General comments

The paper is well written and comprehensible, and the research question is scientifically interesting and relevant. However, the statistical methodology presented in Section 3.4 requires more rigorous justification before publication. In particular, the assumptions underlying the combination of data from multiple turbines need to be carefully examined, as they directly affect the estimated return periods. Without sufficient validation of these assumptions, the main conclusion regarding the 50-year extreme event cannot be considered adequately supported.

Specific comments

S1. Page 2, lines 25-26:

The citation of Hansen et al. (2021) should be clarified. Since this reference was published after the IEC standard, it is unclear whether it contains the original data used to derive the ECD or a later compilation of these data.

S2. Page 3, lines 64-65:

It is not entirely clear why using wind turbines as anemometers is necessarily an advantage. The wind speed reconstruction may introduce uncertainties or smooth out certain features of the original wind field. Although the validation of the wind speed estimator is discussed later, its potential limitations should already be acknowledged here. Furthermore, it should be clarified that the wind direction measurements are obtained from nacelle anemometers rather than from the wind speed estimator.

S3. Page 5, lines 122-124:

The assumption that the events are independently and identically distributed (i.i.d.) requires further justification. As discussed in my comments on Section 3.4, this assumption is not sufficiently established by the proposed statistical analysis. The statement should therefore be revised or supported by additional evidence.

S4. Page 7, lines 135 and 148:

The selection of the 1000 largest values appears somewhat arbitrary. Why was a fixed number chosen instead of a threshold-based criterion? The same question applies to retaining only the top 10% of events. A brief justification of these choices, including their potential influence on the detected event statistics, would be helpful.

S5. Page 7, line 150:

The word “excluded” appears to be incorrect, as it contradicts the subsequent explanation that the analysis focuses on the most extreme events. Presumably, the authors mean “selected” or “retained”.

S6. Page 9, lines 202-204:

The argument concerning the converse of the Poisson convergence theorem requires clarification. Observing Poisson-distributed event counts does not, in general, establish that the underlying events are i.i.d. In particular, the same meteorological event may affect several turbines, potentially introducing spatial dependencies. Furthermore, a Poisson distribution of event counts does not establish that the gust characteristics are identically distributed across sites. The authors should clarify which assumptions are required for the cited converse theorem and demonstrate their applicability to the present dataset.

S7. Pages 17-19, Section 3.4:

Since the assumptions underlying the combination of data from multiple turbines are essential for the extreme value extrapolation and the estimated return periods, I believe that they require a more rigorous justification.

(a) **Poisson fit and the i.i.d. assumption:** As discussed above, a good fit of the event counts to a Poisson distribution does not establish that the underlying events are independent and identically distributed. In particular, potential spatial dependencies between turbines and differences in the distributions of gust characteristics across sites should be considered. How do the authors justify treating the combined observations as i.i.d.?

(b) **Exclusion of sites with zero events:** The authors report a poor Poisson fit when all sites are included, but obtain an improved fit after excluding sites with zero events. Although the use of a zero-truncated Poisson distribution is mathematically valid, this does not necessarily resolve the potential model misspecification indicated by the original poor fit. The authors should justify this approach more thoroughly and investigate how it affects the estimated effective observation time and the resulting return periods.

(c) **Goodness of fit and statistical uncertainty:** The conclusion that the event counts are adequately described by a Poisson distribution appears to rely primarily on the reported RMSE. Given the small number of observed events, particularly in Argentina, it is unclear whether the available data provide sufficient evidence to support this model. The authors should provide a more rigorous goodness-of-fit assessment, accounting for the small sample sizes, and quantify the uncertainty in the estimated Poisson parameters.

(d) **Uncertainty in the effective observation time:** Since the effective observation time T_m directly enters the calculation of the exceedance probability in Eq. (6), uncertainties in the estimated number of sites S_T and the Poisson parameter λ may affect the resulting return periods. The authors should clarify whether these uncertainties are propagated through the extreme value extrapolation.

Overall, I believe these points should be addressed before the validity of the regional extreme value extrapolation can be adequately assessed.

S8. Page 17, lines 342-344:

The authors set the off-diagonal elements of the correlation matrix to zero because no statistically significant correlations were detected. However, the absence of a statistically significant correlation does not establish statistical independence, particularly given the small sample sizes. The authors should clarify how statistical significance was assessed and discuss the sensitivity of the return surface to this assumption.

S9. Page 23, lines 426-430:

The conclusions regarding the effective observation time and the i.i.d. assumption should be reconsidered in light of the comments on Sections 2.4 and 3.4.

Technical corrections

T1. Page 3, Eq. (3):

V_{hub} should be explicitly defined.

T2. Pages 3-4, lines 84-86:

The definitions of $V(z)$ and V_r could be clarified.

T3. Page 5, lines 110-112:

The use of nacelle-mounted sonic anemometers for measuring wind direction should be introduced earlier, alongside the description of the wind speed estimator.

T4. Page 6, line 130:

There appears to be an extra space between "Table" and "2".

T5. Page 12, Table 3 and lines 266-270:

The event counts reported in Table 3 are inconsistent with those given in the text. Furthermore, the counts in Table 3 sum to 358 rather than the reported total of 356 events.

T6. Page 14, Figure 5:

The markers are difficult to distinguish. Increasing their size or using more contrasting colours would improve readability.

T7. Page 19, Table 6:

The fitting error is labelled "MSE", whereas the text refers to "RMSE".

T8. Page 19, Table 6:

For Germany, the reported values $S_T = 22$ and $T_m = 29$ years are inconsistent with Eq. (12), which gives $T_m = 1.5 \times 22 = 33$ years. Please clarify or correct this discrepancy.

T9. Page 21, line 386:

"On other hand" should be replaced with "On the other hand".

T10. Page 23, lines 414-416:

The event counts reported here are inconsistent with those in Table 3. Please ensure consistency throughout the manuscript.

T11. Page 29, Table C1:

For the United States, the reported minimum return period is larger than the reported maximum return period. These values appear to have been interchanged and should be checked.