

## Referee comment to

**Title: Long-term uncertainty quantification in WRF-modeled offshore wind resource off the US Atlantic coast**

Author(s): Nicola Bodini and Simon Castagneri

### General comments:

Overall, the manuscript is well written and structured in an understandable way. The introduction provides enough background and references to connect the article to the current state-of-the-art and the methodology section provides a high degree of details to understand and follow the workflow. Some elements in the methodology and data descriptions however require, in my opinion, additional description or additional references. While the analysis in general is presented concise and in a convincing way, I see the need for further investigation & provision of statistics with respect to the claimed generalization of the machine learning approach (see specific comment #2). Some additional words on limitations of the methodologies range of application would be appreciated as well. While I would suggest also some minor changes in the text and data presentation, my recommendation is publication after the questions/comments below have been addressed adequately.

### Specific comments:

#### general remarks:

#### #1 Usage of the term “boundary condition and parametric uncertainty”

The authors introduce above mentioned term to describe the share of uncertainty that can be described by NWP ensemble runs appearing several times in the manuscript. This term, however, can be highly misleading and misinterpreted especially in the context of regional NWP where the term “boundary conditions” and boundary condition “uncertainties” are used in a different context. I would suggest a term like “Ensemble-derived uncertainty” or something along those lines to avoid misinterpretation.

#### #2 Generalization of the machine learning approach

Section 5 describes the details of the validation of the machine learning approach. While the round-robin cross-validation approach shows promising behavior, the authors acknowledge the impact of spatial correlation due to the close vicinity of locations of validation. Here, I think it very crucial to quantify this impact by e.g. calculating mutual correlation coefficients of the Lidar time series to fully understand how independent the training and validation data actually are (maybe presented in a similar manner to Fig.4). This would also provide more details to the generalization skill and to what extent the high generalization skill is achieved purely due to high correlation of training and validation data. In this context, it would be also valuable if the mutual distance between the Lidars could be stated.

#### #3 More details in description of numerical WRF setup

While the setup is described fairly detailed, the following information is missing:

- Nudging is mentioned in l. 72, but it remains unspecified if grid or spectral nudging has been used. Please specify and provide details on parameter settings if they differ from the default settings in WRF.
- Land surface model, Microphysics, Longwave/Shortwave radiation, topographic data base and land use data in Table 1 lack references. Please add them for completeness in line with the other specifications in Table 1.

#### #4 Limitations of the proposed methodology

In the current version of the manuscript, very little is talked about the limitations of this methodology. While there is a statement made about the atmospheric conditions at the buoy north of Cape Cod (l. 244), a more in-depth critical elaboration is required in my opinion (maybe as additional paragraph in Sect. 6). This concerns especially the reliability / trustworthiness of the method when the random forest is applied to locations that are very different from the training data (geographic location, distance to training data, atmospheric conditions).

remarks addressing specific lines or sections:

l. 8: Since it is the abstract, please be specific about your used method (random forest) instead of the generic term “machine learning technique”.

l. 22: The stated reference for the currently installed offshore wind farms in the US is around 7 years old, please update with newer reference (maybe GEWCs Global Wind Report) to confirm and to use more up-to-date data.

l. 31: The reference (Skamarock et al. 2008) points towards Version 3 of the WRF model, but your WRF version seems to be 4.2.1 (Table 1). Is there a particular reason why the Version 3 reference is used here and not Version 4? Otherwise, please update.

l. 125/126: I would recommend to mention here again that the variables used for training are coming from the Lidar to avoid any ambiguity about the input for the training process.

l. 135 (footnote): The part “[...] which are both needed because each value of sine only (or cosine only) is linked to two different values of the cyclical variable” is unclear to me. What does it mean? Please consider elaboration or reformulation.

l. 149: What exactly do you mean by “typical single-site uncertainty”? Is this the averaged standard deviation of the residual time series for a particular location or something else? Please elaborate.

l. 270: Replace “wind energy” with “wind turbine power production”

**Technical corrections:**

Language corrections:

l. 3 [...] heavier relative weight [...] → [...] heavier is the relative weight [...]

l. 144 Then, to assess the uncertainty → To assess the uncertainty

l. 161 Introduction → introduction

l. 256 [...] to numerically model [...] → [...] to model [...]

Figures:

Fig. 1: For completeness, please state in the figure caption which markers indicate Lidar locations and which markers indicate buoy locations. This is currently unclear by looking at the figure only.

Fig. 6: I would suggest to transform this figure to a table since the bars for the parameters in the middle do not convey much information. For completeness, I would also suggest adding the explanation of “SST” to the caption similar to “WS” and “WD”.