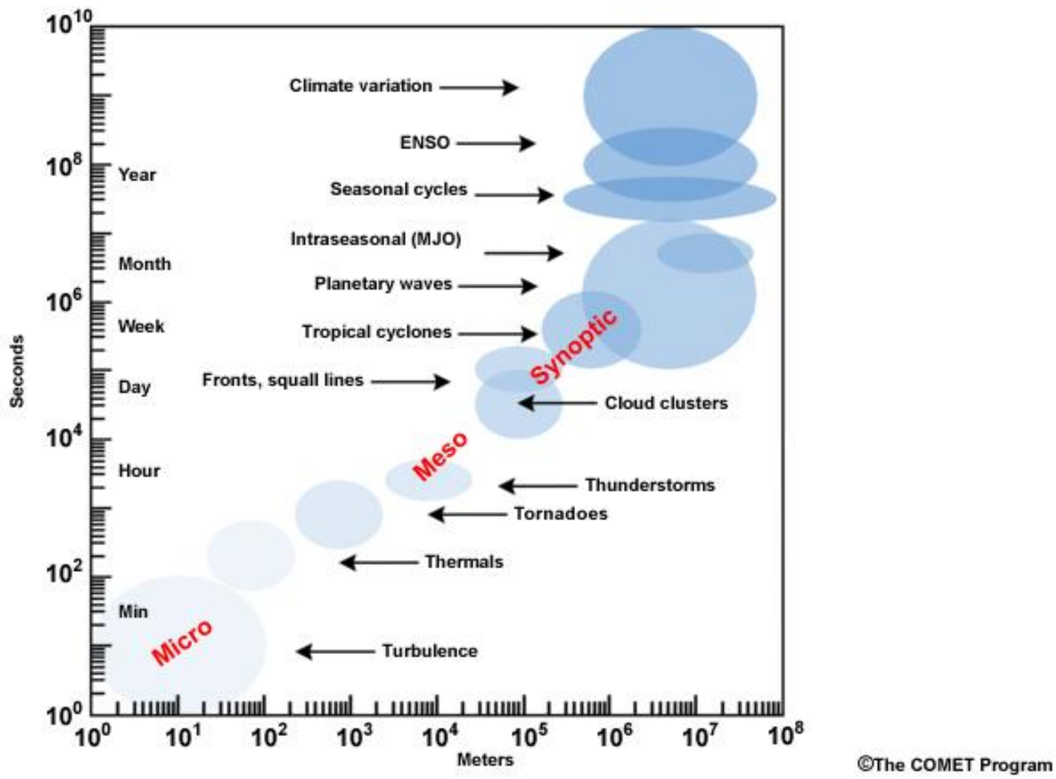
Reviewer 1:

General comments

1. *I really like the attempt to separate the 'mesoscale' and 'synoptic-scale' winds. However, I'm not really sure what physical processes are being picked up by this approach. The 'mesoscale wind', is designed to align with the effective resolution of the model, rather than any physical arguments. The 'synoptic-scale wind', calculated over a 45kmX45km grid box, is still well within the stated length-scale of 100km for mesoscale weather systems. For example, a sea-breeze disturbance could cover the whole larger box, as could a large thunderstorm complex. Can the authors try to quantify/justify this choice of length-scales? Some kind of spectral analysis or spatial filtering with known spectral properties could be helpful here. Either the length-scales should be clearly justified by physical arguments about the scale of mesoscale weather systems, or by practical arguments about the scale of variability that is of relevance to large offshore wind farms.*

The size of the large window is chosen to be smaller than the length scale of the synoptic weather systems in the figure placed hereunder. It is true that the MSVI as we use it has some selectivity towards variability at a certain spatial scale, but this is also the scale of variability that can result in sudden power ramping events for wind farms.



2. *When discussing the variability that the authors call long time-scales (6-12 hours), the focus is on diurnal processes over the land influencing the wind over the sea. However, there can also be a diurnal cycle over the water, especially in shallow areas where the water can warm or cool more*

rapidly. The authors did not state how the SSTs were specified in the model, nor whether they were being updated on daily or sub-daily time-scales.

The sea surface temperatures used in COSMO are inferred from the ERA5 reanalysis dataset. The SSTs do change in our simulation on a sub-daily timescale, but the diurnal cycle of these SSTs is at 1 to 2 K relatively limited. This cycle is comparable to the diurnal cycle found in potential sea surface temperature of the Baltic Sea Physics Analysis and Forecast provided by CMEMS. As our simulation set up does not include an oceanic circulation model the complex structure of eddies found in the Baltic Sea Physics Analysis and Forecast is not present in the COSMO simulation.

In the model setup section we have added a statement regarding the SST's being provided by ERA5, and we have added a section to the appendix discussing our findings from a test run we did with a constant SST.

2.1 Model setup

The model we use is the COnsortium for Small-scale MOdelling-CLimate Mode (COSMO-CLM) non-hydrostatic limited-area atmospheric model (Rockel et al., 2008). COSMO-CLM is a community model which is continuously maintained and developed by its users, under the coordination of the German Weather Service (DWD). The dynamical core of this model solves the primitive thermo-hydrodynamical equations describing a compressible flow in a moist atmosphere (Doms and Baldauf, 2018). The COSMO-model uses an Arakawa C-grid and a staggered Lorenz vertical grid with terrain following Gal-Chen coordinates. The horizontal grid is mapped out in rotated coordinates with a spacing of 0.0135◦. This corresponds to a horizontal distance of approximately 1.5 km which allows for explicit representation of deep convection in the model. For shallow convection the dynamical core of the model is expanded with a shallow convection parametrisation of Tiedtke (1989). Subgrid-scale turbulence is parametrised by a one-dimensional diagnostic level 2.5 closure scheme based on a prognostic TKE equation (Raschendorfer, 2001; Schulz, 2008a, b). Further parametrisations are present in the model that take subgrid-scale processes regarding micro-physical cloud processes and radiative transfer (Ritter and Geleyn, 1992) into account. COSMO-CLM has proven to be an adequate tool for long-term convection-permitting simulations, allowing for a statistical analysis of mesoscale weather systems (Brisson et al., 2016; Thiery et al., 2015; Van de Walle et al., 2020). It has also shown its value for studying wind speed metrics (Nolan et al., 2014; Wiese et al., 2019; Akhtar et al., 2021; Petrik et al., 2021). The model is directly driven by the 31 km resolution ERA5 reanalysis data (Hersbach et al., 2018). The sea surface temperature is also provided by ERA5 and updated on an hourly basis. More information about the nesting strategy can be found in appendix A. Some deficiencies in ERA5, like meridional variability of surface winds, and moist convection (Belmonte Rivas and Stoffelen, 2019) are better represented at the kilometre scale resolution. For practical reasons the calculation of the 10-year simulation has been divided into smaller periods, but using the restart files generated by COSMO these periods were initialised with a warm start. To account for the relaxation and spin up of the forcing data 20 grid points from every side of the domain are excluded from the analysis domain. The setup for this simulation has been used before to investigate the effect of wind farms on the regional climate in the German Bight (Chatterjee and van Lipzig, 2020).

(Appendix) We used some test runs to get to an optimal simulation setup, and in one of these test runs the sea surface temperature (sst) was kept constant at the initial (January) level. When comparing the spectrum for both runs we find on longer timescales no difference the intensity. On the shorter timescales there is however a higher intensity for the simulation with changing sst. As the sst does not change significantly over these timescales this higher intensity on short timescales might be because of the average sst in winter being higher than the sst from the first of January. In summer we find a higher intensity for the longer timescales for the simulation with a constant January sst. An enhanced contrast between land and sea surface temperature might thus result in more wind speed variability on long timescales.

3. *The authors looked at both temporal and spatial variability, but it would have been nice to see a greater attempt to relate these results. In particular, given that the temporal and spatial analysis should be capturing the same thing, why was it necessary to work with the spatial methods that place limitations on coastal areas? Why could the integrated periodograms over 'mesoscale' and 'synoptic-scale' periods have been compared, in a similar way to the MSVI? This would have given well-resolved maps of where the mesoscale variability was playing an important role?*

Mesoscale systems are by definition bound in both space and time, and we tried to examine mesoscale variability in both space and time. In the spatial analysis there was a limitation placed on coastal areas in order to limit the influence of the wind speed-up when transitioning from land to sea. The results of the MSVI metric become harder to interpretate when one of the windows contains a disproportionate amount of land pixels. The following statement on the complementarity that the spatial and temporal analysis have regarding one another is added to the methods section

The Welch method and the MSVI metric complement each other. The Welch method produces a spectrum with information about wind speed variations over different timescales for every pixel. Bundling this information in a spectrum does however remove the temporal resolution of that time series. The MSVI on the other hand aggregates spatial information into a metric, and in doing so gives up some spatial resolution. The temporal resolution in this method stays intact. As mesoscale systems are by definition bound in both space and time these two methods together offer a more complete view on offshore mesoscale wind speed variability.

Specific comments

1. *Page 3, line 41: "The effects of turbulence can be taken into account in Large Eddy Simulations" -> I think it should say "partly taken into account", since LES models only capture the larger part of the turbulence.*

This is changed in the manuscript.

2. *Page 3, line 53: "the majority of research is based on the onshore extent of the systems" -> There are a few references that look at the offshore part of land-sea breeze circulations that might be missing here. For example, Short et al. (2019) and Gille et al. (2005). In this context, the authors should also mention the land-breeze, which may be more relevant for offshore winds.*

These two references are indeed very interesting and are added to the manuscript

3. *Page 2, line 38: There are more recent versions of the wind speed spectrum that you could refrence - eg. Kang et al (2016).*

The data used in Kang et al. (2016) is taken from a continental location. For our paper we therefore opted to include a spectrum from a more coastal climate, such as Van der Hoven (1957) and Larsén et al. (2016).

4. *Page 3, line 45: Change first sentence to "Less is known about the impact of mesoscale weather systems on wind variability, for example in organised convection"*

This whole paragraph has been restructured and does no longer include the sentence: "*Less is known about mesoscale weather systems, for example in organised convection.*"

5. *Page 3, line 55: Add reference to Trombe et al. (2014).*

This interesting paper is quite relevant and is added to the paper.

6. *Page 3, lines 45-60: I think this section is missing discussion of a major source of mesoscale variability, which is from organised thunderstorms or MCS.*

This is added to the manuscript.

7. *Page 4, line 100: What height is the ERA5 wind speed accuracy quoted for?*

As suggested by the other reviewer, this statement is removed from the paper.

8. *Page 5, line 109: 'aggregated' - clarify - is this interpolated, or averaged?*

We have clarified in the manuscript that we mean averaged here.

9. *Page 5, line 130-133: The authors state that the time series is cut in overlapping sections. Later, it says that a 'Hann window is used to cut the signal into sections'. Is this duplication?*

This paragraph is slightly restructured to remove the duplication

10. *Page 5, line 134: 'fast natural variability' - is it 'fast', or just removing the noise?*

We can't really call it noise, as noise is considered an alternation to the original signal due to measurement imperfections. Our simulation should in principle not introduce something that can be classified as noise.

11. *Page 6, line 135: 'good estimate' -> how do you know it's 'good'?*

Good was meant to indicate that we consider the number of periodograms that are averaged to be large enough to converge to the underlying spectrum of wind speed variations. As this is never proven in the paper this statement of it being good is confusing, and "good" is removed from the sentence

12. *Page 6, line 137: 'integrated over a 3-month interval' - what does that mean?*

This is indeed a confusing statement so we removed it.

13. *Page 6, line 144: 'period' -> can the authors choose a different word? This could mean 'a period of time' or 'periods from the spectrum'.*

We have used "time interval" where possible to alleviate confusion for the reader.

14. *Page 7, line 169: suggest changing to "As the small window is contained within the large window'*

This is changed in the manuscript.

15. *Page 7, line 170: 'everywhere' -> 'everywhere else'*

This is changed in the manuscript.

16. *Page 10, line 220: The comparison with lidar data at high frequencies is mentioned, but as far as I can see, this is not shown in the figures.*

A reference to the figures of the comparison of the model periodogram and the lidar data in the appendix is added to the manuscript.

17. *Page 10, line 225-227: Why would the diurnal effects only show up on the 12h time-scale, and not the 24-hour time-scale?*

Masouleh et al. (2019) show that quite a large portion of sea breeze events have a duration of less than 12 hours. That is the reason why there is not one single peak in the spectrum shown in figure 5. While we believe that the temperature difference between land and sea is a driver for wind speed variability at these time scales (since the variability on these timescales is larger in summer compared to winter) the systems that this temperature contrast induces are apparently not always following this diurnal cycle.

18. *Page 11, line 230: It would be useful to contrast/compare the results to Vincent (2011).*

Indeed, the results of Vincent et al. 2011 are quite similar and a comparison is added to the manuscript.

19. *Page 11, line 240: The authors mention the issues of spin-up around the edges of the domain. This is an interesting problem, but raises the question of whether the boundary removed from the edges was sufficient. Can we really trust the results, give these effects?*

The appendix on the nesting strategy includes a small discussion on the tests we performed with a larger simulation domain. A reference to this discussion has been added to the model setup section.

20. *Page 11, line 246: Can the authors show the periodogram for power, as well as the integrated maps? The power time-series presumably has some constant sections where the wind speed is greater than 15 m/s or less than 3 m/s, and possibly some sudden jumps due to the cut-out speed being reached. What impact do these shocks have on the periodogram?*

These shocks elevate the spectrum quite a lot compared to the wind spectrum, but as these shocks don't necessarily happen at specific time intervals there are no clear peaks in this spectrum. The shape of the spectrum remains similar to the wind spectrum.

21. *Page 14, figure 9: I don't find this graph very useful. What is it supposed to be showing? Could the authors show it as an average annual cycle, averaged over the 10-year period? Or overlay a smoothed version so that the curve is more obvious?*

This graph is supposed to make clear to the reader that the MSVI results in a 1-D time series. This point is clarified in the text, and the figure is removed from the paper.

22. *Page 15, line 269: Why are the 'mesoscale winds' always more than the 'synoptic wind speed'? This is attributed to the land-sea breeze circulation and other mesoscale phenomena. In some cases, this might not be the case, since if the sea-breeze opposes the background flow, then it will weaken the wind overall. I agree that usually, mesoscale phenomena would lead to more windy conditions, but it should not be assumed that this is the case.*

It is true that the MSVI detects a locally elevated wind speed in each time step. Some mesoscale systems do indeed generate lower wind speeds, but due to the local nature of mesoscale weather systems a place with elevated wind speeds should be found in the vicinity of that mesoscale system. While our method is definitely not perfect the MSVI has the potential to capture this variability in wind speed created over these relatively small spatial scales.

23. *Page 15, lines 270-272: This section lacks references*

References have been added to the manuscript