https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Physics-Guided Diagnosis of CYGNSS Wind Mismatch under Tropical-Cyclone Wind–Wave Disequilibrium

S M Injamamul Haque Masum[1], Ahmed Wasiu Akande[2], Dongkai Yang[2], Mohammad Shohidul Islam[3], and Syed Shahid Shah[1]

[1]School of Electronic and Information Engineering, Hangzhou International Innovation Institute of Beihang University, Hangzhou 311115, China
[2]International School, Hangzhou International Innovation Institute of Beihang University, Hangzhou 311115, China
[3]Atmospheric Research Division, Bangladesh Space Research and Remote Sensing Organization (SPARRSO), Dhaka, Bangladesh

Correspondence: S M Injamamul Haque Masum (mrmasum005@gmail.com)

Abstract.

The Cyclone Global Navigation Satellite System (CYGNSS) provides extensive ocean-surface wind observations within tropical cyclones, yet its reference-relative mismatch can vary systematically with storm structure and the evolving sea state. This study develops a physics-guided diagnostic framework using 212,697 unique CYGNSS–tropical-cyclone matchups from 147 storms during 2022–2024 across the western North Pacific, Bay of Bengal, and North Atlantic. CYGNSS observations are combined with ERA5 atmospheric and wave fields and IBTrACS best-track information to construct six descriptors of derived wave-height departure, young-wave state, bulk wave steepness, transient wind forcing, transient sea-state response, and storm-relative proximity. A supervised Wind–Wave Disequilibrium Index (WWDI*) is derived using training-only robust normalization and leave-one-component-out evidence weighting, with all parameters frozen before temporal and geographical evaluation.

Across four storm-independent and observation-independent cross-basin transfers, WWDI* retains positive associations with absolute CYGNSS–IBTrACS mismatch. Observation-level Spearman correlations range from 0.219 to 0.381, whereas 10-bin regime-scale R^2 values range from 0.639 to 0.753. Wave steepness receives the largest training-derived weight in both development basins, followed by young-wave state and transient significant-wave-height response. Permutation testing, storm-wise bootstrap resampling, and sensitivity experiments confirm the persistence of the regime-scale relationships. After controlling for ranked tropical-cyclone intensity, storm-relative radius, their quadratic terms, and their interaction, the residual association reverses sign in all four transfers, with residual Spearman coefficients from -0.645 to -0.543. This shows that the positive unconditional relationship is embedded within the joint storm-intensity–radius structure rather than representing an independent monotonic effect.

Independent comparison with 5,771 CMEMS significant-wave-height matchups gives an ERA5 RMSE of 0.631 m and Pearson correlation of 0.942. Diagnostic machine-learning experiments further show that local atmospheric wind provides substantial transferable information beyond tropical-cyclone intensity and radial position, with additional benefit from bulk wave-state variables in three of the four transfers. A training-derived WWDI*-conditioned adjustment reduces raw external

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

RMSE by 29.8–40.4 %; relative to a simpler global training-bias correction, the additional point-estimate improvement is 1.7–13.0 %, with storm-wise bootstrap support for the reciprocal WP–BOB transfers.

These results identify transferable tropical-cyclone wind–wave regimes associated with systematic changes in reference-relative CYGNSS mismatch. WWDI* is therefore interpreted as a supervised physical diagnostic of mismatch-prone storm environments rather than as a causal index or an operational wind-retrieval correction.

Keywords: tropical cyclones; wind–wave interaction; wind–wave disequilibrium; CYGNSS; GNSS reflectometry; ocean surface roughness; ocean waves; air–sea coupling

1 Introduction

Tropical cyclones (TCs) produce some of the most rapidly changing wind and wave conditions over the ocean. Intense surface winds generate locally forced waves, while storm motion, pre-existing swell, and the finite adjustment time of the wave field create a sea state that can differ substantially from that expected under steady local wind forcing. The ocean surface beneath and around a TC is therefore frequently characterized by strong spatial and temporal departures from wind–wave equilibrium. Such conditions affect sea-surface roughness and are consequently important for the interpretation of microwave and reflected-signal observations during extreme marine events (Emanuel, 2003; Prakash et al., 2018; Lima et al., 2022).

Global Navigation Satellite System reflectometry (GNSS-R) observes the ocean by receiving navigation signals after reflection from the sea surface. The measured scattering response depends on the statistical roughness and slope distribution of the reflecting surface and therefore contains information about both ocean wind and wave conditions (Foti et al., 2015; Soisuvarn et al., 2016; Gleason et al., 2016; Bu et al., 2024). The Cyclone Global Navigation Satellite System (CYGNSS) applies this measurement principle using a multi-satellite constellation designed for frequent sampling of tropical oceans, providing valuable observations in regions affected by tropical cyclones (Ruf et al., 2016, 2019a).

Substantial progress has been made in converting CYGNSS measurements into ocean-surface wind estimates. Early retrieval algorithms established empirical relationships between GNSS-R observables and surface wind speed, while subsequent studies improved calibration, geophysical model functions, statistical retrieval procedures, and uncertainty characterization (Clarizia and Ruf, 2016; Gleason et al., 2019; Ruf and Balasubramaniam, 2019; Ruf et al., 2019b; Clarizia and Ruf, 2020; Asgarimehr et al., 2020). Continued validation and product assessment have demonstrated the usefulness of CYGNSS wind observations across the tropics and within TC environments, while also showing that retrieval performance varies with wind regime, calibration approach, product formulation, and reference dataset (Asharaf et al., 2021; Ricciardulli et al., 2021; Ruf et al., 2024). Recent developments in physics-informed and multi-GNSS retrieval methods further illustrate the continuing evolution of GNSS-R ocean-wind estimation (Qiao and Huang, 2025; Li et al., 2026).

Despite these advances, a CYGNSS wind estimate and an external reference do not necessarily represent identical geophysical quantities. This distinction is especially important in tropical cyclones, where strong wind gradients, rapid storm evolution, and spatial variability occur over relatively short distances and time intervals. Differences between CYGNSS and a reference wind product can therefore contain contributions from retrieval uncertainty, sampling differences, spatial representativeness,

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

temporal mismatch, and storm-scale reference characteristics. For this reason, the present study treats the analysed CYGNSS–reference difference as a *reference-relative mismatch* rather than assuming that it represents a pure pointwise retrieval error (Ricciardulli et al., 2021; Knaff et al., 2021; Ruf et al., 2019b; Powell and Ruf, 2025).

Sea state introduces an additional source of variability because GNSS-R responds to the physical ocean surface rather than directly to atmospheric wind. Previous work has demonstrated that GNSS-R scattering responds to both wind and waves and that significant wave height can provide useful information for improving CYGNSS wind estimation (Soisuvarn et al., 2016; Clarizia and Ruf, 2017; Pascual et al., 2021). Hurricane observations have also shown that GNSS-R scattering characteristics can vary with storm-relative geometry and the directional structure of the surface (Balasubramaniam and Ruf, 2020). These findings suggest that part of the observed CYGNSS wind mismatch may be organized by the evolving state of the wave field rather than by wind speed alone.

The physical state of a TC sea can be described from several complementary perspectives. Wave age provides information on the maturity of the wave field relative to wind forcing, whereas wave steepness characterizes the geometric relationship between wave height and wavelength (Rascle and Ardhuin, 2013; Wiese et al., 2018). During rapidly evolving storms, atmospheric forcing may change faster than the bulk wave field can adjust, producing young, steep, or otherwise transient sea states (Prakash et al., 2018; Bethel et al., 2022; Lima et al., 2022). No single bulk variable can represent the full directional wave spectrum, but wave maturity, steepness, dimensionless wave-height departure, and temporal changes in wind and significant wave height provide physically interpretable descriptors of the degree to which the local wind and wave fields are adjusted to one another. Here, these conditions are referred to collectively as *wind–wave disequilibrium*.

A second challenge concerns how such relationships are evaluated. TC observations are strongly clustered by storm, geographical region, and time. Random observation-level splitting can therefore place closely related samples in both development and evaluation subsets, producing performance estimates that do not necessarily represent transfer to independent storms or regions. Structured cross-validation is preferable when observations contain strong temporal or spatial dependence (Roberts et al., 2016; Varoquaux, 2018). Consequently, a transferable diagnostic of CYGNSS mismatch should be constructed without using information from the external evaluation basin and should subsequently be tested under strict geographical separation.

Motivated by these considerations, this study develops a physics-guided Wind–Wave Disequilibrium Index framework to examine whether reference-relative CYGNSS wind mismatch is systematically related to the evolving TC wind–wave state. CYGNSS observations are combined with ERA5 atmospheric and wave fields and IBTrACS tropical-cyclone information across the western North Pacific (WP), Bay of Bengal (BOB), and North Atlantic (ATL). Independent Copernicus Marine Service significant-wave-height observations are used separately to assess the ERA5 wave representation and do not contribute to construction or training of the proposed index.

The framework contains two related indices. First, an equal-weight diagnostic TC-WWDI combines six components representing dimensionless wave-height departure, young-wave state, wave steepness, transient wind forcing, transient sea-state response, and storm-relative context. Second, a supervised WWDI* is constructed using normalization parameters and leave-one-component-out evidence weights estimated exclusively from the training basin. All training-derived parameters are frozen before temporal or external-basin evaluation.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

The study therefore addresses three main objectives. First, it examines whether physically interpretable wind–wave descriptors organize CYGNSS–reference mismatch into coherent TC sea-state regimes. Second, it tests the transferability of the training-derived WWDI* through strict WP→BOB, BOB→WP, WP→ATL, and BOB→ATL experiments, supported by confounder-aware diagnostics, permutation testing, storm-wise bootstrap resampling, and sensitivity analysis. Third, it evaluates the additional diagnostic information contained in the wind–wave variables through independent wave validation, machine-learning benchmarks, and a training-only WWDI*-conditioned mismatch adjustment. The objective is not to replace the operational CYGNSS wind retrieval, but to identify physically interpretable TC regimes in which larger reference-relative wind mismatch becomes more likely.

2 Data

The analysis combines CYGNSS ocean-surface wind observations, ERA5 atmospheric and wave fields, and IBTrACS tropical-cyclone information to construct a multi-basin dataset for examining reference-relative CYGNSS wind mismatch under tropical-cyclone conditions. Satellite-altimeter significant wave height from the Copernicus Marine Service (CMEMS) is used separately to evaluate the ERA5 wave field. The CMEMS observations are kept outside the construction, normalization, and supervised weighting of the wind–wave disequilibrium indices and are not used for machine-learning model development or calibration of the mismatch-adjustment experiments.

The study covers 2022–2024 and includes three tropical-cyclone-active regions: the Bay of Bengal (BOB), western North Pacific (WP), and North Atlantic (ATL). The three domains differ substantially in storm frequency, geographical extent, observation density, and prevailing wind–wave conditions. They therefore provide distinct regional environments for testing whether relationships identified in one development basin remain detectable when transferred geographically.

After CYGNSS product-level quality screening, tropical-cyclone association, ERA5 collocation, physical plausibility filtering, and removal of duplicate physical observations, the final dataset contains 212,697 unique CYGNSS–TC matchups associated with 147 tropical cyclones. The WP contains 113,444 observations from 76 storms, the BOB contains 6,830 observations from 14 storms, and the ATL contains 92,423 observations from 57 storms. The regional domains and final sample composition are summarized in Table 1.

Table 1. Final unique CYGNSS–tropical-cyclone matchup dataset used in the analysis.

Basin	Abbreviation	Longitude range	Latitude range	Unique matchups	Storms
Bay of Bengal	BOB	78°E–100°E	0°–25°N	6,830	14
Western North Pacific	WP	100°E–180°E	0°–45°N	113,444	76
North Atlantic	ATL	90°W–20°E	11°–45°N	92,423	57
Total	–	–	–	212,697	147

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

2.1 CYGNSS ocean-surface wind observations

The Cyclone Global Navigation Satellite System (CYGNSS) observes navigation signals after reflection from the Earth's surface and provides frequent sampling over the tropical oceans (Ruf et al., 2016; Foti et al., 2015; Gleason et al., 2016). Over the ocean, the reflected GNSS signal responds to the statistical roughness and slope distribution of the sea surface, whose scattering characteristics vary with wind forcing and wave conditions (Soisuvarn et al., 2016). Ocean-surface wind speed can consequently be estimated through calibrated relationships between GNSS-R observables and near-surface wind speed (Clarizia and Ruf, 2016; Ruf and Balasubramaniam, 2019).

The primary CYGNSS dataset used in this study is the NOAA CYGNSS Level-2 Science Wind Speed Product (SWSP), version 1.2, with nominal 25-km along-track sampling (NOAA/NESDIS/STAR/SOCD, 2022). Source observations from 2022–2024 are considered. The archived Level-2 wind estimate is denoted U_{CYGNSS} and is used without alteration in the subsequent analysis. Neither the geophysical retrieval model nor the archived wind product is retrained or recalibrated in this study.

This choice allows the analysis to focus on the environmental and storm-relative conditions associated with variation in the mismatch of an existing CYGNSS wind product. Published assessments have shown that CYGNSS wind characteristics depend on retrieval formulation, calibration strategy, wind regime, and reference dataset (Asharaf et al., 2021; Ricciardulli et al., 2021; Ruf et al., 2024). The NOAA product is therefore treated as the observational wind record being diagnosed, rather than as an error-free representation of the local atmospheric wind.

For each source record, the processing workflow retains the acquisition time, specular-point coordinates, archived Level-2 wind speed, and the measurement or retrieval fields required for product-level screening and subsequent processing.

Quality control is applied using the CYGNSS `sample_flags` bit mask. A record is retained only when bit 0 is not set:

$$(\texttt{sample\_flags} \,\&\, 1) = 0. \tag{1}$$

Additional physical screening is performed after association with ERA5 and IBTrACS.

Throughout the study, differences between U_{CYGNSS} and an external wind reference are interpreted cautiously. CYGNSS retrieval uncertainty varies with observational and environmental conditions, while different reference products can represent different spatial and temporal wind quantities (Ruf et al., 2019b; Ricciardulli et al., 2021). The resulting differences are therefore described as reference-relative mismatch unless a specific validation experiment supports a stronger interpretation.

2.2 ERA5 atmospheric and wave fields

Atmospheric forcing and bulk sea-state conditions are represented using the ERA5 global reanalysis produced by the European Centre for Medium-Range Weather Forecasts (Hersbach et al., 2020; Copernicus Climate Change Service (C3S), 2018). Four hourly variables are retained: 10-m zonal wind u_{10}, 10-m meridional wind v_{10}, significant wave height H_s, and mean wave period T_m. The atmospheric single-level fields are represented on a regular $0.25^\circ \times 0.25^\circ$ latitude–longitude grid, whereas the ERA5 ocean-wave fields are represented on a $0.5^\circ \times 0.5^\circ$ grid.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

ERA5 10-m wind speed is calculated from the horizontal wind components as

$$U_{\mathrm{ERA5}} = \sqrt{u_{10}^2 + v_{10}^2}, \tag{2}$$

where u_{10} and v_{10} are the ERA5 components associated with the CYGNSS observation.

The atmospheric and wave fields are collocated to the CYGNSS specular-point position and acquisition time through space–time interpolation. For each CYGNSS observation, the ERA5 time levels bracketing the acquisition time are identified. Spatial interpolation is carried out on the corresponding native latitude–longitude grid, followed by linear interpolation between the bracketing temporal fields. Longitude coordinates are first expressed using a common convention to ensure consistency across the three study regions.

Observations lacking finite values of the required ERA5 fields are removed during subsequent screening. The final physical matchup records therefore contain finite U_{ERA5}, H_s, and T_m.

Significant wave height and mean wave period are used as bulk descriptors of the contemporaneous sea state. They provide the basis for the characteristic wave scale, effective wave age, bulk wave steepness, and dimensionless wave-height quantities defined in Sect. 3.3. Mean wave direction is not used in either TC-WWDI or WWDI*.

ERA5 variables serve as environmental descriptors and do not replace or modify the archived CYGNSS wind. U_{ERA5} provides local atmospheric-wind context, while H_s and T_m describe the bulk wave field associated with the CYGNSS observation. The use of wind and wave properties together is consistent with established evidence that GNSS-R ocean scattering is sensitive to both wind forcing and the state of the reflecting wave field (Soisuvarn et al., 2016; Clarizia and Ruf, 2017; Pascual et al., 2021).

2.3 IBTrACS tropical-cyclone information

Tropical-cyclone position and intensity information is obtained from the International Best Track Archive for Climate Stewardship (IBTrACS), version 4r01 (Knapp et al., 2010; Gahtan et al., 2024). IBTrACS provides best-track records describing tropical-cyclone location, intensity, and temporal evolution.

A common intensity field is used across all three basins in order to avoid introducing an additional source of inter-basin inconsistency through different agency wind variables. The tropical-cyclone intensity reference, denoted U_{TC}, is therefore obtained exclusively from the IBTrACS `usa_wind` field. No fallback to `wmo_wind` or another agency-specific wind field is used. The retained field represents the USA-agency 1-min maximum sustained wind available through the IBTrACS record.

For each retained CYGNSS–TC association, the processing dataset includes the storm identifier, selected best-track time, storm-centre latitude and longitude, maximum sustained wind intensity, temporal separation from the CYGNSS observation, and great-circle distance between the selected storm centre and the CYGNSS specular point.

IBTrACS wind intensity is converted from knots to metres per second using

$$U_{\mathrm{TC}}[\mathrm{m\,s^{-1}}] = 0.514444\, U_{\mathrm{TC}}[\mathrm{kt}]. \tag{3}$$

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

The converted value is used consistently in subsequent intensity, storm-relative, mismatch, confounder, and diagnostic analyses.

Importantly, U_{TC} is not a pointwise measurement of wind speed at the CYGNSS specular point. It represents storm-scale best-track maximum sustained intensity, whereas the CYGNSS observation represents a satellite retrieval at a particular location and time. The two quantities therefore differ in spatial support and physical meaning (Knapp et al., 2010; Knaff et al., 2021).

Consequently, the difference between U_{CYGNSS} and U_{TC} can contain contributions from cyclone radial structure, rapid storm evolution, spatial and temporal sampling differences, CYGNSS retrieval uncertainty, and representativeness or uncertainty in the best-track reference (Ruf et al., 2019b; Ricciardulli et al., 2021; Powell and Ruf, 2025). The CYGNSS–IBTrACS quantity is therefore described throughout the manuscript as a reference-relative mismatch rather than as a direct pointwise retrieval error.

2.4 CYGNSS–IBTrACS association and physical screening

CYGNSS observations are associated with IBTrACS records using a distance-prioritized selection procedure within a prescribed temporal window. For each CYGNSS acquisition, candidate best-track records are first restricted to

$$|\Delta t_{\mathrm{TC}}| \leq 3\ \mathrm{h}, \tag{4}$$

where Δt_{TC} is the temporal separation between the CYGNSS observation and an IBTrACS candidate record.

Great-circle distance is then calculated from the CYGNSS specular point to the storm-centre position of every temporally admissible candidate. The candidate with the minimum spatial separation is selected and retained only when

$$R_{\mathrm{TC}} \leq 300\ \mathrm{km}, \tag{5}$$

where R_{TC} is the great-circle distance between the CYGNSS specular point and the selected best-track storm centre.

The $\pm$3-h condition therefore defines the admissible temporal search window, whereas spatial proximity determines which discrete IBTrACS record is assigned within that window. The storm position and intensity used in the analysis are taken directly from the selected best-track record. No temporal interpolation of IBTrACS storm position or intensity is applied.

The 300-km threshold defines the tropical-cyclone analysis domain used in this work and subsequently provides the reference scale for the storm-proximity descriptor. It is not interpreted as a universal physical radius of a tropical cyclone.

Following product-level CYGNSS screening, IBTrACS association, and ERA5 collocation, additional filters remove non-finite or physically implausible records. The retained wind quantities satisfy

$$0 < U_{\mathrm{CYGNSS}} \leq 90\ \mathrm{m\,s^{-1}}, \tag{6}$$

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$0 < U_{\mathrm{TC}} \leq 95\ \mathrm{m\,s^{-1}}, \tag{7}$$

and

$$0 < U_{\mathrm{ERA5}} \leq 80\ \mathrm{m\,s^{-1}}. \tag{8}$$

The retained bulk-wave quantities satisfy

$$0.03 \leq H_s \leq 25\ \mathrm{m}, \qquad 1 \leq T_m \leq 30\ \mathrm{s}. \tag{9}$$

These limits function only as quality-control and physical-plausibility screens. They do not define the wind-speed, sea-state, or disequilibrium regimes examined later. Observations containing non-finite quantities required for construction of the physical descriptors are removed before the diagnostic indices are calculated.

Duplicate physical observations are removed before the basin-specific master datasets are finalized. The counts in Table 1 therefore represent unique physical CYGNSS–TC matchups. A retained observation may subsequently enter more than one methodological experiment when its basin assumes different transfer roles, but it occurs only once in its regional master dataset.

2.5 Independent CMEMS significant-wave-height observations

An independent satellite-altimeter dataset is used to evaluate the ERA5 significant wave height employed in the physical diagnostics. Level-3 Sentinel-3A significant-wave-height observations are obtained from the Copernicus Marine Service (E.U. Copernicus Marine Service Information, 2026). Satellite-altimeter wave-height observations provide an observational basis for independent assessment of model or reanalysis wave fields (Schulz-Stellenfleth and Staneva, 2019).

Quality screening is performed at the CMEMS source-data stage. Only Sentinel-3A records satisfying

`valueQc=1`

are admitted to the collocation procedure. The original CMEMS `valueQc` field is not retained as an independent column in the final ERA5–CMEMS matchup table. The resulting validation records therefore represent observations that have already passed the prescribed source-level quality screening.

The quality-screened Sentinel-3A observations are associated independently with ERA5 significant wave height. The primary validation configuration requires a spatial separation satisfying

$$\Delta r_{\mathrm{CMEMS}} \leq 100\ \mathrm{km}, \tag{10}$$

and an absolute temporal separation satisfying

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$|\Delta t_{\mathrm{CMEMS}}| \leq 3\ \mathrm{h}. \tag{11}$$

Records containing non-finite CMEMS SWH, ERA5 SWH, spatial separation, or temporal separation are excluded before calculation of the validation statistics.

The primary validation dataset contains 5,771 ERA5–CMEMS SWH matchups: 2,657 in the ATL, 2,948 in the WP, and 166 in the BOB. The median spatial separation is 43.0 km and the median absolute temporal separation is 1.38 h. These values refer specifically to the primary 100-km/3-h configuration. Sensitivity to more restrictive spatial and temporal criteria is evaluated separately in Sect. 3.15.

The CMEMS comparison is isolated from development of the proposed wind–wave diagnostic framework. CMEMS observations are not used to construct or normalize TC-WWDI or WWDI*, estimate supervised component weights, define the temporal or geographical transfer partitions, select or fit the diagnostic machine-learning models, or calibrate the global-bias or WWDI*-conditioned adjustments. The altimeter observations therefore serve exclusively as an independent assessment of the ERA5 H_s field used to characterize the bulk sea state.

2.6 Final multi-basin dataset

Following CYGNSS quality control, IBTrACS association, ERA5 collocation, physical screening, and duplicate removal, the retained observations are assembled into basin-specific master datasets. Each physical record includes the CYGNSS acquisition time and specular-point coordinates, archived CYGNSS Level-2 wind estimate, matched IBTrACS storm identifier and best-track intensity, storm-relative distance, CYGNSS–IBTrACS temporal separation, and collocated ERA5 10-m wind speed, significant wave height, and mean wave period.

The resulting multi-basin dataset comprises 212,697 unique physical CYGNSS–TC matchups associated with 147 tropical cyclones during 2022–2024. The WP contains 113,444 observations from 76 storms, the BOB contains 6,830 observations from 14 storms, and the ATL contains 92,423 observations from 57 storms (Table 1). These basin-level counts describe the underlying regional datasets and are distinct from the later transfer-specific sample counts obtained when each basin is assigned a training, same-basin internal-validation, or external-evaluation role.

The characteristic wave variables, normalized physical components, diagnostic TC-WWDI, and supervised WWDI* are derived subsequently from these retained physical observations according to Sect. 3.1. They are therefore derived diagnostic quantities, not additional independent observations.

The geographical distribution of the final records is shown in Fig. 1. Sampling is strongly unequal among the three regions, particularly because the BOB contains substantially fewer observations and tropical cyclones than WP or ATL. No artificial observation-level balancing or resampling is applied to the external datasets. The cross-basin experiments therefore preserve the naturally available observation and storm composition of each evaluation region.

For every transfer experiment, normalization limits, empirical training-basin relationships, component weights, and other data-derived parameters required by the formulation are estimated using the designated training subset before application to

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

internal-validation or external data. The resulting quantities are then frozen according to Sect. 3.10. Consequently, the external basin does not contribute to construction of WWDI* or calibration of the subsequent diagnostic analyses.

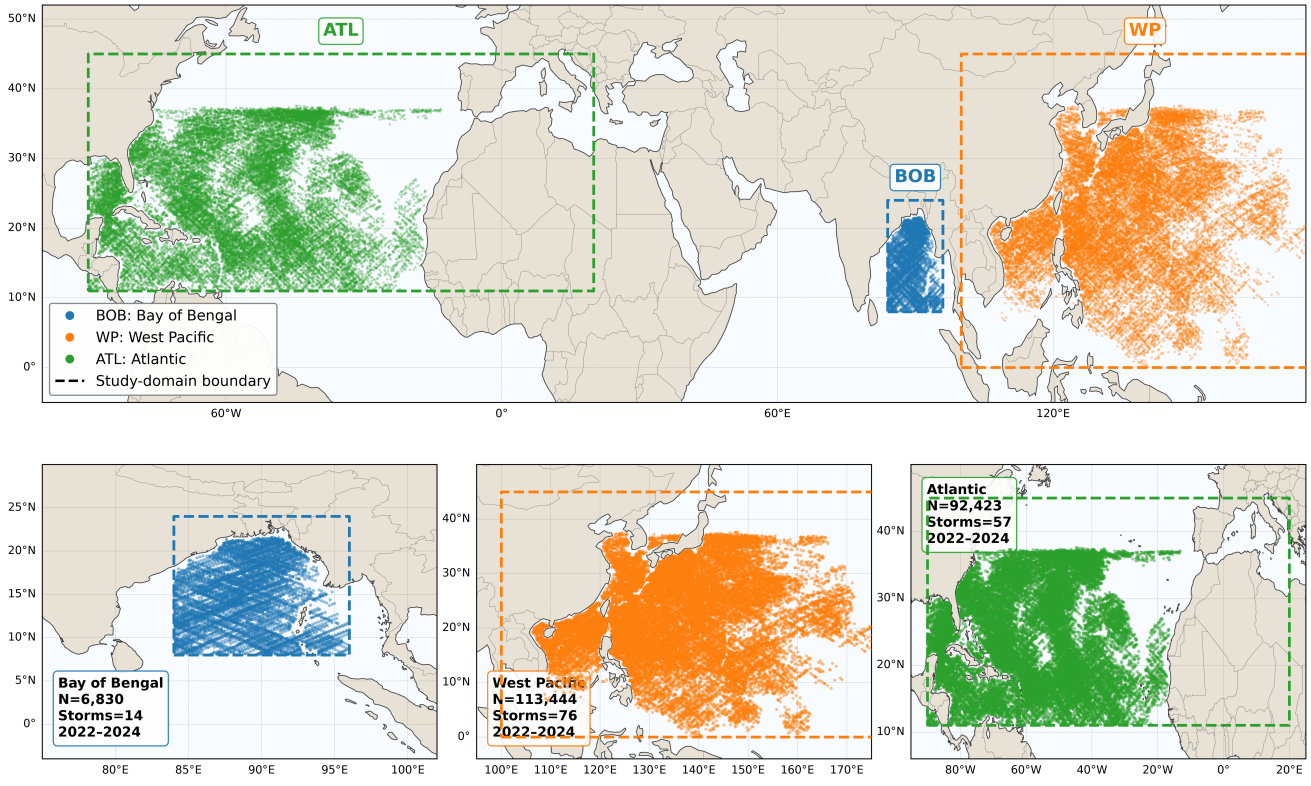


Figure 1. Study regions and spatial distribution of the final unique CYGNSS–tropical-cyclone matchups. The Bay of Bengal (BOB), western North Pacific (WP), and North Atlantic (ATL) domains are shown together with the retained CYGNSS specular-point locations. The final multi-basin dataset contains 212,697 unique physical matchups associated with 147 tropical cyclones during 2022–2024. Geographic base-map features are obtained from Natural Earth and rendered using Cartopy.

3 Methodology

3.1 Analysis framework

The analysis is designed to determine whether reference-relative CYGNSS wind mismatch during tropical cyclones is systematically organized by the contemporaneous wind–wave environment. The archived CYGNSS Level-2 wind estimate is retained without modification throughout the principal analysis. The framework is therefore diagnostic rather than retrieval oriented: it seeks to identify physically interpretable combinations of bulk sea state, storm evolution, and storm-relative context associated

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

with changes in the magnitude and structure of CYGNSS–reference mismatch (Clarizia and Ruf, 2017; Ruf et al., 2019b; Pascual et al., 2021; Powell and Ruf, 2025).

The methodological sequence is summarized in Fig. 2. Starting from the collocated CYGNSS–ERA5–IBTrACS observations described in Sect. 2, bulk descriptors are constructed for wave maturity, wave steepness, wind-normalized wave height, transient atmospheric forcing, transient sea-state response, and storm-relative position.

Two related indices serve different purposes. The first is the equal-weight Tropical-Cyclone Wind–Wave Disequilibrium Index (TC-WWDI), which retains the complete six-component candidate set and is used as a descriptive representation of the combined storm and sea-state environment. The second is WWDI*, a supervised physics-guided index whose normalization limits, empirical wave-height scaling coefficients, and component weights are estimated exclusively from the designated 2022–2023 training subset. All such quantities are frozen before application to either the same-basin 2024 holdout or an external basin.

Random observation-level partitioning is not used to define the training, internal-validation, or external-evaluation datasets. This restriction is important because observations from the same tropical cyclone are strongly clustered in space and time, and distributing closely related samples across development and evaluation sets can yield optimistic estimates of generalization (Roberts et al., 2016; Varoquaux, 2018). The training-only validation fraction used internally by the multilayer perceptron for early stopping is an optimization device confined to the development data and does not define a transfer-evaluation partition.

Transferability is examined using predefined temporal and geographical splits. External relationships are evaluated at both observation and regime scales and are further examined through confounder-aware residual analysis, component-collinearity diagnostics, permutation testing, storm-wise bootstrap resampling, methodological sensitivity tests, independent ERA5–CMEMS wave-height validation, and a diagnostic machine-learning benchmark. A final training-derived WWDI*-conditioned adjustment is compared with a simpler global training-bias correction to distinguish general debiasing from additional regime-dependent information.

3.2 Reference-relative CYGNSS wind mismatch

IBTrACS provides a storm-scale best-track intensity estimate rather than a point measurement of the local surface wind at an individual CYGNSS specular point. Differences between CYGNSS and IBTrACS are therefore treated throughout as reference-relative mismatch rather than as direct measurements of CYGNSS retrieval error (Knapp et al., 2010; Ruf et al., 2019b; Ricciardulli et al., 2021; Knaff et al., 2021; Powell and Ruf, 2025).

The signed CYGNSS–IBTrACS mismatch is defined as

$$e_{\mathrm{TC}} = U_{\mathrm{CYGNSS}} - U_{\mathrm{TC}}, \tag{12}$$

where U_{CYGNSS} is the archived CYGNSS Level-2 wind speed and U_{TC} is the matched IBTrACS storm-intensity reference. The corresponding absolute mismatch is

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

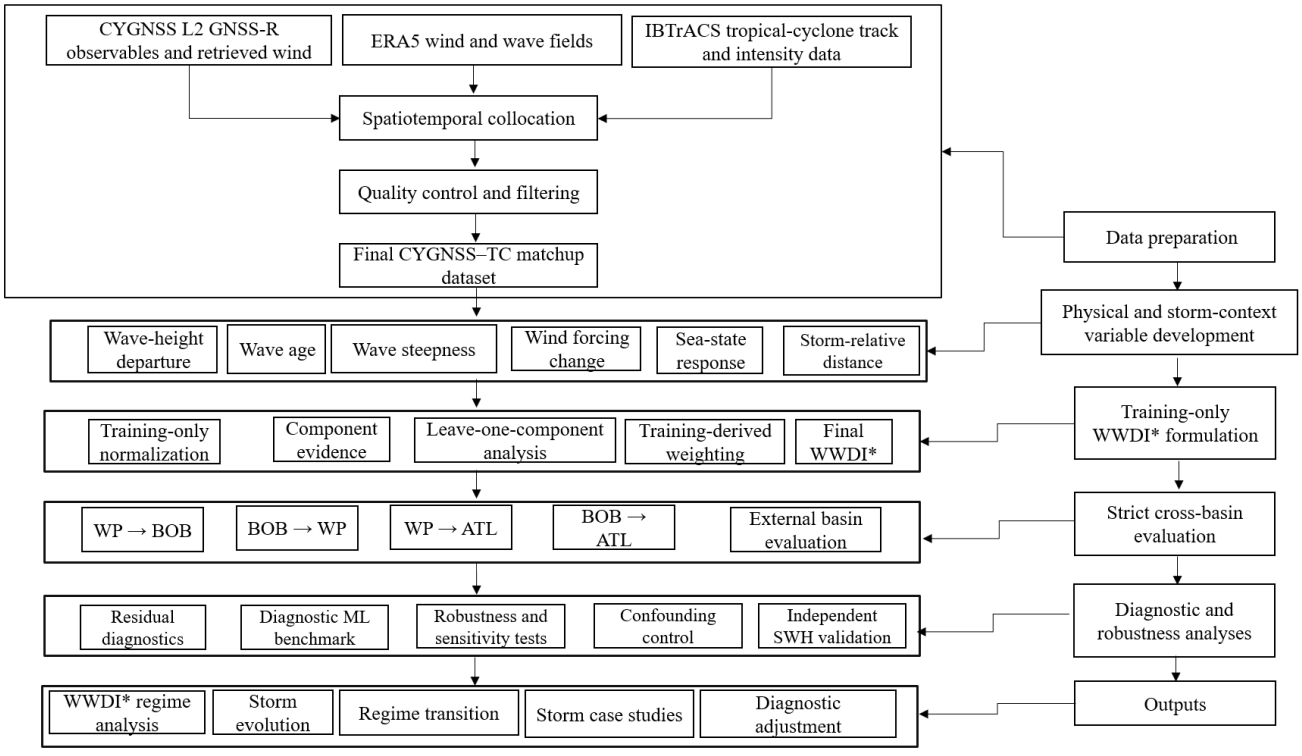


Figure 2. Methodological workflow of the physics-guided tropical-cyclone wind–wave disequilibrium framework. CYGNSS observations are combined with ERA5 atmospheric and wave fields and IBTrACS tropical-cyclone information. Physical and storm-context descriptors are derived from the collocated observations and normalized using training-only statistics. The complete component set is first examined through the equal-weight diagnostic TC-WWDI, after which the supervised WWDI* formulation is constructed from training-derived evidence weights. Evaluation comprises strict temporal and cross-basin transfer, confounder and multicollinearity diagnostics, permutation testing, storm-wise bootstrap resampling, sensitivity analysis, independent ERA5–CMEMS wave validation, diagnostic machine-learning experiments, and training-calibrated mismatch adjustment.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$E_{\mathrm{TC}} = |U_{\mathrm{CYGNSS}} - U_{\mathrm{TC}}| . \tag{13}$$

E_{TC} is the principal response variable for the association analysis, training-stage WWDI* evidence weighting, robustness experiments, confounder analysis, and diagnostic machine-learning benchmark. The signed quantity e_{TC} is retained for bias analysis and for the final training-calibrated diagnostic adjustments.

For comparison with the collocated atmospheric reanalysis, the absolute CYGNSS–ERA5 difference is

$$E_{\mathrm{ERA5}} = |U_{\mathrm{CYGNSS}} - U_{\mathrm{ERA5}}| , \tag{14}$$

where U_{ERA5} denotes the collocated ERA5 10-m wind speed (Hersbach et al., 2020).

3.3 Wind–wave physical variables

3.3.1 Low-wind regularization

Several wind-normalized wave quantities contain wind speed in the denominator and can become numerically unstable under weak-wind conditions. A smooth effective-wind formulation is therefore used in preference to a hard denominator threshold:

$$U_{x,\mathrm{eff}} = \sqrt{U_x^2 + U_0^2}, \qquad x \in \{\mathrm{ERA5}, \mathrm{TC}\}, \tag{15}$$

with the primary regularization value

$$U_0 = 2\ \mathrm{m\,s^{-1}}. \tag{16}$$

At moderate and high wind speeds, $U_{x,\mathrm{eff}}$ approaches the corresponding unregularized wind, whereas the regularized form prevents singular behaviour under weak winds. Sensitivity is evaluated using

$$U_0 \in \{1, 2, 3, 5\}\ \mathrm{m\,s^{-1}}. \tag{17}$$

The wind–wave components entering TC-WWDI and WWDI* are constructed using $U_{\mathrm{ERA5,eff}}$. Although $U_{\mathrm{TC,eff}}$ is retained as an auxiliary derived quantity, it does not enter the final diagnostic machine-learning feature groups.

3.3.2 Characteristic wave scale, wave age, and wave steepness

The bulk sea-state descriptors are constructed from ERA5 significant wave height H_s and mean wave period T_m. Characteristic wavelength and phase-speed scales are obtained from T_m using the deep-water dispersion approximation. These quantities are

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

used as internally consistent bulk diagnostic scales rather than as direct estimates of spectral-peak wavelength, finite-depth wave properties, or the complete directional wave spectrum.

The characteristic wavelength scale is

$$L_{T_m} = \frac{gT_m^2}{2\pi}, \tag{18}$$

and the corresponding phase-speed scale is

$$c_{T_m} = \frac{gT_m}{2\pi}, \tag{19}$$

with

$$g = 9.80665\ \mathrm{m\,s^{-2}}. \tag{20}$$

Because T_m is the ERA5 mean wave period rather than the spectral peak period, L_{T_m} and c_{T_m} are interpreted as characteristic bulk-wave scales (Rascle and Ardhuin, 2013; Soisuvarn et al., 2016).

Effective bulk wave age is defined as

$$A_w = \frac{c_{T_m}}{U_{\mathrm{ERA5,eff}}}, \tag{21}$$

where lower A_w denotes a younger bulk sea state relative to the local wind forcing.

Bulk wave steepness is defined as

$$S_w = \frac{H_s}{L_{T_m}}, \tag{22}$$

and a wind-normalized dimensionless wave-height scale is

$$H_w^* = \frac{gH_s}{U_{\mathrm{ERA5,eff}}^2}. \tag{23}$$

A_w, S_w, and H_w^* provide different representations of the same underlying bulk wind–wave state. Wave age characterizes maturity relative to local forcing, S_w describes geometric steepness, and H_w^* expresses significant wave height relative to the effective wind scale (Rascle and Ardhuin, 2013; Clarizia and Ruf, 2017; Pascual et al., 2021). They are not independent physical degrees of freedom because they are constructed from the common variables H_s, T_m, and $U_{\mathrm{ERA5,eff}}$. Direct substitution gives the exact identity

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$H_w^* = 2\pi S_w A_w^2. \tag{24}$$

The three quantities are consequently treated as physically interpretable alternative representations of a common bulk sea state rather than as independent dynamical variables. Their separate inclusion in the diagnostic development allows the analysis to determine which representation most effectively organizes the observed mismatch.

The resulting dimensionless scaling and weak-wind behaviour are summarized in Fig. 3.

Figure 3. Dimensionless wave scaling and low-wind regularization diagnostics

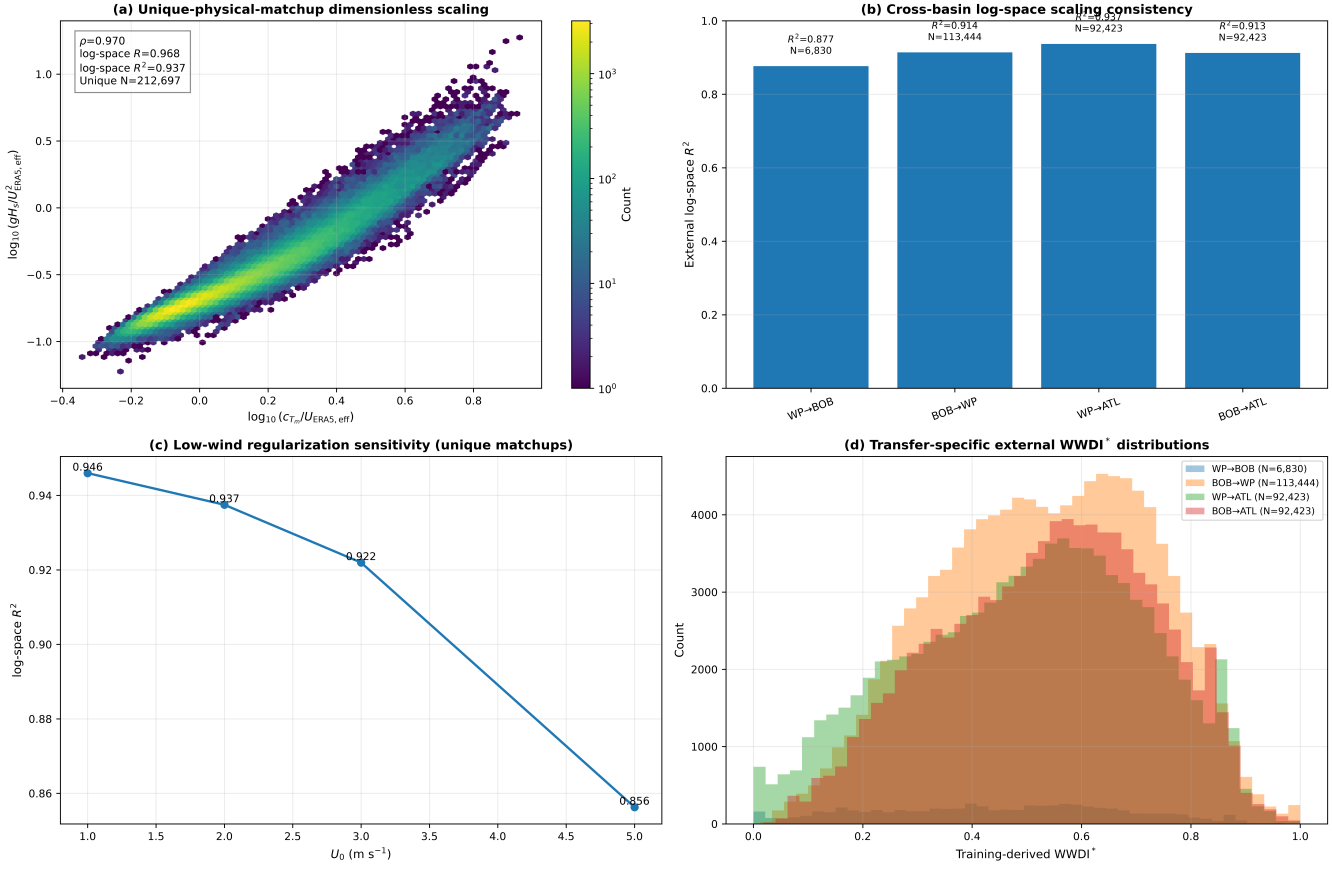


Figure 3. Dimensionless wave scaling and low-wind regularization diagnostics. Panel (a) shows the log-space relationship between effective wave age $A_w = c_{T_m}/U_{\mathrm{ERA5,eff}}$ and dimensionless wave height $H_w^* = gH_s/U_{\mathrm{ERA5,eff}}^2$. Panel (b) shows cross-basin consistency of the training-derived log-space scaling. Panel (c) summarizes sensitivity to the low-wind regularization parameter U_0, and panel (d) shows external distributions of the frozen training-derived WWDI* formulations.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

3.4 Transient wind forcing and sea-state response

Tropical-cyclone atmospheric forcing may evolve more rapidly than the accompanying bulk wave field. Two temporal descriptors are therefore used to represent the relative rate of change of wind forcing and significant wave height. Direct differencing of consecutive CYGNSS observations is avoided because satellite sampling is irregular and very short observation separations can produce unstable finite-difference estimates (Prakash et al., 2018; Bethel et al., 2022; Lima et al., 2022).

Observations belonging to each cyclone are grouped into 3-h storm-time bins. Median ERA5 wind speed, effective ERA5 wind speed, and significant wave height are calculated within each bin. Changes between successive valid bins are estimated as

$$\frac{\Delta U_{\mathrm{ERA5}}}{\Delta t} = \frac{U_{\mathrm{ERA5},b} - U_{\mathrm{ERA5},b-1}}{t_b - t_{b-1}}, \tag{25}$$

and

$$\frac{\Delta H_s}{\Delta t} = \frac{H_{s,b} - H_{s,b-1}}{t_b - t_{b-1}}, \tag{26}$$

where b denotes the storm-time bin.

Only consecutive differences satisfying

$$1\ \mathrm{h} \le \Delta t \le 12\ \mathrm{h} \tag{27}$$

are retained.

The raw transient wind-forcing descriptor is

$$T_{U,\mathrm{raw}} = \frac{|\Delta U_{\mathrm{ERA5}}/\Delta t|}{U_{\mathrm{ERA5,eff}} + \epsilon_U}, \tag{28}$$

with

$$\epsilon_U = 10^{-9}\ \mathrm{m\,s^{-1}}, \tag{29}$$

and the corresponding raw transient sea-state descriptor is

$$T_{H,\mathrm{raw}} = \frac{|\Delta H_s/\Delta t|}{H_s + \epsilon_H}, \tag{30}$$

where

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$\epsilon_H = 10^{-9}\ \mathrm{m}. \tag{31}$$

The small constants protect only against division by zero. Before robust normalization, both transient descriptors are restricted to

$$0 \le T_{U,\mathrm{raw}}, T_{H,\mathrm{raw}} \le 2\ \mathrm{h}^{-1}. \tag{32}$$

The bin-level values are assigned to the retained CYGNSS observations falling within the corresponding cyclone and 3-h interval. These quantities are used as compact descriptors of environmental evolution rather than as a complete spectral wave-growth model.

3.5 Storm-relative position and storm evolution

Storm-relative position is represented by the great-circle distance R_{TC} between the CYGNSS specular point and the matched IBTrACS storm centre. A bounded radial-proximity descriptor is defined as

$$P_{R,\mathrm{raw}} = 1 - \mathrm{clip}\left(\frac{R_{\mathrm{TC}}}{R_0}, 0, 1\right), \tag{33}$$

where

$$R_0 = 300\ \mathrm{km}. \tag{34}$$

$P_{R,\mathrm{raw}}$ approaches one near the matched storm centre and decreases toward zero at the outer boundary of the retained TC domain. The 300-km scale is therefore a contextual normalization tied to the analysis domain and is not interpreted as a universal physical cyclone radius (Knaff et al., 2021; Mayers et al., 2023).

Storm evolution is referenced to the time of maximum collocated U_{TC} for each cyclone. The observations are divided into five relative phases:

- far before peak: $t < -24$ h;
- before peak: $-24 \le t < -6$ h;
- near peak: $-6 \le t \le +6$ h;
- after peak: $+6 < t \le +24$ h;
- far after peak: $t > +24$ h.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

These phase categories and continuous hours from the intensity peak are used for storm-evolution interpretation and as auxiliary inputs to the extended diagnostic machine-learning feature group. They do not contribute to training-stage estimation of the principal WWDI* component weights.

3.6 Training-only robust normalization

For each transfer experiment, candidate components are normalized using statistics estimated exclusively from the corresponding 2022–2023 training subset. Let X_j denote a raw component and $Q_{j,2}$ and $Q_{j,98}$ denote the 2nd and 98th percentiles of its finite training values. The normalized quantity is

$$\widetilde{X}_j = \mathrm{clip}\left(\frac{X_j - Q_{j,2}}{Q_{j,98} - Q_{j,2}}, 0, 1\right). \tag{35}$$

The percentile limits estimated from the training observations are frozen and subsequently applied without modification to the same-basin 2024 internal-validation data and to the external basin. Neither evaluation subset contributes to estimation of the normalization parameters.

A component is treated as invalid for a transfer if

$$N_{j,\mathrm{finite}} < 10 \quad \text{or} \quad Q_{j,98} \le Q_{j,2}. \tag{36}$$

A neutral zero contribution is assigned in such cases.

Because lower wave age corresponds to a younger bulk sea state, its orientation is reversed:

$$D_A = 1 - \widetilde{A}_w. \tag{37}$$

The naturally increasing normalized terms are

$$D_S = \widetilde{S}_w, \qquad T_U = \widetilde{T}_{U,\mathrm{raw}}, \qquad T_H = \widetilde{T}_{H,\mathrm{raw}}. \tag{38}$$

The radial-proximity term is normalized using the same training-only procedure to obtain P_R. The six candidate physical and storm-context components are summarized in Table 2.

3.7 Dimensionless wave-height departure

A derived wave-height-departure component is included to characterize deviation from the training-basin relationship between effective wave age and dimensionless wave height. Because A_w, S_w, and H_w^* satisfy the exact relationship in Eq. 24, D_H does not represent an additional independent physical degree of freedom. It is instead interpreted as a consistency descriptor describing departure from the empirical training relationship linking bulk wave maturity and wind-normalized wave height.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2. Physical and storm-context components used in the wind–wave disequilibrium framework. Final components are bounded to $[0,1]$ using parameters estimated exclusively from the corresponding training subset.

Component	Interpretation	Raw quantity	Larger normalized value indicates
D_H	Derived wave-height departure	$\lvert r_H \rvert$	Larger departure
D_A	Young-wave state	$c_{T_m}/U_{\mathrm{ERA5,eff}}$	Younger sea
D_S	Bulk wave steepness	H_s/L_{T_m}	Steeper sea
T_U	Transient wind forcing	$\lvert \Delta U_{\mathrm{ERA5}}/\Delta t \rvert / U_{\mathrm{ERA5,eff}}$	Faster wind change
T_H	Transient sea-state response	$\lvert \Delta H_s/\Delta t \rvert / H_s$	Faster SWH response
P_R	Storm-relative proximity	$1-\mathrm{clip}(R_{\mathrm{TC}}/300,0,1)$	Closer to storm centre

Using finite positive observations from the designated training subset, the log-space relationship is fitted as

$$\log_{10} H_w^* = \alpha_H + \beta_H \log_{10} A_w + r_H, \tag{39}$$

where α_H and β_H are estimated exclusively from the training subset and r_H is the residual.

The raw departure magnitude is

$$D_{H,\mathrm{raw}} = |r_H| = \left|\log_{10} H_w^* - (\alpha_H + \beta_H \log_{10} A_w)\right|. \tag{40}$$

Larger $D_{H,\mathrm{raw}}$ therefore identifies observations lying farther from the wave-age–dimensionless-height relationship characteristic of the training basin. Because of the exact algebraic dependence among the underlying bulk quantities, such departures are interpreted in terms of changes in the coupled age–steepness structure rather than as an independent wave-height mechanism.

After estimation, α_H and β_H are frozen and applied unchanged to internal-validation and external data. D_H is obtained by applying the training-only robust normalization in Eq. 35 to $D_{H,\mathrm{raw}}$ (Rascle and Ardhuin, 2013; Clarizia and Ruf, 2017; Pascual et al., 2021).

3.8 Diagnostic TC-WWDI

Before supervised weighting is introduced, the complete normalized component set is retained:

$$\mathcal{C} = \{D_H, D_A, D_S, T_U, T_H, P_R\}. \tag{41}$$

The equal-weight diagnostic TC-WWDI is

$$\mathrm{TC{-}WWDI} = \frac{1}{6}\left(D_H + D_A + D_S + T_U + T_H + P_R\right). \tag{42}$$

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

TC-WWDI is not the final supervised transferable formulation. Its purpose is to examine whether the complete candidate component set organizes observations into coherent differences in wave maturity, bulk steepness, derived wave-height departure, transient environmental evolution, storm-relative position, and reference-relative CYGNSS mismatch.

3.9 Training-derived evidence-weighted WWDI*

WWDI* is constructed entirely from the designated 2022–2023 training subset. Its weighting therefore depends only on training-stage evidence and is not optimized against either the same-basin 2024 holdout or an external dataset.

An equal-weight candidate containing all six components is first calculated:

$$I_{\mathrm{full}} = \frac{1}{6}\sum_{j=1}^{6} C_j, \qquad C_j \in \mathcal{C}. \tag{43}$$

Its training-stage Spearman correlation with E_{TC} is denoted ρ_{full}.

Each component is then omitted individually:

$$I_{-j} = \frac{1}{5}\sum_{\substack{k=1\\k\neq j}}^{6} C_k, \tag{44}$$

with corresponding training correlation ρ_{-j}. The incremental training evidence associated with component j is

$$\Delta\rho_j = \rho_{\mathrm{full}} - \rho_{-j}. \tag{45}$$

Only positive evidence increments are retained:

$$q_j = \max(\Delta\rho_j, 0). \tag{46}$$

The normalized component weight is

$$w_j = \frac{q_j}{\sum_{k=1}^{6} q_k}. \tag{47}$$

If every q_j is zero, equal six-component weights are used as a predefined fallback. Otherwise, components with non-positive incremental training evidence receive zero weight.

The final supervised index is

$$\mathrm{WWDI}^* = \sum_{j=1}^{6} w_j C_j. \tag{48}$$

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

All normalization limits, wave-height scaling coefficients, and WWDI* component weights are estimated before either evaluation subset is examined. Because E_{TC} enters explicitly into the training-stage evidence calculation, WWDI* is interpreted as a supervised physics-guided diagnostic index rather than as an independently derived causal measure of wind–wave disequilibrium.

3.10 Strict temporal and cross-basin evaluation

The evaluation strategy separates parameter development, same-basin temporal assessment, and external geographical transfer. Random observation-level splitting is not used to define these partitions because such a procedure could distribute observations from the same cyclone or closely related storm environment across development and evaluation samples (Roberts et al., 2016; Varoquaux, 2018).

For WP-based development, observations from 2022–2023 form the training subset, containing 73,322 observations from 48 tropical cyclones. WP observations from 2024 form a distinct same-basin internal-validation subset containing 40,122 observations from 28 storms. The two subsets share neither storm identifiers nor physical CYGNSS observations.

For BOB-based development, the 2022–2023 training subset contains 5,802 observations from 11 storms. The corresponding BOB-2024 internal-validation subset contains 1,028 observations from three storms. These subsets likewise share no storm identifier or physical CYGNSS observation.

Geographical evaluation uses the complete 2022–2024 record from an independent basin. Four predefined transfers are considered: WP→BOB, BOB→WP, WP→ATL, and BOB→ATL. The BOB external dataset contains 6,830 observations from 14 storms, the WP external dataset contains 113,444 observations from 76 storms, and the ATL external dataset contains 92,423 observations from 57 storms. The evaluation design is summarized in Table 3.

Table 3. Strict temporal and cross-basin evaluation design. Training is restricted to 2022–2023 in the development basin, same-basin 2024 observations form the internal temporal holdout, and the complete 2022–2024 record of an independent basin is used for external evaluation. Values in parentheses denote numbers of tropical cyclones.

Transfer	Training set	Training N	Internal validation	Internal N	External set	External N
WP→BOB	WP, 2022–2023	73,322 (48)	WP, 2024	40,122 (28)	BOB, 2022–2024	6,830 (14)
BOB→WP	BOB, 2022–2023	5,802 (11)	BOB, 2024	1,028 (3)	WP, 2022–2024	113,444 (76)
WP→ATL	WP, 2022–2023	73,322 (48)	WP, 2024	40,122 (28)	ATL, 2022–2024	92,423 (57)
BOB→ATL	BOB, 2022–2023	5,802 (11)	BOB, 2024	1,028 (3)	ATL, 2022–2024	92,423 (57)

Partition independence is verified directly from the final basin master datasets. No storm identifier or physical CYGNSS observation is shared between training, internal-validation, and external stages in any transfer. Pairwise comparison of the WP, BOB, and ATL master datasets likewise identifies no shared storm identifier or duplicate physical CYGNSS observation. External evaluation is therefore independent at both storm and physical-observation levels.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

The external datasets span 2022–2024 and consequently overlap the development data in calendar year. They are not claimed to be temporally independent of the training period. The calendar overlap does not constitute storm or observation leakage because the external observations originate from geographically separate basins and contain neither cyclone identities nor physical CYGNSS records represented in the corresponding development basin. Same-basin 2024 observations provide the explicit temporal holdouts, whereas the external datasets provide geographical, storm-independent, and observation-independent evaluation.

All quantities defining the physical diagnostic formulation are derived exclusively from the 2022–2023 training subset. These include normalization limits, empirical wave-height scaling coefficients, WWDI* component weights, and calibration parameters for the final diagnostic adjustments.

Candidate machine-learning algorithms are also fitted exclusively using the 2022–2023 development data. The predefined same-basin 2024 subset is used only to select among candidate machine-learning algorithms for each feature group, as described in Sect. 3.16; it is not used to re-estimate WWDI* normalization parameters, scaling coefficients, or component weights.

The external basin is never used for fitting, normalization, component weighting, algorithm selection, hyperparameter specification, or adjustment calibration. The ATL experiments therefore provide a third geographically independent basin in which separately developed WP- and BOB-based formulations are tested against the same external regional dataset.

3.11 Association and error metrics

Observation-level monotonic association between WWDI* and E_{TC} is quantified using Spearman rank correlation. Pearson correlation is additionally used where linear association is of interest.

Because WWDI* is intended primarily as a regime diagnostic, external observations are ordered by WWDI* and divided into 10 approximately equal-count quantile bins. Median WWDI* and median E_{TC} are calculated within each retained bin. A linear regression is then fitted to the resulting pairs of bin medians, and the coefficient of determination of this median-level regression is reported as the binned R^2.

For each external transfer, a bin is retained only when it contains at least the larger of 30 observations or 0.2 % of the complete external sample. The predefined 10-bin configuration satisfies this criterion in all four principal transfers.

The binned R^2 is therefore interpreted as a descriptive measure of regime-scale organization and not as an observation-level predictive coefficient of determination. Uncertainty in reported bin medians is represented using distribution-free 95 % confidence intervals based on binomial order statistics, retaining separate lower and upper limits.

For signed differences e_i, mean bias is

$$\mathrm{Bias} = \frac{1}{N}\sum_{i=1}^{N} e_i, \tag{49}$$

mean absolute error is

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|e_i|, \tag{50}$$

and root mean square error is

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}e_i^2}. \tag{51}$$

3.12 Confounder control and multicollinearity

Tropical-cyclone intensity and storm-relative radius can influence both the sea-state descriptors and the reference-relative CYGNSS mismatch. Their contribution is therefore examined using a rank-polynomial residual analysis (Knaff et al., 2021; Mayers et al., 2023).

For the N finite observations entering a given external-transfer analysis, fractional ranks of TC intensity and storm-relative distance are defined as $r_U = \mathrm{rank}(U_{\mathrm{TC}})/(N+1)$ and $r_R = \mathrm{rank}(R_{\mathrm{TC}})/(N+1)$, respectively. The nonlinear control basis is then

$$\mathbf{Z}_r = \left\{r_U, r_U^2, r_R, r_R^2, r_U r_R\right\}. \tag{52}$$

Constructing the polynomial terms from the two base fractional ranks preserves distinct linear, quadratic, and interaction terms in rank space. The same fractional-rank transformation is applied separately to WWDI* and E_{TC}. Each ranked response is then residualized against $\mathbf{Z}_r$ using ordinary least-squares regression with an intercept, producing residual vectors r_W and r_E.

The residual monotonic association is

$$\rho_{\mathrm{partial}} = \rho_s(r_W, r_E), \tag{53}$$

where ρ_s is Spearman rank correlation.

For complementary reporting, the rank-residual coefficient of determination is

$$R^2_{\mathrm{partial}} = \left[\mathrm{corr}(r_W, r_E)\right]^2, \tag{54}$$

where corr is the Pearson correlation between the two residual vectors. R^2_{partial} is therefore not the square of the reported residual Spearman coefficient. Both quantities are treated as conditional diagnostics and are not assigned a causal interpretation.

Linear dependence among the six normalized candidate components is examined separately using the conventional variance-inflation factor

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

$$\mathrm{VIF}_j = \frac{1}{1 - R_j^2}, \tag{55}$$

where R_j^2 is obtained from an ordinary least-squares regression with an intercept in which component j is regressed on the remaining candidate components. WWDI* itself is excluded from the VIF calculation because it is explicitly constructed from the same component set.

3.13 Permutation testing and storm-wise bootstrap uncertainty

Permutation testing and storm-wise bootstrap resampling address distinct aspects of uncertainty in the external regime relationships.

For the permutation analysis, the observed WWDI* values are retained while E_{TC} is randomly reassigned among individual observations. The complete 10-bin regime calculation is repeated for every permutation. For B permutations, the finite-sample empirical probability is

$$p_{\mathrm{perm}} = \frac{1 + \sum_{b=1}^{B} \mathbb{1}\left(R^2_{b,\mathrm{perm}} \geq R^2_{\mathrm{obs}}\right)}{B + 1}, \tag{56}$$

with

$$B = 500. \tag{57}$$

The finite-sample correction prevents a zero Monte Carlo probability. Because reassignment occurs at the observation level, this test evaluates whether the observed regime organization exceeds that obtained under random pairing of WWDI* and mismatch observations.

Storm dependence is examined independently through cluster bootstrap resampling. Complete storm identities are sampled with replacement, and all observations belonging to each sampled cyclone are retained. The 10-bin regime statistic is recalculated for 500 storm-wise bootstrap realizations and summarized using the median and percentile 95 % confidence interval. Resampling whole storms preserves the within-cyclone clustering of CYGNSS observations (Roberts et al., 2016; Varoquaux, 2018).

A fixed base random seed of 42 is used to derive deterministic, procedure-specific random-number streams for the stochastic analyses. Consequently, permutation and bootstrap results do not depend on execution order and can be reproduced from the same input datasets and software environment.

3.14 Sensitivity experiments

Sensitivity experiments are performed without relaxing the training-only construction of WWDI*.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

First, the primary 2nd–98th percentile normalization interval is compared with 1st–99th and 5th–95th percentile alternatives. For every configuration, normalization limits and WWDI* component weights are re-estimated exclusively from the corresponding training subset before external application.

Second, weak-wind regularization is examined using

$$U_0 \in \{1, 2, 3, 5\}\ \mathrm{m\,s^{-1}}, \tag{58}$$

while retaining the primary normalization procedure.

Third, the specification of D_H is examined because the dimensionless-height quantity from which it is constructed is algebraically linked to wave age and bulk wave steepness. The equal-weight diagnostic TC-WWDI is recomputed after exclusion of D_H, and the resulting external associations are compared with those from the six-component formulation. The supervised WWDI* formulation is also reconstructed after explicit removal of D_H to determine whether the principal transferable index depends numerically on this derived descriptor.

Finally, sensitivity of the regime-scale statistic to bin count is examined using 5, 8, 10, 12, and 15 approximately equal-count WWDI* bins. The 10-bin configuration remains the predefined primary analysis; the alternative bin counts are used only to assess dependence of the reported regime relationship on discretization.

3.15 Independent ERA5–CMEMS SWH validation

Because ERA5 significant wave height H_s enters several of the wind–wave descriptors, its representation of the observed sea state is evaluated independently against Sentinel-3A altimeter observations from the Copernicus Marine Service described in Sect. 2.5.

CMEMS quality screening is applied during source-data preprocessing. Only Sentinel-3A observations satisfying `valueQc=1` are admitted to the collocation procedure. The source quality flag is not retained as a separate field in the final ERA5–CMEMS matchup table; the final table therefore contains observations that have already passed the prescribed source-data quality screening.

The primary validation requires a maximum spatial separation of

$$\Delta r_{\mathrm{CMEMS}} \leq 100\ \mathrm{km}, \tag{59}$$

and a maximum absolute temporal separation of

$$|\Delta t_{\mathrm{CMEMS}}| \leq 3\ \mathrm{h}. \tag{60}$$

Records containing non-finite ERA5 SWH, CMEMS SWH, spatial separation, or temporal separation are excluded before calculation of validation statistics. The resulting observations are used solely to assess the ERA5 wave field and remain separate from construction or calibration of the wind–wave disequilibrium framework.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

CMEMS observations do not contribute to TC-WWDI or WWDI* normalization, training-derived component weights, machine-learning selection or fitting, transfer-partition definition, or calibration of either diagnostic adjustment. The comparison therefore provides an independent observational assessment of the ERA5 H_s input without introducing altimeter measurements into the subsequent diagnostic analysis (Hersbach et al., 2020; Wiese et al., 2018).

For each retained matchup, the signed SWH difference is

$$e_{H_s} = H_{s,\mathrm{ERA5}} - H_{s,\mathrm{CMEMS}}, \tag{61}$$

such that positive values denote higher ERA5 SWH and negative values denote lower ERA5 SWH relative to the associated CMEMS observation.

Agreement is quantified using bias, MAE, RMSE, coefficient of determination, Pearson correlation, and Spearman rank correlation. The reported R^2 uses the conventional residual-sum-of-squares definition with CMEMS SWH treated as the observed variable and ERA5 SWH as the estimate; it is therefore calculated independently of squared Pearson correlation.

The primary validation uses the 100-km/3-h collocation window. Sensitivity to collocation proximity is evaluated by tightening only the spatial threshold to 50 km, tightening only the temporal threshold to 1 h, and applying both restrictions simultaneously. These configurations are robustness tests and do not replace the primary validation sample.

The validation dataset is additionally inspected for repeated use of physical observations. Duplicate CYGNSS physical observations and repeated associations involving the same CMEMS altimeter observation are identified separately. A supplementary sampling-structure test retains one representative association per unique CMEMS observation and recalculates the aggregate ERA5–CMEMS statistics. This analysis determines whether repeated association of an altimeter observation with multiple admissible CYGNSS records materially affects the principal validation conclusion.

The ERA5–CMEMS experiment is intended to assess whether the reanalysis captures the dominant significant-wave-height variability required for regime-scale characterization. Agreement with Sentinel-3A is not interpreted as evidence that ERA5 is free from local or event-specific uncertainty during rapidly evolving tropical cyclones.

3.16 Diagnostic machine-learning benchmark

The diagnostic machine-learning benchmark is used solely to examine the incremental information carried by local atmospheric and wave-state variables. The response variable is E_{TC}, the absolute CYGNSS–IBTrACS reference-relative mismatch; ocean-surface wind speed itself is not predicted. The experiment therefore does not constitute a new CYGNSS wind-retrieval model (Asgarimehr et al., 2020; Qiao and Huang, 2025).

Five predefined feature groups are evaluated:

- **A: TC–radial context**: U_{TC} and R_{TC};
- **B: wind context**: Group A plus U_{ERA5};
- **C: bulk wind–wave**: Group B plus H_s and T_m;

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.



- **D: WWDI***: Group B plus the training-derived WWDI*;
- **E: extended**: Group B plus H_s, T_m, WWDI*, T_U, T_H, continuous hours from the storm-intensity peak, and categorical storm-phase code.

The feature groups are deliberately constructed to avoid unnecessary duplication of deterministic transformations of the same physical information. In particular, $U_{\mathrm{TC,eff}}$, the radial-proximity transformation P_R, and the complete algebraically related wave-age–steepness–dimensionless-height triplet are not simultaneously added when the corresponding underlying information is already represented by the selected features.

Six candidate algorithms are considered: Ridge regression, Huber regression, histogram gradient boosting, random-forest regression, extremely randomized trees, and a multilayer perceptron. Their configurations are specified before external evaluation rather than tuned against the external datasets (Asgarimehr et al., 2020; Roberts et al., 2016; Varoquaux, 2018; Qiao and Huang, 2025).

Ridge regression uses $\alpha = 1$ following robust feature scaling. Huber regression uses $\epsilon = 1.75$, $\alpha = 10^{-3}$, and a maximum of 1000 iterations, also following robust scaling.

Histogram gradient boosting uses 500 maximum boosting iterations, a learning rate of 0.035, 15 maximum leaf nodes, a minimum leaf size of 30, and L_2 regularization of 0.05.

Random forest and extremely randomized trees each use 220 estimators, maximum tree depth 18, and minimum leaf size 5. Deterministic random seeds are used for both algorithms.

The multilayer perceptron contains hidden layers of 48 and 24 units, ReLU activation, and the Adam optimizer. Training uses a maximum of 500 iterations, an initial learning rate of 5×10^{-4}, an L_2 penalty of 10^{-3}, automatic batch size, shuffling, and early stopping. The early-stopping validation fraction is 0.15 and the no-improvement criterion is 25 iterations. Input features are scaled using a robust 5th–95th percentile transformation, while the response variable is standardized within the training procedure. Random operations use deterministic seeds derived from a fixed base seed of 42.

A deterministic upper training-sample safeguard of 80,000 observations is defined for random forest, extremely randomized trees, and the multilayer perceptron. The safeguard is inactive in the final experiments because the largest development subset contains 73,322 observations.

Candidate algorithms are fitted exclusively to the 2022–2023 training subset of the relevant development basin. For each feature group, the algorithm giving the lowest RMSE on the corresponding same-basin 2024 internal-validation subset is selected. Internal-validation MAE and then algorithm name are used as deterministic secondary tie-breakers. The selected algorithm is subsequently frozen and applied to the external basin. External metrics play no role in algorithm selection or hyperparameter specification.

Before model fitting, observations containing non-finite values in the required feature set or E_{TC} are excluded separately for the corresponding feature group. No transformation or imputation parameter is estimated from the external basin. Where scaling is required, the transformation is fitted using training observations only.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Because E_{TC} is non-negative, predicted values are constrained to a non-negative range. The upper prediction bound is determined exclusively from the fitted training-response distribution as 1.5 times its 99.5th percentile, subject to predefined lower and upper limits of 20 and 90 $\mathrm{m\,s^{-1}}$, respectively. The same frozen bound is used for internal-validation and external predictions.

The benchmark is interpreted strictly as an information-content experiment. Because U_{TC} contributes both to the storm-reference context and to the definition of E_{TC}, the reported performance does not constitute independent operational wind-retrieval skill.

3.17 Outcome-independent case-study selection

A single Western Pacific tropical cyclone is used to illustrate within-storm evolution of the diagnostic variables. To avoid selecting a storm because it exhibits an unusually favourable WWDI*–mismatch relationship, selection is performed without reference to WWDI*, E_{TC}, wave age, wave steepness, or their subsequent correlations.

Candidate cyclones are drawn from the BOB→WP external evaluation and are required to contain at least 120 unique physical matchups. Among eligible storms, the cyclone with the largest observed IBTrACS maximum sustained wind is selected. Matchup count is used only as a deterministic tie-breaker if required.

The time of maximum collocated storm intensity is used solely as a temporal reference for visualization. The case study is illustrative and is not treated as independent causal validation of WWDI*.

3.18 Training-only WWDI*-conditioned diagnostic adjustment

The final experiment examines whether the training-derived WWDI* regimes contain transferable information about systematic signed CYGNSS–IBTrACS mismatch and whether any resulting improvement extends beyond removal of a single mean training-basin offset.

Two training-calibrated adjustment strategies are evaluated. The first is a global-bias baseline. For each transfer, the mean signed e_{TC} is calculated over the complete 2022–2023 training subset and subtracted from every external CYGNSS wind estimate. This baseline quantifies how much external mismatch can be reduced by removal of one training-derived mean offset without using WWDI* or any regime information.

The second strategy conditions the training residual on WWDI*. For each transfer, the training-basin WWDI* distribution is divided into 10 approximately equal-count quantile bins. Both the bin boundaries and the corresponding signed mismatch statistics are estimated exclusively from the training observations and are then frozen.

For training bin b, the mean signed mismatch is

$$\bar{e}_b = \frac{1}{N_b} \sum_{i \in b} \left(U_{\mathrm{CYGNSS},i} - U_{\mathrm{TC},i} \right), \tag{62}$$

where N_b is the number of training observations assigned to that bin.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

An external observation is assigned using the frozen training-derived WWDI* boundaries. Its regime-conditioned diagnostic adjustment is

$$U_{\mathrm{CYGNSS,corr},i} = U_{\mathrm{CYGNSS},i} - \bar{e}_{b(i)}, \tag{63}$$

where $b(i)$ is the corresponding training-defined WWDI* bin.

The resulting signed reference-relative mismatch is

$$e_{\mathrm{TC,corr},i} = U_{\mathrm{CYGNSS,corr},i} - U_{\mathrm{TC},i}. \tag{64}$$

No external residual is used to estimate, modify, select, or tune either adjustment. Performance is evaluated over the complete external dataset, with RMSE used as the principal comparison and bias and MAE retained as secondary error diagnostics.

The total improvement of the WWDI*-conditioned adjustment relative to the raw mismatch is reported separately from its incremental improvement over the global training-bias baseline. The latter is quantified through the difference between the globally corrected and WWDI*-conditioned RMSE values and by the corresponding percentage reduction relative to the globally corrected RMSE.

Uncertainty in this incremental comparison is evaluated using paired storm-wise bootstrap resampling with 500 realizations. Complete external storm identities are sampled with replacement, and both adjustment methods are evaluated on exactly the same resampled cyclones in every realization. The distribution of the paired RMSE difference is summarized using its median and percentile 95 % confidence interval. This paired design tests whether WWDI* conditioning provides additional storm-level improvement beyond global debiasing rather than attributing the complete raw-to-adjusted reduction to the wind–wave regime.

The low-, medium-, and high-WWDI* categories displayed in the Results are external reporting tertiles only. They are not calibration intervals: the diagnostic adjustment always uses the 10 quantile bins determined exclusively from the training subset.

The experiment is therefore interpreted as an assessment of transferable, regime-dependent mismatch reducibility relative to IBTrACS. Neither the global-bias correction nor the WWDI*-conditioned adjustment establishes a new geophysical model function, a universally valid pointwise correction, or an operational CYGNSS wind-retrieval algorithm (Ruf et al., 2019b; Pascual et al., 2021; Powell and Ruf, 2025).

4 Results

4.1 Physical behaviour of diagnostic TC-WWDI

The physical behaviour of the diagnostic TC-WWDI was first examined under the four predefined external-transfer configurations. Each transfer was analysed separately so that differences in storm population, regional sampling, and sea-state conditions remained explicit rather than being masked through basin pooling.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Figure 4 shows a systematic reorganization of the bulk wind–wave state with increasing TC-WWDI. Wave steepness generally increases toward the higher-index regimes, whereas wave age decreases, indicating progressively younger and steeper bulk sea states. Dimensionless wave height also varies across the diagnosed TC-WWDI range, although the magnitude and detailed form of this response differ among the four external transfers.

The two reference-relative CYGNSS mismatch quantities exhibit a related regime structure. Absolute CYGNSS–IBTrACS mismatch generally becomes larger toward the higher TC-WWDI regimes, while absolute CYGNSS–ERA5 difference shows a broadly comparable tendency. These mismatch responses are less regular at the individual-observation level than the accompanying changes in the bulk wave variables. Diagnostic TC-WWDI is therefore treated as a descriptor of physically distinct storm and sea-state regimes rather than as a deterministic predictor of individual mismatch values.

The points in Fig. 4 represent quantile-bin medians. Where uncertainty is displayed, the error bars denote asymmetric, distribution-free 95 % confidence intervals for the corresponding median.

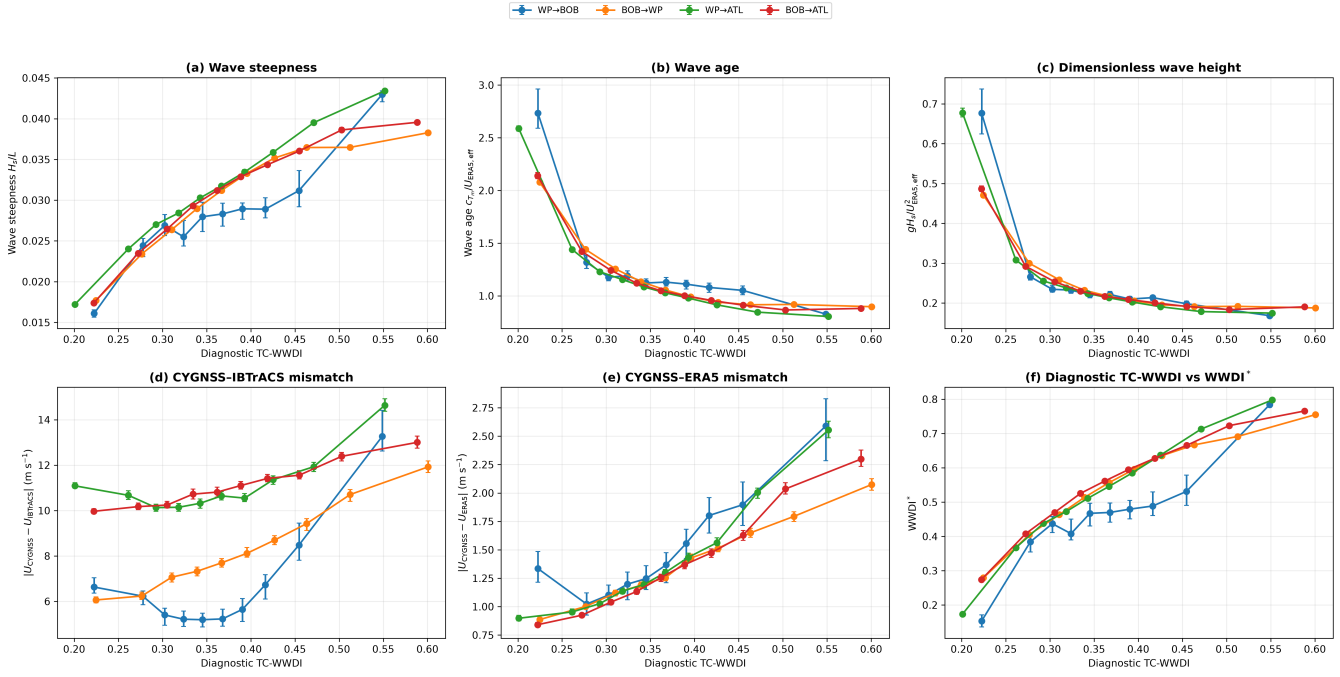


Figure 4. Physical regime behaviour of diagnostic TC-WWDI across the four strict external-transfer experiments. Panels show the regime evolution of (a) wave steepness, (b) wave age, (c) dimensionless wave height, (d) absolute CYGNSS–IBTrACS mismatch, (e) absolute CYGNSS–ERA5 mismatch, and (f) WWDI*. Each transfer is evaluated independently. Points denote quantile-bin medians, and error bars represent asymmetric distribution-free 95 % confidence intervals for the median.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

4.2 Training-derived WWDI* and strict external transfer

WWDI* was constructed independently from the WP and BOB training subsets using the leave-one-component-out evidence procedure described in Sect. 3.9. Components were assigned non-zero weight only when their omission produced a positive training-stage evidence increment.

The resulting contributions are summarized in Table 4. For the WP-trained formulation, D_A, D_S, and T_H receive weights of 0.2207, 0.6397, and 0.1396, respectively. The corresponding BOB-trained weights are 0.2051, 0.6408, and 0.1541. Wave-steepness disequilibrium therefore accounts for the largest training-derived contribution in both configurations. Young-wave state provides the second-largest contribution, while transient sea-state response contributes a smaller but consistently positive fraction.

The remaining candidate components receive zero weight because their leave-one-component-out evidence increments are non-positive. This outcome is specific to the supervised weighting procedure and does not imply that those variables are physically unimportant within tropical-cyclone environments.

Table 4. Training-only evidence used to construct WWDI*. Only components with positive leave-one-component-out evidence increments receive non-zero weights.

Training configuration	Component	$\Delta\rho$	Weight
WP training	D_A	0.0442	0.2207
	D_S	0.1280	0.6397
	T_H	0.0279	0.1396
BOB training	D_A	0.0413	0.2051
	D_S	0.1291	0.6408
	T_H	0.0310	0.1541

The two training-derived formulations were then transferred without re-estimation to their respective external basins. Figure 5 and Table 5 summarize the external WWDI*–mismatch relationships.

Positive observation-level Spearman associations are retained in all four experiments. The coefficients are 0.381 for WP→BOB, 0.292 for BOB→WP, 0.219 for WP→ATL, and 0.223 for BOB→ATL. The corresponding 10-bin regime-scale R^2 values are 0.639, 0.733, 0.714, and 0.753, respectively.

The observation-level correlations are therefore moderate, whereas the relationships between the bin medians show considerably stronger regime-scale organization. This difference is consistent with the intended role of WWDI* as a diagnostic of mismatch-prone wind–wave regimes rather than an observation-by-observation mismatch predictor.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Table 5. External-basin performance of training-derived WWDI* under strict cross-basin evaluation. Binned R^2 is calculated from 10 approximately equal-count WWDI* bins.

Transfer	External N	Storms	Spearman ρ	Binned R^2
WP→BOB	6,830	14	0.381	0.639
BOB→WP	113,444	76	0.292	0.733
WP→ATL	92,423	57	0.219	0.714
BOB→ATL	92,423	57	0.223	0.753

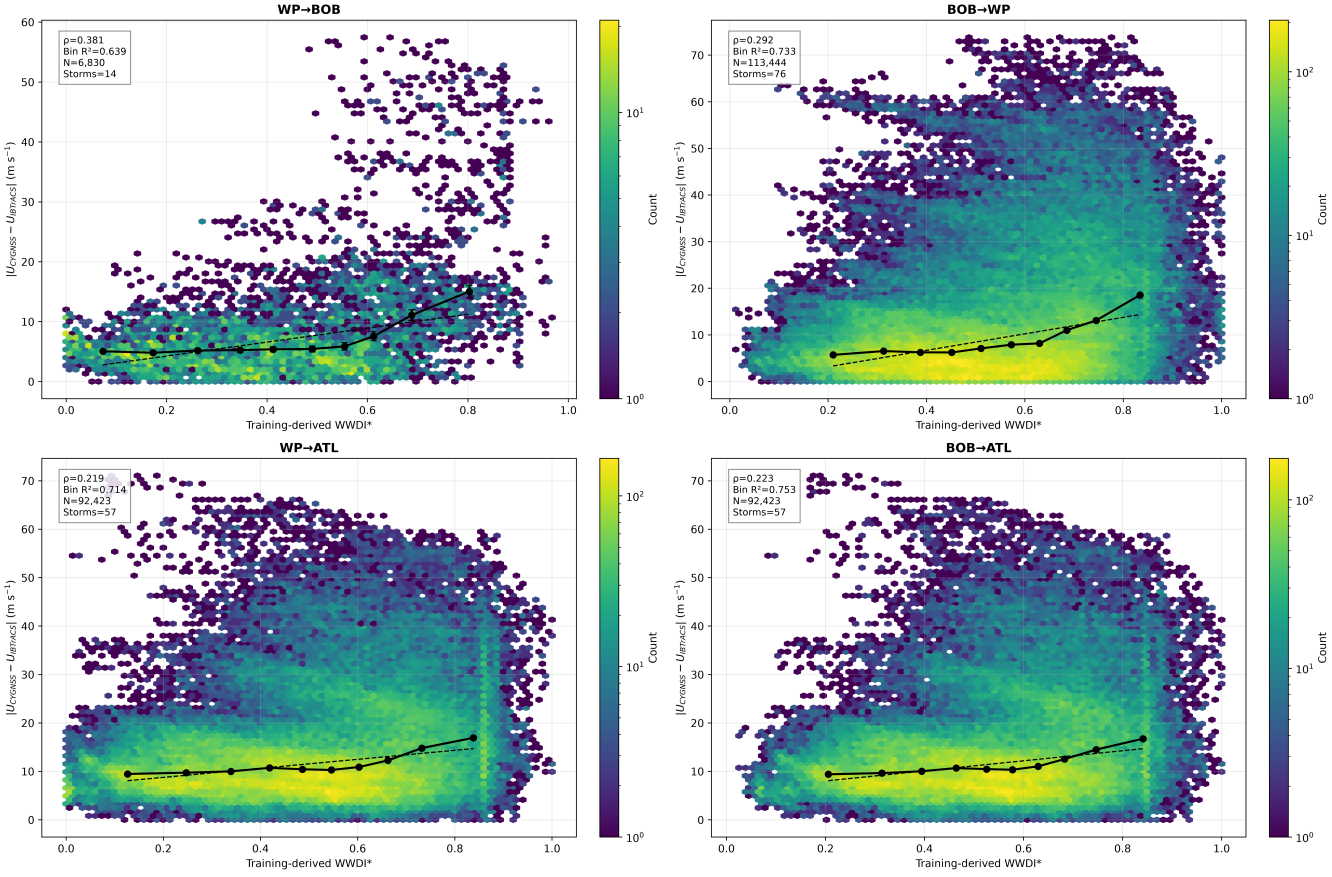


Figure 5. Strict cross-basin external evaluation of WWDI*. Each panel shows the relationship between the frozen training-derived WWDI* and absolute CYGNSS–IBTrACS mismatch. Background density represents individual external observations, while the overlaid points show the 10-bin regime-scale response.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

4.3 Confounder-aware analysis

Tropical-cyclone intensity and storm-relative position evolve jointly with the sea state and can influence both the wind–wave descriptors and the reference-relative CYGNSS mismatch. Their contribution was therefore examined using the rank-polynomial residual analysis described in Sect. 3.12.

The conventional component-level variance-inflation factors remain below 5 in all four external-transfer experiments. The largest value is 4.830, indicating that severe linear multicollinearity among the six normalized candidate components is not evident from this diagnostic.

After controlling for ranked TC intensity, ranked storm-relative radius, their quadratic terms, and their interaction, the residual Spearman correlations between WWDI* and E_{TC} are -0.543, -0.645, -0.622, and -0.611 for WP→BOB, BOB→WP, WP→ATL, and BOB→ATL, respectively. The corresponding partial-rank R^2 values are 0.280, 0.366, 0.355, and 0.344 (Table 6).

These negative conditional associations contrast with the positive unconditional relationships reported in Sect. 4.2. The sign reversal persists after use of the non-redundant rank-polynomial control basis, indicating that the positive external WWDI*–mismatch association is strongly embedded within the joint intensity–radius structure of the sampled tropical cyclones. Once these dominant dependencies are removed, the residual rank association is negative in every transfer.

The conditional sign reversal does not invalidate the positive unconditional cross-basin relationship because the two analyses address different aspects of the storm environment. The unconditional relationship describes the co-variation of WWDI* and mismatch across the complete external cyclone population, whereas the residual analysis characterizes their association after the principal intensity and radial dependencies have been removed. The conditional result therefore limits interpretation of WWDI* as an independent monotonic driver of mismatch and instead supports its use as a storm-regime diagnostic whose external behaviour is coupled to tropical-cyclone structure.

The component-collinearity, conditional-association, and rank-residual diagnostics are shown in Fig. 6.

Table 6. Rank-polynomial confounding-control results after accounting for tropical-cyclone intensity, storm-relative radius, their quadratic rank terms, and their interaction. Partial-rank R^2 is calculated from the Pearson correlation between the two rank-residual vectors and is not the square of the reported residual Spearman coefficient.

Transfer	N	Partial Spearman ρ	Partial-rank R^2
WP→BOB	6,830	-0.543	0.280
BOB→WP	113,444	-0.645	0.366
WP→ATL	92,423	-0.622	0.355
BOB→ATL	92,423	-0.611	0.344

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

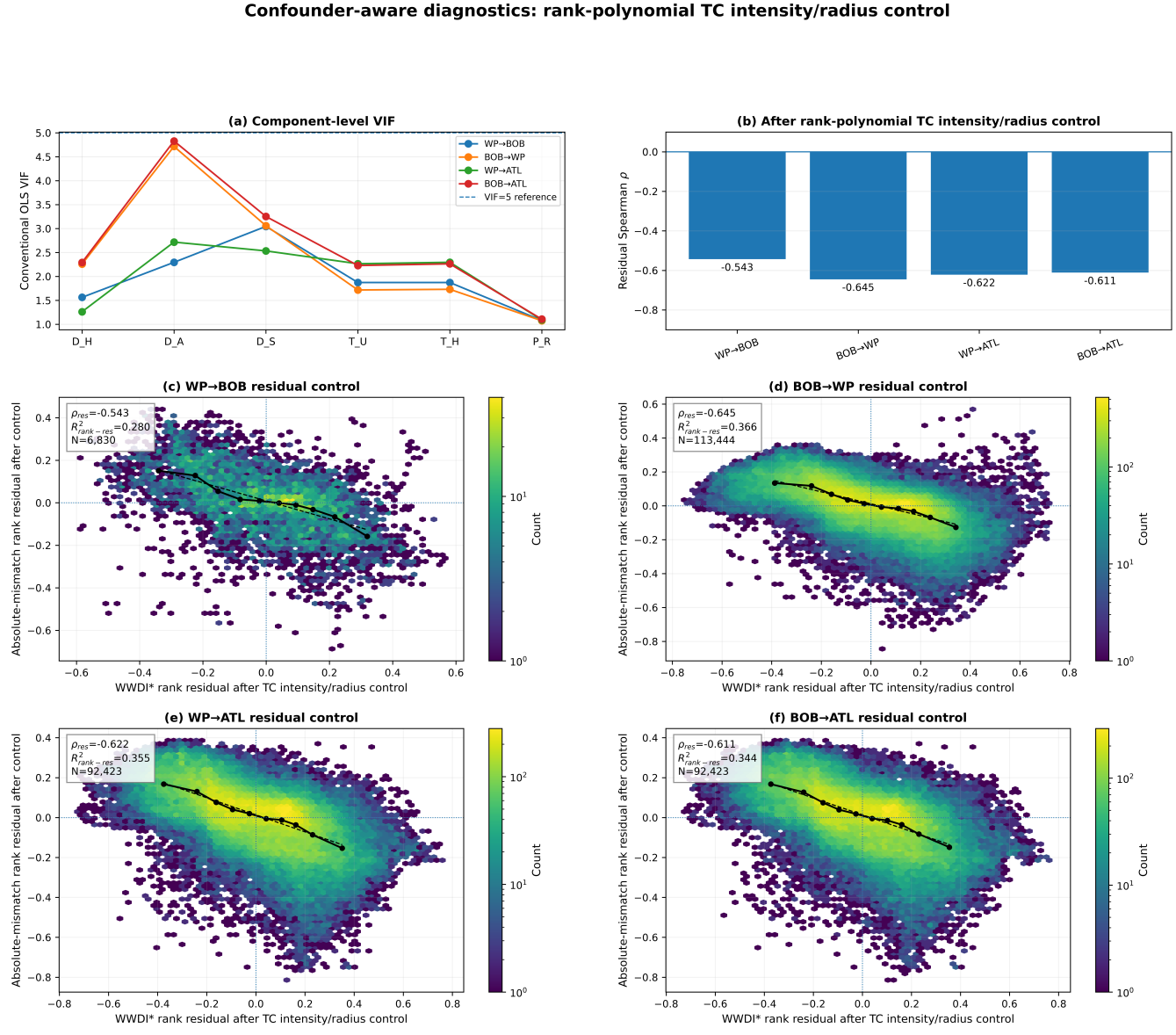


Figure 6. Confounder-aware diagnostic analysis across the four strict external-transfer experiments. Panel (a) shows conventional OLS variance-inflation factors for the six normalized candidate components. Panel (b) shows the residual Spearman association between WWDI* and E_{TC} after controlling for ranked TC intensity, ranked storm-relative radius, their quadratic terms, and their interaction. Panels (c–f) show the corresponding rank-residual relationships. Negative conditional associations persist in all four transfers and are interpreted as conditional diagnostics rather than causal effects.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

4.4 Cross-basin consistency of the physical relationships

The external behaviour of the principal sea-state descriptors was examined separately to determine whether the transferred WWDI* response is accompanied by recurring structure in its constituent physical representations. Because several of these descriptors are algebraically related, this analysis is not intended to establish statistical independence among them. Instead, it assesses whether comparable physical tendencies occur across geographically distinct external datasets.

Figure 7 summarizes the external Spearman associations involving the principal wave-state descriptors and the two reference-relative CYGNSS mismatch measures. The magnitude of the associations varies among transfers, reflecting differences in storm population, regional wave conditions, geographical sampling, and sample size. Several recurring features nevertheless remain evident.

Bulk wave steepness exhibits one of the strongest and most persistent relationships across the four transfer configurations. The young-wave-state descriptor also retains a clear external association, showing that the transferred WWDI* response repeatedly occurs with shifts toward younger and steeper bulk sea states. The derived wave-height-departure term shows weaker and more transfer-dependent behaviour, consistent with its role as a derived consistency descriptor rather than an independent physical degree of freedom.

The reference-relative mismatch quantities also retain recurring external associations, although these are generally weaker than those involving the dominant wave-state descriptors. This difference is consistent with a regime-level interpretation: the composite diagnostic organizes systematic changes in bulk wind–wave conditions more clearly than it predicts individual mismatch values.

Taken together, the results show that the transferred WWDI* response is accompanied by recurring changes in the underlying wind–wave state. However, the variation in association strength among basins also demonstrates that the expression of the diagnosed disequilibrium is not geographically invariant. The external response is therefore interpreted as a recurring regime structure with basin-dependent magnitude rather than as a universal one-to-one relationship between any individual physical descriptor and CYGNSS mismatch.

4.5 Permutation and storm-wise bootstrap robustness

The external WWDI*–mismatch relationships were examined using two complementary resampling procedures. Observation-level permutation testing evaluates the random-pairing null, whereas storm-wise bootstrap resampling examines the sensitivity of the regime-scale relationship to the sampled cyclone population.

For all four transfers, the observed 10-bin R^2 values lie well above the centres of their corresponding permutation distributions (Fig. 8). The observed values are 0.639, 0.733, 0.714, and 0.753 for WP→BOB, BOB→WP, WP→ATL, and BOB→ATL, respectively. Median values under random permutation are 0.064, 0.067, 0.058, and 0.062, and the corresponding finite-sample permutation probabilities are 0.008, 0.010, 0.006, and 0.004. The observed regime organization is therefore substantially stronger than that obtained after random reassignment of the mismatch observations.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

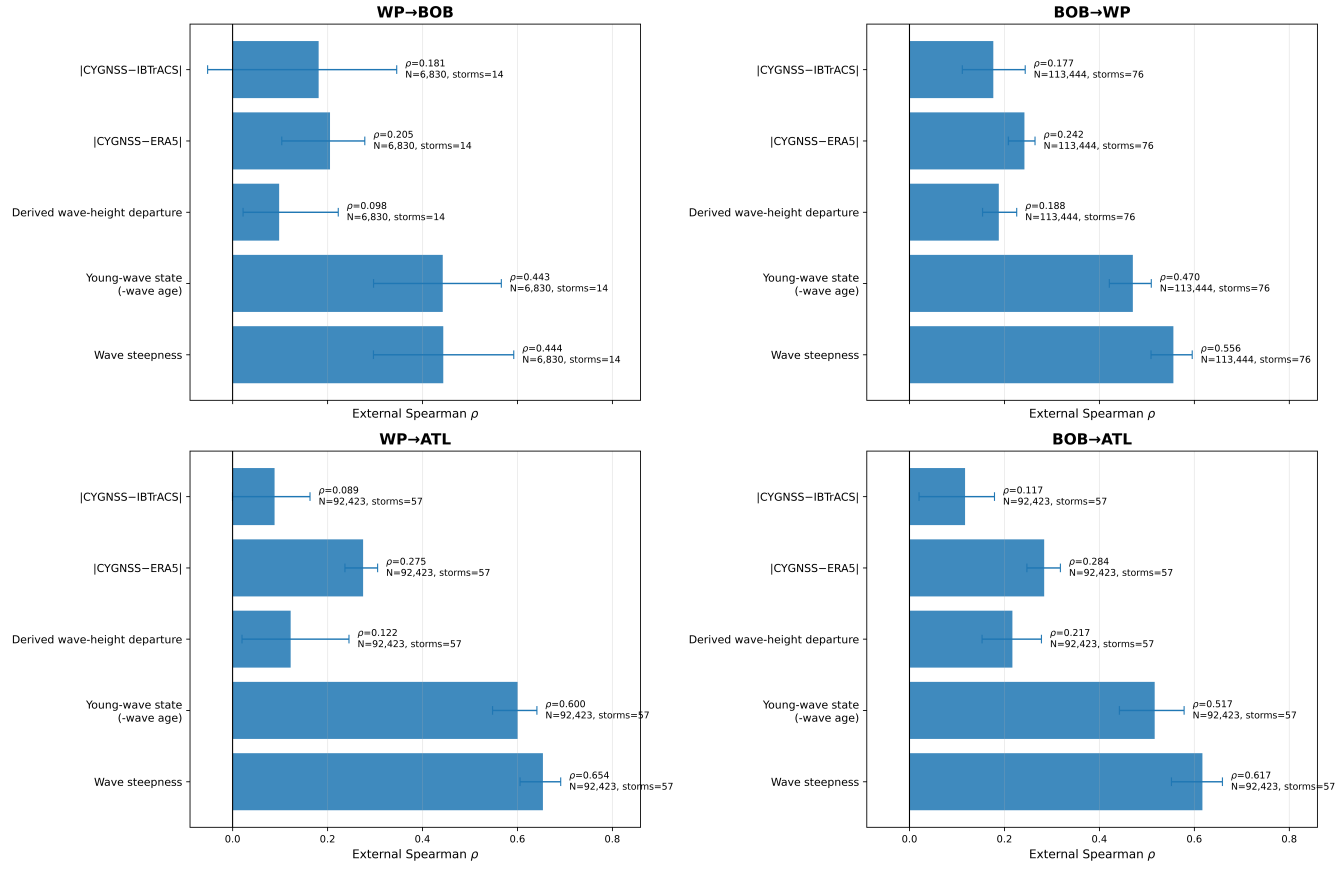


Figure 7. Cross-basin consistency of the external physical relationships. Bars show Spearman rank associations for the two reference-relative CYGNSS mismatch measures and the principal wind–wave descriptors under each strict external-transfer configuration. Wave steepness and young-wave state show the most persistent cross-basin structure, whereas the derived wave-height-departure term exhibits weaker and more transfer-dependent behaviour. Each transfer is evaluated independently.

Storm-wise bootstrap resampling produces median R^2 values of 0.601 for WP→BOB, 0.738 for BOB→WP, 0.695 for WP→ATL, and 0.718 for BOB→ATL. The corresponding 95 % percentile intervals are [0.090, 0.878], [0.501, 0.883], [0.246, 0.916], and [0.309, 0.920].

The width of the bootstrap distributions differs appreciably among the four transfers. WP→BOB and WP→ATL exhibit greater storm-to-storm uncertainty, whereas BOB→WP and BOB→ATL retain more concentrated distributions around high regime-scale R^2. Nevertheless, the bootstrap medians remain close to the corresponding observed values. The resampling analyses therefore show that the external regime structure is not reproduced by random pairing and remains evident when uncertainty is evaluated at the cyclone level, while also demonstrating that its estimated strength is not equally precise across all transfers.

The complete resampling results are summarized in Table 7.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

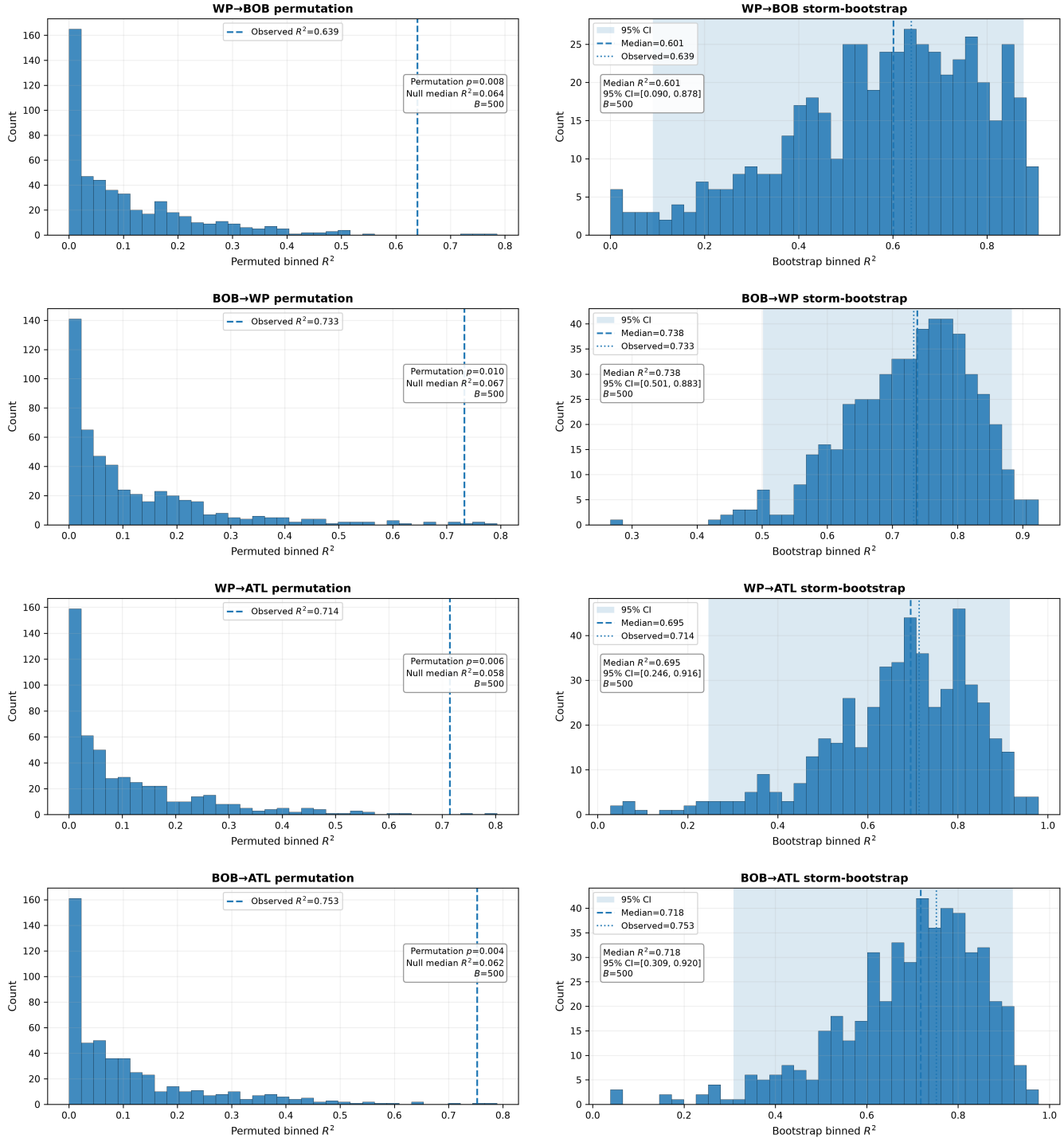


Figure 8. Permutation and storm-wise bootstrap assessment of the external WWDI* regime relationships. Left-column panels show null distributions of the 10-bin R^2 obtained from 500 random permutations, with the observed external value marked for comparison. Right-column panels show the corresponding distributions from 500 storm-wise bootstrap resamples together with the bootstrap median, observed R^2, and 95 % percentile interval.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Table 7. Permutation and storm-wise bootstrap assessment of the external WWDI* regime relationships. The permutation probability is the finite-sample empirical probability obtained from 500 random permutations. Bootstrap statistics are based on 500 storm-wise resamples.

Transfer	Observed R^2	Permutation p	Permutation median R^2	Bootstrap median R^2	95 % CI
WP→BOB	0.639	0.008	0.064	0.601	[0.090, 0.878]
BOB→WP	0.733	0.010	0.067	0.738	[0.501, 0.883]
WP→ATL	0.714	0.006	0.058	0.695	[0.246, 0.916]
BOB→ATL	0.753	0.004	0.062	0.718	[0.309, 0.920]

4.6 Sensitivity to normalization, low-wind regularization, derived-component specification, and binning

The stability of the external WWDI* response was examined with respect to methodological choices that could influence either construction of the index or calculation of its regime-scale summary. The experiments considered the training-derived normalization interval, weak-wind regularization, inclusion of the derived wave-height-departure component D_H, and the number of equal-count bins used for the regime analysis.

Changing the primary 2nd–98th percentile normalization to 1st–99th or 5th–95th percentile alternatives produces only small changes in the external observation-level associations. Spearman coefficients remain within 0.377–0.382 for WP→BOB, 0.292–0.294 for BOB→WP, 0.217–0.220 for WP→ATL, and 0.219–0.225 for BOB→ATL. The corresponding binned R^2 ranges are 0.613–0.656, 0.726–0.751, 0.679–0.717, and 0.752–0.775. The direction of the external WWDI*–mismatch relationship remains unchanged in every configuration.

Weak-wind regularization was assessed using $U_0 = 1$, 2, 3, and 5 ms^{-1}. Across these settings, the external Spearman coefficient ranges from 0.372 to 0.384 for WP→BOB, 0.283 to 0.296 for BOB→WP, 0.210 to 0.222 for WP→ATL, and 0.214 to 0.226 for BOB→ATL. The corresponding binned R^2 ranges are 0.624–0.659, 0.709–0.751, 0.687–0.733, and 0.718–0.777. The principal cross-basin relationships therefore persist across the tested regularization range.

A separate specification experiment was conducted for D_H because the dimensionless wave-height quantity from which it is derived is algebraically related to wave age and bulk wave steepness. The underlying dimensionless identity is reproduced numerically to machine precision, with a maximum absolute discrepancy of 5.33×10^{-15}. D_H is consequently interpreted as a derived consistency or departure descriptor rather than an independent wave-state degree of freedom.

For the equal-weight diagnostic TC-WWDI, removing D_H changes the Spearman association with absolute CYGNSS–IBTrACS mismatch from 0.181 to 0.278 for WP→BOB, from 0.177 to 0.166 for BOB→WP, from 0.089 to 0.112 for WP→ATL, and from 0.117 to 0.099 for BOB→ATL. The corresponding differences are +0.097, -0.011, +0.023, and -0.019. Despite these changes, the six- and five-component diagnostic indices remain strongly rank-correlated, with external correlations ranging from 0.849 to 0.951.

The paired storm-wise bootstrap analysis shows no common direction of the D_H-exclusion effect across all four transfers. The median change in Spearman correlation is 0.092, with a 95 % interval of [-0.041, 0.180], for WP→BOB; -0.010 [-0.025,

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

0.004] for BOB→WP; and 0.021 [-0.007, 0.055] for WP→ATL. For BOB→ATL, the median change is -0.018 with an interval of [-0.033, -0.003].

This specification issue does not alter the supervised WWDI* used in the principal transfer analysis. The training-derived evidence procedure assigns zero weight to D_H in both formulations, with training-stage $\Delta\rho$ values of -0.0169 for WP and -0.0336 for BOB. Explicit removal of D_H therefore leaves the final WWDI* values and all four principal external-transfer results unchanged.

Sensitivity of the regime-scale statistic to bin count was also evaluated using 5, 8, 10, 12, and 15 equal-count WWDI* bins. The independently reconstructed 10-bin configuration reproduces the principal binned R^2 values exactly: 0.639 for WP→BOB, 0.733 for BOB→WP, 0.714 for WP→ATL, and 0.753 for BOB→ATL.

Across the tested configurations, WP→BOB produces binned R^2 values between 0.636 and 0.658, corresponding to a total span of 0.022. BOB→WP varies from 0.716 to 0.766, giving the largest span, 0.050. The corresponding ranges are 0.705–0.719 for WP→ATL and 0.745–0.756 for BOB→ATL, with spans of 0.014 and 0.011. The largest absolute departures from the principal 10-bin values are 0.019, 0.033, 0.009, and 0.008, respectively.

Taken together, these experiments show that the principal WWDI* transfer relationships are stable to reasonable changes in normalization, weak-wind regularization, and regime-bin count. The derived D_H component influences the equal-weight diagnostic TC-WWDI to a limited and transfer-dependent extent but has no numerical effect on the final supervised WWDI* formulation.

4.7 Independent validation of ERA5 significant wave height

ERA5 significant wave height was evaluated independently against CMEMS satellite-altimeter observations because H_s contributes directly to several of the wind–wave descriptors. CMEMS observations are excluded from WWDI* construction, training-derived normalization and weighting, diagnostic machine-learning development, and calibration of the final training-derived mismatch adjustments. The comparison therefore provides an independent assessment of the ERA5 wave field rather than an internal validation of WWDI*.

The primary validation sample contains 5,771 ERA5–CMEMS SWH matchups satisfying the prescribed quality-control and collocation criteria. Across the complete sample, ERA5 gives a bias of -0.139 m, an MAE of 0.440 m, and an RMSE of 0.631 m. The corresponding R^2, Pearson correlation, and Spearman correlation are 0.878, 0.942, and 0.939, respectively. The two datasets therefore show strong correspondence in bulk SWH variability while retaining non-negligible error at the individual matchup level.

Regional statistics are summarized in Table 8. The ATL sample contains 2,657 matchups and yields a bias of -0.117 m, an RMSE of 0.655 m, and a Pearson correlation of 0.944. The WP sample contains 2,948 matchups, with a bias of -0.168 m, an RMSE of 0.616 m, and a Pearson correlation of 0.940. The BOB sample is substantially smaller ($N = 166$) and yields a bias of 0.026 m, an RMSE of 0.470 m, and a Pearson correlation of 0.872. The considerably smaller BOB validation population therefore requires greater caution when comparing its basin-specific statistics with those from ATL and WP.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Sensitivity to the adopted space–time collocation window was examined independently. Tightening only the spatial threshold from 100 to 50 km retains 3,378 matchups and reduces the overall RMSE from 0.631 to 0.589 m, while R^2 increases from 0.878 to 0.893. Restricting only the absolute temporal separation from 3 to 1 h retains 2,007 matchups and gives an RMSE of 0.605 m and $R^2 = 0.897$.

The strictest configuration simultaneously requires a separation not exceeding 50 km and an absolute time difference not exceeding 1 h. This subset contains 1,167 matchups, approximately 20.2 % of the primary validation sample. Its bias is -0.112 m, MAE is 0.382 m, and RMSE is 0.527 m. The corresponding R^2, Pearson correlation, and Spearman correlation are 0.919, 0.962, and 0.950.

Relative to the primary 100-km/3-h configuration, the absolute RMSE difference under the strict 50-km/1-h requirement is 0.104 m. The corresponding absolute changes in R^2, Pearson correlation, and Spearman correlation are 0.041, 0.021, and 0.011. Agreement therefore remains strong and generally improves as the matchup constraints are tightened.

The validation sample was also examined for repeated physical observations. No duplicate CYGNSS physical observation occurs in the tested collocation configurations. Individual CMEMS observations can, however, contribute to more than one admissible association when several CYGNSS records occur within the same space–time neighbourhood. Under the primary criteria, 22.5 % of matchup rows involve a CMEMS observation appearing in more than one association. Restricting the analysis to one representative association per unique CMEMS observation does not materially alter the overall agreement statistics.

Figure 9 summarizes the ERA5–CMEMS relationship, SWH-difference distribution, regional performance, and principal collocation characteristics. Collectively, these results support the use of ERA5 H_s for regime-scale characterization of the bulk wave environment while not implying negligible reanalysis uncertainty at individual locations or during rapidly evolving tropical-cyclone conditions.

Table 8. Independent ERA5–CMEMS significant-wave-height validation under the primary 100-km/3-h collocation criteria.

Region	N	Bias (m)	MAE (m)	RMSE (m)	R^2	Pearson R
All	5,771	-0.139	0.440	0.631	0.878	0.942
ATL	2,657	-0.117	0.457	0.655	0.878	0.944
BOB	166	0.026	0.324	0.470	0.732	0.872
WP	2,948	-0.168	0.431	0.616	0.874	0.940

4.8 Diagnostic machine-learning benchmark

The diagnostic machine-learning benchmark evaluates the incremental information carried by local atmospheric and wave-state variables for explaining E_{TC}. The response variable is therefore the absolute reference-relative CYGNSS–IBTrACS mismatch rather than ocean-surface wind speed, and the analysis is treated as an information-content experiment rather than as an alternative CYGNSS retrieval model.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

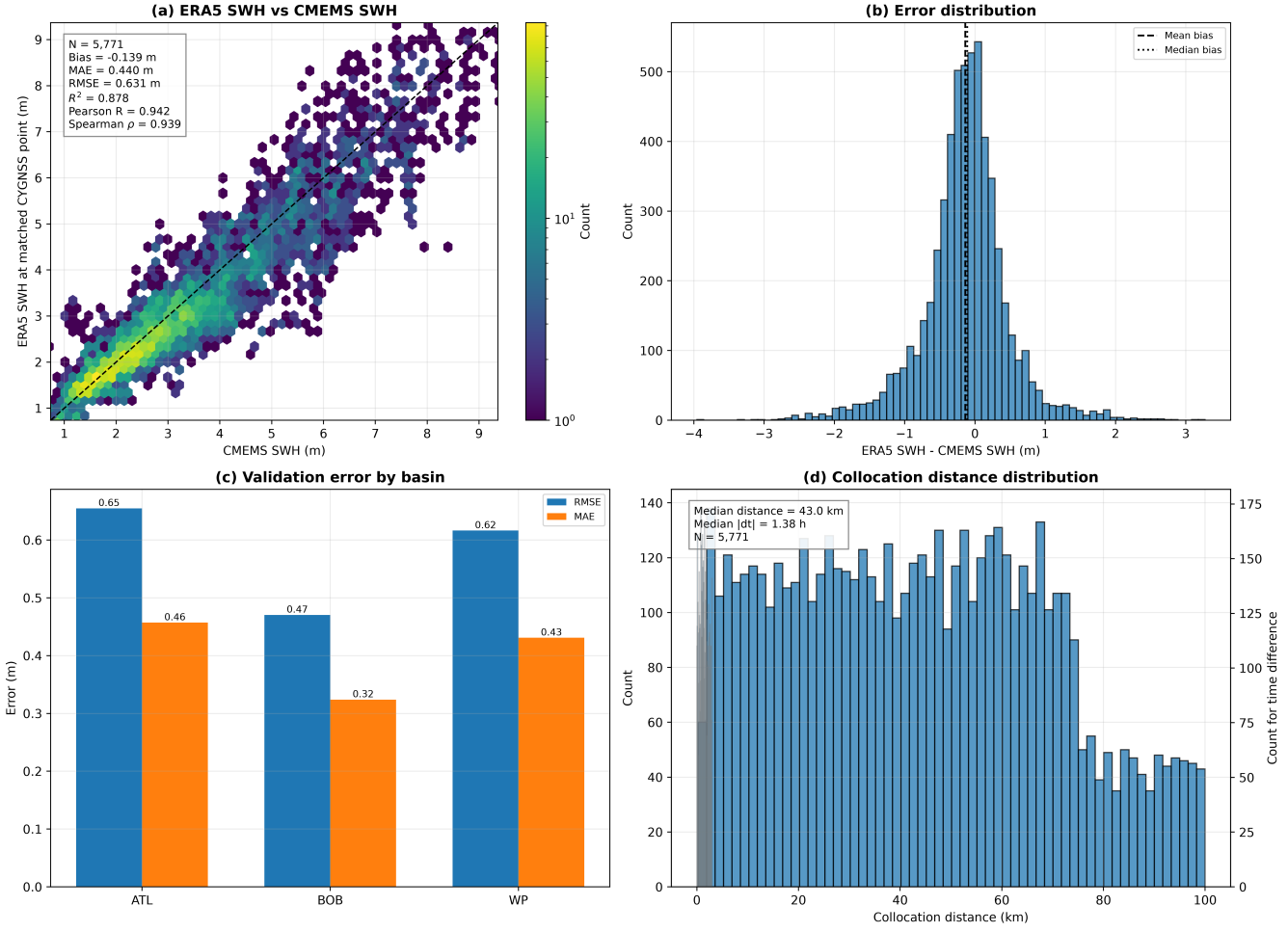


Figure 9. Independent validation of ERA5 significant wave height against CMEMS satellite-altimeter observations under the primary 100-km/3-h collocation criteria. Panels summarize the ERA5–CMEMS SWH relationship, SWH-difference distribution, regional validation statistics, and collocation-distance characteristics. The temporal separation of the retained matchups is reported together with the collocation diagnostics. Sensitivity to stricter space–time criteria is evaluated separately in the accompanying text.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Candidate algorithms were fitted using only the 2022–2023 portion of the corresponding development basin. For each predefined feature group, algorithm selection was performed exclusively using the same-basin 2024 internal-validation subset. The selected algorithm was then frozen and applied to the external basin without further selection or tuning. The complete external results are summarized in Table 9.

Five feature groups were evaluated. Group A contains TC intensity and storm-relative radial distance. Group B adds local ERA5 wind speed. Group C augments this wind-context representation with significant wave height and mean wave period. Group D instead supplements the wind-context baseline with WWDI*. Group E combines local atmospheric wind, bulk wave state, WWDI*, transient terms, and storm-evolution information.

For WP→BOB, the TC–radial reference gives an RMSE of 3.212 ms^{-1} and $R^2 = 0.878$. Adding local ERA5 wind reduces RMSE to 2.562 ms^{-1} and increases R^2 to 0.923. The bulk wind–wave and WWDI* groups yield RMSE values of 2.569 and 2.575 ms^{-1}, respectively, while the extended configuration gives 2.690 ms^{-1}. Most of the improvement over the TC–radial reference is therefore already captured by local atmospheric wind context for this transfer.

For BOB→WP, the TC–radial configuration yields an RMSE of 3.890 ms^{-1}. Adding ERA5 wind reduces this value to 3.059 ms^{-1}. Inclusion of bulk wave information produces a further decrease to 2.631 ms^{-1}, with $R^2 = 0.958$, whereas the WWDI* representation yields 2.863 ms^{-1}. The lowest RMSE for this transfer, 2.605 ms^{-1}, is obtained from the extended feature set.

For WP→ATL, the TC–radial reference gives an RMSE of 4.019 ms^{-1} and $R^2 = 0.862$. The wind-context configuration reduces RMSE to 2.460 ms^{-1}, while adding the bulk wave variables reduces it further to 2.407 ms^{-1}, with $R^2 = 0.950$. The WWDI* and extended representations give RMSE values of 2.452 and 2.509 ms^{-1}, respectively.

For BOB→ATL, the TC–radial configuration gives an RMSE of 4.108 ms^{-1}. Local ERA5 wind reduces the value to 3.057 ms^{-1}. The bulk wind–wave group produces the lowest RMSE, 2.700 ms^{-1}, with $R^2 = 0.938$. The WWDI* and extended groups yield 2.929 and 2.771 ms^{-1}, respectively.

Across the four transfers, local atmospheric wind context provides a substantial improvement relative to TC intensity and radial position alone. Direct bulk wave information provides an additional reduction for BOB→WP, WP→ATL, and BOB→ATL, but not for WP→BOB. WWDI* also carries transferable information, although it does not consistently outperform either the local wind-context baseline or direct bulk wave variables. Its machine-learning contribution is therefore consistent with its intended role as a compact physical diagnostic rather than as a universally optimal predictive feature.

4.9 Storm-centred evolution

The temporal evolution of WWDI* and the associated wave-state variables was examined relative to the time of maximum collocated tropical-cyclone intensity.

Figure 10 shows that median WWDI* generally increases from the earlier storm stages toward the vicinity of peak intensity. Wave steepness follows a broadly comparable evolution, whereas wave age decreases toward the higher-intensity period. The combined response is consistent with a transition toward younger and steeper bulk sea states as the sampled cyclone environment intensifies.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Table 9. Diagnostic machine-learning performance under strict external-basin transfer. Candidate algorithms were fitted using the 2022–2023 training subset and selected using only the corresponding 2024 same-basin internal-validation subset. The selected model was then frozen and evaluated on the external basin. MAE and RMSE are in $\mathrm{m\,s^{-1}}$.

Transfer	Feature group	Selected model	MAE	RMSE	R^2
WP→BOB	A: TC–radial	MLP	2.291	3.212	0.878
	B: Wind context	MLP	1.724	2.562	0.923
	C: Bulk wind–wave	MLP	1.769	2.569	0.922
	D: WWDI*	MLP	1.728	2.575	0.922
	E: Extended	MLP	1.846	2.690	0.915
BOB→WP	A: TC–radial	MLP	2.942	3.890	0.907
	B: Wind context	MLP	1.998	3.059	0.943
	C: Bulk wind–wave	Huber	1.873	2.631	0.958
	D: WWDI*	MLP	1.926	2.863	0.950
	E: Extended	Huber	1.845	2.605	0.958
WP→ATL	A: TC–radial	MLP	3.116	4.019	0.862
	B: Wind context	MLP	1.755	2.460	0.948
	C: Bulk wind–wave	MLP	1.715	2.407	0.950
	D: WWDI*	MLP	1.753	2.452	0.949
	E: Extended	MLP	1.815	2.509	0.946
BOB→ATL	A: TC–radial	MLP	3.164	4.108	0.856
	B: Wind context	MLP	2.063	3.057	0.920
	C: Bulk wind–wave	Huber	1.984	2.700	0.938
	D: WWDI*	MLP	2.025	2.929	0.927
	E: Extended	Huber	2.015	2.771	0.934

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

The evolution of absolute CYGNSS–IBTrACS mismatch differs among the four transfers, reflecting regional and storm-to-storm variability. Mismatch is nevertheless frequently elevated near or after the higher-intensity stages relative to the earlier storm period. This phase dependence is interpreted descriptively and does not imply that storm phase alone determines the observed mismatch.

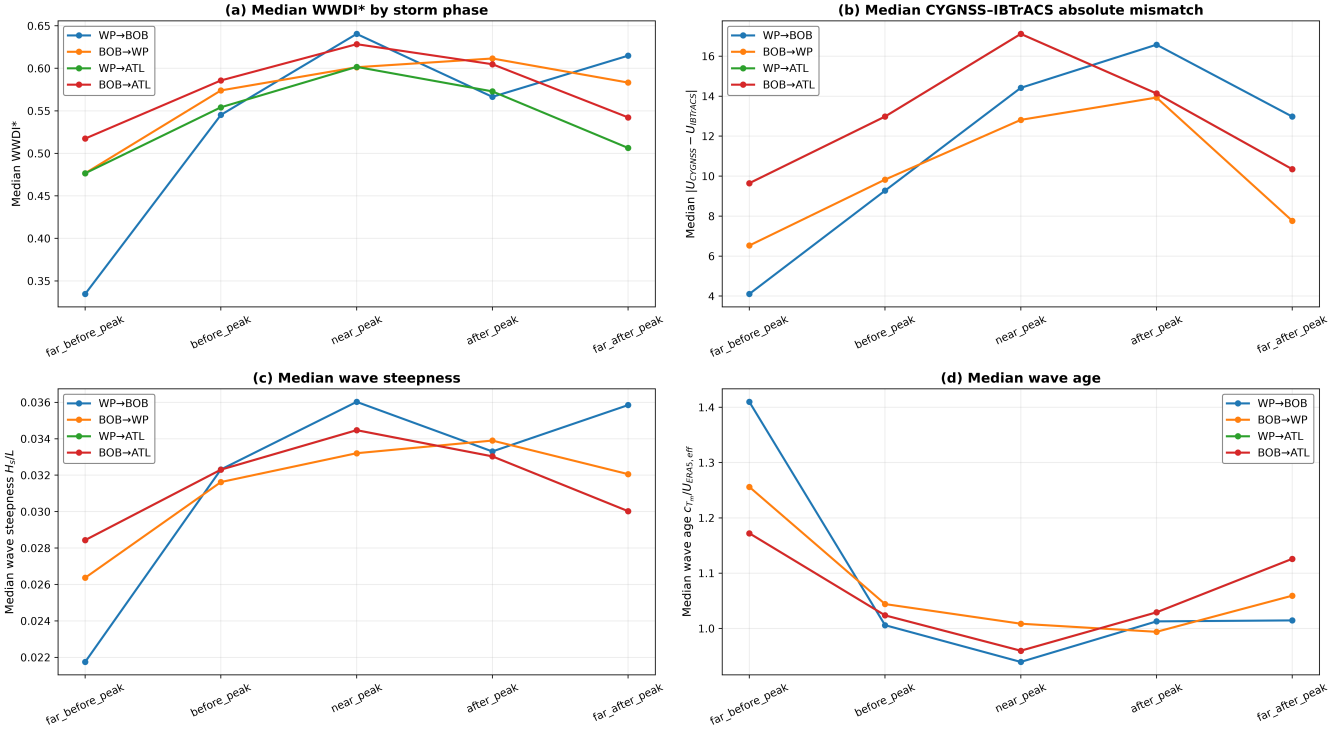


Figure 10. Storm-centred evolution of WWDI*, absolute CYGNSS–IBTrACS mismatch, wave steepness, and wave age relative to the tropical-cyclone intensity peak.

4.10 Regime transition across diagnostic TC-WWDI

The physical environment was further examined by organizing external observations into fixed diagnostic TC-WWDI intervals. Unlike the storm-centred analysis, this representation groups observations according to the diagnosed wind–wave state rather than temporal position relative to the storm-intensity peak.

Figure 11 shows that higher TC-WWDI is generally associated with greater TC intensity and smaller storm-relative distance. Wave age decreases substantially from the lower-index regimes, while wave steepness increases toward higher TC-WWDI.

The reference-relative mismatch quantities show a corresponding broad transition. Absolute CYGNSS–ERA5 difference generally increases with TC-WWDI, while absolute CYGNSS–IBTrACS mismatch becomes larger in the moderate-to-high

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

index regimes. The response is not strictly monotonic through every low-index interval or every transfer, but the highest sampled TC-WWDI regimes consistently exhibit greater mismatch than the lowest regimes.

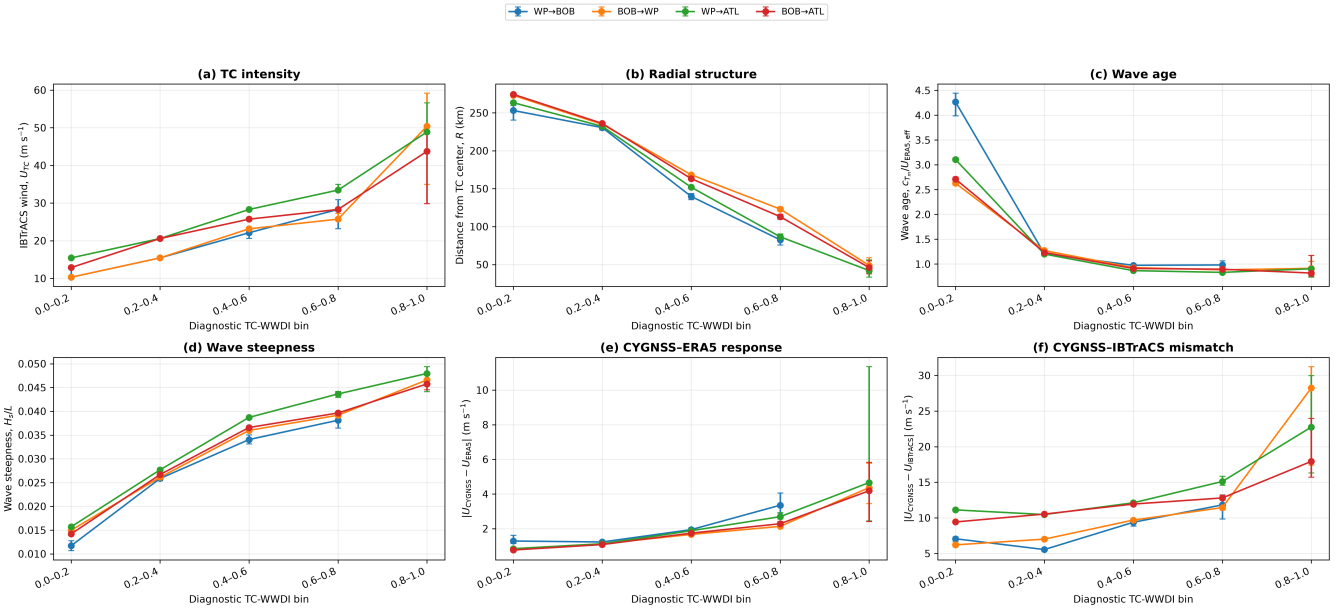


Figure 11. Physical regime transition across fixed diagnostic TC-WWDI intervals. Points denote bin medians, and error bars represent asymmetric distribution-free 95 % confidence intervals for the median. Each strict transfer experiment is shown independently.

4.11 Outcome-independent Western Pacific case study

A Western Pacific tropical cyclone was selected to illustrate the within-storm behaviour of the diagnostic variables. Case selection was performed without using WWDI*, reference-relative mismatch, wave-state correlations, or subsequent diagnostic performance.

Among storms in the BOB→WP external dataset containing at least 120 unique physical matchups, the cyclone with the largest observed IBTrACS maximum wind was selected. The resulting case, `WP_2023138N05151`, contains 2,833 unique CYGNSS–TC matchups between 18 May and 2 June 2023 and reaches a maximum IBTrACS wind of 84.883 ms^{-1}.

Figure 12 shows pronounced spatial and temporal variability in WWDI*, wave steepness, wave age, ERA5 significant wave height, and CYGNSS–IBTrACS mismatch. Across the complete storm period, the Spearman correlation between WWDI* and absolute mismatch is 0.167. The corresponding correlation for wave steepness is 0.173, whereas wave age has a correlation of -0.259.

These within-storm relationships are modest and substantially weaker than would be expected from a deterministic pointwise association. They are therefore consistent with the interpretation of WWDI* as a regime-level diagnostic. The storm-intensity peak is used only as a temporal reference for visualization and is not interpreted as evidence of a causal relationship.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

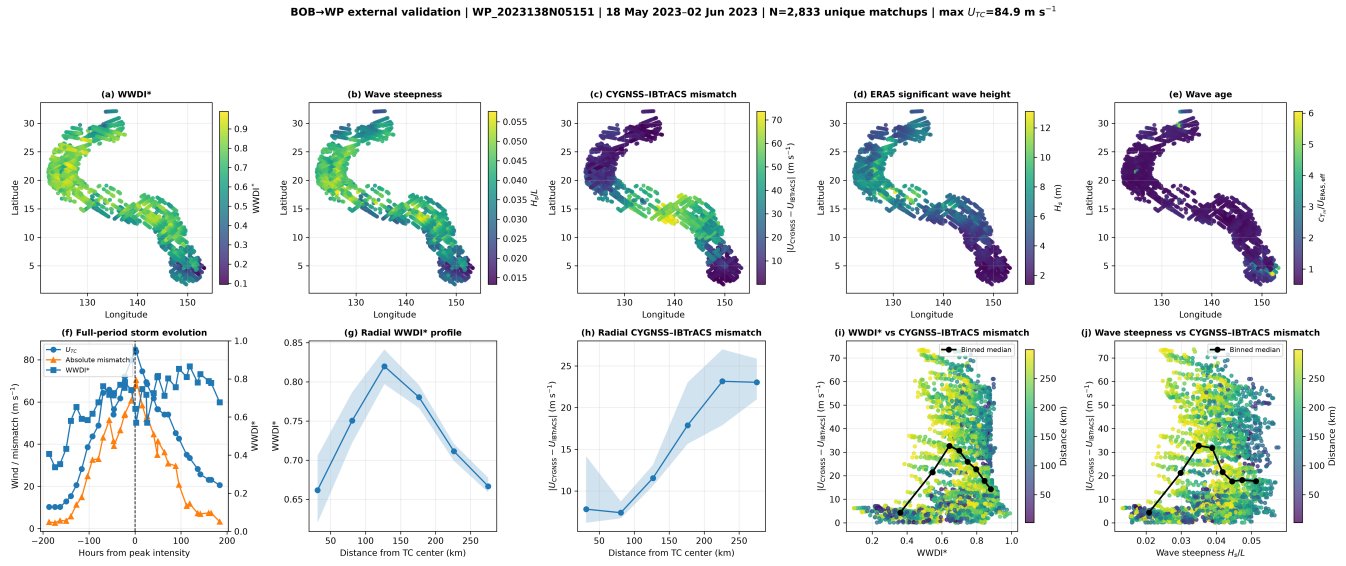


Figure 12. Outcome-independent Western Pacific case study for WP_2023138N05151. The figure summarizes the spatial distribution and storm-time evolution of WWDI*, wave-state descriptors, and reference-relative CYGNSS mismatch. Case selection uses maximum IBTrACS wind after application of the minimum-matchup requirement and does not use the diagnostic outcome variables.

4.12 WWDI*-conditioned diagnostic adjustment

The final experiment examines whether WWDI* contains transferable information about systematic signed CYGNSS–IBTrACS mismatch beyond that captured by a single basin-wide training-bias correction. The comparison therefore distinguishes general debiasing from additional regime-dependent adjustment.

The WWDI*-conditioned adjustment reduces RMSE relative to the raw external mismatch in every transfer (Table 10). For WP→BOB, RMSE decreases from 12.907 to 8.899 $\mathrm{m\,s^{-1}}$, corresponding to a 31.1 % reduction. For BOB→WP, RMSE decreases from 18.012 to 12.642 $\mathrm{m\,s^{-1}}$, a reduction of 29.8 %. The corresponding reductions are 40.4 % for WP→ATL and 35.6 % for BOB→ATL.

A substantial fraction of these raw-to-adjusted improvements is also obtained using the simpler global training-bias correction. The global baseline reduces RMSE to 10.232, 13.277, 10.962, and 11.917 $\mathrm{m\,s^{-1}}$ for WP→BOB, BOB→WP, WP→ATL, and BOB→ATL, respectively. These correspond to reductions of 20.7 %, 26.3 %, 39.4 %, and 34.1 % relative to the raw mismatch.

The WWDI*-conditioned adjustment nevertheless gives a lower point-estimate RMSE than global debiasing in all four external datasets. Relative to the globally corrected result, the additional reductions are 13.0 % for WP→BOB, 4.8 % for BOB→WP, 1.7 % for WP→ATL, and 2.2 % for BOB→ATL. The largest incremental contribution occurs for WP→BOB, where RMSE decreases from 10.232 to 8.899 $\mathrm{m\,s^{-1}}$ when the single global adjustment is replaced by WWDI*-conditioned calibration.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

The storm-level robustness of this incremental difference was evaluated using paired bootstrap resampling. For WP→BOB, the median RMSE advantage of WWDI* conditioning over global debiasing is 1.310 ms^{-1}, with a 95 % interval of [0.338, 2.090] ms^{-1}. For BOB→WP, the corresponding median advantage is 0.610 ms^{-1}, with an interval of [0.149, 1.118] ms^{-1}. Because both intervals remain above zero, the additional storm-level improvement is supported for the reciprocal WP–BOB transfers.

For the Atlantic evaluations, the point estimates also favour WWDI* conditioning, but the incremental differences are smaller. The bootstrap median advantages are 0.184 ms^{-1} for WP→ATL and 0.254 ms^{-1} for BOB→ATL. Their 95 % intervals are [-0.291, 0.586] and [-0.378, 0.758] ms^{-1}, respectively. Because both intervals include zero, the additional storm-level improvement over global debiasing is not regarded as robust for the Atlantic transfers.

Figure 13 shows the raw, global-bias-corrected, and WWDI*-conditioned RMSE together with the performance across the external low-, medium-, and high-WWDI* reporting regimes. The 29.8–40.4 % raw-to-adjusted reductions should not be attributed entirely to WWDI* conditioning because a substantial proportion is explained by removal of the mean training-basin bias. The additional regime-dependent contribution is strongest for WP→BOB, remains supported for BOB→WP, and is substantially smaller and less stable for the two Atlantic evaluations.

The adjustment experiment therefore provides evidence that the diagnosed wind–wave regimes contain transferable information about systematic reference-relative mismatch, while not establishing an independently validated operational CYGNSS wind correction.

Table 10. Comparison of raw, global training-bias-corrected, and WWDI*-conditioned CYGNSS–IBTrACS mismatch under strict external-basin transfer. Both adjustments are calibrated exclusively from the training basin. Extra WWDI* gain is the percentage reduction in RMSE relative to the globally debiased result. RMSE is in ms^{-1}.

Transfer	Raw RMSE	Global RMSE	WWDI* RMSE	Global reduction (%)	WWDI* reduction (%)	Extra WWDI* gain (%)
WP→BOB	12.907	10.232	8.899	20.7	31.1	13.0
BOB→WP	18.012	13.277	12.642	26.3	29.8	4.8
WP→ATL	18.093	10.962	10.774	39.4	40.4	1.7
BOB→ATL	18.093	11.917	11.660	34.1	35.6	2.2

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

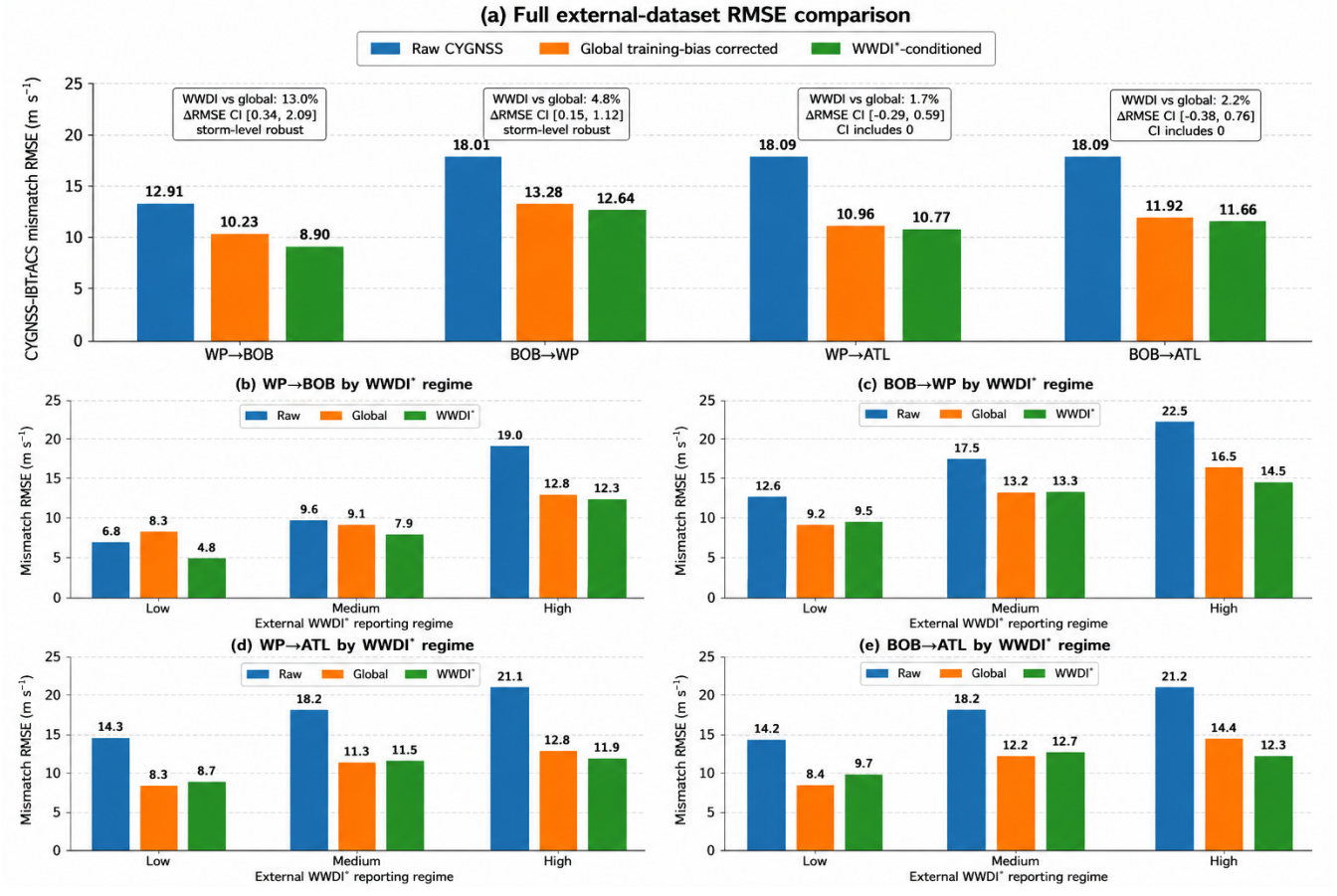


Figure 13. Training-only diagnostic mismatch adjustment under strict cross-basin transfer. Panel (a) compares external RMSE for the raw CYGNSS–IBTrACS mismatch, the global training-bias correction, and the WWDI*-conditioned adjustment. Panels (b–e) summarize performance across low-, medium-, and high-WWDI* external reporting regimes. The global correction uses a single mean signed mismatch estimated from the training basin, whereas the WWDI* adjustment uses mean signed mismatch within 10 training-defined WWDI* quantile bins. All calibration quantities are frozen before external evaluation. The low-, medium-, and high-WWDI* categories are used only for visualization and reporting and do not define the adjustment itself.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

5 Discussion

The present results show that reference-relative CYGNSS wind mismatch during tropical cyclones is systematically organized by the evolving storm and bulk sea-state environment. The strongest training-derived contribution to WWDI* is associated with wave steepness in both development basins, while young-wave state and transient significant-wave-height response provide smaller but consistently positive contributions. The close agreement in this ordering between the WP- and BOB-derived formulations is notable because the two training datasets differ substantially in storm population and sample size. The result suggests that changes in the bulk surface state repeatedly emerge as informative descriptors of CYGNSS–reference mismatch across distinct tropical-cyclone environments.

This behaviour is physically consistent with the nature of GNSS-R ocean observations. CYGNSS does not measure atmospheric wind directly; the reflected signal responds to the statistical roughness and slope distribution of the ocean surface, which in turn reflects the interaction between local wind forcing and the evolving wave field (Soisuvarn et al., 2016; Clarizia and Ruf, 2017; Pascual et al., 2021). During a tropical cyclone, atmospheric forcing can change rapidly over time and space, whereas the bulk wave field may respond on a different timescale. Similar instantaneous wind conditions can therefore occur over sea states with different maturity, steepness, and transient response. The present results are consistent with this physical picture, although they do not establish a causal pathway from any individual wave descriptor to CYGNSS mismatch.

The role of the individual wave-state variables also requires careful interpretation because several are constructed from the same underlying ERA5 quantities. Wave age, bulk wave steepness, and dimensionless wave height are not independent physical degrees of freedom. In particular, H_w^*, S_w, and A_w satisfy an exact algebraic relationship. The derived wave-height-departure component D_H should therefore be viewed as a consistency descriptor within the coupled wind–wave state, rather than as an additional independent process. This interpretation is supported empirically by the training-derived weighting: D_H receives zero weight in both final WWDI* formulations. Explicit removal of D_H consequently leaves the principal WWDI* transfer results unchanged.

The four strict external-transfer experiments provide the principal evidence for geographical transferability. WWDI* remains positively associated with absolute CYGNSS–IBTrACS mismatch in every external basin, despite all normalization parameters, empirical scaling coefficients, and supervised weights being estimated before evaluation of the external observations. Observation-level Spearman correlations range from 0.219 to 0.381, whereas the corresponding 10-bin regime-scale R^2 values range from 0.639 to 0.753. The marked difference between these two scales is important. The moderate pointwise correlations indicate that WWDI* is not a deterministic predictor of mismatch for individual CYGNSS observations. Instead, the stronger bin-median relationships show that the index organizes broader environmental regimes in which the distribution of reference-relative mismatch changes systematically.

This regime-level behaviour is consistent with previous evidence that CYGNSS wind uncertainty is influenced by environmental conditions and by the state of the ocean surface (Ruf et al., 2019b; Ricciardulli et al., 2021; Pascual et al., 2021). The present analysis extends that interpretation by explicitly examining cross-basin transfer under storm- and observation-

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

independent external evaluation. Nevertheless, the binned R^2 values should remain interpreted as descriptive measures of regime organization. They are not equivalent to observation-level predictive coefficients of determination.

The resampling experiments provide additional support for this interpretation. In all four transfers, the observed regime-scale relationships are substantially stronger than those generated after random reassignment of the mismatch observations. Storm-wise bootstrap resampling also retains high median R^2 values when complete cyclone identities are resampled, indicating that the observed structure is not produced solely by a small subset of individual observations. The width of the bootstrap intervals nevertheless differs among transfers. In particular, the WP→BOB and WP→ATL configurations show greater storm-to-storm uncertainty than the reciprocal BOB→WP and BOB→ATL evaluations. The external relationship is therefore reproducible at the regime scale, but its precision is not uniform across regional storm populations.

The corrected confounder analysis places an important limit on the physical interpretation of the positive unconditional associations. After controlling for ranked tropical-cyclone intensity, ranked storm-relative radius, their quadratic terms, and their interaction, the residual Spearman coefficients are -0.543, -0.645, -0.622, and -0.611 for WP→BOB, BOB→WP, WP→ATL, and BOB→ATL, respectively. The corresponding partial-rank R^2 values are 0.280, 0.366, 0.355, and 0.344. The conditional sign reversal therefore persists after use of the non-redundant rank-polynomial control basis.

The negative conditional associations do not contradict the positive unconditional transfer relationships. The two analyses describe different features of the sampled tropical-cyclone environment. The unconditional relationship reflects the joint evolution of storm intensity, radial position, sea state, and mismatch across the complete external population. The residual analysis asks how WWDI* and mismatch co-vary after the dominant intensity–radius structure has been removed. The sign reversal therefore shows that the positive external association should not be interpreted as evidence that WWDI* acts as an independent monotonic driver of mismatch. Rather, WWDI* characterizes a broader storm regime whose behaviour is embedded within the coupled evolution of tropical-cyclone intensity, radial structure, and sea state.

The component-collinearity analysis does not indicate severe linear multicollinearity within the six normalized candidate components. The maximum conventional OLS variance-inflation factor is 4.830 across the four external configurations. This result supports numerical use of the component set while not implying physical independence among the descriptors. The weights assigned by the supervised formulation should therefore continue to be interpreted as training-derived evidence weights rather than as independent causal coefficients.

The methodological sensitivity experiments further show that the principal external relationships are not dependent on a single numerical choice. Changing the training normalization interval and the weak-wind regularization parameter preserves the direction and general magnitude of the transferred WWDI* associations. The binned R^2 also remains stable when the number of equal-count bins is varied from 5 to 15. This result is particularly relevant because the binned statistic is central to the regime-scale interpretation. Its persistence across alternative bin counts reduces the likelihood that the reported high R^2 values arise solely from the predefined 10-bin representation.

Independent comparison with CMEMS provides an observational check on the ERA5 significant-wave-height information used to construct the wind–wave descriptors. Across the 5,771 primary matchups, ERA5 yields an RMSE of 0.631 m, $R^2 = 0.878$, and a Pearson correlation of 0.942 relative to the Sentinel-3A observations. Agreement remains strong when the

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

space–time collocation criteria are tightened: the 50-km/1-h subset yields an RMSE of 0.527 m, $R^2 = 0.919$, and Pearson correlation of 0.962. The overall validation conclusion is also insensitive to reducing repeated CMEMS associations to a single representative matchup per altimeter observation.

These results support the use of ERA5 H_s for bulk regime-scale characterization, but they should not be interpreted as evidence that ERA5 accurately resolves every local wave state during an individual tropical cyclone. Reanalysis uncertainty remains relevant during rapidly evolving and extreme conditions (Hersbach et al., 2020; Wiese et al., 2018). Moreover, the physical framework uses bulk significant wave height and mean wave period rather than the complete directional wave spectrum. Directional wave propagation, swell–wind-sea partitioning, and finite-depth effects are therefore not explicitly represented.

The diagnostic machine-learning benchmark provides a complementary assessment of the information content carried by the environmental variables. Adding local ERA5 wind to tropical-cyclone intensity and radial position produces a substantial reduction in external mismatch-prediction error in all four transfers. Direct bulk wave information provides an additional improvement in BOB→WP, WP→ATL, and BOB→ATL, while the gain is negligible for WP→BOB. WWDI* also improves on the TC–radial reference in all four transfers, but does not consistently outperform either the local wind-context representation or the direct bulk wave variables.

The machine-learning results therefore do not support interpretation of WWDI* as a universally optimal predictive feature. Its value lies instead in providing a compact and physically interpretable summary of part of the environmental information contained in the underlying atmospheric and wave variables. This distinction is important because the purpose of the framework is diagnostic understanding rather than maximization of predictive performance.

The final training-derived adjustment experiment provides further evidence that the diagnosed regimes contain transferable information about systematic signed mismatch, while also showing why the magnitude of the improvement must be interpreted cautiously. Relative to the raw external CYGNSS–IBTrACS mismatch, the WWDI*-conditioned adjustment reduces RMSE by 29.8–40.4 %. A considerable fraction of that reduction, however, is obtained by subtracting a single mean bias estimated from the training basin. Relative to this global training-bias baseline, the additional point-estimate improvement associated with WWDI* conditioning is 13.0 % for WP→BOB, 4.8 % for BOB→WP, 1.7 % for WP→ATL, and 2.2 % for BOB→ATL.

Paired storm-wise bootstrap resampling further distinguishes the strength of this incremental contribution. The WWDI*-conditioned adjustment retains a positive storm-level RMSE advantage over global debiasing for both reciprocal WP–BOB transfers, with the corresponding confidence intervals remaining above zero. In contrast, the bootstrap intervals include zero for both Atlantic evaluations. The full 29.8–40.4 % raw-to-adjusted reduction therefore cannot be attributed exclusively to the WWDI*-conditioned regime correction. Evidence for a distinct regime-dependent contribution is strongest for WP→BOB, remains supported for BOB→WP, and is weaker for transfer to the Atlantic.

Several limitations follow directly from the study design. First, IBTrACS is a storm-scale best-track maximum sustained wind estimate rather than a pointwise measurement of the local wind at the CYGNSS specular location. Consequently, E_{TC} combines CYGNSS retrieval uncertainty with differences arising from storm spatial structure, temporal separation, radial sampling, and reference representativeness (Ruf et al., 2019b; Ricciardulli et al., 2021; Powell and Ruf, 2025). The present

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

framework therefore diagnoses reference-relative mismatch and does not provide a direct decomposition of CYGNSS retrieval error.

Second, the physical descriptors are derived from bulk ERA5 wind, significant wave height, and mean wave period. They do not represent the complete directional wave spectrum or explicitly separate swell from locally forced wind sea. The characteristic wavelength and phase-speed quantities are consequently diagnostic scales rather than complete descriptions of tropical-cyclone wave dynamics.

Third, the amount of development data differs strongly between WP and BOB. This imbalance is most consequential for the machine-learning benchmark. The BOB-2024 internal-validation subset contains only 1,028 observations from three tropical cyclones, compared with 40,122 observations from 28 storms in the WP-2024 holdout. Algorithm selection for BOB-trained configurations is therefore based on a considerably smaller storm population, even though the selected models are subsequently tested on large independent external datasets.

Finally, WWDI* is explicitly supervised because its component weights are derived from training-stage association with E_{TC}. The final regime-conditioned adjustment is likewise calibrated against training-basin CYGNSS–IBTrACS mismatch. Neither should therefore be interpreted as an independently derived geophysical law or as an operational retrieval correction. Future work should examine the identified regime structure using additional pointwise wind and wave observations, particularly at extreme tropical-cyclone wind speeds, and assess whether physically informed environmental dependence can be incorporated into uncertainty-aware GNSS-R retrieval systems while preserving strict storm-level and geographical independence.

6 Conclusions

This study developed a physics-guided framework for diagnosing reference-relative CYGNSS wind mismatch under tropical-cyclone wind–wave conditions using 212,697 unique matchups from 147 storms across the western North Pacific, Bay of Bengal, and North Atlantic. The archived CYGNSS Level-2 wind retrieval was retained unchanged, allowing the analysis to focus on how its mismatch relative to IBTrACS varies with bulk sea-state development, transient environmental response, and storm-relative context.

The two independently trained WWDI* formulations show a consistent component hierarchy. Wave steepness receives the largest supervised weight in both WP- and BOB-based development, followed by young-wave state and transient significant-wave-height response. The derived wave-height departure D_H receives zero weight in both final formulations, and its explicit removal leaves the principal WWDI* transfer results unchanged.

Under four strict storm-independent and observation-independent external transfers, WWDI* remains positively associated with absolute CYGNSS–IBTrACS mismatch. Observation-level Spearman correlations range from 0.219 to 0.381, whereas the corresponding 10-bin regime-scale R^2 values range from 0.639 to 0.753. The stronger bin-level organization indicates that WWDI* is most informative as a descriptor of mismatch-prone storm regimes rather than as a deterministic predictor of individual observations.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Permutation testing and storm-wise bootstrap resampling show that the external regime relationships are substantially stronger than those expected from random pairing and remain evident when uncertainty is evaluated at the cyclone level. Sensitivity experiments further demonstrate that the principal relationships persist under alternative normalization limits, weak-wind regularization values, derived-component specification, and regime-bin counts.

The corrected confounder analysis provides an important qualification to the unconditional transfer relationships. After controlling for ranked tropical-cyclone intensity, ranked storm-relative radius, their quadratic terms, and their interaction, the residual WWDI*–mismatch association is negative in all four transfers, with residual Spearman coefficients from -0.645 to -0.543. The positive unconditional relationship therefore does not establish an independent monotonic or causal effect of wind–wave disequilibrium. Instead, WWDI* identifies a broader storm regime in which tropical-cyclone intensity, radial structure, evolving sea state, and reference-relative mismatch vary jointly.

Independent Sentinel-3A observations support the use of ERA5 significant wave height for bulk regime characterization. The primary 5,771-matchup comparison yields an RMSE of 0.631 m and a Pearson correlation of 0.942, with agreement remaining strong under the stricter 50-km/1-h collocation configuration.

The diagnostic machine-learning benchmark further shows that environmental context contains transferable information about E_{TC}. Local ERA5 wind substantially improves performance relative to tropical-cyclone intensity and radial position alone, while direct bulk wave information provides additional improvement in three of the four external transfers. WWDI* carries transferable information but does not consistently outperform the direct atmospheric and wave-state variables, supporting its use as a compact diagnostic representation rather than as a universally optimal predictive feature.

The training-derived WWDI*-conditioned adjustment reduces raw external RMSE by 29.8–40.4 %. A substantial fraction of this reduction is shared with a simpler global training-bias correction. Relative to that baseline, the additional point-estimate improvement is 1.7–13.0 %, with paired storm-wise bootstrap support for the reciprocal WP–BOB transfers but not for the two Atlantic transfers. The complete raw-to-adjusted improvement should therefore not be attributed solely to WWDI* conditioning.

Overall, the results support WWDI* as a supervised, physically interpretable diagnostic of tropical-cyclone environments associated with systematic changes in reference-relative CYGNSS wind mismatch. It should not be interpreted as a universal geophysical relationship, a causal wind–wave index, or an operational CYGNSS retrieval correction. The framework instead provides a basis for identifying environmental regimes in which CYGNSS reference-relative mismatch becomes systematically larger and for developing future uncertainty-aware GNSS-R analyses using additional independent wind and wave observations.

Code availability. The Python and Jupyter-based analysis code developed for this study is available from the corresponding author upon reasonable request.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Data availability. The processed CYGNSS–ERA5–IBTrACS matchup datasets and independent ERA5–CMEMS significant-wave-height validation datasets underlying this study are publicly available on Zenodo (Masum, 2026). The original CYGNSS, ERA5, IBTrACS, and CMEMS source datasets are available from their respective data providers.

Author contributions. SMIHM developed the study concept and methodological framework, performed the data processing and analysis, developed and implemented the computational workflow, prepared the figures and tables, interpreted the results, and wrote the original manuscript. AWA contributed to scientific discussion, methodological evaluation, and manuscript review. DY provided the primary research supervision, contributed to scientific interpretation and funding acquisition, and reviewed the manuscript. MSI contributed to the interpretation of tropical-cyclone and atmospheric processes and to manuscript review. SSS contributed to data processing and analysis, scientific discussion, and preparation of the original manuscript. All authors reviewed and approved the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors acknowledge NASA and the CYGNSS Science Team for providing the CYGNSS ocean-surface wind observations, the European Centre for Medium-Range Weather Forecasts and the Copernicus Climate Change Service for providing the ERA5 reanalysis, NOAA's National Centers for Environmental Information for maintaining the IBTrACS tropical-cyclone archive, and the Copernicus Marine Service for providing the independent satellite significant-wave-height observations used in this study.

The authors also gratefully acknowledge the support of the Hangzhou International Innovation Institute, Beihang University, Hangzhou 311115, China, in conducting this research.

Financial support. This work was supported by Beihang University of China under Grant KQ24016.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

References

Asgarimehr, M., Zhelavskaya, I., Foti, G., Reich, S., and Wickert, J.: A GNSS-R Geophysical Model Function: Machine Learning for Wind Speed Retrievals, IEEE Geoscience and Remote Sensing Letters, 17, 1333–1337, https://doi.org/10.1109/LGRS.2019.2948566, 2020.

Asharaf, S., Waliser, D. E., Posselt, D. J., Ruf, C. S., Zhang, C., and Putra, A. W.: CYGNSS Ocean Surface Wind Validation in the Tropics, Journal of Atmospheric and Oceanic Technology, 38, 711–724, https://doi.org/10.1175/JTECH-D-20-0079.1, 2021.

Balasubramaniam, R. and Ruf, C. S.: Azimuthal Dependence of GNSS-R Scattering Cross-Section in Hurricanes, Journal of Geophysical Research: Oceans, 125, e2020JC016 167, https://doi.org/10.1029/2020JC016167, 2020.

Bethel, B. J., Sun, W., Dong, C., and Wang, D.: Forecasting Hurricane-Forced Significant Wave Heights Using a Long Short-Term Memory Network in the Caribbean Sea, Ocean Science, 18, 419–436, https://doi.org/10.5194/os-18-419-2022, 2022.

Bu, J., Liu, X., Wang, Q., Li, L., Zuo, X., Yu, K., and Huang, W.: Ocean Remote Sensing Using Spaceborne GNSS-Reflectometry: A Review, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 13 047–13 076, https://doi.org/10.1109/JSTARS.2024.3418429, 2024.

Clarizia, M. P. and Ruf, C. S.: Wind Speed Retrieval Algorithm for the Cyclone Global Navigation Satellite System (CYGNSS) Mission, IEEE Transactions on Geoscience and Remote Sensing, 54, 4419–4432, https://doi.org/10.1109/TGRS.2016.2541343, 2016.

Clarizia, M. P. and Ruf, C. S.: Bayesian Wind Speed Estimation Conditioned on Significant Wave Height for GNSS-R Ocean Observations, Journal of Atmospheric and Oceanic Technology, 34, 1193–1202, https://doi.org/10.1175/JTECH-D-16-0196.1, 2017.

Clarizia, M. P. and Ruf, C. S.: Statistical Derivation of Wind Speeds from CYGNSS Data, IEEE Transactions on Geoscience and Remote Sensing, 58, 3955–3964, https://doi.org/10.1109/TGRS.2019.2959715, 2020.

Copernicus Climate Change Service (C3S): ERA5 hourly data on single levels from 1940 to present, Climate Data Store [data set], https://doi.org/10.24381/cds.adbb2d47, 2018.

Emanuel, K.: Tropical Cyclones, Annual Review of Earth and Planetary Sciences, 31, 75–104, https://doi.org/10.1146/annurev.earth.31.100901.141259, 2003.

E.U. Copernicus Marine Service Information: Global Ocean L3 Significant Wave Height from Near-Real-Time Satellite Measurements, Copernicus Marine Service [data set], https://doi.org/10.48670/moi-00179, 2026.

Foti, G., Gommenginger, C., Jales, P., Unwin, M., Shaw, A., Robertson, C., and Rosello, J.: Spaceborne GNSS-Reflectometry for Ocean Winds: First Results from the UK TechDemoSat-1 Mission, Geophysical Research Letters, 42, 5435–5441, https://doi.org/10.1002/2015GL064204, 2015.

Gahtan, J., Knapp, K. R., Schreck, C. J., Diamond, H. J., Kossin, J. P., and Kruk, M. C.: International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4r01, NOAA National Centers for Environmental Information [data set], https://doi.org/10.25921/82ty-9e16, 2024.

Gleason, S., Ruf, C. S., Clarizia, M. P., and O'Brien, A. J.: Calibration and Unwrapping of the Normalized Scattering Cross Section for the Cyclone Global Navigation Satellite System, IEEE Transactions on Geoscience and Remote Sensing, 54, 2495–2509, https://doi.org/10.1109/TGRS.2015.2502245, 2016.

Gleason, S., Ruf, C. S., O'Brien, A. J., and McKague, D. S.: The CYGNSS Level 1 Calibration Algorithm and Error Analysis Based on On-Orbit Measurements, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12, 37–49, https://doi.org/10.1109/JSTARS.2018.2832981, 2019.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., Munoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., et al.: The ERA5 Global Reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Knaff, J. A., Sampson, C. R., Kucas, M. E., Slocum, C. J., Brennan, M. J., Meissner, T., Ricciardulli, L., Mouche, A., Reul, N., Morris, M., Chirokova, G., and Caroff, P.: Estimating Tropical Cyclone Surface Winds: Current Status, Emerging Technologies, Historical Evolution, and a Look to the Future, Tropical Cyclone Research and Review, 10, 125–150, https://doi.org/10.1016/j.tcrr.2021.09.002, 2021.

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., and Neumann, C. J.: The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying Tropical Cyclone Data, Bulletin of the American Meteorological Society, 91, 363–376, https://doi.org/10.1175/2009BAMS2755.1, 2010.

Li, Z., Guo, F., Zhang, X., Yang, W., Zhu, Y., Zhang, Z., Chen, P., Liu, H., and Song, X.: Global Sea Surface Wind Speed Retrieval Using FengYun-3F Multi-GNSS Reflections, Advances in Space Research, 77, 11 556–11 565, https://doi.org/10.1016/j.asr.2026.03.082, 2026.

Lima, M. M., Gouveia, C. M., and Trigo, R. M.: Upper-Ocean Response to the Passage of Tropical Cyclones in the Azores Region, Ocean Science, 18, 1419–1430, https://doi.org/10.5194/os-18-1419-2022, 2022.

Masum, S. M. I. H.: Processed CYGNSS–ERA5–IBTrACS Tropical-Cyclone Wind–Wave Disequilibrium Dataset (2022–2024), Zenodo [data set], https://doi.org/10.5281/zenodo.22168565, 2026.

Mayers, D. R., Ruf, C. S., and Warnock, A. M.: CYGNSS Storm-Centric Tropical Cyclone Gridded Wind Speed Product, Journal of Applied Meteorology and Climatology, 62, 329–339, https://doi.org/10.1175/JAMC-D-22-0054.1, 2023.

NOAA/NESDIS/STAR/SOCD: NOAA CYGNSS Level 2 Science Wind Speed 25-km Product Version 1.2, PO.DAAC [data set], https://doi.org/10.5067/CYGNN-22512, 2022.

Pascual, D., Clarizia, M. P., and Ruf, C. S.: Improved CYGNSS Wind Speed Retrieval Using Significant Wave Height Correction, Remote Sensing, 13, 4313, https://doi.org/10.3390/rs13214313, 2021.

Powell, C. E. and Ruf, C. S.: A Method for Estimating and Partitioning the Covariance of Wind Speed Observation Error from GNSS Reflectometry, Journal of Atmospheric and Oceanic Technology, 42, 1027–1040, https://doi.org/10.1175/JTECH-D-24-0031.1, 2025.

Prakash, K. R., Nigam, T., and Pant, V.: Estimation of Oceanic Subsurface Mixing under a Severe Cyclonic Storm Using a Coupled Atmosphere–Ocean–Wave Model, Ocean Science, 14, 259–272, https://doi.org/10.5194/os-14-259-2018, 2018.

Qiao, X. and Huang, W.: Ocean Surface Wind Speed Estimation from GNSS-R Data Using Physics-Informed Attention-Aided Convolutional Neural Network, IEEE Transactions on Geoscience and Remote Sensing, 63, 1–16, https://doi.org/10.1109/TGRS.2025.3594633, 2025.

Rascle, N. and Ardhuin, F.: A Global Wave Parameter Database for Geophysical Applications. Part 2: Model Validation with Improved Source Term Parameterization, Ocean Modelling, 70, 174–188, https://doi.org/10.1016/j.ocemod.2012.12.001, 2013.

Ricciardulli, L., Mears, C., Manaster, A., and Meissner, T.: Assessment of CYGNSS Wind Speed Retrievals in Tropical Cyclones, Remote Sensing, 13, 5110, https://doi.org/10.3390/rs13245110, 2021.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schroder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F.: Cross-validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure, Ecography, 40, 913–929, https://doi.org/10.1111/ecog.02881, 2016.

Ruf, C., Al-Khaldi, M., Asharaf, S., Balasubramaniam, R., McKague, D., Pascual, D., Russel, A., Twigg, D., and Warnock, A.: Characterization of CYGNSS Ocean Surface Wind Speed Products, Remote Sensing, 16, 4341, https://doi.org/10.3390/rs16224341, 2024.

Ruf, C. S. and Balasubramaniam, R.: Development of the CYGNSS Geophysical Model Function for Wind Speed, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12, 66–77, https://doi.org/10.1109/JSTARS.2018.2833075, 2019.

https://doi.org/10.5194/egusphere-2026-5300
Preprint. Discussion started: 11 September 2026
© Author(s) 2026. CC BY 4.0 License.

Ruf, C. S., Atlas, R., Chang, P. S., Clarizia, M. P., Garrison, J. L., Gleason, S., Katzberg, S. J., Jelenak, Z., Johnson, J. T., Majumdar, S. J., O'Brien, A., Posselt, D. J., Ridley, A. J., Rose, R. J., and Zavorotny, V. U.: New Ocean Winds Satellite Mission to Probe Hurricanes and Tropical Convection, Bulletin of the American Meteorological Society, 97, 385–395, https://doi.org/10.1175/BAMS-D-14-00218.1, 2016.

Ruf, C. S., Asharaf, S., Balasubramaniam, R., Gleason, S., Lang, T., McKague, D., Twigg, D., and Waliser, D.: In-Orbit Performance of the Constellation of CYGNSS Hurricane Satellites, Bulletin of the American Meteorological Society, 100, 2009–2023, https://doi.org/10.1175/BAMS-D-18-0337.1, 2019a.

Ruf, C. S., Gleason, S., and McKague, D. S.: Assessment of CYGNSS Wind Speed Retrieval Uncertainty, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12, 87–97, https://doi.org/10.1109/JSTARS.2018.2825948, 2019b.

Schulz-Stellenfleth, J. and Staneva, J.: A Multi-collocation Method for Coastal Zone Observations with Applications to Sentinel-3A Altimeter Wave Height Data, Ocean Science, 15, 249–268, https://doi.org/10.5194/os-15-249-2019, 2019.

Soisuvarn, S., Jelenak, Z., Said, F., Chang, P. S., and Egido, A.: The GNSS Reflectometry Response to the Ocean Surface Winds and Waves, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9, 4678–4699, https://doi.org/10.1109/JSTARS.2016.2602703, 2016.

Varoquaux, G.: Cross-validation Failure: Small Sample Sizes Lead to Large Error Bars, NeuroImage, 180, 68–77, https://doi.org/10.1016/j.neuroimage.2017.06.061, 2018.

Wiese, A., Staneva, J., Schulz-Stellenfleth, J., Behrens, A., Fenoglio-Marc, L., and Bidlot, J.-R.: Synergy of Wind Wave Model Simulations and Satellite Observations during Extreme Events, Ocean Science, 14, 1503–1521, https://doi.org/10.5194/os-14-1503-2018, 2018.