



SCS-TCWave: an event-based hindcast dataset of typhoon waves and reconstructed wind fields for the South China Sea (1979–2025)

Feifan Fang^{1,2}, Yifeng Yang^{1,2}, Peng Jin^{1,2}, and Binzhen Zhou^{1,2}

¹State Key Laboratory of Subtropical Building and Urban Science, South China University of Technology, Guangzhou 510641, China

²School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China

Correspondence: Binzhen Zhou (zhoubinzhen@scut.edu.cn)

Abstract. Extreme-wave estimates for the South China Sea (SCS) depend on the peak sea states of tropical cyclones, yet reanalysis winds smooth the storm core and no openly archived, event-resolved typhoon wave record exists for the basin. We present SCS-TCWave, an event-based hindcast of the 480 tropical cyclones that reached tropical storm intensity inside the SCS during 1979–2025. Each event is forced by 10-m winds reconstructed from best-track records (a Holland-type vortex with motion asymmetry and far-field gale-radius anchoring, blended into ERA5) and simulated with SWAN on one-way nested $0.25^\circ/0.05^\circ$ grids under a frozen, fully audited configuration. Each event ships three CF-1.8 NetCDF layers: hourly wave fields, the reconstructed wind forcing, and an extreme-value catalogue of per-cell event maxima. Validation against 64,529 calibrated altimeter super-observations from 187 events shows near-zero bias in background sea states and a single-signed core overestimation that grows with intensity (+1.15 to +2.99 m within 100 km of the centre). A distance-binned multiplicative correction annex with bootstrap confidence intervals centres the per-event peak distribution: the median relative error of per-event maxima moves from +9.5 % raw to −1.4 % corrected, with 78 % of events within ± 25 %. At the 9.4 m Pearl River estuary anchor during Typhoon Mangkhut, the product stays within ± 25 % where an identically configured ERA5-forced run does not; beyond 400 km ERA5-based products remain preferable, and this distance-based usage rule ships with the data. Method tiers, quality flags, the frozen configuration, and the full calibration record accompany the archive (<https://doi.org/10.57760/sciencedb.011ul>; Fang et al., 2026).

1 Introduction

Typhoon waves dominate the extreme wave climate of the South China Sea (SCS). The basin hosts one of the highest concentrations of tropical cyclone activity and coastal exposure in the world, and the design of its offshore wind farms, platforms, ports, and coastal defences rests on estimates of extreme significant wave height (Xu et al., 2017; Shao et al., 2019; Sun et al., 2020). Extreme-value methods for this purpose consume a specific quantity: the peak sea state of each storm event (Jonathan and Ewans, 2013). Long, homogeneous, event-resolved wave records are therefore the raw material of typhoon wave hazard assessment, and they are exactly what the region lacks.

Observations alone cannot supply them. Buoy records in the SCS are sparse, short, and largely held by the institutions that operate them (Xu et al., 2017; Zhai et al., 2023). Satellite altimetry provides accurate, well-calibrated wave heights (Ribal and



Table 1. Attribute comparison of wave products covering the South China Sea. The availability column follows each product’s own data availability statement or documentation.

Product	Period	Resolution	TC-core wind reconstruction	Event-resolved	Availability	Per-event extremes
ERA5 ocean waves (Hersbach et al., 2020)	1940–present	0.5°	no	no	open (Copernicus CDS)	not provided
CAWCR wave hindcast (Smith et al., 2021)	1979–present	0.4° global 10′/4′ nests	no	no	open (CSIRO DAP)	not provided
SCS coupled hindcast (Bagnasco et al., 2025)	1970–2022	unstructured, to 0.008° nearshore	no	no	on request ^a	not provided
SCS wave-power hindcast (Sun et al., 2020)	1991–2016	0.067°	no	no	none stated ^a	not provided
SCS-TCWave (this work)	1979–2025	0.25°/0.05° waves 0.1° wind	yes (Sect. 2.2)	yes (480 events)	open (Science Data Bank, CC BY 4.0)	per-event maxima catalogue

^a As stated in the respective article: “on request” quotes the article’s data availability statement; “none stated” means no data availability statement or public archive was located (checked 2 September 2026).

25 Young, 2019; Dodet et al., 2020), but a nadir track crosses a moving storm core only rarely, so altimeters characterise the wave climate while systematically undersampling the eyewall peaks that govern design (Tamizi and Young, 2020). Global model products fill the temporal gaps and introduce a different problem. Reanalysis winds smooth the tropical cyclone core: ERA5 underestimates observed annual-maximum winds over the SCS by up to 38 % (Zhai et al., 2023), and the wave extremes forced by those winds are biased low in proportion. ERA5 also assimilates altimeter wave heights (Hersbach et al., 2020), which keeps
30 its background sea states close to the very observations later used to evaluate it; agreement at ordinary conditions therefore says little about the unassimilated, undersampled tail where typhoon design conditions live.

Regional hindcasting addresses the core problem directly, and the SCS has a substantial hindcast literature, including multi-decadal simulations at high nearshore resolution (Bagnasco et al., 2025; Sun et al., 2020). The resulting fields, however, remain with their producers: availability statements of the “data available on request” type are the norm, and empirical studies across
35 disciplines show that such statements are honoured only in a small minority of cases (Stodden et al., 2018; Tedersoo et al., 2021; Gabelica et al., 2022). For the SCS there is, to our knowledge, no openly deposited, long-record typhoon wave product whose wind forcing reconstructs the storm core. Each group that needs one builds one, calibrates it against whatever observations it holds, and the community’s collective effort accumulates in inaccessible form. An event-based, openly archived hurricane surge-and-wave database for the United States Atlantic and Gulf coasts (Deb et al., 2026), following the precedent of open
40 storm surge archives (Needham and Keim, 2012), has recently demonstrated the value of the product form; the western North Pacific side has no counterpart. Table 1 summarises the attributes of the main wave products available for the basin alongside SCS-TCWave.



SCS-TCWave is built to close this gap. It is an event-based hindcast of 480 tropical cyclones affecting the SCS over 1979–2025, forced by winds reconstructed from best-track records with a parametric vortex blended into the ERA5 background, simulated with SWAN on nested grids, and released in full: hourly wave fields, the reconstructed wind forcing, and a per-event extreme-value catalogue, together with the frozen model configuration, the calibration record, and a bias-correction annex with confidence intervals. Validation is two-tiered by design. Basin-wide skill is measured against 64,529 altimeter matchups, and extreme skill is measured at fixed points that recorded eyewall passages directly, with an ERA5-forced control run of identical configuration quantifying the increment attributable to the wind reconstruction throughout. The dataset ships uncorrected, with every known limitation documented and a distance-based usage rule that states where its skill lies and where existing products serve better.

The remainder of this article describes the data and methods (Sect. 2), the validation and the correction annex (Sect. 3), the uncertainty and known limitations (Sect. 4), and usage guidance (Sect. 5); the data availability statement closes the article.

2 Data and methods

2.1 Scope and event selection

SCS-TCWave is an event-based hindcast: each product unit is one tropical cyclone (TC) simulated over its own time window, rather than a continuous multi-decadal run. This design follows the needs of extreme-value and engineering applications, which consume per-event process fields and per-event maxima, and it concentrates computational effort on the sea states that global products resolve least well.

Candidate events were drawn from the China Meteorological Administration (CMA) best-track dataset (Ying et al., 2014; Lu et al., 2021), cross-referenced with IBTrACS v04r01 (Knapp et al., 2010) for wind structure parameters. A TC enters the catalogue if its best-track centre lies inside the South China Sea selection box (105–121° E, 5–24° N; Fig. 1) at tropical storm intensity or above ($V_{\max} \geq 17.2 \text{ m s}^{-1}$) at any record during 1979–2025. This yields 480 events. For each event the simulation window extends from 48 h before the first qualifying record to 24 h after the last one, clipped to the track lifetime, giving a mean window of 4.7 days. The 48 h lead is set by a spin-up sensitivity test on six storms rerun with 24–96 h leads (Table B1): from 48 h upward the event-maximum H_s changes by at most 0.5 %, peak timing by at most 1 h, and peak location by at most one outer-grid cell, while a 24 h lead leaves the 200–400 km ring about 1.5 m short of far-field swell in the strongest case; 48 h is the shortest converged lead. Figure 2 summarises the full production chain from these inputs to the released product.

2.2 Wind field reconstruction

Surface winds are the dominant source of skill and of error in TC wave hindcasting (Cavaleri, 2009), so the forcing is reconstructed rather than taken from a reanalysis, a standard practice in tropical cyclone wave modelling (e.g., Phadke et al., 2003). Each wind field combines a parametric vortex (e.g., Vickery et al., 2009) with an ERA5 background (Hersbach et al., 2020).

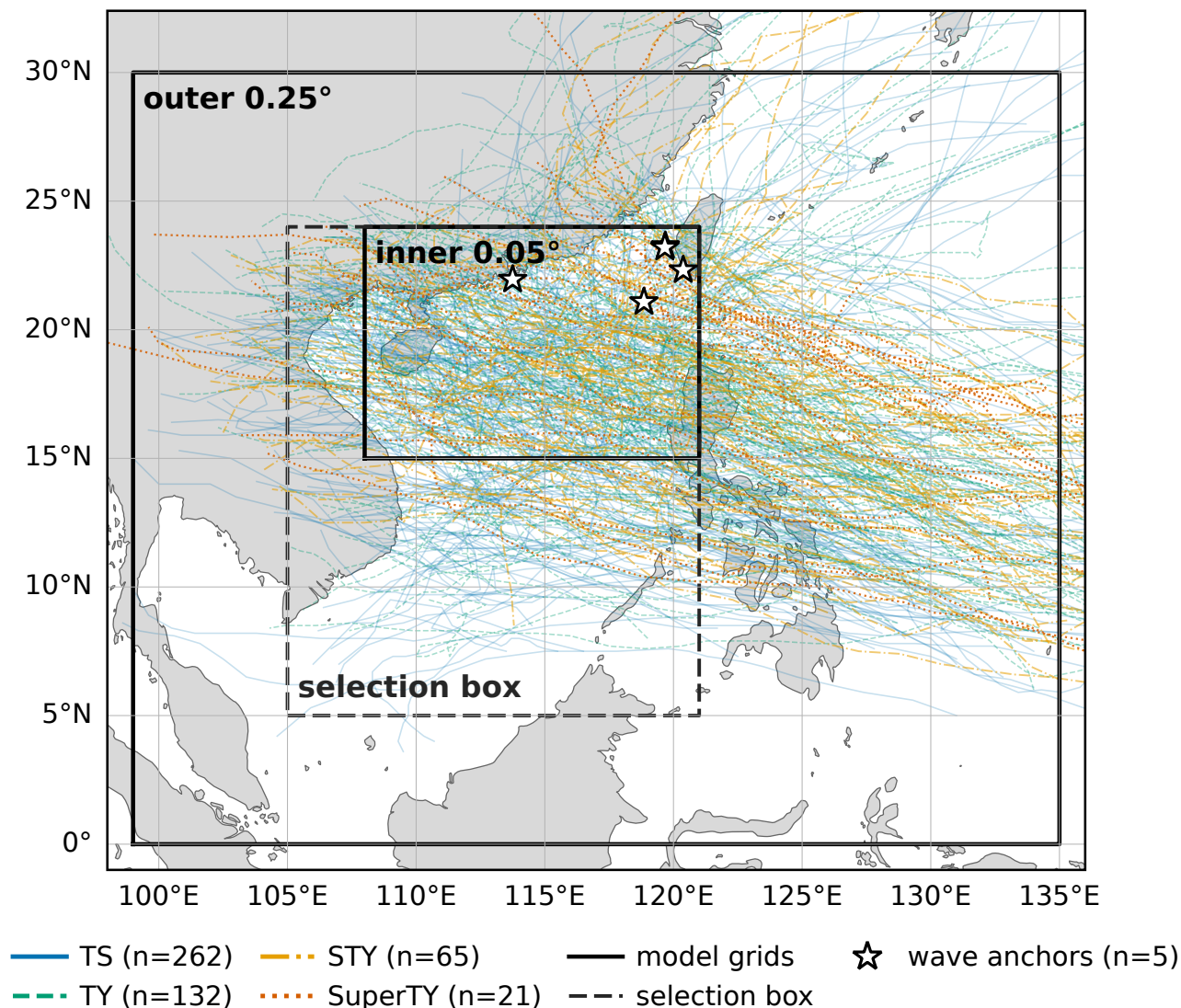


Figure 1. Study domain and the 480 hindcast events (1979–2025). Solid black frames mark the outer 0.25° SWAN grid (99–135° E, 0–30° N) and the inner 0.05° grid (108–121° E, 15–24° N); the dashed frame marks the event-selection box (105–121° E, 5–24° N), and stars mark the five fixed-point wave anchors of the validation (Sect. 3.5). Best tracks are distinguished by colour and line style according to the peak CMA intensity class inside the selection box: tropical storm (TS, $n = 262$), typhoon (TY, $n = 132$), severe typhoon (STY, $n = 65$), and super typhoon (SuperTY, $n = 21$). Coastline data: Natural Earth (public domain).

The vortex follows the Holland (1980) radial profile, with the shape parameter diagnosed from the recorded wind–pressure pair (Holland, 2008; Holland et al., 2010; Vickery and Wadhera, 2008). Its inputs per best-track time step are the centre position, the central pressure P_c , the maximum wind speed $V_{\max}$ (a 2 min mean in the CMA records; agency averaging conventions

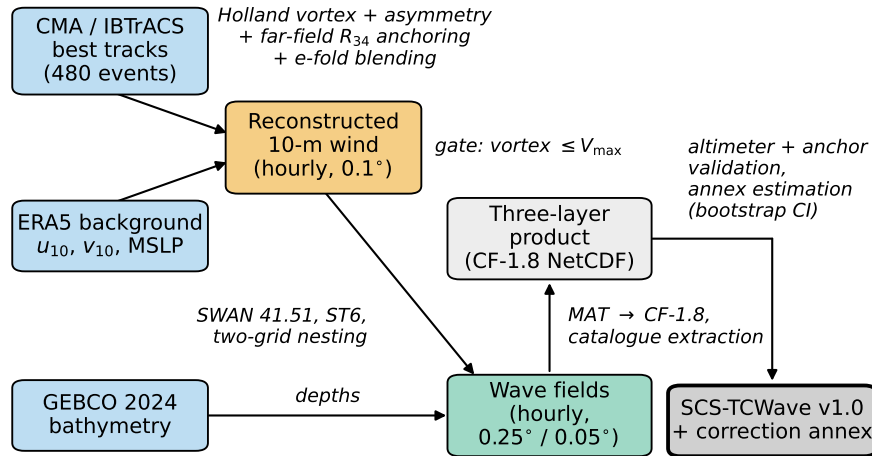


Figure 2. The SCS-TCWave production pipeline. Boxes are data states (blue: external inputs; orange: wind stage; green: wave stage; grey: released products); arrows carry the actions that transform them. Every event passes the reconstruction gate before simulation, and the frozen configuration (v1.0-calib) fixes all physics choices and parameters across the catalogue.

differ, Harper et al., 2010), and the radius of maximum wind $R_{\max}$. $R_{\max}$ is taken from the IBTrACS USA_RMW record where available and estimated from an intensity–latitude climatology (Willoughby and Rahn, 2004) otherwise; this distinction contributes to the method tiers described in Sect. 2.5. Before profile fitting, the translation contribution is removed from the recorded maximum wind, $V_{\text{fit}} = V_{\max} - k V_t$ with $k = 0.6$ (Lin and Chavas, 2012), because best-track $V_{\max}$ already contains the translation asymmetry. Motion asymmetry is then reintroduced with the Georgiou (1985) formulation, which folds the translation effect into the gradient-wind balance and confines it to the inner core. Direct verification of the reconstructed winds against coastal station records (Sect. 2.4) localised the remaining profile error to the far field, not the core, consistent with the outer wind field varying largely independently of the inner core (Chavas et al., 2015). The frozen configuration therefore anchors the outer profile to the observed gale radius: beyond 200 km the Holland profile is blended, with a weight of 0.5 applied over a 50 km ramp, toward the outer exponential of Willoughby et al. (2006), whose decay length is fitted to each best-track record by requiring the profile to pass through 17.49 m s^{-1} (34 kt) at the IBTrACS R_{34} radius (quadrant mean). Records without R_{34} fall back to the climatological decay length; this fallback is the second component of the method tiers. The composite vortex merges into the ERA5 background through an e-folding weight function beyond a core radius of 300 km, with an e-folding length of 150 km, and a wind speed reduction factor of 0.85 converts the gradient-level profile to 10-m winds (Kepert, 2001).

Every reconstructed field passes an automated gate before it reaches the wave model: the vortex component may not exceed the recorded $V_{\max}$ while the centre is at sea and the storm is at tropical storm intensity or above. Gate violations during weak



or post-landfall stages, when the monsoon background legitimately exceeds a decaying vortex, are logged as warnings. In 98 events the wind–pressure records are mutually inconsistent: the Holland shape parameter reaches its physical lower bound and the vortex amplitude is renormalised down to the recorded $V_{\max}$. These events carry the quality flag described in Sect. 2.5.

2.3 Wave model configuration

Waves are simulated with SWAN 41.51 (Booij et al., 1999) on two one-way nested regular grids: an outer grid at 0.25° covering $99\text{--}135^\circ$ E, $0\text{--}30^\circ$ N, whose western Pacific buffer supplies swell entering through the Luzon Strait, and an inner grid at 0.05° covering the northern South China Sea ($108\text{--}121^\circ$ E, $15\text{--}24^\circ$ N); a single nested wave model suffices at these basin scales. Bathymetry is GEBCO 2024 (GEBCO Compilation Group, 2024). The spectral space is discretised into 36 directions and 32 frequencies spanning $0.0345\text{--}0.6$ Hz.

Source terms use the ST6 physics package (Rogers et al., 2012; Zieger et al., 2015; Liu et al., 2019), the observation-based package shared with WAVEWATCH III, with the wind-input drag scaling reduced to 0.90, the calibrated response to the wave-growth overshoot at high winds identified in the two-sided validation (Sect. 2.4), consistent with the observed saturation of the surface drag coefficient at extreme wind speeds (Powell et al., 2003; Donelan et al., 2004; Hwang, 2011; Holthuijsen et al., 2012). Negative wind input and the observation-based swell dissipation of Zieger et al. (2015) are active. Depth-induced breaking (default Battjes–Janssen formulation), quadruplets, and Madsen bottom friction ($k_n = 0.05$ m) are active on both grids; triad interactions are not. Propagation uses the first-order BSBT scheme in nonstationary mode with two iterations per time step of 20 min (outer) and 10 min (inner). Each event runs as an independent serial process; the full catalogue was produced by process-level parallelism (8 concurrent events) on a single workstation, with per-event integrity checks and automatic retry. The complete configuration is frozen in versioned form (v1.0-calib) and distributed with the dataset.

2.4 Calibration

The frozen configuration is the outcome of a staged calibration against independent observations, conducted on four storms chosen for their altimeter core coverage (Mangkhet 2018, Hato 2017, Krathon 2024, Yinxing 2024). The stages were: (i) single-factor scans of drag cap, blending radius, and wind reduction factor; (ii) direct verification of the reconstructed winds against coastal station records digitised from Hong Kong Observatory tropical cyclone reports, which localised the remaining error to quadrant asymmetry and far-field decay; and (iii) a combined wind-side and wave-side adjustment (far-field gale-radius anchoring, translation deduction, ST6 input scaling) selected by ablation. The full scan matrix (85 complete model chains) and every decision are distributed with the archive as the calibration record, so each parameter choice can be traced to the evidence that motivated it. A pure-ERA5-forced control run of the calibration storms is retained throughout as a reference configuration.

2.5 Production, method tiers, and quality flags

The catalogue carries two orthogonal quality dimensions. *Method tier* records the wind-input provenance: tier 1 events (239 storms, predominantly 2002–2025) use observed vortex structure ($R_{\max}$ and R_{34}); tier 2 events (241 storms, predominantly



Table 2. Data records. Every NetCDF file is CF-1.8 with standard names, coordinate attributes, and global provenance attributes (title, institution, source, history, references, version); the archive licence is CC BY 4.0. Grid shapes are fixed across events; the time dimension t follows each event window (hourly).

File (per event <id>/)	Grid	Dimensions	Variables (units)
<id>_waves_inner.nc	inner 0.05° (108–121° E, 15–24° N)	$t \times 181 \times 261$	hs (m), tp (s), tm01 (s), dir (degree), depth (m), u10, v10 (m s^{-1})
<id>_waves_outer.nc	outer 0.25° (99–135° E, 0–30° N)	$t \times 121 \times 145$	as above
<id>_wind.nc	0.1° (99–135° E, 0–30° N)	$t \times 301 \times 361$	u10, v10 (m s^{-1})
<id>_catalog.nc	inner grid	181×261	hs_max (m), peak_hour (h from window start), wind_max (m s^{-1})
Repository root files:			
event_catalog_index.csv	one row per event: window, peak H_s and its location and time, maximum wind, method_tier, qc_flag		
CORRECTION_ANNEX.csv	multiplicative factors per distance ring with bootstrap 95 % CI and sample sizes (Table 6)		
MANIFEST_v1.0.csv	SHA-256 checksum of every file		
documentation set	quick-start README, product documentation (PRODUCT_DOCUMENTATION_v1.0.md), calibration record (CALIBRATION_RECORD_v1.0.md), data dictionary (DATA_DICTIONARY_v1.0.md)		
configuration snapshots	frozen v1.0-calib wind and SWAN configurations		
scs_tcwave_usage_example.ipynb	worked notebook: reading the three layers, applying the annex by distance ring		

1979–2001) use empirical estimates. A cross-era consistency experiment, in which 30 tier-1 storms were re-simulated with the tier-2 method (27 completed pairs; the reconstruction gate stopped 3, an expected outcome of the degraded pathway), found no systematic offset in per-event peak significant wave height (+1.0 %, 95 % CI $[-1.7, +4.2]$ %) and an added per-event scatter of 7.8 % (SD), which is propagated into the tier-2 uncertainty. *Quality flag* marks 98 events (marginal) whose vortex reconstruction required amplitude clamping (Sect. 2.2); retaining these events with an explicit flag avoids the selection bias that outright exclusion of weak-vortex, monsoon-season storms would introduce.

130 2.6 Dataset structure

Each event ships three product layers in CF-1.8 NetCDF (Table 2): (i) hourly fields of significant wave height, peak and mean period, mean direction, depth, and forcing wind on both grids; (ii) the reconstructed 10-m wind forcing on a 0.1° grid; and (iii) an event catalogue entry containing the per-grid-point event maximum, its timing, and the footprint of maximum significant wave height. A global index, the frozen configuration, the calibration record, and the bias-correction annex (Sect. 3) accompany the archive. The dataset itself is uncorrected model output; the annex provides optional multiplicative correction factors with bootstrap confidence intervals for extreme-value applications.



3 Validation

3.1 Verification framework

The dataset is verified at two complementary levels. Basin-scale skill is assessed against the calibrated multi-mission altimeter database of Ribal and Young (2019), built on established multi-mission calibration and cross-validation practice (Young et al., 2017; Ribal and Young, 2020), which samples the full product domain but rarely crosses a storm core at peak intensity. Extreme-sea-state skill is therefore assessed separately against fixed-point anchors: coastal and open-water wave measurements recorded during direct TC passages, drawn from the Pearl River estuary buoy record published for Typhoon Mangkhut (Luo et al., 2021) and from Central Weather Administration (CWA) buoys around Taiwan for Typhoons Hato, Mangkhut, and Krathon. A pure-ERA5-forced simulation of the anchor storms, run on the identical grid and physics, serves as the reference configuration throughout. SCS-TCWave does not assimilate altimeter wave observations, and its configuration was frozen before the production runs: the 52 first-batch storms outside the calibration set reproduce the calibration error structure at eight times the calibration sample size. ERA5 assimilates altimeter significant wave heights, so the same comparison is partly circular for the reference and is read with that asymmetry in mind (Hersbach et al., 2020).

3.2 Wind product verification

Wave skill is inherited from the forcing, so the wind product is verified before the waves. Across the 187 matchup storms, 64,507 altimeter wind super-observations pair with the reconstruction and with ERA5 on the same geometry as the wave validation (Fig. 3). Within 200 km of the centre the reconstruction stands above the censored observations ($+4.60 \text{ m s}^{-1}$ at 0–100 km, $+2.02$ at 100–200 km) where ERA5 stands below or near them (-2.92 and -0.14); beyond 200 km the two forcings are close (-1.11 versus -0.33 at 200–400 km, -0.93 versus -0.54 beyond 400 km). Altimeter winds saturate near 25 m s^{-1} (280 of the 64,507 pairs, 59 of them in the core ring, are excluded); above that limit the verification rests on the wave-side comparison.

3.3 Basin-scale skill against altimetry

Across 187 directly sampled events, 64,529 altimeter super-observations (1-Hz records clustered into 25 km along-track segments, following standard altimeter validation practice; Queffelecoulou, 2004) pair with the simulated fields within 30 min (Fig. 4). Figure 5 illustrates the sampling geometry and the core comparison for Typhoon Mangkhut (2018). Background sea states dominate the sample: below a model wind of 15 m s^{-1} (61,846 super-observations) the comparison gives a bias of -0.019 m , an RMSE of 0.894 m , and a correlation of 0.753 .

Beyond 200 km from the storm centre the fields are essentially unbiased ($+0.085 \text{ m}$ for 200–400 km, -0.030 m beyond 400 km; Table 3). Within the core, the bias grows monotonically with storm intensity, from $+1.15 \text{ m}$ for tropical storms to $+2.99 \text{ m}$ for super typhoons in the 0–100 km ring (Table 4; Fig. 6a). This regular, single-signed structure is the residual of the frozen configuration (Sect. 2.4); its concentration in a known distance band and its monotonicity are what make the correction

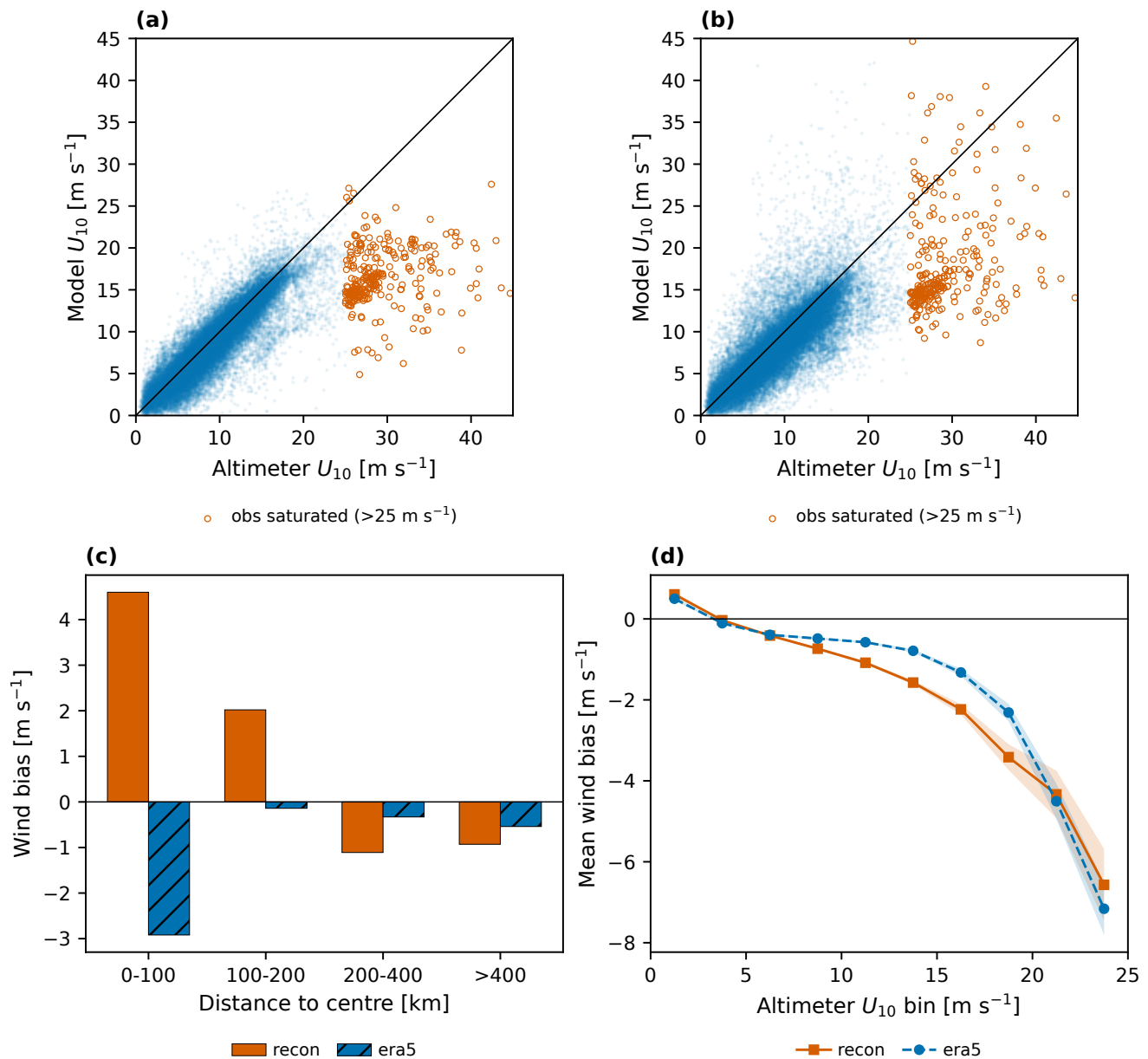


Figure 3. Verification of the wind forcing against calibrated altimeter winds across the 187 matchup storms (64,507 pairs): modelled versus observed 10-m wind speed for (a) the ERA5 background and (b) the reconstruction, with observations above the altimeter saturation limit ($> 25 \text{ m s}^{-1}$) shown hollow and excluded from the statistics; (c) wind bias by distance ring over the unsaturated pairs; and (d) mean wind bias by observed wind-speed bin (2.5 m s^{-1} bins, shading: 95 % confidence interval of the bin mean).

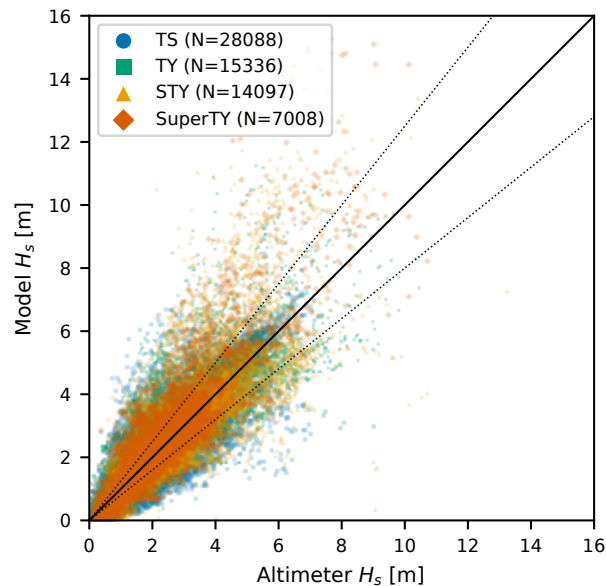


Figure 4. Altimeter validation of significant wave height: 64,529 super-observations from the 187 directly sampled events, by intensity class (TS $N = 28,088$; TY $N = 15,336$; STY $N = 14,097$; SuperTY $N = 7,008$). The solid line is 1:1; dotted lines mark model-to-altimeter ratios of 1.25 and 0.8. Stratified statistics are given in Tables 3 and 4.

annex possible. The same stratification, resolved by distance ring, wind-speed band, intensity class, and storm quadrant, is distributed with the dataset.

170 3.4 Per-event peak skill and the correction annex

Extreme-value applications consume one number per event, the process maximum of significant wave height, so peak skill is reported separately (Table 5; Fig. 7). An event enters this comparison when its matchup set holds at least 30 super-observations, the threshold for the sampled maximum to stand in for the event maximum; 185 of the 187 matchup events qualify. For these 185 events, the raw median relative error of the per-event maximum is +9.5 %, with 111 of 185 events (60 %) within ± 25 %.

175 The overestimation concentrates where the core bias sits: severe typhoons show a raw median of +33.6 %.

The correction annex converts this systematic component into an optional, transparent adjustment. The 185 altimeter-sampled events are split into an estimation set, the 56 first-batch storms whose 24,319 super-observations set the factors, and a validation set of the remaining 129 events. The correction multiplies each value by its ring factor, $H_{s,\text{corr}} = f_r H_{s,\text{raw}}$, with f_r estimated as the ratio of summed observed to summed modelled heights over the ring's super-observations in the estimation set. The 95 % confidence intervals come from a storm-level bootstrap that resamples whole storms (2,000 draws; Table 6): 0.704 [0.661, 0.752] for 0–100 km, 0.836 [0.792, 0.877] for 100–200 km, 0.971 [0.915, 1.023] for 200–400 km, and 1.050 [1.013, 1.095] beyond 400 km; the 200–400 km interval spans unity, so no correction is established for that ring at the

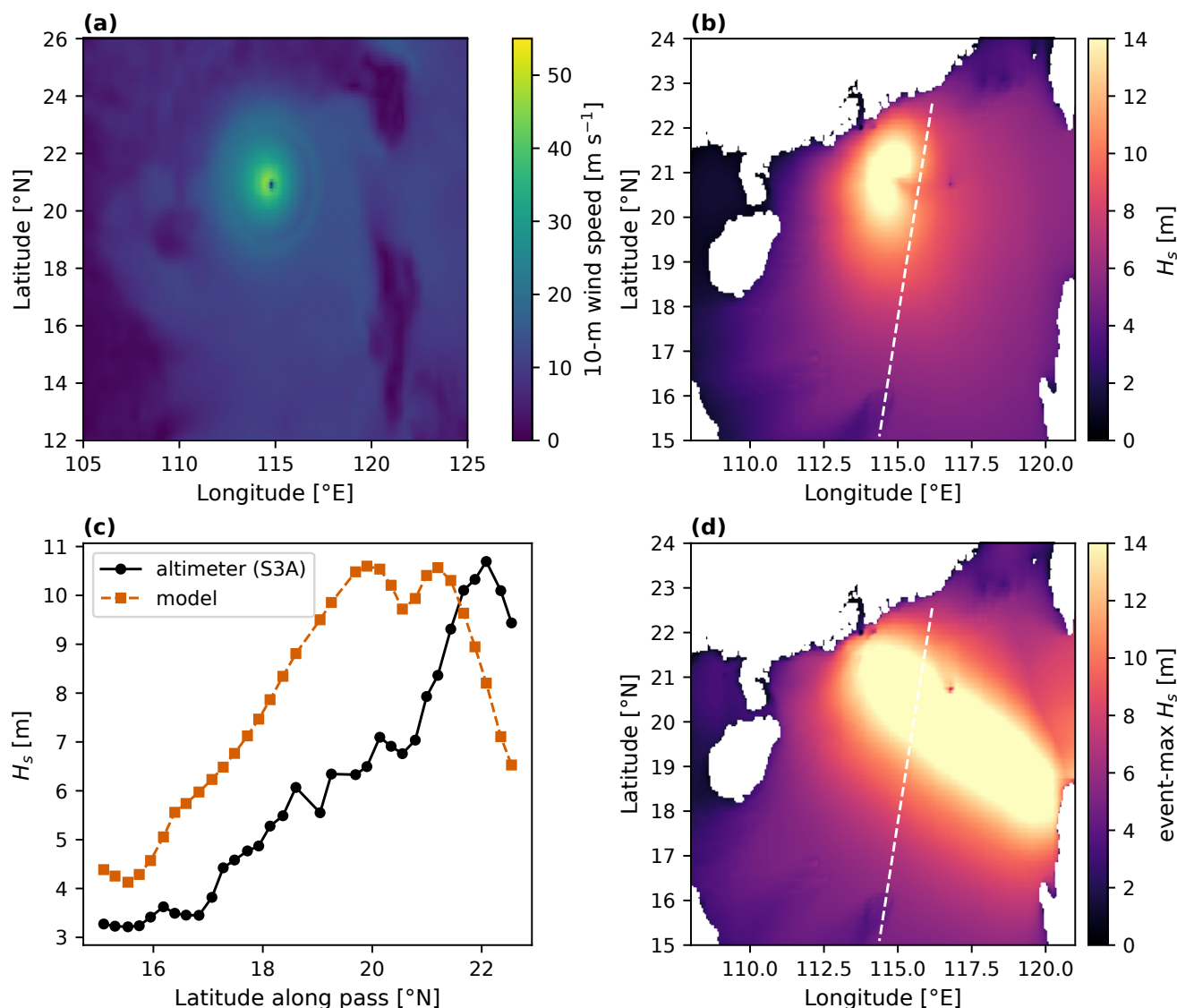


Figure 5. Typhoon Mangkhut (2018) and the three product layers: (a) reconstructed 10-m wind forcing at 02:00 UTC on 16 September 2018 (wind layer), (b) simulated significant wave height at the same time with the Sentinel-3A ground track (dashed; hourly wave-field layer), (c) the core transect at 02:30 UTC, model against altimeter super-observations along the pass, and (d) the event-maximum H_s footprint with the same pass overlaid (event-catalogue layer), on the colour scale of panel (b). The land mask in panels (b) and (d) is derived from GEBCO 2024 (GEBCO Compilation Group, 2024).

95 % level. On the validation set the corrected median relative error is -1.1% with 101 of 129 events (78 %) within $\pm 25\%$; on the estimation set it is -2.9% (44 of 56), and over all 185 events -1.4% (145 of 185, 78 %). Over the full sample severe ty-



Table 3. Full-dataset pooled validation matrix (64,529 super-observations): stratification by model wind band, distance to the storm centre, and storm quadrant relative to the motion heading (RF/RR/LR/LF = right-front, right-rear, left-rear, left-front, for points within 400 km of the centre).

Dimension	Bin	Super-obs	Bias (m)	RMSE (m)	SI	r
wind bin (m/s)	U<15	61846	-0.02	0.89	0.39	0.75
wind bin (m/s)	15-25	2370	1.06	1.85	0.40	0.56
wind bin (m/s)	U>=25	192	4.28	4.84	0.74	0.40
distance (km)	0-100	607	1.83	2.81	0.69	0.77
distance (km)	100-200	2170	0.89	1.83	0.51	0.73
distance (km)	200-400	8013	0.09	1.21	0.41	0.72
distance (km)	400-inf	53739	-0.03	0.84	0.38	0.78
quadrant	RF	2555	0.03	1.49	0.45	0.74
quadrant	RR	2750	-0.14	1.21	0.34	0.77
quadrant	LR	2977	0.86	1.67	0.57	0.73
quadrant	LF	2508	0.59	1.52	0.55	0.78

The wind-band rows cover 64,408 super-observations; the remaining 121 lack a valid model wind value at the matchup point and enter the distance and quadrant rows only.

185 phoons move to +3.1 % (22 of 33 within ± 25 %) and super typhoons to -0.1 %. The dataset itself remains uncorrected model output; the annex ships as a separate table and script so that users can apply, adapt, or decline the adjustment and propagate its confidence intervals into their own uncertainty budgets.

3.5 Eyewall anchor verification

190 Fixed-point anchors probe the sea states that altimeter tracks miss. Anchors are classified by the distance between the station and the best-track centre at the time of the observed peak: near-core anchors ($r \leq 200$ km) test eyewall skill, and far-field stations ($r > 200$ km) document swell propagation. All 12 station–storm combinations (10 A-grade, 2 B-grade) are reported (Table 7; far-field stations in Table A1); none are excluded.

At the 5 near-core anchors, SCS-TCWave reproduces the observed peaks within ± 25 % at 3 of 5 before correction. The open-water anchor at Dongsha Island during Typhoon Krathon ($r = 61$ km at the observed peak) illustrates the annex operating in
 195 its design zone: the raw error of +50.5 % reduces to +6.0 % under the 0–100 km factor. The Pearl River estuary anchor during Mangkhut shows the opposite case: the raw error of -24.5 % deepens to -36.9 % under correction, because the station is depth-limited and the annex encodes an open-water radial bias. The annex documentation therefore restricts its application to open-water points. At the same estuary anchor, the ERA5-forced reference underestimates the observed peak by 29.7 %; the reference also matches the A-grade far-field stations well (4 of 5 within ± 25 %), where SCS-TCWave runs systematically low



Table 4. Distance-binned validation of significant wave height against altimeter super-observations, by CMA intensity class (peak intensity inside the selection box). Bias and RMSE in metres; SI is the scatter index and r the correlation coefficient.

Class	r bin (km)	Events	Super-obs	Bias (m)	RMSE (m)	SI	r
TS_STS	0-100	45	298	1.15	1.68	0.63	0.41
TS_STS	100-200	59	961	0.55	1.33	0.51	0.47
TS_STS	200-400	78	3382	-0.10	0.92	0.42	0.56
TS_STS	400-inf	95	23447	-0.10	0.47	0.23	0.92
TY	0-100	16	90	1.98	2.67	0.62	0.80
TY	100-200	30	512	1.15	1.85	0.49	0.78
TY	200-400	40	1899	0.17	1.20	0.35	0.73
TY	400-inf	43	12835	0.02	0.56	0.24	0.90
STY	0-100	21	145	2.56	3.82	0.64	0.40
STY	100-200	28	492	1.07	2.26	0.47	0.48
STY	200-400	33	1948	0.13	1.55	0.42	0.59
STY	400-inf	33	11512	-0.04	1.53	0.62	0.52
SuperTY	0-100	9	74	2.99	4.02	0.70	0.85
SuperTY	100-200	13	205	1.35	2.50	0.53	0.74
SuperTY	200-400	16	784	0.55	1.36	0.43	0.79
SuperTY	400-inf	16	5945	0.13	0.61	0.31	0.87

Table 5. Per-event peak significant wave height for the 185 altimeter-sampled events, raw and annex-corrected, by intensity stratum: mean peak bias, median relative error, and the count of events within $\pm 25\%$.

Stratum	Events	Bias raw (m)	Med RE raw (%)	$ RE \leq 25$ raw	Bias corr (m)	Med RE corr (%)	$ RE \leq 25$ corr
ALL	185	0.75	9.50	111/185	-0.11	-1.40	145/185
TS/STS	93	0.04	1.70	68/93	-0.36	-8.20	78/93
TY	43	0.98	20.90	24/43	0.26	7.30	33/43
STY	33	1.58	33.60	11/33	-0.30	3.10	22/33
SuperTY	16	2.59	19.70	8/16	0.71	-0.10	12/16

200 (-22.5 to -62.9% , 1 of 5 within $\pm 25\%$) under the combined effect of far-field wind deficit and one-way nesting. The two B-grade stations lie beyond 1100 km and are biased high in both runs.

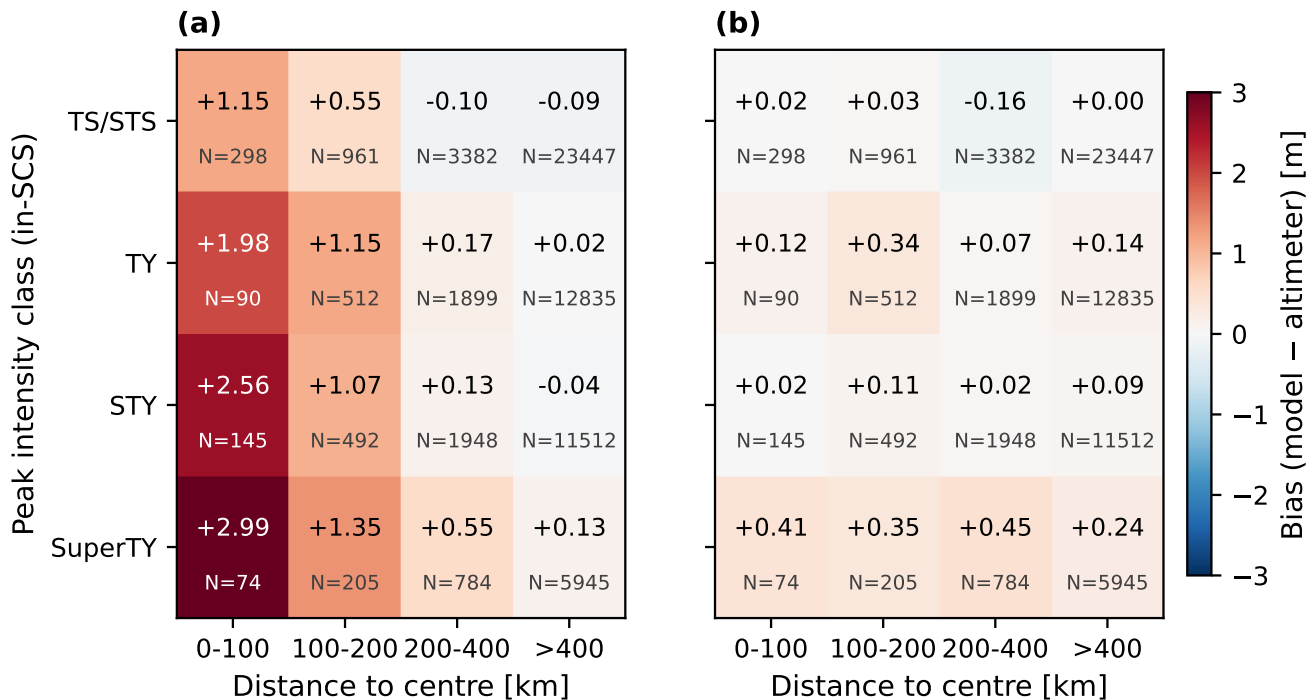


Figure 6. Residual structure of significant wave height against altimeter super-observations, stratified by distance ring and in-SCS peak intensity class: (a) raw model output and (b) with the annex factors applied to every matchup point. Cell annotations give the bias and the sample size; both panels share the colour scale. The raw panel reproduces the values of Table 4.

Table 6. Multiplicative correction annex coefficients by distance ring, with 95 % confidence intervals from a storm-level bootstrap (2,000 draws resampling whole storms) and the sample they are estimated from (first-batch super-observation comparison).

Distance ring (km)	Factor	95 % CI	Super-obs	Storms
0–100	0.704	[0.661, 0.752]	454	52
100–200	0.836	[0.792, 0.877]	1390	56
200–400	0.971	[0.915, 1.023]	4103	56
>400	1.050	[1.013, 1.095]	18372	56

3.6 Complementarity and recommended use

The two verification levels draw a consistent map, and the dividing line sits in the extreme tail. At the 9.4 m Pearl River estuary anchor during Typhoon Mangkhut, SCS-TCWave stays within $\pm 25\%$ and the ERA5-forced reference does not. The remaining near-core anchors are comparable between the two configurations, and the correction annex returns open-water near-core points

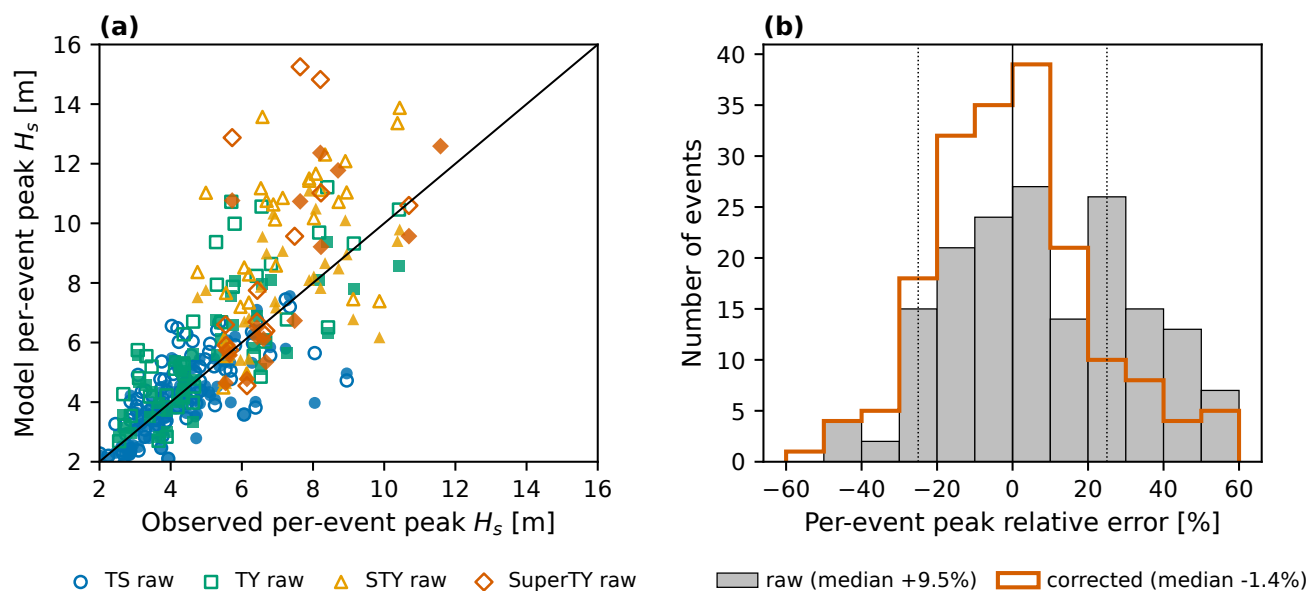


Figure 7. Per-event peak significant wave height for the 185 events with altimeter sampling near the peak. (a) Modelled versus observed peaks, raw (open symbols) and annex-corrected (filled), by intensity class. (b) Relative-error distributions before (raw, median +9.5 %) and after correction (median −1.4 %); dotted lines mark $\pm 25\%$.

Table 7. Near-core ($r \leq 200$ km at the observed peak) anchor verification: observed and modelled peak significant wave height, relative error before and after the correction annex, and the ERA5-forced reference. All five combinations are A-grade records.

Station	Storm	r (km)	Obs H_s (m)	Model H_s (m)	RE raw (%)	RE corr (%)	ERA5 RE (%)
Dongsha Island buoy	Hato (2017)	79	8.52	6.16	-27.7	-49.1	-16.1
Dongsha Island buoy	Krathon (2024)	61	9.67	14.56	+50.5	+6.0	-0.7
Xiaoliuqi buoy	Krathon (2024)	106	9.05	9.86	+9.0	-8.9	-12.9
Qimei buoy	Krathon (2024)	114	6.46	7.31	+13.1	-5.4	+17.3
Dawanshan buoy (PRE mouth)	Mangkhut (2018)	163	9.40	7.09	-24.5	-36.9	-29.7

such as Dongsha to within $\pm 25\%$ while centring the per-event peak distribution of the full catalogue. At the far-field stations the reference is clearly superior, and beyond roughly 400 km the smooth reanalysis background is already adequate: there the reanalysis wave fields are the natural choice. The distributed product documentation encodes this map as a recommended-use layer: the primary product with the annex inside 400 km, and ERA5 wave fields beyond it.



210 4 Uncertainty and known limitations

The validation in Sect. 3 defines where the dataset is strong. This section states, in one place, where it is not, so that users inherit the boundaries together with the data.

Core overestimation. The uncorrected fields overestimate significant wave height within 200 km of the storm centre, and the bias grows with storm intensity (Sect. 3.3; +1.15 m for tropical storms to +2.99 m for super typhoons in the innermost ring).
 215 The correction annex removes the systematic part over open water; a per-event scatter remains after correction (interquartile range 25.5 percentage points, from −14.8 to +10.7 %; the standard deviation of 23.1 % is inflated by eight events beyond ±50 %) and should enter the uncertainty budget of any extreme-value application built on the catalogue.

Far-field swell deficit. Beyond about 400 km from the centre the product underestimates swell, by 22.5 to 62.9 % at the Taiwan-side reference points of Typhoons Mangkhut and Hato (Sect. 3.5). Two design choices cause this: the reconstructed
 220 wind decays toward the ERA5 background outside the vortex, and the one-way nesting supplies no wave energy generated outside the outer grid. The recommended-use rule in Sect. 5 routes far-field applications to ERA5 accordingly.

Method tiers and the pre-2002 era. Tier-2 events (241 storms, predominantly 1979–2001) rely on empirical $R_{\max}$ and climatological profile parameters because observed wind-structure records do not exist for that era. The cross-era experiment (Sect. 2.5) bounds the method-switch effect (peak offset +1.0 %, 95 % CI [−1.7, +4.2] %, added scatter 7.8 % SD), and no
 225 direct altimeter matchups constrain tier-2 storms in the present archive (the local altimeter holdings start in 2008). Tier-2 fields therefore carry wider, explicitly labelled uncertainty, and the reconstruction-gate trigger rate during production was higher for tier-2 storms (27.0 % versus 13.8 %), consistent with noisier wind–pressure records in the early era.

No coupling to water levels or currents. The simulations use still-water bathymetry: no astronomical tide, no storm surge, no wave–current interaction. At depth-limited coastal points the peak wave height depends on the instantaneous water level,
 230 which is one reason the Dawanshan case in Sect. 3.5 is underpredicted and why the correction annex must not be applied there. Nearshore engineering applications should treat the fields as offshore boundary conditions rather than as design values at the structure.

Spatial resolution at complex coasts. The 0.05° inner grid does not resolve sheltering by small islands and headlands; the sheltered inshore wave recorders of the West Lamma Channel during Mangkhut are the documented example. Grid points
 235 adjacent to complex coastline geometry inherit this limit.

Peak period. Altimeters measure wave height, not period. T_p carries no systematic observational verification in this release: the fixed-point anchor records were evaluated for wave height, and no spectral comparison has been performed. Its error statistics are correspondingly weaker than those of H_s .

Background sea states. The event-window design targets typhoon waves. Monsoon and background conditions inside the
 240 windows are retained for continuity but are not the product’s purpose, and their skill is lower (Sect. 3.3). Users who need year-round climatology should combine the catalogue with a continuous product, as described in Sect. 5.



Flagged events. 98 events carry the `marginal` flag: their best-track wind and pressure records are mutually inconsistent, and the vortex was reconstructed with the shape parameter clamped at its physical bounds. They are retained to avoid selection bias against weak-vortex, monsoon-season storms; users performing sensitive statistics can exclude them with one filter.

245 A planned v1.1 will extend direct altimeter validation back to about 1992 using the ERS-1/2, TOPEX, and GFO missions, which addresses the tier-2 validation gap without changing the model chain.

5 Usage notes

5.1 Reading the products

Each event directory contains the three layers described in Sect. 2.6: hourly wave fields, the reconstructed wind forcing,
250 and the event-catalogue entry. All files are CF-1.8 NetCDF and open with standard tools. A worked example notebook, `scs_tcwave_usage_example.ipynb`, distributed with the archive, demonstrates reading the three layers and applying the correction annex by distance ring. The global index lists every event with its dates, track summary, method tier, and quality flag; the catalogue layer provides per-grid-point event maxima and their timing without touching the hourly fields.

5.2 Extreme-value applications

255 The event catalogue is the intended input for extreme-value analysis: one process maximum per storm per grid point, which feeds peaks-over-threshold and compound Poisson frameworks directly (Jonathan and Ewans, 2013). For open-water points, apply the correction annex factors by distance ring before fitting, and propagate the annex confidence intervals together with the post-correction scatter into the return-level uncertainty. Do not apply the annex at depth-limited coastal points. The catalogue records the peak-time distance to the storm centre, so the rule can be applied programmatically.

260 5.3 Recommended use by distance band

Within 200 km of the storm centre, use SCS-TCWave, with the annex over open water. Between 200 and 400 km, SCS-TCWave is recommended; the annex factor may be applied for consistency, noting that its 95 % confidence interval includes unity. Beyond 400 km, ERA5-based wave products serve background and remote swell better than this dataset does. For total wave hazard, combine the two: SCS-TCWave supplies the tropical cyclone component and a continuous product supplies the back-
265 ground climate, which mirrors how mixed-population extreme-value methods separate typhoon and monsoon contributions (Jonathan and Ewans, 2013; Shao et al., 2019).

5.4 Quality filtering

Two orthogonal attributes support filtering. `method_tier` separates observed-parameter events (tier 1) from empirical-parameter events (tier 2); trend and climate studies that span both eras should either work within tier 1 or propagate the



270 documented tier-2 scatter. `qc_flag` marks the 98 clamped reconstructions. Both attributes appear in every file's global meta-data and in the index.

5.5 Citation

Users of the dataset should cite the present article and the dataset itself (<https://doi.org/10.57760/sciencedb.011ul>; Fang et al., 2026). The repository README repeats this guidance.



275 Appendix A: Far-field anchor verification

Table A1 lists the seven far-field ($r > 200$ km at the observed peak) station–storm combinations, completing the anchor set of Sect. 3.5.

Table A1. Far-field ($r > 200$ km at the observed peak) anchor verification, columns as in Table 7 plus the record grade (A: quantitative peer-reviewed or agency record; B: reduced-confidence record).

Station	Storm	Grade	r (km)	Obs H_s (m)	Model H_s (m)	RE raw (%)	RE corr (%)	ERA5 RE (%)
Xiaoliuqi buoy	Hato (2017)	A	472	4.22	1.99	-52.8	-50.5	-14.7
Qimei buoy	Hato (2017)	A	378	5.41	2.01	-62.9	-63.9	-33.9
Dongjidaο wave station	Hato (2017)	B	1388	1.54	1.95	+26.8	+33.0	+122.6
Dongsha Island buoy	Mangkhut (2018)	A	248	12.30	8.52	-30.7	-32.7	+5.8
Xiaoliuqi buoy	Mangkhut (2018)	A	549	5.61	4.35	-22.5	-18.6	+3.1
Qimei buoy	Mangkhut (2018)	A	526	6.95	3.96	-43.1	-40.2	-24.4
Dongjidaο wave station	Mangkhut (2018)	B	1164	2.80	3.75	+33.8	+40.6	+72.8

Appendix B: Spin-up sensitivity

280 Table B1 lists the event-maximum significant wave height of the six storms rerun with nominal spin-up leads of 24–96 h, the basis of the 48 h lead in Sect. 2.1.

Table B1. Spin-up sensitivity: event-maximum H_s for six storms rerun with nominal leads of 24–96 h. [†] effective lead limited to 48 h by genesis inside the domain; dashes: the 72 and 96 h chains of storm 2015_0010 were not produced. At the 24 h lead the 200–400 km ring of storm 2024_0020 is 1.5 m below its 72 h value.

Storm	Event-maximum H_s (m) by nominal lead			
	24 h	48 h	72 h	96 h
2015_0010	10.98	10.98	–	–
2017_0009	4.75	4.75	4.75 [†]	4.75 [†]
2018_0026	18.09	18.14	18.17	18.17
2022_0003	11.42	11.43	11.43 [†]	11.43 [†]
2023_0005	11.83	11.82	11.82 [†]	11.82 [†]
2024_0020	17.70	17.71	17.73	17.72



Code availability

The SWAN wave model (version 41.51) is openly distributed at <https://swanmodel.sourceforge.io>. The frozen model configuration, the complete calibration decision record, and a usage-example notebook that reads the products and applies the correction annex are distributed with the dataset (<https://doi.org/10.57760/sciencedb.011ul>). Additional pre- and post-processing scripts
285 are available from the corresponding author upon reasonable request.

Data availability

SCS-TCWave v1.0 is archived as a single record at the Science Data Bank under a CC BY 4.0 licence and is available at <https://doi.org/10.57760/sciencedb.011ul> (Fang et al., 2026), which is the citable identifier for the dataset. The repository record contains 47 annual archives of the event products for 1979–2025, together with the documentation and data dictionary,
290 event-catalogue index, validation tables, correction annex, frozen model configuration and calibration record, and a usage-example notebook (`scs_tcwave_usage_example.ipynb`). The best-track data used to construct SCS-TCWave were obtained from the CMA tropical cyclone database (Ying et al., 2014; Lu et al., 2021) and IBTrACS v04r01 (Knapp et al., 2010). ERA5 data were obtained from the Copernicus Climate Data Store (Hersbach et al., 2020), bathymetric data from GEBCO 2024 (GEBCO Compilation Group, 2024), and calibrated altimeter data from the Australian Ocean Data Network
295 (Ribal and Young, 2019).

Author contributions. FF: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing (original draft preparation), and visualization. YY: validation, writing (review and editing), and project administration. PJ: validation, writing (review and editing), and project administration. BZ: resources, writing (review and editing), supervision, project administration, and funding acquisition.

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

Acknowledgements. We thank the China Meteorological Administration for providing the tropical cyclone best-track database, NOAA's National Centers for Environmental Information (NCEI) for the International Best Track Archive for Climate Stewardship (IBTrACS), the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Copernicus Climate Data Store for ERA5, the GEBCO Compilation Group for bathymetric data, the Australian Ocean Data Network (AODN) and the Integrated Marine Observing System (IMOS)
305 for the calibrated altimeter database, the Central Weather Administration (Taiwan) for buoy observations, and the Hong Kong Observatory for tropical cyclone reports.



Claude (Anthropic) was used for programming assistance and for language refinement of author-prepared manuscript text. All scientific decisions, methodological choices, analyses, interpretations, and conclusions were made and verified by the authors, who take full responsibility for the manuscript and dataset.

310 *Financial support.* This work is supported by the National Natural Science Foundation of China (52571291), Fundamental Research Funds for the Central Universities (2025ZYGXZR026) and Project of State Key Laboratory of Subtropical Building and Urban Science (2023ZB14).



References

- Bagnasco, T., Stocchino, A., Vousdoukas, M. I., and Wang, J.: A two-way coupled wave–current high resolution hindcast for the South China Sea, *Applied Ocean Research*, 161, 104 672, <https://doi.org/10.1016/j.apor.2025.104672>, 2025.
- 315 Booij, N., Ris, R. C., and Holthuijsen, L. H.: A third-generation wave model for coastal regions: 1. Model description and validation, *Journal of Geophysical Research: Oceans*, 104, 7649–7666, <https://doi.org/10.1029/98JC02622>, 1999.
- Cavaleri, L.: Wave modeling—missing the peaks, *Journal of Physical Oceanography*, 39, 2757–2778, <https://doi.org/10.1175/2009JPO4067.1>, 2009.
- Chavas, D. R., Lin, N., and Emanuel, K.: A model for the complete radial structure of the tropical cyclone wind field. Part I: Comparison
 320 with observed structure, *Journal of the Atmospheric Sciences*, 72, 3647–3662, <https://doi.org/10.1175/JAS-D-15-0014.1>, 2015.
- Deb, M., Balaguru, K., Rice, J., McPherson, T., and Daniel, W. B.: Tropical cyclone-driven storm surge and wave database for the US North Atlantic and Gulf coastlines, *Scientific Data*, <https://doi.org/10.1038/s41597-026-07635-z>, 2026.
- Dodet, G., Piolle, J.-F., Quilfen, Y., Abdalla, S., Accensi, M., Arduin, F., et al.: The Sea State CCI dataset v1: towards a sea state climate data record based on satellite observations, *Earth System Science Data*, 12, 1929–1951, <https://doi.org/10.5194/essd-12-1929-2020>, 2020.
- 325 Donelan, M. A., Haus, B. K., Reul, N., Plant, W. J., Stiassnie, M., Graber, H. C., Brown, O. B., and Saltzman, E. S.: On the limiting aerodynamic roughness of the ocean in very strong winds, *Geophysical Research Letters*, 31, L18 306, <https://doi.org/10.1029/2004GL019460>, 2004.
- Fang, F., Yang, Y., Jin, P., and Zhou, B.: SCS-TCWave: an event-based hindcast dataset of typhoon waves and reconstructed wind fields for the South China Sea (1979–2025), *Science Data Bank [data set]*, <https://doi.org/10.57760/sciencedb.011ul>, 2026.
- 330 Gabelica, M., Bojčić, R., and Puljak, L.: Many researchers were not compliant with their published data sharing statement: mixed-methods study, *Journal of Clinical Epidemiology*, 150, 33–41, <https://doi.org/10.1016/j.jclinepi.2022.05.019>, 2022.
- GEBCO Compilation Group: GEBCO 2024 Grid, <https://doi.org/10.5285/1c44ce99-0a0d-5f4f-e063-7086abc0ea0f>, 2024.
- Georgiou, P. N.: Design wind speeds in tropical cyclone-prone regions, Ph.D. thesis, University of Western Ontario, 1985.
- Harper, B. A., Kepert, J. D., and Ginger, J. D.: Guidelines for converting between various wind averaging periods in tropical cyclone
 335 conditions, Tech. Rep. WMO/TD-1555, World Meteorological Organization, 2010.
- Hersbach, H., Bell, B., Berrisford, P., et al.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Holland, G. J.: An analytic model of the wind and pressure profiles in hurricanes, *Monthly Weather Review*, 108, 1212–1218, [https://doi.org/10.1175/1520-0493\(1980\)108<1212:AAMOTW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1980)108<1212:AAMOTW>2.0.CO;2), 1980.
- 340 Holland, G. J.: A revised hurricane pressure–wind model, *Monthly Weather Review*, 136, 3432–3445, <https://doi.org/10.1175/2008MWR2395.1>, 2008.
- Holland, G. J., Belanger, J. I., and Fritz, A.: A revised model for radial profiles of hurricane winds, *Monthly Weather Review*, 138, 4393–4401, <https://doi.org/10.1175/2010MWR3317.1>, 2010.
- Holthuijsen, L. H., Powell, M. D., and Pietrzak, J. D.: Wind and waves in extreme hurricanes, *Journal of Geophysical Research: Oceans*,
 345 117, C09 003, <https://doi.org/10.1029/2012JC007983>, 2012.
- Hwang, P. A.: A note on the ocean surface roughness spectrum, *Journal of Atmospheric and Oceanic Technology*, 28, 436–443, <https://doi.org/10.1175/2010JTECHO812.1>, 2011.



- Jonathan, P. and Ewans, K.: Statistical modelling of extreme ocean environments for marine design: A review, *Ocean Engineering*, 62, 91–109, <https://doi.org/10.1016/j.oceaneng.2013.01.004>, 2013.
- 350 Keptert, J. D.: The dynamics of boundary layer jets within the tropical cyclone core. Part I: Linear theory, *Journal of the Atmospheric Sciences*, 58, 2469–2484, [https://doi.org/10.1175/1520-0469\(2001\)058<2469:TDOBLJ>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<2469:TDOBLJ>2.0.CO;2), 2001.
- 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), *Bulletin of the American Meteorological Society*, 91, 363–376, <https://doi.org/10.1175/2009BAMS2755.1>, 2010.
- Lin, N. and Chavas, D.: On hurricane parametric wind and applications in storm surge modeling, *Journal of Geophysical Research: Atmospheres*, 117, D09 120, <https://doi.org/10.1029/2011JD017126>, 2012.
- 355 Liu, Q., Rogers, W. E., Babanin, A. V., Young, I. R., Romero, L., Zieger, S., Qiao, F., and Guan, C.: Observation-based source terms in the third-generation wave model WAVEWATCH III: Updates and verification, *Journal of Physical Oceanography*, 49, 489–517, <https://doi.org/10.1175/JPO-D-18-0137.1>, 2019.
- Lu, X.-Q., Yu, H., Ying, M., Zhao, B.-K., Zhang, S., Lin, L.-M., Bai, L.-N., and Wan, R.-J.: Western North Pacific tropical cyclone database created by the China Meteorological Administration, *Advances in Atmospheric Sciences*, 38, 690–699, [https://doi.org/10.1007/s00376-](https://doi.org/10.1007/s00376-020-0211-7)
- 360 [020-0211-7](https://doi.org/10.1007/s00376-020-0211-7), 2021.
- Luo, Z., Huang, B., Chen, X., Tan, C., Qiu, J., and Huang, G.: Effects of wave–current interaction on storm surge in the Pearl River Estuary: A case study of Super Typhoon Mangkhut, *Frontiers in Marine Science*, 8, 692 359, <https://doi.org/10.3389/fmars.2021.692359>, 2021.
- Needham, H. F. and Keim, B. D.: A storm surge database for the US Gulf Coast, *International Journal of Climatology*, 32, 2108–2123, <https://doi.org/10.1002/joc.2425>, 2012.
- 365 Phadke, A. C., Martino, C. D., Cheung, K. F., and Houston, S. H.: Modeling of tropical cyclone winds and waves for emergency management, *Ocean Engineering*, 30, 553–578, [https://doi.org/10.1016/S0029-8018\(02\)00033-1](https://doi.org/10.1016/S0029-8018(02)00033-1), 2003.
- Powell, M. D., Vickery, P. J., and Reinhold, T. A.: Reduced drag coefficient for high wind speeds in tropical cyclones, *Nature*, 422, 279–283, <https://doi.org/10.1038/nature01481>, 2003.
- 370 Queffeuilou, P.: Long-term validation of wave height measurements from altimeters, *Marine Geodesy*, 27, 495–510, <https://doi.org/10.1080/01490410490883478>, 2004.
- Ribal, A. and Young, I. R.: 33 years of globally calibrated wave height and wind speed data based on altimeter observations, *Scientific Data*, 6, 77, <https://doi.org/10.1038/s41597-019-0083-9>, 2019.
- Ribal, A. and Young, I. R.: Global calibration and error estimation of altimeter, scatterometer, and radiometer wind speed using triple collocation, *Remote Sensing*, 12, 1997, <https://doi.org/10.3390/rs12121997>, 2020.
- 375 Rogers, W. E., Babanin, A. V., and Wang, D. W.: Observation-consistent input and whitecapping dissipation in a model for wind-generated surface waves, *Journal of Atmospheric and Oceanic Technology*, 29, 1329–1346, <https://doi.org/10.1175/JTECH-D-11-00092.1>, 2012.
- Shao, Z., Liang, B., Li, H., Li, P., and Lee, D.: Extreme significant wave height of tropical cyclone waves in the South China Sea, *Natural Hazards and Earth System Sciences*, 19, 2067–2077, <https://doi.org/10.5194/nhess-19-2067-2019>, 2019.
- 380 Smith, G. A., Hemer, M., Greenslade, D., Trenham, C., Zieger, S., and Durrant, T.: Global wave hindcast with Australian and Pacific Island focus: From past to present, *Geoscience Data Journal*, 8, 24–33, <https://doi.org/10.1002/gdj3.104>, 2021.
- Stodden, V., Seiler, J., and Ma, Z.: An empirical analysis of journal policy effectiveness for computational reproducibility, *Proceedings of the National Academy of Sciences*, 115, 2584–2589, <https://doi.org/10.1073/pnas.1708290115>, 2018.
- Sun, Z., Zhang, H., Xu, D., Liu, X., and Ding, J.: Assessment of wave power in the South China Sea based on 26-year high-resolution hindcast data, *Energy*, 197, 117 218, <https://doi.org/10.1016/j.energy.2020.117218>, 2020.
- 385



- Tamizi, A. and Young, I. R.: The spatial distribution of ocean waves in tropical cyclones, *Journal of Physical Oceanography*, 50, 2123–2139, <https://doi.org/10.1175/JPO-D-20-0020.1>, 2020.
- Tedersoo, L., Küngas, R., Oras, E., et al.: Data sharing practices and data availability upon request differ across scientific disciplines, *Scientific Data*, 8, 192, <https://doi.org/10.1038/s41597-021-00981-0>, 2021.
- 390 Vickery, P. J. and Wadhera, D.: Statistical models of Holland pressure profile parameter and radius to maximum winds of hurricanes from flight-level pressure and H*Wind data, *Journal of Applied Meteorology and Climatology*, 47, 2497–2517, <https://doi.org/10.1175/2008JAMC1837.1>, 2008.
- Vickery, P. J., Masters, F. J., Powell, M. D., and Wadhera, D.: Hurricane hazard modeling: The past, present, and future, *Journal of Wind Engineering and Industrial Aerodynamics*, 97, 392–405, <https://doi.org/10.1016/j.jweia.2009.05.005>, 2009.
- 395 Willoughby, H. E. and Rahn, M. E.: Parametric representation of the primary hurricane vortex. Part I, *Monthly Weather Review*, 132, 3033–3048, <https://doi.org/10.1175/MWR2831.1>, 2004.
- Willoughby, H. E., Darling, R. W. R., and Rahn, M. E.: Parametric representation of the primary hurricane vortex. Part II: A new family of sectionally continuous profiles, *Monthly Weather Review*, 134, 1102–1120, <https://doi.org/10.1175/MWR3106.1>, 2006.
- Xu, Y., He, H., Song, J., Hou, Y., and Li, F.: Observations and modeling of typhoon waves in the South China Sea, *Journal of Physical Oceanography*, 47, 1307–1324, <https://doi.org/10.1175/JPO-D-16-0174.1>, 2017.
- 400 Ying, M., Zhang, W., Yu, H., et al.: An overview of the China Meteorological Administration tropical cyclone database, *Journal of Atmospheric and Oceanic Technology*, 31, 287–301, <https://doi.org/10.1175/JTECH-D-12-00119.1>, 2014.
- Young, I. R., Sanina, E., and Babanin, A. V.: Calibration and cross validation of a global wind and wave database of altimeter, radiometer, and scatterometer measurements, *Journal of Atmospheric and Oceanic Technology*, 34, 1285–1306, <https://doi.org/10.1175/JTECH-D-16-0145.1>, 2017.
- 405 Zhai, R., Huang, C., Yang, W., et al.: Applicability evaluation of ERA5 wind and wave reanalysis data in the South China Sea, *Journal of Oceanology and Limnology*, 41, 495–517, <https://doi.org/10.1007/s00343-022-2047-8>, 2023.
- Zieger, S., Babanin, A. V., Rogers, W. E., and Young, I. R.: Observation-based source terms in the third-generation wave model WAVEWATCH, *Ocean Modelling*, 96, 2–25, <https://doi.org/10.1016/j.ocemod.2015.07.014>, 2015.