware wind forecasts.

Comparing the baseline reference **ALO4_CP** with the most performant implementation **ALO4_WFP_NPN** (see Figures 5 and 6), which features a wind farm parameterization and a neural network postprocessing layer, we find that the bias can be reduced from **5% to a virtually unbiased power forecast**, whereas mean absolute error improves by approximately **15% to 20%**, depending on lead time and wind direction.

Beyond deterministic forecasting, the data-driven nowcasting framework introduced by VUB (Ally et al., 2025) marks a significant step forward. By leveraging LSTM architectures and Gaussian-process

uncertainty layers, this approach not only improves RMSE by up to 25% but also provides probabilistic forecasts critical for risk-aware grid operations.

Together, these developments pave the way for more reliable, uncertainty-aware forecasting systems, supporting efficient renewable integration and enhancing grid resilience.

5. Dissemination

Conference presentation: Smet, G., Van den Bleeken, D., Van den Bergh, J., Dehmous, I., Degrauwe, D., Van Ginderachter, M., and Deckmyn, A.: Improving wind power forecasts in the Belgian North Sea, EMS Annual Meeting 2024, Barcelona, Spain, 1–6 Sep 2024, **https://doi.org/10.5194/ems2024-872** .

Journal publication: Van den Bleeken, D., Smet, G., Van den Bergh, J., Dehmous, I., Degrauwe, D., Van Ginderachter, M., and Deckmyn, A. (2025). Improving wind power forecasts in the Belgian North Sea with a wind farm parameterization and a neural network. *Adv. Sci. Res., 22*, 59–67. **https://doi.org/10.5194/asr-22-59-2025**

Journal publication: Ally, S., Verstraeten, T., Daems, P.J., Nowé, A. and Helsen, J., 2024. Modular deep learning approach for wind farm power forecasting and wake loss prediction. *Wind Energy Science Discussions*, *2024*, pp.1-33. **https://doi.org/10.5194/wes-10-779-2025**

6. References

Ally, S., Verstraeten, T., Daems, P.J., Nowé, A. and Helsen, J., 2024. Modular deep learning approach for wind farm power forecasting and wake loss prediction. *Wind Energy Science Discussions*, *2024*, pp.1-33. **https://doi.org/10.5194/wes-10-779-2025**

Bastankhah M. and Porté-Agel F. (2014). A new analytical model for wind-turbine wakes. J. Renew. Energy. 2014;70:116-23.

Jensen, N. O. (1983). A note on wind generator interaction. Risø National Laboratory.

Mads M. Pedersen, Alexander Meyer Forsting, Paul van der Laan, Riccardo Riva, Leonardo A. Alcayaga Romàn, Javier Criado Risco, Mikkel Friis-Møller, Julian Quick, Jens Peter Schøler Christiansen, Rafael Valotta Rodrigues, Bjarke Tobias Olsen and Pierre-Elouan Réthoré. (2023, February). PyWake 2.5.0: An open-source wind farm simulation tool. DTU Wind, Technical University of Denmark. https://gitlab.windenergy.dtu.dk/TOPFARM/PyWake

Ma, Yulong, Cristina L. Archer, and Ahmad Vasel-Be-Hagh. 2022. "Comparison of Individual versus Ensemble Wind Farm Parameterizations Inclusive of Sub-grid Wakes for the WRF Model." *Wind Energy* 25 (9): 1573–95. https://doi.org/10.1002/we.2758.

MUMM, 2025. https://odnature.naturalsciences.be/mumm/en/windfarms/

Phipps, K., Lerch, S., Andersson, M., Mikut, R., Hagenmeyer, V., & Ludwig, N. (2022). Evaluating ensemble post-processing for wind power forecasts. Wind Energy, 25(8), 1379-1405. https://doi.org/10.1002/we.2736

Ott, S. and Nielsen, M. (2014). Developments of the offshore wind turbine wake model Fuga. DTU Wind Energy. DTU Wind Energy E No. 0046.

Skamarock, C., Bogumiła Klemp, Jimy Dudhia, O. Gill, Zhiquan Liu, Judith Berner, Wei Wang, G. Powers, Greg Duda, Dale Melvyn Barker and Xiangyu Huang. “A Description of the Advanced Research WRF Model Version 4.” (2019). https://doi.org/10.5065/1DFH-6P97

Smet, G., J. Van den Bergh, and P. Termonia, 2019: Probabilistic storm forecasts for wind farms in the North Sea. ALADIN-HIRLAM Newsletter, 12, 17–24.

Van den Bleeken, D., Smet, G., Van den Bergh, J., Dehmous, I., Degrauwe, D., Van Ginderachter, M., and Deckmyn, A. (2025). Improving wind power forecasts in the Belgian North Sea with a wind farm parameterization and a neural network. *Adv. Sci. Res., 22*, 59–67. **https://doi.org/10.5194/asr-22-59-2025**