Review comments
Responses to Reviewer 1

We thank Reviewer 1 for the valuable comments and suggestions, which helped us greatly in improving the quality of the paper. The following shows each reviewer comment in blue followed by the author response in black.

The method is clear and results are useful. A few points need attention:

Annual maxima sampling: Using the single yearly max wind speed for the Gumbel distribution is standard but may miss years with several strong typhoons. The second-strongest wind in an active year might be greater than the annual max in a quiet year. This could affect the U50 estimate. Please discuss this limitation.

We completely agree, the Gumbel Distribution, while it is very useful, is not without its limitations. Therefore, in Section '2.6 The Gumbel Distribution' we have added the following statement addressing the limitations of the Gumbel distribution: Using the yearly maximum for the Gumbel fit can result in the discarding of other large storms that occur in the same active year. This, in turn, could lead to the tail retaining less information than other approaches. Since U_{50} is derived in the Gumbel distribution by extrapolating the fitted tail to an exceedance probability, the estimate could carry additional uncertainty. A peak-over-threshold approach using the Generalised Pareto distribution was tested, but the grid resolution produced too few exceedances per point. This lead to unstable and often negative shape parameters. The Gumbel distribution remains a standard choice but for this procedure, the resulting U_{50} represents the wind level associated with exceedance of the yearly maximum, rather than the full set of extreme storms that may occur within a year.

Comparison request: Please compare your 100m U50 results with those from Imberger et al. (2024). This will help show consistency with other recent methods.

Yes, these comparisons can absolutely be included. Their omission in the first submission was primarily due to the differences in the analysed heights and the limited visibility of the spatial distributions across the regions in the Imberger analysis. However, the data made available from the Imberger analysis made it possible to compare the studies more. The following statements have been added to the Results sections:

1. 3.1 Taiwan: While Imberger et al. (2024) presents results at 50 m and 150 m using CFSR, MERRA 2, and ERA5, the resultant dataset from the project has also been made available. Examining the 100 m results, the spatial distribution appears most similar to the CFSR output, with the peak located to the east of the Philippines. Whereas for the ERA5 and MERRA 2 results, the peak is located further north. The maximum U_{50} at 100 m in this analysis is 84.3 ms^{-1} which is closest to the ERA5 estimate of U_{50} of 86.7 ms^{-1}.

2. 3.2 Japan: The results presented here differ spatially from Imberger et al. (2024). In Imberger_2024, peak wind speeds occur primarily below 35°N and up to 140°E, with a north-eastward trailing feature. In contrast, this analysis shows a less pronounced decrease beyond 140°E and does not capture this north-eastward feature. This could be due to the difference in sampling size at high latitudes, as the dataset used in Imberger et al. (2024) does not exhibit the same reduction in data availability with increasing latitude. The maximum U_{50} in this study is lower than all values used in Imberger et al. (2024), however the ERA5 and MERRA 2 maxima are shifted further north into the Japan region, where as the CFSR results are more consistent with those presented here, aside from a single pronounced peak in the Imberger et al. (2024) analysis.
3. 3.3 East Coast of the United States: Imberger et al. (2024) identifies a Caribbean hotspot across all three datasets. In addition, the CFSR results show a separate, lower-latitude hotspot in the Atlantic Ocean, which is not evident in the ERA5 or MERRA 2 outputs which this analysis is more in line with. The CFSR and MERRA 2 datasets otherwise place their primary peaks above 35°N latitude. The maximum wind speed in this analysis (74.9 ms^{-1}) is closest in magnitude to the CFSR value of 72.0 ms^{-1}, although the locations of these maxima differ substantially.

Minor corrections:

Table 4: Longitude units for Taiwan/Japan should be "E", not "W". Please check the whole text.

Suggestion taken and the longitude units have been corrected**.**

Figure 2: Add latitude/longitude coordinate labels to the maps for clarity.

Suggestion taken and coordinate units have been added to Figure 2.

Responses to Reviewer 2

We thank Reviewer 2 for the valuable comments and suggestions, which helped us greatly in improving the quality of the paper. The following shows each reviewer comment in blue followed by the author response in black.

This paper presents a practical computational method for estimating 50-year return wind speeds at turbine hub heights across three tropical cyclone-prone regions: Taiwan, Japan, and the East Coast of the United States. The approach integrates IBTracks, the Holland model and Gumbel extreme value distribution, propagating uncertainties through Monte Carlo simulations. This is a standard industry practice, but usually performed at a particular offshore wind site, so I see the innovation on expanding the study to cover entire sea basins. The method performs robustly in Taiwan and Japan but results over US east coast exhibit spatial fragmentation and unexpected latitudinal patterns coming from the low data coverage. The uncertainty analysis effectively identifies input parameters and the sensitivity of each component.

Overall Assessment: The paper is well written and provides new insights for design of offshore wind turbines in regions where 50-year winds approach the limit for IEC Class S turbines. The contribution of this methodology is incremental but substantive, with limitations that are properly discussed. I have minor revisions throughout the paper that need to be addressed before publication. See below:

Line 6 – "method performs well in Taiwan and Japan". What do you mean by this? Be specific in terms of statistics and concrete numbers

Line 8 – "…performs less well". Same comment as above.

The reviewer is correct, the abstract has been updated to reflect the true meaning of these statements: The method aligns with previous studies through the spatial representation of wind speeds and maximum 50-year return wind speeds in Taiwan and Japan which can be attributed to the large sample size of data points located within a limited spatial area. The east coast of the United States exhibits spatial fragmentation and only partially aligns to the spatial representation of 50 year return wind speeds from previous studies, which conversely, is due to the smaller sample size and wider spatial region of which they cover.

Line 106 – I suggest you already explain here the reason for the choice of time frame for each region. It is strange to see such short period for ECUS, for example.

Thank you for the suggestion. A statement has now been included within Section 2.1 so that the reader is not caught off guard by the shorter time period: The time frame of ECUS is limited by the availability of the RMW parameter and is further explained in Section 2.4.

Figure 2- Improve the plots by adding lat/long coordinates and scale.

Thank you for the suggestion. The plots have been improved by adding the lat/lon coordinates and a scale (km).

Table 4 – this table is not needed if you properly make the plots in Figure 2.

Thank you for the suggestion. The advice from the previous suggestions has been followed and therefore this table has been removed as suggested.

Figure 3 – Same comment. Improve the quality and geo referencing of this plot, showing lat/long coordinates in the axis and a scale for reference.

Thank you for your comment. Coordinates and a scale have been added to this figure as well.

Line 174-175 – The methodology to calibrate the geostrophic drag law by tuning z0 is indeed non-realistic. How did you arrive to a single z0 value for the entire region? I suggest to plot the wind speed comparisons between the downscaled winds with the IBTracks, so one could see what are the spatial differences even with the calibrated z0.

Thank you for your comment and the review is right to be apprehensive about the calculation of z_0. The z_0 parameter was challenging to define for this paper and below we will try to outline how we arrived on a single value for z_0 for the entire region. There are several ways to calculate z_0 , each with its own challenges. We will outline what each potential avenue is, what the challenges where, whether these challenges could be overcome or if we moved on to another avenue.

First, we could calculate z_0 at every grid point and timestep for each tropical cyclone. Using the calculated surface friction (from the gradient wind speed formula), this could be input into the Charnock relation. We tested this approach, but when comparing the maximum wind speed from the scaled Holland model to the IBTrACS 10 m maximum wind speed, the Holland model significantly underestimates the winds, with the distribution centred around −10% $\left(\Delta u = 100 \times \left(\frac{\mathrm{U}_{holland} - \mathrm{U}_{\mathrm{bt}}}{\mathrm{U}_{bt}}\right)\right)$.

Since this study focuses on extreme wind speeds, using a z_0 that produces a 10% deficit is problematic, even if it is a standard approach – the validity of the Charnock formulation is questionable during tropical cyclone conditions. The maximum wind speeds are the most important in this context, as they are most likely to be selected in the annual maxima used for the Gumbel distribution.

For this reason, we treat z_0 as a surface correction parameter rather than a strictly physical roughness length, to match the Best Track wind magnitude. This can be justified by looking at the Holland model, which is itself empirical and therefore the framework already departs from a purely physical representation. In this formulation, z_0 does not represent a physical surface property but instead a scalar which maps from gradient wind (given by the Holland model) to 10 m wind speeds.

Given that IBTrACS provides 10 m maximum wind speeds, z_0 can be determined such that the scaled Holland model matches these values on average. This calibration can only be performed at locations where the maximum wind speed exists (this is the only wind speed within IBTrACS). Extending z_0 spatially using additional assumptions could introduce further uncertainty. Combined with the fact that the maximum wind speed is the most critical quantity, this leads to using a spatially uniform z_0 at each timestep for a given cyclone.

At this stage, a further decision is required: whether z_0 should vary by timestep, by cyclone, or be constant across all cyclones and timesteps. A single regional value is chosen. Estimating z_0 per timestep or per cyclone would effectively fit it to individual events rather than surface characteristics and would risk absorbing errors from both the Holland model and IBTrACS, and could lead to unstable values.

Therefore, a constant z_0 is used for the entire region. Once z_0 is fixed in space and time, uncertainty can be assessed. IBTrACS wind speeds have uncertainties of approximately ±7 kt in the North Atlantic and ±10 kt since year 2000, corresponding to about ±5–7% for category 5 tropical cyclones and ±9–12% for category 1 tropical cyclones in the North Atlantic and western Pacific respectively. Approximately 99% of the modelled values fall within 10% of IBTrACS, which is consistent with this level of observational uncertainty.

The overall goal is to produce reliable wind fields for estimating 50-year return levels. Using a regional mean z_0 provides a stable and consistent transformation, ensuring that variability in the results is driven by storm statistics rather than parameter fitting.

We hope that this outlines the different options available and why we have made the choices that we have made. We do not think that this is a perfect solution but we do think it is a reasonable solution that provides a conservative estimate for wind speed, which is suitable for offshore applications.

Following this, we have updated the text to reflect that it is a surface correction parameter rather than the actual roughness length within the Method section and further added an appendix explaining the choices taken to calibrate z_0.

Figure 4 – The 95th quantile of Gumbell does not seem to be the best way to illustrate the spatial uncertainty in these results. I would expect the 50-year winds to be much worse where we have lower data samples. Can you clarify why you chose this parameter? Also, since you have chosen an annual maxima to select your peaks, I think it is worth mentioning how many points are used for the Gumbel fit for each region.

Thank you for your comment. We understand that it can seem that the 95th quantile does not seem the best way to illustrate the spatial uncertainty of the results, however, it speaks to the uncertainty within the Gumbel distribution itself given the dataset. The objective was to show that even without analysing all the other uncertainties, this method itself inherently has uncertainty. In regards to mentioning the number of points used to fit the Gumbel for each region, we think there is a small amount of confusion here. The number of data points at each grid point is shown in Table 3, and that is because at each grid point, there is the exact same number of data points (within a given region). The reason for this is that all IBTrACS datapoints go through the Holland model, and so once all results from the Holland Model are compiled together, each grid point now has the same amount of data points associated to it. Where the difference lies is that some areas might have many more values closer to zero than others. Following this, this is why Figure 5c, 6c and 7c were developed, as a way to try and show where the data points were that had substantial wind speed values.

Within Section 3.1 we have included the following: In this context, the 95% confidence interval reflects the variability associated with the fitted Gumbel distribution. It does not represent the uncertainty of all input parameters. Uncertainties relating to the entire process are discussed in Section 4. Given that the Holland model is applied to every IBTrACS data point that fits the criteria, all grid points have the same number of data points which are listed in Table 2. To illustrate which areas in each region contain the highest wind speeds Figure 4c has bee developed. It shows the number of data points above 19m/s from the Holland model at each grid point. … The same reasoning holds true in Section 3.2 and 3.3.

Line 275 – For the uncertainty analysis, I understand that you assume a Gaussian distribution of all parameters for your Monte Carlo simulations. However, according to IEC 610400-1, Annex J, these parameters might actually have a Weibull distribution. Can you justify why you chose to stick with a Gaussian? What is the impact of this choice on your results?

The review is right that the distribution of the parameters could be Weibull or lognormal as the IEC standard states. However, within this analysis, we are approximating the distribution of the error to be Gaussian rather than the parameter itself. When running the Monte Carlo simulations, the wind speed is not selected from a wind speed distribution – as is done in the suggested methodology from the IEC standard -, but rather the wind speed that was previously defined given the values from IBTrACS and then it is varied within its error bounds. Given that these errors are given to us as a uniform distribution, we approximate them to be Gaussian rather than Weibull or lognormal because we have no information about the spread or skew so Gaussian is the most reasonable approximation.

The following has been added to Section 4.1.1: Note that while the IEC standard outlines that parameters such as wind speed may follow a Weibull distribution, the present analysis assumes a Gaussian for the associated error rather than the parameter itself.

Discussion – One uncertainty component not taken into account nor discussed is the fact that IBTracks dataset has a coarse spatial resolution, usually providing track data every six hours or so. How do you tackle the coarse spatial resolution of this underlying dataset? I know your method is grid-based, so it compiles all data within each grid cell. However, there is no spatial interpolation or uncertainty quantification related to the spatial resolution. Please add some lines explaining if your results account for this effect and how.

Thank you for the suggestion, the reviewer is right that we do not discuss the issue with the IBTrACS data being provided every 3 hours. We think that there are two components to this discussion. The first issue related to the sampling rate of observations. The track is sampled every 3 hours which could lead to under-representation of peak winds. This is a clear limitation that the reviewer has pointed out which has not been addressed.

The second issue relates to the temporal resolution of the data. The wind speed/observations which have a sampling rate of every 3 hours have an averaging period of of 1, 2, or 10 minutes depending on the agency that provides it to IBTrACS. Within the analysis here, they are all converted to 10-minute average. The method presented is not using the observations from the storm to build a time series but rather at each grid point, one value is taken per year (the annual maxima) and that is fed into the Gumbel fit. Therefore the temporal resolution of the data is determined by the averaging period of the wind speed (10-minute avg) rather than the sampling rate (3-hourly). Therefore, we would suggest that this is primarily a sampling issue rather than a resolution issue.

The following statements have been included in the Uncertainty section and the Discussion to address the limitations relating of the sampling rate. It elaborates that while it involves uncetainty itself, attemping to fix this problem then risks increasing the uncertainty.

The following has been included in the start of the Uncertainty section: The uncertainty associated with the sampling rate of IBTrACS is not quantified in this section. The dataset provides 3-hourly observations, and no interpolation to a finer resolution was performed to avoid introducing additional uncertainty. The aim is to evaluate the method using the raw IBTrACS data.

The following has been included in the discussion: An uncertainty which has not been quantified is the results of the sampling resolution of IBTrACS. When a TC is positioned between two consecutive track points, there is no information available and therefore there are spatial gaps in the track. This could lead to under-representation of peak winds, particularly where storms move quickly or data is sparse. As such, this is more likely to affect the results within the ECUS region than Taiwan or Japan regions. Interpolating the track to a finer temporal resolution could be attempted, however, it would require assumptions about the TC evolution between observations. This in turn which would risk the introduction of greater uncertainty than that of the sampling gap.

Line 471 – I suggest you explicitly mention some other uncertainty components that are not taken into account. For example, the impact of climate change on the track database when looking at 50-year winds.

Thank you for the suggestion. The reviewer is correct that there is no mention of climate variability within this analysis and it is an important factor to touch upon. However, it is a difficult uncertainty to quantify. An attempt to verify the impact on the analysis from climate variability was made however it became apparent that the uncertainty related to the Gumbel distribution dominated of the climate variability. We have included the following within Section 4, outlining the analysis: A second source of uncertainty that has not been extensively evaluated is the climatological variability of the underlying dataset. An attempt was made to quantify this by splitting each regional dataset into two equal halves and fitting the Gumbel distribution independently to each half. The difference between the two half-period estimates was tested for statistical significance by identifying differences between the two results that fell outside of the combined Gumbel confidence interval. Almost no grid points showed a statistically significant difference between the two periods. This indicates that the Gumbel uncertainty dominates over any detectable shift in the extreme wind climate between the two halves.

The following has been included in the discussion: A secondary uncertainty which has not been addressed is the potential impact of climate variability on the results. While all available data has been used, it may not represent the long-term TC climate if the period coincides with an active or in-active phase of natural variability. A period of enhanced TC activity could inflate the annual maxima which is fed into the Gumbel distribution which could return a higher U_{50}. A suppressed period would do the opposite. Although the split-half analysis revealed that the Gumbel uncertainty dominates over any detectable shift in the extreme wind climate between sub-periods, climatological variability remains an unquantified source of uncertainty.

Line 475 – Another methodology to generate synthetic datasets is clearly mentioned and suggested by IEC 61400-1 in Annex J. I suggest this is also mentioned as a potential solution to increase sample sizes and compute extremes from synthetic tropical cyclones from Monte Carlo simulations. This is particularly of interest since higher return periods are also needed for design.

Thank you for the suggestion, we had not previously included this method. The following statement in the description has been added to the discussion: A third approach is to generate a synthetic tropical cyclone database following the methodology outlined in Annex J of the IEC 61400-1 standard.

Conclusions – I suggest you expand your conclusions to summarize the potential for future studies and more detailed analysis with variations of Holand, synthetic tracks, other extreme value analysis methods, etc.. This section should not be just a bullet list of your findings.

Thank you for you comment. We have now updated the conclusion to summarise potential future studies and variations of the current method. The following has been included as the final paragraph within the conclusion: Future work could explore the use of alternative parametric models to better capture regional variations in TC characteristics and specifically improve performance within the ECUS region. In addition, applying the existing methodology while using larger datasets may reduce uncertainty and could provide reasonable results within the ECUS region. Using the existing methodology while comparing results using different variations of datasets would enable a systematic comparison between historical TC records and synthetic databases which represent a broader range of plausible, but unobserved, events. Finally, incorporating the influence of climate variability into the uncertainty framework is an important extension, as the current U_{50} does not account for this.