Thank you to both reviewers for the helpful comments. We have provided a point-wise response to both reviewers in this document. We copied the useful annotations into this document so that we could provide a clearer response.

Reviewer Comment 1

The paper investigates the wind resource around south-east Australia during peak electricity demand associated with abnormally cold and hot days. For this purpose the authors use three years of consumption data and 29 years of reanalysis data. They estimate the high demand days using a random forest model trained for which both consumption and reanalysis data was available. They conclude that particular weather patterns are associated with peak demand days.

The paper is structured well, is well written and provides the necessary amount of information, only the quality of the figures could be improved as well as the captions. The conclusions are clear and provide a nice overview of the potential of offshore wind being able to provide the necessary power supply on peak days.

More detailed comments are given in the attached PDF.

Thank you very much for the constructive comments. We have addressed all the comments and we think that the manuscript is now improved relative to the initial version. We have copied the annotations from the PDF to this document so that we could provide a point-by-point response.

Line 97: Query about errors in wind perturbations in ACCESS – R1: "I guess this is due to the resolution of 12km"

This is a good point. Actually the ACCESS version used in Brown et al (2017) had a resolution of 4km, but still had some errors in the wind perturbations. That is still probably partly to the resolution of the coastline, but also due to under-resolved mesoscale processes – for example, the strength of the land-sea breeze relative to the background wind will be affected by the land-sea temperature difference, or by boundary layer processes. We have clarified that we're talking about a different resolution of the same model. We have very few instances of the ACCESS model being validated relative to low-level winds, so this is still a key reference.

Line 114: R1: "difference turbine types would also be interesting, maybe some with lower rated wind speed or those that use smooth cut-out (most turbines these days deregulated smoothly from 25 to 30 m/s to avoid strong cut-outs of entire wind farms). Maybe something that could be included already."

Thank you for the suggestion, respectfully however it would be more suited for a follow up paper. Due to the wide range of possible wind turbines and site configurations (and subsequent expansion of the text required), we settled on computing the capacity factor using a single wind turbine power curve (Figure 2) to focus on the meteorology driving wind electricity generation. We note that differences from fine tuning the power curve were not sufficient to radically change the meteorology, and our BARRA wind speed data rarely exceeded the 25 m/s needed to be affected by the smooth cut-out. We *take this as a limitation of the research and have added the following text to the closing paragraph of the manuscript: "Furthermore, the single turbine assumption that underpins the capacity factor calculation could be improved by further incorporating common wind farm configurations and design typologies."*

Figure 2: R1: It is important to mention that the capacity factor was computed without wake effects (or am I misunderstanding something here). One could apply a reduction in capacity factor due to wakes for modern wind farms to account for this.
Also did you include density effects?

Wake effects were not included, due to the coarse resolution data used. Also, since we deal mostly with the relative impacts or anomalies of capacity factor, it unlikely that including wake effects would make much difference to the overall results. For similar reasons, density effects were not considered. We have also added extra text in several places to highlight the single-turbine assumption, that does not include wake effects.

There might be some interesting follow-up studies to be done here, considering the patterns in prevailing meteorology on high-demand days, that might introduce systematic differences in wake impacts on these days. This is, however, beyond the scope of this current study.

Line 119: R1: "But post-Covid consumption is again similar to the 2015-2018 period?"

We agree that it is a little unclear which period to use for the study, as there was still a long period of adjustment (such as high levels of working and studying from home) in 2022 and 2023. Therefore we chose to use the 2015-2018 period to be completely clear of this. We have added a caveat that we actually don't know how consumption patterns will evolve in the coming decade, which is also marked by societal factors such as uptake of EVs, level of electrified heating/cooking, and population growth. These factors place a limitation on this study, but we still think that similar types of synoptic patterns will continue to induce high electricity demand in the future.

Line 161: R1: "Any ideas as to why the RF is outperforming the LR on cold days? They seem to be on-par on hot days."

This is a very good question. We think this is related to the fact that the high-demand cold days are split between the two synoptic patterns shown in figure 10. This means that there are two distinct sets of predictors leading to high demand. The RF model is effective for handling this situation.

This is exemplified in fig R.1. Demand on both CDI and HDI days is strongly linked to temperature, with a negative correlation for HDI days and a positive correlation for CDI days, as expected. However, the higher CDI days also tend to be clustered around lower wind speeds, making wind a useful predictor. On the other hand, there is almost no relationship with wind speed on HDI days, because there can be both cold windy days and cold calm days.

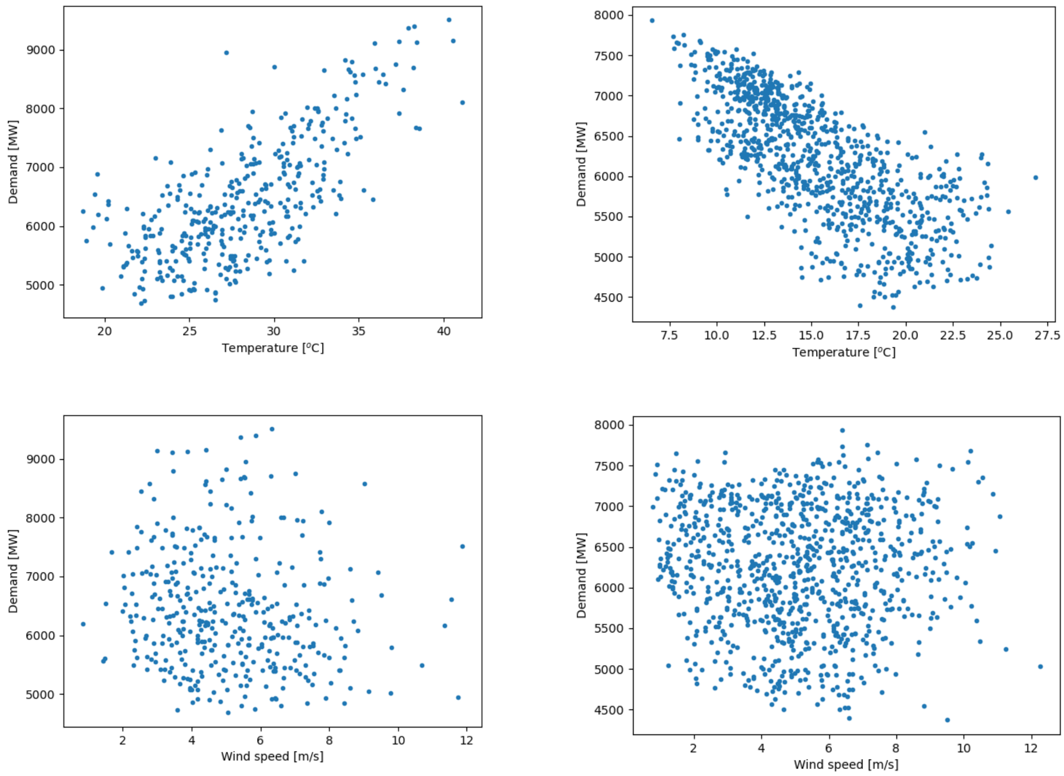

Fig R.1. Scatter plots of demand against temperature and wind speed for CDI days (left) and HDI days (right).

Moreover, if we consider the Linear-regression predicted versus observed demand for the two cases (fig R.2), colour coded by true positives (red), misses (black), false alarms (blue) and true negatives (green), we see that while the fit is reasonably good overall for the HDI days, it is clustered towards the middle of the distribution where there are many datapoints, and does not capture the high extremes very well.

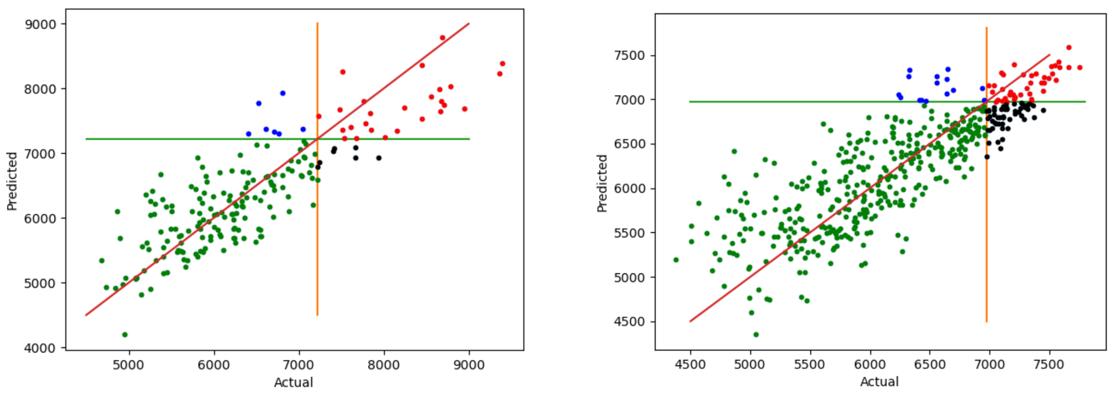

Fig R.2. Scatter plots predicted versus actual demand (testing data) for CDI days (left) and HDI days (right). Vertical red lines and horizontal green lines show the 80th percentile threshold of demand for each set of days. Points are colour coded according to true positives (red), misses (black), false alarms (blue) and true negatives (green).

A comment has also been added about this in the text.

Table1: Typo in table heading

Thanks. This has been fixed.

Line 184-185: R1: "It would be good to add a topographic map for those unfamiliar."

This figure has been updated to include topographic contours, and the caption has been updated accordingly (figure R.3). We have also deleted figure 5, as this was superfluous once the areas had been added to figure 6.

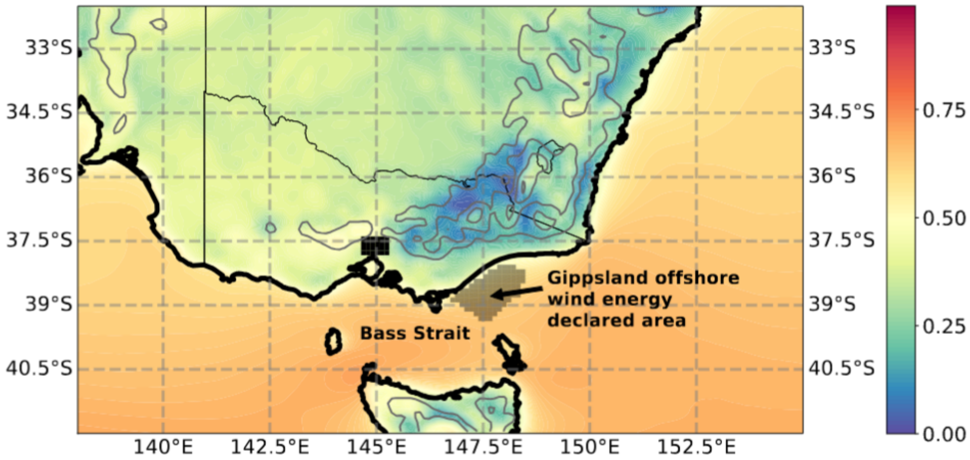

Figure R.3: "Add an outline of the offshore wind area, same applies to the anomalies maps"

This has been done for figure 6 (now figure 5). We decided not to add the area to the capacity factor anomaly maps, as this section does not focus on the Gippsland area.

Figure 8: "(left), (right) should be added"

Extra labels have been added to figures 7, 8 and 9.

Figure 11: "extremely large figure, rescale"

Figures 10 and 11 have been rescaled.

Line 294: "It would also be good to mention that due to the low resolution some local patterns and around complex terrain might change."

A comment to this effect has been added in the concluding statements.

Reviewer Comment 2

This study examines the wind resource across southeastern Australia during high electricity demand days, including both warm and cold days. The analysis combines three years of energy consumption data with 29 years of reanalysis data.

The study concludes that wind resources and energy consumption data align well on both warm and cold high-demand days.

Overall, I find the study being well-structured and a significant step toward making more informed decisions about offshore wind farm placement. That said, there is room for improvement, and my specific comments are outlined below. Also, as the authors noted at some point, I would be interested in seeing how higher spatial and temporal resolution data—capable of capturing complex topography, coastline features, and daily flow variations—could complement the findings presented in this study.

Thank you very much for the constructive comments. We have responded to all the comments and now have an improved version of the manuscript. Relating to the comment about the role of higher spatial and temporal data, we agree that this is a very important area. We are currently pursuing some extended studies using higher-resolution datasets now available and targeted modelling for wind energy areas. This did not really fit within the scope of this current paper.

More specific comments and suggestions:

Line 34: "This emphasizes that the variation in wind power is a combination of synoptic-scale processes and local-scale processes"

How are these local processes represented in the dataset?

Thanks for this comment. We changed the sentence to: "This emphasises that the variation in wind power is a combination of synoptic-scale processes and local-scale processes, some of which will exhibit a pronounced diurnal cycle.", since this paragraph is actually about the diurnal-scale variations. Later, in lins34-35, we add "We use a moderately high resolution reanalysis over Australia with a horizontal grid spacing of 12 km, which will partially resolve local-scale variations in wind speed associated with topographic and coastal processes."

Line 97: "some errors in the direction of the perturbations around areas of complex coastline."

Are these errors reflected in the patterns of any variables used in this study? If so, how?

It's a good question, and we don't know. The study of Brown et al (2017) was over a different domain to the one used here, and we have scarce data available to evaluate the seabreezes. This is an area that we are closely looking into in ongoing work. In the text, we have added the following: "Errors in the fine-scale structure of the seabreeze are likely present in the data that we use here too, although we note that the Gippsland offshore wind energy declared area is adjacent to a relatively straight coastline with the complex topography set back from the coast."

I recommend including a high-resolution regional map that shows both topography and coastline features, so readers can better understand the regional characteristics influencing the findings.

Topography contours have been added to figure 6 (now figure 5).

Line 149: "False Alarm Rate (FAR) and Probability of Detection (POD)"
Please include a reference or justification for selecting these two metrics.

A reference to Wilks (2011) has been added.

Line 151: "The predictors for the best performing model for high-demand hot days were CDI, Tmax, Tav, RH, WS, and DOW, where all predictors were significant at the 5% level except for RH, and Tav, while the predictors for the best performing model for high-demand cold days were HDI, Tmax, RH, Tmaxlag1, and DOW, where all predictors were significant at the 5% level."

It would be helpful to include the significance level for each variable in the final models. Which variables had the strongest statistical significance?

p-values have been added to the text.

Line 156: "Although there was some variation in the order of the most important predictors…"

Again, please include the significance levels for each variable. If there is a variation in the order of predictor importance, where might this variation stem from? How do you explain it?

The average importance of each of the selected variables has now been included in the text. There was some variation in the ordering of the most important variables due to the random seeding of the RF model. We specified 1000 random seeds and ran 100 realisations of the model, and selected the variables that consistently were identified as important predictors. Note that the relative importances of the top predictors were quite close to one another – for example, for the Hot80 days, the relative importance of the predictors was 26%, 22%, 29% and 10%. This means that it is not too surprising that the order of these first three predictors varied between realisations of the model, due to the randomness of the selected data and variables in each level of the tree.

Line 161: There appears to be a typo in Table 1 (should "Hot80 RF" be "Cold80 RF"?). Additionally, how do you explain the performance of the Cold80 LR model?

Thanks for picking up the typo. This has been fixed.

The performance of the Cold80 LR model can be explained by the fact that the HDI80 days are attributable to two broad sets of synoptic conditions, as shown in figure 10 (now figure 9). This means that there are distinct sets of predictors that both lead to cold days. We have discussed this in more detail in relation to a similar comment from reviewer 1. A comment has also been added about this in the text.

Line 183: "Average wind capacity factors"
Please clarify how this factor is calculated or assessed.

The following text has been added:

The average capacity factors are calculated by applying the wind turbine power curve (Fig. 2) to all wind speed data, and then averaging over the 29 year dataset.

Figures 8, 9, and 10: The plots should have clear titles or labels indicating which represent Hot80 and which represent Cold80 days.

These labels have been added.

Line 281: "While not examined in this work, we also note that certain synoptic patterns may be associated with concurrent hazards…"

While this is an important point, it feels somewhat disconnected from the rest of the discussion. Consider linking it more clearly to the implications for demand/management during extreme weather events.

We agree with this statement, and have slightly expanded the discussion around this point:

We also note that while certain scenarios may be associated with a favourable wind resource, this resource may be undermined by the co-occurrence of hazards - for example, the high-demand hot days are similar to the synoptic setup for heat-wave conditions (Henderson et. al. 2023) or when combined with an approaching cold front, fire weather risk (Reeder et. al. 2015) in SE Australia. Similarly the high-demand cold days are likely to overlap with hazards associated with cold, wet and windy conditions (Ashcroft et. al. 2009).