2026-02-13

The quality of the revised manuscript has been improved significantly. The authors have addressed the previous comments with care.

A few remaining comments are listed below for potential further improvement:

- In the revised abstract, the primary motivation for developing the data-driven surrogate model is described as the need to perform *numerous* simulations to account for wake effects from multiple neighbouring farms. It is not clear why such a large number of simulations is required in this context. I recommend clarifying this point in the introduction.
- Page 25 appears to be extremely large in terms of tile size (MB). Consider compressing the figures on this page to make the manuscript more manageable.
- In Table 2, error metrics have been added for a "Naïve free stream baseline". I'm surprised by these really bad metrics, for example, for MAE [ms^{-1}]. In reality, the wind speed reduction by the wake is typically much lower than 11.87 ms^{-1}. How can the MAE be so large for a naïve baseline which assumes a wind speed reduction of 0 ms^{-1}.
- Line 487: As you noted correctly in the introduction, in other works GNNs have been deployed for modelling multiple parameters at turbine locations. It could be interesting to briefly discuss how your model compares with these GNN approaches. If a quantitative comparison is not feasible within the current scope, a qualitative comparison of your observations may be valuable.
- In the conclusion is stated: "The GNO is less accurate than PyWake, as expected, since a machine learning surrogate cannot fully replicate the fidelity of its source model." I recommend reformulating this sentence or removing it. Although this information would be very interesting, in the manuscript, the accuracy of the GNO and the accuracy of PyWake (or more precisely, the selected wake model) have not been compared, as PyWake has been taken as the ground-truth. Only if a more accurate ground-truth were used, such as a high-fidelity model or wind measurements in the field, could the accuracy of the two methods be compared. It is not excluded that the accuracy of the GNO could be better than PyWake, similar to data-driven methods trained with measurement data, which manage to predict wake losses in a wind farm more accurately than low-fidelity physics-based models.