

# Response to the comments from reviewer 2

Thanks for your comments on our manuscript. Here is our response to each of your comments. The response is in black color.

## General Comments

This study investigated future changes in extreme winds over Northern Europe by analyzing CMIP6 model results. The authors used a method called “spectral correction” to obtain wind speed data at 10-minute interval using 6-hourly model outputs, and then calculated the 50-year wind. By comparing historical simulations and future projections, they discussed possible changes in extreme winds in the region. This research covers an important topic, and it has some merits. However, I have a few major concerns about the methodology and the statistical significance of the results. I recommend a major revision and have listed my specific comments below.

## Specific Comments

1. Influences from natural climate variability. The authors compared the CMIP6 historical simulations with reanalysis data over the period 1980–2009 to assess the model performance. However, the observational results during 30-year period are likely affected by both anthropogenic forcing and decadal to inter-decadal natural climate variability. But phases of natural climate variabilities are randomly distributed in the models and therefore not synchronized with observations. In particular, given the relatively small sample size (6 models selected), comparisons between the model and observational results may be strongly affected by low-frequency natural climate variabilities (such as Pacific Decadal Oscillation, Atlantic Multidecadal Oscillation). Similarly, comparing future projections during 2020-2049 with historical simulations may not necessarily isolate the climate change effect, either.

We will break the above comment into the following questions:

- 1) what is the impact of natural climate variability on the extreme wind?
- 2) how is the low-frequency natural climate variability relevant for comparing the 6 selected models and observational results?
- 3) Can 2020-2049 vs historical simulations provide information on climate change effect?

Indeed, the natural climate variability was not examined in the calculation of the extreme wind. Probably it needs to be pointed out that U50 is obtained from a sample of annual maxima with a Gumbel distribution fit. This means that with 30-year data, we obtain ONE value of U50. Uncertainty related to the “low-frequency” variability will be expected in connection with such a calculation (with 30-year data) when the occurrence period of this “low-frequency” event is longer than 30 years. Indeed, if such effect is present, it is difficult to separate if it is the multidecadal variability or climate change.

It is such an interesting subject and we will include the discussion in the Discussion session and at the same time, we will look into and analyze the model data to see if such variabilities are present and if so, what is the effect of using 30-year data instead of a longer period.

On the other hand, the focus of this study is to investigate the change in U50 from historic period to the future period, and therefore it is of secondary importance to find out what caused the differences.

We also would like to point it out that when comparing with measurements, we used 18 (not 6) models (e.g., Fig.4). There are different groups of models formed based on some criteria and the purpose was also to see how this grouping affect the results. However, the results seem to be in general consistent. Apparently we need to improve the corresponding text.

2. Since CMIP models still exhibit biases in simulating many aspects of the global climate system, it is important to select the ones that can better simulate the subject of the research. However, the criteria used to evaluate the CMIP6 model performance (L161-168) seem a bit subjective to me. I would suggest to try using spatial correlation coefficient to assess the model-data agreement, and I wonder if how this may affect the conclusions in this study.

It is a relevant question whether, or how, these biases in the different aspects of the global climate system as in the CMIP models affect our conclusions.

The criteria from L161 – 168 are not enough to evaluate the CMIP6 model performance, as this is also reflected by the comparison with measurements from the three FINO stations (Fig. 4). We did not try to find out which model is best; the criteria and the corresponding grouping rather serve as sensitivity test how the conclusions depend on such grouping. We will make this clear in the revised version.

The reviewer suggests to use spatial correlation coefficient to assess the model-data agreement. The authors are not sure how to bring added values to this subject: it is quite difficult to find out which model can best represent the climate change impact regarding the change in the extreme wind – that’s the reason we use an ensemble of models and tried different groupings, as also mentioned in the previous paragraph. Another question is how to transfer the agreement for the historic period to the future period, as the subject is on the climate change effect.

3. I find that there is overall lack of statistical significance test in the results. For instance, L204-209 uses relative changes in the wind speed to determine whether the changes are significant. But even if a change is small in magnitude, it still can be significant as long as the signal is large compared to the noise. I would suggest to perform either a student’s t test or show model agreement in the figures to better demonstrate what signals are significant, which may provide more useful information.

We will consider using a student’s t test and some other tests to add more analysis on the statistical significance in the revised version.

4. L179-190: Here the authors compared their results with FINO masts, but I think more information about the data is needed to help the readers better understand the results, such as locations, observational periods etc.

More information about the FINO masts will be added. We will also compare the CMIP model results to reanalysis as written in the response to reviewer #1.

Best regards,  
The authors