In this document, the editor's comments are in black, the authors' responses are in red.

We thank the editor for her thoughtful and productive comments.

Dear Dr Pronk and co-authors,

I have read your answers to the reviewer's comments, and I find that you have not entirely addressed their suggestions. The manuscript is partly a repetition of what others have done on many sites for only two locations in the USA.

We would like to explain our choice of the two US locations selected for the analysis, which are indeed limited in number, but they represent sites of primary importance for current and future wind energy development in the US:

1) the offshore location is where the vast majority of the near-future offshore wind energy development will occur;
2) the land-based location is surrounded by many wind plants, a long-term atmospheric observatory, has hosted a field campaign (LABLE) in the last decade, and will host an international field campaign (AWAKEN) starting next year.

We have edited the text to better communicate the reasons behind our choice.

You display the results, but there is little attempt to explain or discuss their implications.

We have created a separate section to present our analysis of the wind plant wake impacts at SGP, which we think contribute to the WRF overestimation of nighttime winds at the site. We think that having a separate section, with slightly stronger text, will help the reader follow our attempt to explain the results presented in the sections before.

Also, we have added comments from the analysis of the surface heat fluxes at SGP, as suggested in a later comment.

The request of comparing the wind direction distributions for the sites to the ones simulated by ERA5 and WRF is reasonable.

We have added the following section to describe how the considered error metrics vary as a function of wind direction:

3.3 Impact of wind direction

Different wind direction regimes could also have an impact on the relative performance of WTK-LED and ERA-5. In general, we find that both WTK-LED and ERA-5 are capable of accurately representing the observed wind direction distributions at the considered locations (see histograms in the supplementary materials). At SGP, the wind is mainly from the south or north (wind rose in Figure 2), with a strong seasonal variability as summer months experience mostly southerly flow, whereas in winter a broader range of wind directions is observed. At the offshore location, the winds are mostly from the northwest or the southeast (wind rose in Figure 3), once again with a significant seasonal variability (as described in Bodini et al. (2019)), with summer seeing mostly southwesterly winds, and winter experiencing mostly northwesterlies. Based on these dominant wind regimes, we show in Figure 7 vertical profiles of the five performance metrics with data segregated by wind direction. At SGP,

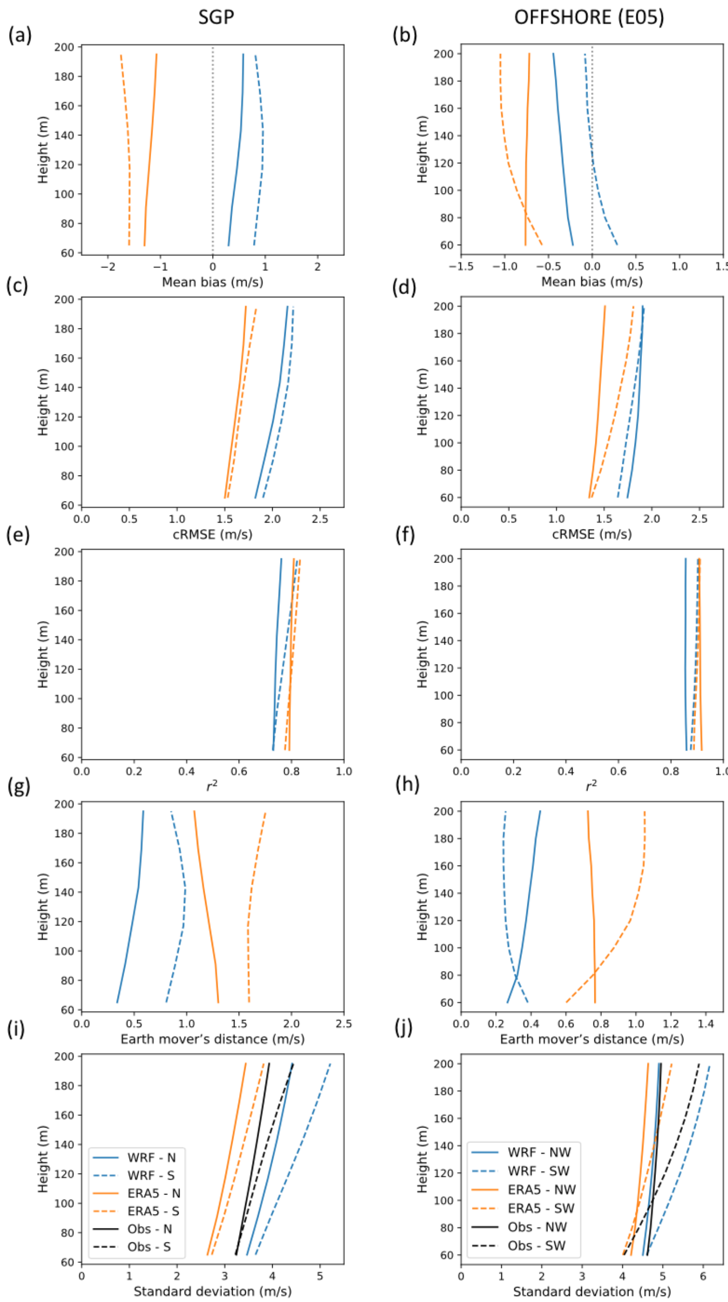

Figure 7. Vertical profiles of performance metrics segregated by dominant wind direction regimes at the SGP C1 site (left) and at the location of the E05 lidar (right). Results from the E06 lidar are included in the Supplementary Materials.

we include northerly (315-45°) and southerly (135-225°) flow, at the E05 lidar we show results for northwesterly (270-360°) and southwesterly (180-270°) winds.

At SGP, both WTK-LED and ERA-5 show a better performance for northerly flow, when looking at bias, cRMSE and EMD. In terms of correlation, southerly flow cases are slightly better for both data sources, but with a very limited difference. As will be detailed in Section 3.5, the worse performance for southerly flow at the site can be connected to potential wind turbine and wind plant wake effects, which impact the SGP observational site for southerly winds.

At the offshore site, WTK-LED shows better skills when modeling southwesterly flow, whereas ERA-5 is more accurate at representing northwesterly flow. This is true across all performance metrics.

Also, in the SI, we have included histograms which compare the distributions of hub-height wind direction from lidar, WRF, and ERA-5.

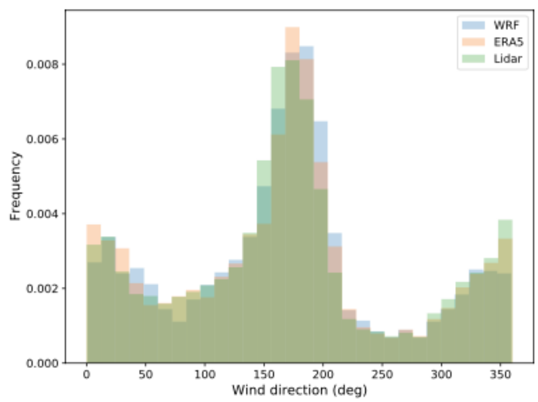

Figure S3. Histogram of 100-m wind direction from lidar data, WRF and ERA-5 at the SGP location.

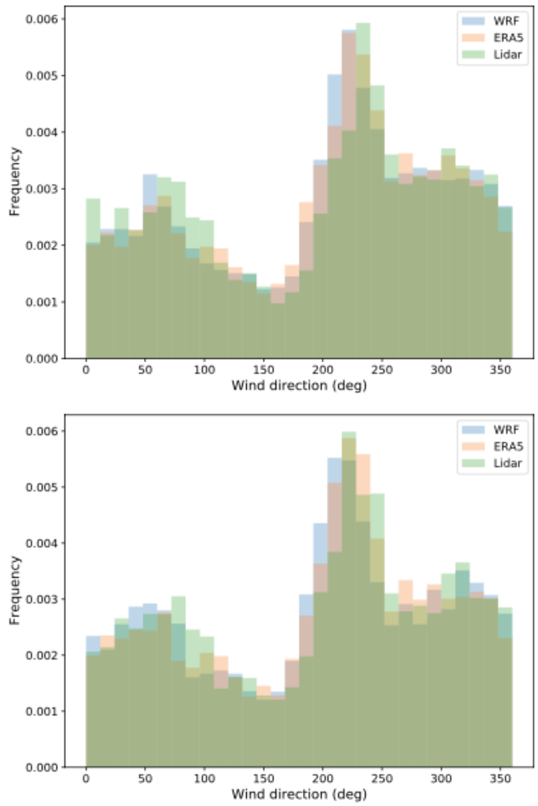

Figure S4. Histogram of 100-m wind direction from lidar data, WRF and ERA-5 at the E05 (top) and E06 (bottom) lidar locations.

Understanding the overprediction of the diurnal cycle over the land location will be very valuable. DTU has done WRF simulations in South Africa (https://orbit.dtu.dk/en/publications/mesoscale-and-microscale-downscaling-for-the-wind-atlas-of-south-) and compared them to the data at 19 62-m masts. We find that the overprediction of the daily cycle occurs in some sites but not others. We have difficulties identifying the physical causes of the errors because of the lack of surface fluxes. In your study, you have such data available at the land site. Perhaps there is a connection to the simulation of the soil moisture and surface fluxes, as suggested by a recent manuscript by Xia et (2021) (DOI: 10.1175/MWR-D-20-0363.1)? I suggest you try further to diagnose the causes of the daily cycle errors and incorporate the results and interpretation in your manuscript.

We have analyzed the sensible and latent heat fluxes from observations and from WRF at the SGP site as suggested, and added the following section to the paper:

3.6 Impact of surface fluxes at land-based site

The ability of WTK-LED of accurately representing soil moisture could also contribute explaining the overestimation of the diurnal cycle at SGP. In Figure 12, we compare the average diurnal cycle in sensible and latent heat fluxes from near-surface observations at SGP and from the WTK-LED. We find that WTK-LED overestimates the sensible heat flux during the day, while at night the WTK-LED prediction is, on average, quite accurate. The diurnal cycle of latent heat flux, on the other hand, is quite accurately modeled by WTK-LED throughout the day, on average, with a slight delay compared to the observed one. These results suggest that the exaggerated strong sensible heat flux by WTK-LED during the day (potentially caused by issues in the WTK-LED land use/land cover) is consistent with the WTK-LED's slight underprediction in hub-height wind speed during the day, as a strong sensible heat flux would increase turbulence and drag. On the other hand, soil moisture seems accurately modeled in WTK-LED given the relatively accurate latent heat flux values, so that the effects on the exaggerated diurnal cycle in hub-height wind speed are likely negligible Xia et al. (2021).

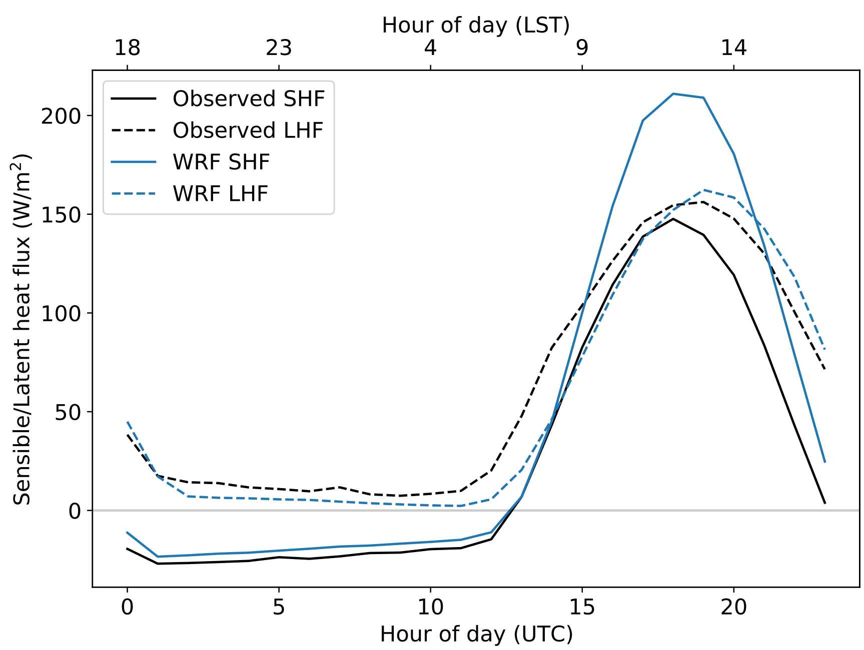

Figure 12. Observed and WTK-LED-modeled average diurnal cycle of sensible and latent heat fluxes at the SGP site.

Moreover, as mentioned above, we think that wake effects are related to the strong WRF overestimation of wind speed at night, as confirmed by the directional results added to the paper. These effects are discussed in Section 3.5.
All these considerations have been added to the Conclusions, too.

There are also a few other details to be sorted:

1. Some references are incomplete, e.g. "Vaterite ˘ı: On the stabilization of the general linear group over a ring." No year, no source. Other references are missing their DOI.
   We have carefully reviewed the references and added missing information where needed. We note that we could not find a DOI for the paper by Babíc et al. 2012.

2. The statement "ERA-5 data are publicly available from the ECMWF's MARS archive." is incorrect. The ERA5 model-level data is available from the MARS system; the pressure data is available from the Copernicus website.
   We have fixed this.