

Response to Referee 2

We sincerely thank Referee 2 for the detailed, thoughtful, and constructive assessment of our manuscript. We particularly appreciate that the referee identified two closely connected interpretive issues that affected several of the original results: first, the use of the difference between a local CYGNSS observation and the IBTrACS storm-maximum wind as the principal mismatch quantity; and second, the possibility that the apparent sea-state relationship was largely inherited from storm intensity and wind-field structure rather than representing an independent sea-state contribution.

In response, we performed a set of additional analyses specifically designed to separate these effects. These included: (i) a signed decomposition separating the local CYGNSS–ERA5 difference from the ERA5-local-wind versus IBTrACS-storm-maximum difference; (ii) expanded storm-structure controls using normalized radius/RMW and storm-relative asymmetry; (iii) a fixed-algorithm nested feature analysis; (iv) a sensitivity analysis removing the ERA5-wind-dependent wave-age contribution; (v) a sampling-geometry sensitivity analysis for the SWH-transience component; and (vi) an evaluation of the storm-maximum-targeted adjustment against the local ERA5 wind.

These additional analyses materially changed the interpretation of the study. Across the four transfer experiments, the RMS magnitude of the local CYGNSS–ERA5 difference is approximately $3.0\text{--}3.2\text{ ms}^{-1}$, whereas the ERA5-local-wind versus IBTrACS-storm-maximum term has an RMS magnitude of approximately $13.7\text{--}19.1\text{ ms}^{-1}$. Thus, the large CYGNSS–IBTrACS difference is governed primarily by the contrast between a local wind quantity and a storm-scale maximum rather than by local CYGNSS–ERA5 disagreement.

After controlling jointly for storm intensity, normalized radius/RMW, and storm-relative asymmetry, the WWDI* partial Spearman association with the absolute local CYGNSS–ERA5 difference becomes approximately zero, whereas its association with the structural ERA5–IBTrACS term remains strong. The fixed-model feature analysis leads to the same conclusion: local atmospheric wind provides the dominant incremental information, bulk wave variables provide smaller and model-dependent contributions, and WWDI* provides only small and inconsistent additional information.

Accordingly, we have substantially revised the scientific framing. The revised manuscript no longer presents WWDI* as an independent local retrieval-error index or wind–wave disequilibrium as the demonstrated primary cause of local CYGNSS disagreement. Instead, storm intensity, local wind, and storm structure are treated as the primary context, while sea-state descriptors are treated as secondary, partly wind-coupled quantities whose incremental contribution must be demonstrated rather than assumed.

Main Issue 1

Reviewer Comment:

The first and main issue is the choice of measuring local wind errors as a difference with IBTrACS maximum wind. As detailed in the comments below, this introduces an immediate local bias in all observations, as errors will grow with distance to the maximum wind location. This general bias induces, in addition, system-scale and TC-case-wise biases, as the structure of the TC (notably profile steepness and wind field asymmetry) will strongly influence the mean errors over one case (a steeper/more intense TC will show a stronger average mismatch). This problem is actually visible in several of the presented results, notably the case study where the dependence of errors on the radial distance is very clear (Fig. 12h). The authors discuss this shortly but dismiss the

problem as a “local issue” that does not prevent their index from being relevant as a “system-scale / bulk state reference mismatch”. However, in my opinion, this judgment is invalid, for the reasons listed above: individual errors propagate on the system scale and thus on general results, due to TC structure. This choice of using the difference with IBTrACS is not really justified by the authors and appears as confusing to me: existing studies all rely on local wind measurements (Powell and Ruf 2025 use ERA5 winds, Ricciardulli et al. compare with other satellite wind fields, Ruf et al. with SFMR, etc.), and ERA5 winds are used in this study, but never as a basis to compute wind speed mismatch, which I find a bit puzzling.

Response:

We agree with the referee. The original manuscript used the CYGNSS–IBTrACS difference too broadly and, at several places, referred to it in language that could be interpreted as a local retrieval error.

IBTrACS maximum sustained wind is not a pointwise local-wind reference. A CYGNSS observation located away from the maximum-wind region can therefore differ strongly from IBTrACS $V_{\max}$ even when neither source is erroneous.

We directly quantified this issue by decomposing the signed CYGNSS–IBTrACS difference into a local CYGNSS–ERA5 term and an ERA5–local-wind versus IBTrACS–storm-maximum term. Across the four transfer experiments, the RMS magnitude of the local CYGNSS–ERA5 term is approximately $3.0\text{--}3.2\text{ m s}^{-1}$, whereas the structural/reference-scale term is approximately $13.7\text{--}19.1\text{ m s}^{-1}$.

We additionally repeated the analysis using normalized radius/RMW and storm-relative asymmetry. After these expanded controls, the WWDI* association with the absolute local CYGNSS–ERA5 discrepancy is approximately zero, whereas the association with the structural ERA5–IBTrACS term remains strong.

We have therefore revised the terminology throughout the manuscript. The CYGNSS–IBTrACS quantity is now described as a storm-maximum/reference-scale mismatch, not as a local retrieval error. The CYGNSS–ERA5 quantity is presented separately as a local product-to-product wind difference. ERA5 is not treated as absolute truth.

The reference-scale distinction has also been moved to a central position in the Results and Discussion rather than being treated only as a limitation near the end of the manuscript.

Main Issue 2

Reviewer Comment:

The second problem, connected to the first, is the choice to ignore or remove the local wind intensity from the construction of the sea state index. While the link between CYGNSS mismatch and intensity is documented (Ricciardulli et al., 2021, notably) and is discussed in some of the presented results in this manuscript (cf. confounder analysis and machine learning benchmark that both show that the WWDI index does not notably improve the correlation compared to intensity and wind structure parameters alone), the wind speed is not explicitly included as part of the designed index. Consequently, the analysis relies only on wave parameters and draws a link between CYGNSS mismatch and wave steepness. One could argue that this link is entirely due to the primary dependence on intensity, which link with wave steepness is strong. As such, the observed link between WWDI variables and CYGNSS mismatch consequently does not allow to describe the dependence of mismatch on sea state variables. This comment is backed up by the analysis of Figure 6, where the error evolution compared with sea state appears as reversed

when the wind contribution is removed. In my opinion, the study should seek to explain this “intensity-free” tendency, rather than only considering WWDI as a “bulk state” index that, as such, loses a lot of meaning and interpretability.

Response:

We agree that the original manuscript gave insufficient priority to local wind intensity and tropical-cyclone structure.

Rather than inserting intensity directly into WWDI* and thereby changing the meaning of an index originally designed from wind–wave descriptors, we performed a fixed-algorithm nested feature analysis. Storm intensity and radius were used as the structural baseline; local ERA5 wind was then added; bulk wave quantities and wave-state descriptors were added subsequently; and WWDI* was evaluated only after these predictors were already available.

The result is clear. Local atmospheric wind provides the dominant incremental information. Bulk wave variables provide more modest and model-dependent improvements, while WWDI* provides only small and inconsistent additional information.

We also strengthened the residual analysis using normalized radius/RMW and storm-relative asymmetry. After these controls, the WWDI* association with the absolute local CYGNSS–ERA5 difference is approximately zero. We therefore interpret the previous reversal in sign as evidence of suppression/confounding among intensity, storm structure, and sea-state variables rather than as evidence of a reversed physical wave effect.

Accordingly, we have revised the manuscript so that storm intensity, local wind, and storm structure are treated explicitly as the primary explanatory context, while sea-state variables are assessed only for their secondary incremental contribution.

General Remark: Figures are not linked clearly enough with the corresponding text

Reviewer Comment:

figures are not linked clearly enough with the corresponding text. Some figures are not even discussed in the text, or barely referenced, and others would benefit from a more explicit and thorough referencing

Response:

We agree.

We have revised the figure sequence, captions, and surrounding Results text so that each figure is introduced where its corresponding result is discussed. The revised text explicitly identifies the relevant panels and quantitative values rather than presenting the interpretation first and referring to the figure only afterward.

Panels that were not necessary to the immediate argument have been separated, relocated, or removed where appropriate.

General Remark: Interpretation of variability among basins and storm phases

Reviewer Comment:

In several results, authors point at a variability in the error/mismatch, without actually inter-

preting, or trying to explain it. For instance, the fact that some basin-to-basin transfers behave differently from others, possibly pointing at basin-specific conditions, are only mentioned without any interpretation. Same with the comparison with maximum intensity location, which shows larger errors for some basins occurring after the main intensity peak. While the main issue of IBTrACs referencing makes any interpretation complicated, the fact that none is proposed to explain results that appear as interesting left me with questions

Response:

We agree that the original manuscript sometimes reported transfer or storm-phase differences without sufficiently distinguishing between an observed pattern and an established physical explanation.

We have revised the discussion so that interpretations are provided where they are supported by the expanded analyses. Where the analyses do not identify a unique mechanism, this is now stated explicitly rather than attributing the differences speculatively to basin geometry, synoptic forcing, coastal configuration, or delayed wave adjustment.

In particular, the additional analyses do not identify a unique cause of the observed post-peak asymmetry. This pattern is therefore now reported descriptively rather than being attributed uniquely to delayed wind–wave adjustment.

Comment on L. 35

Reviewer Comment:

a short definition could be useful here

Response:

Agreed. We have added a concise definition of wind–wave disequilibrium at first use.

In the revised manuscript, wind–wave disequilibrium refers to a departure of the observed bulk wave state from a state adjusted to the contemporaneous local wind forcing, recognizing that wave growth, decay, propagation, and changing wind forcing occur over finite temporal and spatial scales.

We have also avoided implying that every such departure necessarily causes a CYGNSS retrieval difference.

Comment on L. 48–49

Reviewer Comment:

what is the advantage of CYGNSS for TC observation ? Coverage, resolution, frequency (saturation/cloud penetration) ? What are the defaults ? This paragraph states that CYGNSS is useful but does not back up the statement.

Response:

We agree that the original statement was insufficiently supported.

We have clarified that CYGNSS provides frequent tropical-ocean sampling through a multi-satellite GNSS-R constellation and operates using reflected L-band signals, providing complementary all-weather sampling of tropical cyclones.

We have also stated the relevant limitations, including finite spatial support, dependence on an empirical geophysical model function, heterogeneous high-wind scenes, and limited ability to represent compact local extrema.

The revised manuscript therefore describes CYGNSS as a complementary observing system rather than as generically superior to other satellite wind sensors.

Comment on L. 64

Reviewer Comment:

vague... surface waves, swell ?

Response:

Agreed. We have replaced the vague wording with more specific terminology.

Where appropriate, the revised manuscript distinguishes wind-sea and swell contributions and refers explicitly to the bulk variables used in the analysis, including significant wave height, mean wave period, wave age, wave steepness, and the derived departure/transience descriptors.

Comment on L. 74

Reviewer Comment:

Which “conditions” exactly ? The previous sentence lists a set of parameters that describe the sea state. We understand that the disequilibrium is a departure from a steady state of balance between wind and waves, but this should be stated more clearly (maybe explicitly defined above, cf previous comment)

Response:

Agreed. We have stated explicitly that the “conditions” refer to departures from an adjusted relationship between contemporaneous local wind forcing and the bulk wave state, rather than simply to large waves or strong winds.

This clarification has been introduced together with the first definition of wind-wave disequilibrium.

Comment on L. 94

Reviewer Comment:

this might not be clear without acronym explanation. Maybe use an expression like “basin to basin transferability” instead ?

Response:

Agreed. We have used the explicit expression “basin-to-basin transferability” at first occurrence and introduced the WP→BOB, BOB→WP, WP→ATL, and BOB→ATL abbreviations only after the transfer design is defined.

Comment on L. 111–112

Reviewer Comment:

this is very vague at this stage. Maybe reference section 2.4 ?

Response:

Agreed. We have replaced the vague wording with an explicit cross-reference to the subsection in which the corresponding methodology is defined.

Comment on L. 118–122

Reviewer Comment:

the high sensitivity of GPS signal to sea surface conditions might gain to be clarified, here or in the introduction. Most satellite surface wind measurements (microwave scatterometers, SAR) are not as affected by surface wave conditions, so it might be useful to explain clearly that this is a specificity of CYGNSS due to its frequency band, which makes it more sensitive than other usual measurement satellites.

Response:

We agree that the underlying measurement physics required clearer explanation.

We have clarified that CYGNSS uses reflected L-band GNSS signals whose bistatic scattering depends on sea-surface roughness and slope statistics. The empirical wind retrieval therefore relates a roughness-sensitive observable to wind speed, meaning that wind and wave conditions are physically coupled in the measurement environment.

However, we have avoided making the broader claim that GNSS-R is universally more sensitive to wave conditions than scatterometers or SAR, because such a comparison depends on frequency, measurement geometry, observable, and retrieval formulation.

Comment on L. 134

Reviewer Comment:

what does this flag consist of ?

Response:

Thank you. We have clarified this explicitly.

For the NOAA CYGNSS Level-2 SWSP v1.2 product used in this study, `sample_flags` is a bit field and bit 0 represents the overall quality flag. Samples were retained when bit 0 was equal to zero, corresponding to the criterion `(sample_flags & 1) == 0`.

The revised manuscript now states the meaning of the flag directly rather than reporting only the bitwise expression.

Comment on L. 155

Reviewer Comment:

I am not familiar with ERA5 data, but in what conditions would observations “lack finite values” ? Is it near land ? What does it refer to ? Please specify and maybe explain shortly in the text.

Response:

We agree that the wording was too vague.

“Non-finite” refers to a numerical validity check applied after collocation/interpolation. A required ERA5 or derived variable must contain a finite numerical value rather than a missing, masked, NaN, or infinite value.

We have not attributed such cases specifically to coastal proximity because that cause was not demonstrated by the data. The revised text now describes this as a numerical/collocation validity requirement.

Comment on L. 212–213

Reviewer Comment:

How much of the dataset is affected and filtered as physically non plausible ? Specifying at each step of selection how many samples are removed from the dataset could be useful to realize the impact/utility of each filtering step

Response:

We agree that sample attrition should be reported wherever it can be reproduced reliably.

We re-audited the final analysis-ready datasets and confirmed the final sample of 212,697 unique physical matchups. Reapplication of the final numerical and physical validity criteria used in the Ocean Science analysis retained the analysis-ready sample.

However, the historical upstream dataset-building workflow did not preserve a complete and independently auditable filter-by-filter counter for every earlier construction operation. Rather than reconstruct unsupported retrospective numbers, we have reported the sample counts at the major stages for which the counts can be reproduced directly and have clearly distinguished the upstream construction stage from the final analysis-level screening.

Comment on L. 214

Reviewer Comment:

what is a duplicate observation ? Does it mean the exact same observation and if so, why does CYGNSS record duplicates ? How many data does this represent over the entire dataset ?

Response:

We agree that the original wording could incorrectly imply that CYGNSS itself repeatedly records the same physical observation.

Here, a duplicate refers to repeated representations of the same physical CYGNSS matchup arising during file concatenation or processing, rather than to legitimate nearby observations.

We re-audited the final BOB, WP, and ATL master datasets using physical observation keys and found zero duplicate physical observations remaining in the final analysis-ready files.

We have clarified this definition and now state explicitly that the final total of 212,697 represents

unique physical matchups.

Comment on L. 288–292

Reviewer Comment:

this list is very unclear at this stage of the article. enumerating the methods that are described in following sections might be confusing. Authors should. at least, reference the section describing these methods in this sentence.

Response:

Agreed. We have shortened this passage and explicitly referenced the subsections in which the corresponding methods are defined rather than presenting a long unexplained list before the methods are introduced.

Comment on Fig. 2 caption

Reviewer Comment:

the first TC-WWDI index does not appear on the figure. I am guessing it should be on the right in the second box. Maybe add it for clarity ?

Response:

Agreed. We have revised the workflow figure and caption so that the equal-weight TC-WWDI and the subsequently learned WWDI* are shown as distinct but related quantities.

The revised figure now makes clear that TC-WWDI is the equal-weight diagnostic formulation, whereas WWDI* is the training-only evidence-weighted formulation derived from the same component family.

Comment on L. 303–305

Reviewer Comment:

It is unclear to me why ETC is used as the principal response variable when it is a reference-relative mismatch, while EERA5 is a point-wise error. What is the meaning of ETC, given that a mismatch at a given point between CYGNSS and IBTrACs doesn't necessarily mean that there is an error in either of the two sources, but maybe that the observation is located away from the maximum wind location ?

Response:

We agree with the referee.

The quantity previously denoted E_{TC} cannot be interpreted as a local error because IBTrACS provides a storm-maximum reference rather than the wind at the CYGNSS observation location.

We therefore performed the signed decomposition described above. The results show that the structural/reference-scale component is substantially larger than the local CYGNSS–ERA5 product difference.

Accordingly, the revised manuscript describes E_{TC} as a storm-maximum/reference-scale mismatch and distinguishes it explicitly from the local CYGNSS–ERA5 product difference.

Comment on the figure around L. 350

Reviewer Comment:

This figure is not described or interpreted anywhere. The link between the text and the different panels is unclear: “dimensionless scaling” is shown in panel (a), but what exactly is the point of comparing Hw^* and A_w ? The plot is not interpreted at all. “weak wind behaviour” is shown, it seems, in panel C, but again what is the point of the figure ? What is the conclusion ? Panels (b) and (d) regarding transfers are not even mentioned in the text. Why does the WP-BOB transfer show a lower correlation, for instance ? This figure should be referenced in the corresponding paragraphs, along with the corresponding interpretation showing how it backs up/validates your methodology.

Response:

We agree. The original figure combined several methodological diagnostics without explaining sufficiently what each panel was intended to demonstrate.

We have revised the figure and surrounding text so that each panel is discussed explicitly where it is introduced. The revised text states why the dimensionless variables are compared, what the weak-wind sensitivity test assesses, and what conclusions can and cannot be drawn from the transfer panels.

We have not assigned a unique physical explanation to the lower WP-BOB correspondence because the present data do not independently establish such a mechanism.

Comment on the y-axis labels of panels (b) and (c)

Reviewer Comment:

The y-axis labels on panels b and c are not very clear: R^2 regarding to what ? correlation between what and what ? This could be more explicit.

Response:

Agreed.

We have explicitly distinguished between observation-level predictive statistics and the descriptive R^2 calculated from binned median responses. Where the binned statistic is retained, it is now labeled as binned-median R^2 so that it cannot be mistaken for an observation-level predictive coefficient of determination.

Comment on the graphical form of panels (b) and (c)

Reviewer Comment:

The form chosen for panels (b) and (c) is not optimal: The bar plots for panel (b) are all between 0.8 and 1, making the differences hard to see. Maybe 95% intervals could also be added to show the spreading, instead of just the mean, replacing the bar plot with a box plot including mean/median, quantiles and min/max

same remark for panel (c): the line plot for a comparison like this is not the best choice, a box plot would give more information and with more clarity.

Response:

We agree that the original visualization compressed the differences and did not communicate uncertainty sufficiently well.

We have revised these panels so that variability or uncertainty is displayed more clearly using the available resampling information rather than emphasizing small differences through compressed bar or line plots.

Comment on panel (d)

Reviewer Comment:

Panel (d) should probably be separated from this figure, as it shows the $WWDI^*$ index which is introduced much later in the text.

Response:

Agreed. We have separated or relocated this panel so that $WWDI^*$ is no longer shown before it has been formally introduced.

Comment on L. 408

Reviewer Comment:

this variable is not defined

Response:

Agreed. We have defined the variable explicitly at first use and ensured that all symbols appearing in the normalization procedure are introduced before the corresponding conditions are stated.

Comment on L. 409

Reviewer Comment:

the criteria for invalidation are unclear. How can $Q_{j,98}$ be smaller than $Q_{j,2}$, since they are defined to be otherwise ? What is $N_{j,\text{finite}}$? Please give more details.

Response:

Thank you for identifying this ambiguity.

$N_{j,\text{finite}}$ denotes the number of finite training observations available for component j .

The percentile condition is a defensive numerical safeguard. Under a non-degenerate distribution, $Q_{j,98}$ is greater than $Q_{j,2}$. The implementation checks that the upper percentile is strictly greater than the lower percentile because equality can occur for a constant or effectively degenerate component, in which case normalization would involve a zero range.

We have stated this explicitly and simplified the wording in the revised manuscript.

Comment on L. 444–445

Reviewer Comment:

this is equal to TC-WWDI, right ? If so, why not relating WWDI to TC-WWDI more explicitly ?*

Response:

We agree that the relationship needed to be stated more directly.

TC-WWDI and WWDI* are constructed from the same diagnostic component family but are not identical. TC-WWDI is the equal-weight diagnostic index, whereas WWDI* is the training-only evidence-weighted version with weights frozen before external evaluation.

We have made this relationship explicit when WWDI* is introduced.

Comment on L. 468–473

Reviewer Comment:

what about the ATL basin ? If it is only for validation and not a training dataset, it should be specified.

Response:

Correct. ATL is used as an external third-basin evaluation dataset and is not used to learn the WP- or BOB-based normalization parameters or WWDI* weights.

We have stated this explicitly in the revised transfer-design subsection.

Comment on L. 498

Reviewer Comment:

what motivates this choice ? What is the advantage of the Spearman correlation compared to the Pearson correlation ?

Response:

Spearman correlation was selected as the primary association metric because the relationships are expected to be monotonic but not necessarily linear, and because rank correlation is less sensitive to extreme values and nonlinear scaling than Pearson correlation.

We have added this rationale explicitly. Pearson correlation is also reported where it provides useful complementary information.

Comment on L. 504

Reviewer Comment:

I am not sure about the meaning of this expression: “the larger of 30 observations”. Is it a typo ? Does it mean “more than 30 observations” ? If not, larger than what ?

Response:

We agree that the wording was unclear.

The implementation used the maximum of 30 observations and a small fraction of the external sample as a defensive minimum-support safeguard.

We have rewritten this statement unambiguously.

Comment on L. 504–505

Reviewer Comment:

this is unclear to the reviewer. The previous paragraph states that “external observations are divided in 10 equal count bins”, meaning each bin represents 10% of the external dataset. Yet, here, you suggest that a bin could contain less than 0.2% of the external dataset. I most likely misunderstood something... Could you clarify ?

Response:

The referee is correct that the original wording was unnecessarily confusing.

The external observations are divided into approximately equal-count deciles. Therefore, under normal conditions, each populated bin contains approximately 10% of the external observations, making the much smaller fractional support threshold effectively redundant.

We have simplified this section so that the primary procedure is described as equal-count decile binning. The minimum-support condition is now described only as a defensive numerical safeguard.

Comment on L. 506–507

Reviewer Comment:

There are many justifications regarding the interpretation of the meaning of variables (lines 505,460,428,419,140,28,1148,1103,1021,825,788, and more). Is there a specific reason why you insist on that ?

If there isn't, then I suggest avoiding repetitions. In the present example, it is clear based on the previous variable definitions that bin-wise defined variables are “regime-scale” and not point-wise observation level measurements.

Response:

Agreed.

The original manuscript repeated the distinction between regime-scale and pointwise interpretation too often. We have consolidated this explanation at the point where the diagnostic is introduced and retained it elsewhere only where it materially affects interpretation.

Comment on L. 716–717

Reviewer Comment:

this difference among basins doesn't seem specific to dimensionless wave height. The behaviour of wave age is very similar and the wave steepness WP-BOB transfer seems more dissociated from other transfers than for the two other variables. Could you explain why you insist on dimensionless wave height ?

Response:

We agree that the observed transfer differences are not specific to dimensionless wave height.

The revised discussion now compares dimensionless wave height, wave age, and wave steepness consistently rather than emphasizing one variable without sufficient justification.

Comment on L. 720–721

Reviewer Comment:

What justifies this assertion on the figure ? This is unclear to me. Is it because of the error bars variability ? Please specify/detail.

Response:

Agreed. We have explicitly identified the figure features and quantitative statistics supporting the corresponding statements.

Where the figure shows a change in binned median response but does not establish strong pointwise discrimination, this distinction is now stated directly.

Comment on L. 721–722

Reviewer Comment:

maybe expliciting the previous sentence will clarify this, but from the figure it is not obvious why TC-WWDI would not be considered as an individual mismatch diagnostic as well as a regime separator. The error bars are pretty separated from one bin to the next so it seems that the individual errors are basically related to the regime, and thus to TC-WWDI. More explicit links between the figure and your assertions would help identify and understand this distinction.

Response:

We agree that the original figure and text did not establish this distinction clearly enough.

Separated bin medians and confidence intervals show that the conditional mismatch distribution changes across diagnostic regimes, but this does not establish accurate diagnosis of individual observations because substantial within-bin variability remains.

More importantly, the expanded storm-structure-controlled analysis shows that the WWDI* association with the absolute local CYGNSS–ERA5 difference becomes approximately zero.

We have therefore removed the interpretation of the original bin separation as evidence that TC-WWDI or WWDI* provides an independent pointwise local-error diagnostic.

Comment on L. 723–724

Reviewer Comment:

this sentence is unclear. is it a justification for the previous assertion ? what is the signification/interpretation of “asymmetric, distribution-free confidence intervals” ?

Response:

We agree that this terminology required explicit definition.

The plotted intervals are nonparametric confidence intervals for the population median based on binomial order statistics. Because the lower and upper order-statistic bounds need not lie

symmetrically around the sample median, the displayed error-bar distances are asymmetric.

We have replaced the shorthand phrase with this explicit explanation in the Methods and figure caption.

Comment on Fig. 4 caption

Reviewer Comment:

the last panel is not discussed anywhere... It shows a clear correspondence between WWDI and WWDI, that could also feed the discussion in the following paragraph as to why the basic sea state variables receive negative weights : the overlapping correlation between temporal diagnostics and instant wave parameters in the WWDI* weighting procedure explains the proximity between TCWWDI and WWDI**

Response:

We agree that the final panel required explicit discussion.

Both formulations are constructed from overlapping diagnostic components, so substantial correspondence between them is expected.

We have also clarified that the learned WWDI* weights are constrained to be nonnegative. Components can receive zero weight, but they do not receive negative weights. For the reported training experiments, wave steepness receives the largest weight, with smaller contributions from wave age and SWH transience.

The revised manuscript now explains both the weighting structure and the relationship between TC-WWDI and WWDI*.

Comment on L. 742–744

Reviewer Comment:

While this indeed proves the regime-scale character of WWDI, the very large spreading also tends to indicate that the designed variables might not be the proper ones to explain the mismatches. A question that arises is the one of the intensity bias, which is the object of the next subsection. A strong intensity bias of CYGNSS with other sources has been noted in previous papers (cf. fig. 11 of Ricciradulli et al., 2021: <https://www.mdpi.com/2072-4292/13/24/5110>)*

While the GPS frequency indeed points at the importance of the sea state itself in the wind estimate, which justifies the present study, the reason why the authors didn't include intensity itself directly in the TC-WWDI diagnostic is unclear.*

Did the authors conduct a sensitivity study to assess the weight in the WWDI of intensity alone ? If so, could they discuss it and, at least, mention it in the text ?*

cf next comment as well regarding intensity bias treatment

Response:

We agree that this is a central issue.

Rather than inserting intensity directly into WWDI* and thereby changing the physical meaning of a wind–wave–state index, we performed a nested fixed-model analysis designed to address the underlying scientific question directly.

Storm intensity and radius form the structural baseline. Local ERA5 wind is then added, followed by bulk wave quantities, wave-state descriptors, and WWDI*.

The analysis shows that local wind provides the dominant incremental information. Bulk wave quantities provide smaller and model-dependent contributions, and WWDI* provides only small and inconsistent additional information.

Accordingly, the revised manuscript now treats storm intensity, local wind, and structure as primary variables, while sea-state quantities are treated as complementary rather than as the primary basis for mismatch diagnosis.

Comment on L. 745

Reviewer Comment:

This section, as well pointed out by the authors, highlights the clear colinearity/correlation between storm intensity and sea state properties. When dissociated through residual analysis, the dependence of sea state on mismatch is reversed. The meaning of this is unclear to me: does it mean there is no dependence on sea state once the structural dependence is removed, or does it mean that, for a given intensity and structure, the bias decreases with TC-WWDI, i.e., smoother and smaller waves have larger mismatch at equivalent intensity ?*

The efforts of the authors to explicit the impact of intensity on the sea state diagnostic are appreciable. However, as they show that this link is indeed very strong and important, one could ask why intensity was not considered as the primary variable for mismatch evaluation, and why sea state was not considered only as complementary to intensity ? This result questions the choice of WWDI as a mismatch diagnosis, since it does not include intensity itself and choses to approach the mismatch distribution under the angle of sea state variability only. It seems that, seeing these results, a TC-WWDI* including both intensity as a primary variable and intensity-normalized sea state as a secondary driver might help better dissociate the mismatches. Is there a fundamental reason why the authors chose not to include intensity in the analysis ? Did the authors conduct sensitivity analyses including intensity into WWDI* ?*

Response:

We agree that the original sign reversal should not be given a simple physical interpretation such as “smoother waves cause larger mismatch.”

The expanded storm-structure analysis shows that, after intensity, normalized radius/RMW, and asymmetry are accounted for, the association between WWDI* and the absolute local CYGNSS–ERA5 difference is approximately zero, whereas its association with the ERA5–storm-maximum structural term remains strong.

We therefore interpret the earlier reversal primarily as evidence of suppression/confounding among intensity, radial structure, and sea-state variables rather than as evidence for a reversed causal wave effect.

We also agree that intensity should be treated as primary context. The revised analysis does this directly through the fixed-model nested framework rather than by embedding intensity inside WWDI* itself.

Comment on L. 752

Reviewer Comment:

the term “controlling” is quite unclear, even with the methodology paragraph in mind. What exactly does it mean ? How are the scatter plots in Fig.6 obtained ? How do you manipulate the rank distributions to obtain the “controlled” data ? Apologies if I missed something.

Response:

We agree that the procedure was not explained sufficiently clearly.

In the rank-residual analysis, the response variable and WWDI* are first transformed to fractional ranks. Each ranked variable is then separately modeled as a function of the specified control variables. The residuals from these two models are then correlated.

The resulting coefficient therefore measures the remaining rank association after variation represented by the selected controls has been removed.

For the revision, we additionally expanded the structural controls to include normalized radius/RMW and storm-relative asymmetry.

We have explained this procedure explicitly in the revised Methods rather than relying on the single word “controlling.”

Comment on L. 761–765

Reviewer Comment:

while this interpretation seems right, it rises the question asked in the previous comment: then, why use sea state to summarize the whole TC structure, and not use TC structure as a primary driver and normalized sea state separately, in addition ?

Response:

We agree.

The revised manuscript no longer uses sea state as a surrogate summary of the complete tropical-cyclone structure.

Storm intensity, local wind, radial position, and storm-relative structure are now treated explicitly as primary context. Wind-wave variables are then evaluated only for their secondary incremental contribution beyond that context.

Comment on L. 766

Reviewer Comment:

same remark as for previous figures: instead of discussing the results and then referencing the figure separately, the discussion above should be done in conjunction with the figure. Each result discussed section 4.3 is illustrated in Fig. 6, so it would help to reference them within the paragraph. Please modify accordingly

Response:

Agreed. We have referenced the corresponding Fig. 6 panels directly within the paragraphs where the results are interpreted.

Comment on L. 772

Reviewer Comment:

again, even though this time the text thereafter kind of follows the figure, it would be preferable to give explicit links with the visuals: correlation values, location on the plots, etc. to help the reader follow your interpretation of the figure

Response:

Agreed. We have added the relevant quantitative values and explicitly identified the corresponding panels so that the reader can directly connect the numerical results with the visual evidence.

Comment on L. 776–784

Reviewer Comment:

it would be nice to have more details about the actual reasons why these discrepancies are observed. How do you explain that WP transfers in wave height departure have lower correlation with ATL and BOB than ATL and BOB with each other, for instance ? Are there studies of sea state properties that allow to interpret these differences ? Is it due to basin size, synoptic wind regimes, coastal configuration ? Do you have insights in these regards ?

Response:

We agree that the original discussion was incomplete.

The present analyses demonstrate transfer heterogeneity but do not uniquely determine whether the observed differences arise from basin geometry, wave climate, synoptic forcing, coastal configuration, storm-population differences, or sampling differences.

The revised discussion identifies these as possible contributors where appropriate but avoids assigning a unique physical cause that cannot be established from the present data.

The revised conclusions therefore do not depend on explaining the exact ordering of the individual basin-transfer correlations.

Comment on L. 825

Reviewer Comment:

unit ?

Response:

Agreed. We have added the appropriate SI unit and checked the surrounding text, tables, figures, and captions for unit consistency.

Comment on L. 851

Reviewer Comment:

since ERA5 data are used point-wise and not as system-averaged, and since the point of your study is to study how the sea state under tropical cyclones affects the CYGNSS mismatch, it would be interesting to have insights on the representation of sea state by ERA5 in terms of TC

conditions: does ERA5 feature the asymmetry of wave fields under TCs ? Does it correctly feature the dependence on intensity regime ? I guess there is already existing studies on the subject... Maybe reference and discuss them could be useful. But this is just a suggestion, and the results in this section are already clear and satisfying as they are.

Response:

Thank you for this useful suggestion.

We have added a discussion emphasizing that ERA5 wave quantities are model/reanalysis fields and may not reproduce all fine-scale tropical-cyclone wave asymmetry or localized extremes.

Our independent CMEMS comparison provides support for the bulk ERA5 significant-wave-height representation in the available matched sample, but it does not constitute validation of the complete azimuthally asymmetric tropical-cyclone wave field.

The revised manuscript therefore does not treat ERA5 sea-state quantities as perfect truth.

Comment on L. 853–855

Reviewer Comment:

this point is repeated 5 or 6 times throughout the manuscript. It could be removed here

Response:

Agreed. We have removed the repetition and consolidated the corresponding limitation in the main interpretation/Discussion section.

Comment on L. 883–886

Reviewer Comment:

same remark as before, the figure could be referenced within the paragraph to back up your points visually, rather than independently at the end.

Response:

Agreed. We have connected the quantitative statements directly to the corresponding figure panels within the paragraph.

Comment on L. 889

Reviewer Comment:

The choice of the IBTrACs-CYGNSS wind mismatch rather than the ERA5-CYGNSS mismatch is confusing to me. Why prefer a bulk difference rather than a local one ? This introduces additional biases in the analysis: for instance, as intense TCs have a more peaked profile, they necessarily have a stronger mismatch out of the eyewall. Weak TCs are more asymmetric, and consequently also display a mean bias in their mismatch simply due to the average gap between maximum wind and local wind in the TC wind field. This introduces additional sources of bias in the mismatch between intensity categories. This choice is not really justified in the manuscript. It would be nice to have more details about why the authors chose IBTrACs over direct ERA5 comparison.

Response:

We agree that the original rationale was insufficient.

IBTrACS was originally selected deliberately as a storm-scale intensity reference because the study sought to examine how local CYGNSS observations relate to the reported storm maximum.

The problem was not the use of IBTrACS for that storm-scale question; the problem was allowing the resulting difference to be interpreted too broadly as though it represented local retrieval error.

Our new decomposition demonstrates quantitatively that this distinction is fundamental.

The revised manuscript therefore retains IBTrACS for storm-scale intensity and structural context, while the CYGNSS–ERA5 comparison provides the local product-to-product perspective.

ERA5 is not described as ground truth.

Comment on L. 927–929

Reviewer Comment:

Do you have ideas regarding why the mismatch appears to be larger after the peak ? wave steepness and wave age don't show the same tendency.

Response:

The additional analyses do not identify a unique physical cause for the observed post-peak asymmetry.

Wave steepness and wave age do not show the same temporal behavior, and our sampling-geometry analysis shows that the SWH-transience descriptor is itself partly affected by changing spatial sampling.

We have therefore reported the pre-/post-peak difference as a descriptive result and removed the attribution of this pattern uniquely to delayed wind–wave adjustment.

Comment on L. 932

Reviewer Comment:

centred -> centered ?

Response:

Agreed. We have changed “centred” to “centered” for consistency with the manuscript’s English usage.

Comment on Fig. 10 caption

Reviewer Comment:

Why is there no WP->ATL transfer for panels b-d ?

Response:

Thank you for identifying this inconsistency.

WP→ATL is part of the external-validation design. We have revised the figure and caption so that transfer coverage is represented consistently across panels or, where a panel uses a restricted subset, that restriction is now stated explicitly.

Comment on Fig. 11 caption

Reviewer Comment:

This figure shows almost identical behaviours of CYGNSS-IBTrACs mismatch and TC intensity. This is consistent with the comment posted above about the choice of computing differences relative to IBTrACs, which introduces a bias in mismatch due to the intensity itself (higher intensity = steep profile = larger mismatch). Most results in this manuscript can thus be interpreted as being, to the least, influenced by your methodological choice. This is a considerable limitation. It also questions the choice of not including the intensity itself in the diagnostic, as it seems to be an important driver, as said in previous comments (but this would require a prior correction of this methodology-induced bias)

Response:

We agree that this figure illustrates the referee's central concern.

The similarity between storm intensity and the CYGNSS-IBTrACS mismatch is not interpreted in the revised manuscript as independent evidence of a sea-state effect. Instead, it is interpreted as being consistent with the structural consequence of comparing a local wind observation with a storm maximum.

The revised interpretation of this figure has therefore been linked directly to the new decomposition and the expanded storm-structure controls.

Comment on Fig. 12 caption

Reviewer Comment:

This figure and section again highlight the limitations of the methodology. The mismatch and WWDI radial distributions (g,h) clearly highlight the problem of choosing IBTrACs to assess the mismatch: mismatch depends primarily on the distance to the center, rather than on WWDI and wave conditions that show a very different dependence on radius. This should by itself be a sufficient motivation for revising the mismatch definition: the mismatch is induced by its own definition. This pointwise issue also induces a system-scale bias, which makes the authors justification regarding the interpretation of WWDI* as a system-scale diagnostic questionable: Since the radial distance strongly determines the bias locally, the radial steepness of the wind profile will strongly influence the system-scale bias. Panel (c) tends to confirm that, showing very large biases when the TC reaches its peak intensity (although the intensity curve is not shown). An additional panel showing the ERA5 mismatch would be really useful to assess if, yes or no, the bias is more related to the wind distribution or to the wave properties. As such, conclusions cannot be made clearly, in my opinion.*

Other remark: the colormap limits on panel (e) are clearly not adapted, nothing can be seen except the old waves in the beginning of the life cycle.

Response:

We agree that the strong radial dependence is central rather than a secondary local issue.

The expanded RMW/asymmetry-controlled analysis confirms that, once storm context is accounted for, the WWDI* association with the absolute local CYGNSS–ERA5 difference is approximately zero, whereas its association with the ERA5–IBTrACS structural term remains strong.

We have therefore revised the case-study analysis so that the local CYGNSS–ERA5 perspective is shown alongside the storm-maximum mismatch, allowing the reader to distinguish local product disagreement from radial/reference-scale structure directly.

We have also revised the original colormap limits in panel (e) so that the relevant wave-state variability is visible.

Comment on L. 1000–1002

Reviewer Comment:

if the wind wave disequilibrium and the temporal lag between wind and wave adjustment was the reason for the CYGNSS mismatch, the transient response variable should carry a more important weight and be the major source of mismatch. Instead, and as noted several times through the manuscript, the main driver in your WWDI and WWDI indices is consistently wave steepness. Moreover, wave steepness is closely related to the wind speed itself, and in many results the direct link between the CYGNSS-IBTrACs mismatch and the intensity is clearly shown. Hence, it seems that in the present state your study shows more evidence regarding the connection between TC intensity and mismatch, than regarding the role of the sea state. To the least, the evidence about the specificity and importance of wave properties in mismatch characterization is not, in my opinion, sufficiently supported by the results. Even though the hypothesis that sea state explains part of the CYGNSS mismatch is most likely true, the choices made by the authors (mainly, not including the local wind contribution in the WWDI* index, measuring the bias as a system-scale difference with IBTrACs, and treating the sea state index as a systemic diagnostic including TC structure and intensity), do not allow to draw this conclusion properly.*

Response:

We agree with the referee’s interpretation.

If delayed wind–wave adjustment were the dominant independent mechanism controlling local CYGNSS disagreement, a stronger and more robust contribution from the transient term would be expected.

Instead, wave steepness receives the largest learned weight, whereas the SWH-transience term contributes much less. Wave steepness is also strongly coupled to wind forcing and storm intensity.

Our additional sampling analysis further shows that the raw SWH-transience descriptor is influenced by sampling geometry, although the resulting changes in the overall WWDI* associations are small, with maximum absolute changes in Spearman correlation of approximately 0.02.

Combined with the fixed-model and storm-structure analyses, these results do not support the claim that wind–wave disequilibrium has been demonstrated as the principal cause of local CYGNSS disagreement.

We have revised the Discussion and Conclusion accordingly.

Comment on L. 1086–1089

Reviewer Comment:

This limitation is indeed the main and most problematic one. The three studies cited here do not use IBTrACs as a reference for bias correction: Ricciardulli et al 2021 use radiometer/scatterometer data, and Powell and Ruf 2025 use ERA5 winds. The choice of authors to use the IBTrACs error as the main mismatch diagnostic is thus puzzling to me. As discussed in previous comments, the main problem with that choice is that it introduces a mean bias in the analysis, due to the fact that bias will naturally be larger as observations are further from the maximum wind; and that this mean bias induces a systemic bias which is that the wind field asymmetry and the shape of the wind profile will then be primary drivers of the mean mismatch at the scale of a TC acquisition. This invalidates in itself the idea of the authors of diagnosing a “reference-relative mismatch”, since the reference itself induces a system-scale mismatch. What led the authors to chose to use IBTrACs for bias measurement, even though ERA5 local winds were available and have already been used in Powell and Ruf 2025 is one of my main interrogations

Response:

We agree. This issue is now treated as a central result rather than only as a limitation.

IBTrACS was originally used to provide storm-scale intensity context. However, the original manuscript overextended that comparison when discussing the difference as though it diagnosed local retrieval behavior.

The new signed decomposition shows that the local-to-storm-maximum structural difference dominates the large CYGNSS–IBTrACS mismatch.

We have therefore explicitly distinguished between two different questions:

1. How closely does CYGNSS agree with a collocated local wind product?
2. How does a local CYGNSS observation relate to the storm-reported maximum intensity?

These quantities have different physical meanings and are no longer conflated in the revised manuscript.

Comment on L. 1114–1117

Reviewer Comment:

The fact that wave steepness is the main driver, while it is also the most related to local wind speed, indicates in itself the contribution of wind speed to the mismatch.

In my opinion, this is how I am led to interpret your study based on your results:

comparison between CYGNSS and Vmax means that points far from the center will have larger bias, and that bias will increase with intensity, since the difference with Vmax is overall larger as Vmax increases. Since the wind is not included in the analysis, the difference in wind speed will be reflected mostly by the wave steepness: steeper waves $\sim$ higher winds $\sim$ larger mean difference with Vmax

Response:

We agree that this is a plausible interpretation of the original analysis and that the original manuscript did not sufficiently separate these effects.

The new fixed-model analysis, expanded storm-structure controls, and local-versus-structural decomposition all support a more conservative interpretation.

Wave steepness is the strongest of the tested sea-state descriptors in the original weighting, but much of its apparent relationship occurs within regimes organized by wind intensity and storm structure.

We have therefore removed the interpretation of the dominance of wave steepness as evidence of an independent sea-state-specific mechanism.

Comment on L. 1138–1140

Reviewer Comment:

If WWDI does not provide an improvement on the system-scale compared to bulk wind data, then its use as a specific diagnostic of sea state is not justified, even though it is designed as systemic and not local. The point of this work being to prove the importance of including sea state diagnosis in the characterization of CYGNSS wind mismatches is not supported by this result.*

Response:

We agree with the logic of this comment.

The fixed-model analysis does not support the claim that WWDI* provides a large or consistently robust improvement after local wind and storm structure are already known.

The revised manuscript therefore no longer presents WWDI* as evidence that a specific sea-state diagnostic is necessary for local CYGNSS error characterization.

Where WWDI* is retained, it is described as a secondary physics-informed descriptor of covarying wind–wave/storm regimes, and its limited incremental contribution is stated explicitly.

Comment on L. 1148–1150

Reviewer Comment:

what “environmental regimes” were identified here as systematically inducing larger mismatches ? Apart from the very general relation between wave steepness and mean bias, which was discussed above as being a potential reflection of the wind bias induced by the authors definition of mismatch, the authors do not discuss in their figures the existence of specific regimes/combinations of variables leading to higher mismatches. The local variability is ignored in the study (which is a fair choice), but the system-scale changes are not analyzed in depth, as the analysis only provides general error diagnostics and does not draw conjectures on the reason for basin-scale or bin-wise differences in general.

The authors take a lot of precautions to describe their “interpretation” of the result at the end of all sections, but as a result, the designed WWDI appears simply as a “bulk diagnostic” of sea state that contains inherent biases and does not actually bring information on specific conditions leading to mismatches. The rare occurrences showing specific behaviours are not interpreted, such as the reason why the bias is larger after the peak intensity than before, or why some basins some transfers appear to be different from others (cf. comment in section 4.4)*

Response:

We agree that the phrase “environmental regimes” was too broad relative to what the analysis

demonstrated.

We have removed or replaced this terminology and now identify only the structures directly supported by the analysis: storm intensity, local wind, storm-relative position and structure, reference-scale contrast, and associated wind–wave state.

We have also distinguished demonstrated relationships from unresolved descriptive patterns. In particular, the present analyses do not identify a unique physical cause of the post-peak asymmetry or the precise ordering of basin-transfer differences, and this is now stated explicitly rather than assigning unsupported physical explanations.

Overall Response and Revised Interpretation

Taken together, the referee’s comments and the additional analyses led us to substantially revise the scientific interpretation of the manuscript.

The revised manuscript now states that:

The large CYGNSS–IBTrACS differences observed in tropical cyclones are governed primarily by reference-scale and storm-structure mismatch rather than by local CYGNSS retrieval disagreement. Storm intensity, local wind, and storm-relative structure therefore provide the primary context. Wind–wave descriptors covary with these regimes and can provide secondary information, but the present evidence does not support WWDI* as an independent local retrieval-error index or wind–wave disequilibrium as the dominant cause of local CYGNSS disagreement.

The revised manuscript also makes clear that:

An adjustment calibrated against storm-maximum intensity can substantially reduce the CYGNSS–IBTrACS mismatch while simultaneously degrading agreement with a collocated local-wind reference. Reference-relative mismatch reduction therefore must not be interpreted as retrieval-accuracy improvement.

We sincerely thank the referee for the detailed and constructive report. The comments led us not simply to add limitations, but to reorganize the scientific interpretation around a clearer distinction among local product disagreement, storm structure, reference-scale mismatch, sampling geometry, and secondary wind–wave-state variability.