Scientific Data

https://doi.org/10.1038/s41597-026-06859-3

Article in Press

A Forty-year regional-scale dataset of shoreline change and nearshore wave conditions in Southeast Australia

Received: 21 August 2025

Accepted: 9 February 2026

Cite this article as: Mao, Y., Vos, K., Cagigal, L. *et al.* A Forty-year regional-scale dataset of shoreline change and nearshore wave conditions in Southeast Australia. *Sci Data* (2026). https://doi.org/10.1038/s41597-026-06859-3

Yongjing Mao, Kilian Vos, Laura Cagigal, Valentine Bodin, Mitchell D. Harley & Kristen D. Splinter

A Forty-year regional-scale dataset of shoreline change and nearshore wave conditions in Southeast Australia

Yongjing Mao[1], Kilian Vos[2], Laura Cagigal[3], Valentine Bodin[1,4], Mitchell D. Harley[1], and Kristen D. Splinter[1]

[1]Water Research Laboratory, UNSW Sydney, Australia

[2]OHB Digital Services GmbH, Bremen, Germany

[3]Geomatics and Ocean Engineering Group, Departamento de Ciencias y Tecnicas del Agua y del Medio Ambiente, Universidad de Cantabria, Cantabria, Spain

[4]School of Centrale Marseille, Aix-Marseille Université, Marseille, France

Corresponding author: Yongjing Mao (Yongjing.mao@unsw.edu.au)

Abstract

Coastal erosion at wave-dominated beaches, primarily driven by nearshore wave dynamics, poses a substantial challenge for coastal management. While existing datasets from individual beaches have improved our understanding of site-specific coastal morphodynamics, there is a growing demand for regional-scale datasets to understand and predict regional shoreline responses to climate variability. To address this, we present a combined shoreline and nearshore wave dataset for the wave-dominated coast of southeast Australia, comprising over 8,000 cross-shore transects at 100 m spacing for over 300 beaches. For each transect, satellite-derived shoreline positions (1984–2024) and beach-face slopes are provided, alongside hourly nearshore wave parameters (1979–2024) extracted at the 10 m depth contour. Shoreline data have been validated using available field surveys, and wave data have been assessed against offshore and nearshore buoy observations. This dataset provides a valuable resource for developing regional-scale understanding of shoreline variability along wave-dominated and embayed coastlines.

Background & Summary

Understanding and predicting shoreline change are important for managing erosion risks, informing land use decisions, and supporting long-term coastal adaptation strategies [1, 2]. One of the primary drivers of shoreline change on many coastlines is nearshore wave dynamics, which influence sediment transport processes both in the cross-shore and longshore directions [3-5]. As such, historical records of shoreline position and wave conditions are critical for advancing our understanding of coastal morphodynamics and developing models capable of quantitatively forecasting shoreline evolution [6, 7].

Several beach profile datasets from individual beaches worldwide have supported evidence-based coastal research across various environments, from micro- to meso-tidal settings [8-13]. Combined with nearshore wave data, these datasets have provided insights into volume changes driven by storms [14, 15] and seasonal to interannual climate variability (e.g. ENSO) [16]. Shoreline position, derivable from these datasets, serves as a key coastal state indicator for shoreline management [17] and a proxy that correlates with the total beach-dune volume [18]. The shoreline position data also underpins the development of shoreline change models, which have largely been tested at individual sites [19-23]. Model transferability without site specific calibration has been shown to be limited [24-27].

To move beyond beach-specific applications, synthesizing shoreline and wave data across multiple sites enables regional-scale analyses of shoreline-wave dynamics [28-33]. With increasing computational power and advances in reduced-complexity and data-driven models, there is growing potential for regional-scale shoreline models [34-36]. These models rely on extensive regional datasets for calibration, validation, and deployment. Such datasets are also crucial for assessing model transferability across diverse settings, which is key for large-scale coastal planning and early warning systems [37]. In the context of emerging data-driven approaches, training machine learning models, including deep learning architectures (e.g. transformers [38]), on large, diverse datasets can enhance predictive accuracy and generalizability beyond what site-specific models can achieve [39, 40].

Recent advances in satellite derived shoreline (SDS) [41] and efficient wave downscaling methods [42, 43] have made it increasingly feasible to generate accurate shoreline and nearshore wave datasets at regional scales. Herein, we present a comprehensive shoreline and nearshore wave

dataset for the coast of New South Wales (NSW), Australia. The NSW coastline is composed exclusively of wave-dominated sandy beaches from dissipative to reflective states, and exhibits distinct morphodynamics [44, 45]. With wave-driven longshore and cross-shore sediment transport as the dominant processes in shaping the shoreline, NSW beaches are influenced by both short-term storm events and longer-term climate variability associated with ENSO [14, 30]. The region's microtidal conditions further enhance the reliability of satellite-derived shoreline measurements, as they reduce the influence of tidal variability on shoreline detection [41].

The dataset comprises over 8,000 cross-shore transects spaced at 100-meter intervals along the NSW coast. For each transect, satellite-derived shoreline positions are provided at approximately bi-weekly intervals from 1984 to 2024, along with associated beach-face slopes. Hourly nearshore wave parameters—including significant wave height (H_s), peak period (T_p), mean wave period (T_{m02}), as well as peak and mean wave directions (D_p and D_m), are provided at the 10 m depth contour for each transect from 1979 to 2024. The dataset has been validated using field survey data (for shorelines) and offshore and nearshore wave buoys (for waves). A subset of this dataset (e.g., Curl Curl Beach) has been previously used for benchmarking shoreline models in the ShoreShop2.0 initiative [46].

Both the shoreline and nearshore wave data are valuable for coastal process studies and hazard management applications, such as shoreline erosion, wave overtopping, and coastal inundation [47]. The combined dataset also facilitates regional understanding of coastal morphodynamics and provides a robust foundation for developing, validating, and implementing shoreline evolution models, especially hybrid and data-driven ones that benefit from large datasets [31, 32, 36].

Methods

Definition of transects

The transects used in this dataset are a subset of the shore-normal transects defined for wave-dominated sandy beaches across the entire Pacific Rim [30], spaced at 100 m intervals alongshore. In NSW, 8,828 transects were identified, spanning 307 wave-dominated sandy beaches selected using the criteria introduced by Vos et al. [30] (Figure 1a). These transects were used to extract both shoreline positions and nearshore wave conditions.

(Figure1 goes here)

Satellite derived shorelines

The CoastSat open-source Python toolkit V3.2 [48, 49] was used to extract shoreline positions at mean sea level (MSL) from satellite imagery. The process involves four key steps: (1) retrieving satellite images via Google Earth Engine; (2) automatically extracting waterlines from the selected images; (3) intersecting the waterlines with predefined shore-normal transects; and (4) applying tidal correction to obtain the shoreline time series.

This workflow was applied individually to each of the 307 beaches. For each beach, a polygonal region of interest (ROI) encompassing the entire beach was defined. All available images in the Landsat archive (Landsat 5, 7, 8, and 9) between 1984 and 2024 that intersected the ROI were clipped, downloaded and pansharpened to 15 m resolution using the panchromatic band.

During waterline extraction, each image was segmented using a pre-trained model developed across a range of Pacific Rim beaches [30]. The classifier categorized pixels into four classes: sand, water, white water (breaking waves), and other. To improve waterline detection, the Modified Normalized Difference Water Index [MNDWI, 50] was computed using the green and short-wave infrared (SWIR1) bands. A histogram of MNDWI values was constructed only for the sand and water pixels, and Otsu's thresholding algorithm [51] was applied to determine an optimal threshold that maximized inter-class variance. The waterline contour corresponding to this threshold was then extracted at sub-pixel resolution using the Marching Squares algorithm [52].

To eliminate outliers and false detections, a reference shoreline was defined for each beach, and shoreline vertices located beyond a 150 m buffer from this reference were removed. The intersections between valid waterline contours and pre-defined shore-normal transects were calculated. Waterline positions were defined as the distances between the landward ends of shore-normal transects and their intersections with valid waterline contours.

Since every satellite image is taken at a different stage of the tide, the waterline position extracted from the previous step reflects the instantaneous boundary between land and water. To retrieve the shoreline position at MSL, a tidal correction was applied to remove high-frequency waterline dynamics related to tidal variations based on the instantaneous water level and the slope of the intertidal zone. Water levels at the time of each image acquisition were extracted from the nearest

grid point in the FES2022 global tide model [53]. Shoreline position was obtained by adding a tidal related horizontal (shore-normal) shift Δx to the extracted instantaneous waterline position at each transect. By assuming a constant beach face slope m, given the reference elevation z_{ref} (0 m for MSL, defined by FES2022 output) and the local water level at the time of image acquisition z_{wt}, the horizontal shift Δx can be estimated as shown in equation 1. The beach face slope m is estimated as the value that, when used for tidal correction, minimizes waterline variance at the tidal frequency relative to erosion/accretion signals with frequency domain analysis [54].

$$\Delta x = \frac{z_{ref} - z_{wt}}{m} \quad (1)$$

Nearshore wave reconstruction

The nearshore wave data were derived by downscaling offshore wave hindcasts from the Centre for Australian Weather and Climate Research (CAWCR) between 1979 and 2024 [55, 56]. This hindcast was provided hourly and globally at 0.4° resolution with nested grids down to 4 arcminutes (~7 km) for the Australian region. Directional wave spectra were available at a 0.5° resolution, discretised into 29 frequencies exponentially spaced from 0.038 Hz to 0.5 Hz and 24 directions with a constant 15° directional resolution. To improve accuracy, the CAWCR hindcasts were bias-corrected using the IMOS satellite altimetry dataset [57], which provides global wind speed and significant wave height data since 1985. Spectral correction followed the method of Cagigal et al. [42], using partitioned sea and swell components corrected independently based on the directional method of Albuquerque et al. [58]. Only the deep water (depth>300 m) altimetry data was used for the correction.

The BinWaves approach [42] was applied to efficiently propagate the full directional wave spectra to nearshore locations. This method is based on linear wave theory and assumes that the propagation of individual wave components is independent. While this assumption holds well in deep water, its accuracy decreases in shallow coastal zones where non-linear processes such as shoaling and wave breaking become significant. Nonetheless, previous studies have demonstrated that this approach can accurately reconstruct nearshore wave conditions at depths of approximately 10~15 m [59].

BinWaves approach consists of 2 main steps: (1) generate a library of monochromatic pre-runs of a wave propagation model; and (2) reconstruct propagated wave spectra at points of interest given deepwater wave boundary conditions (e.g. CAWCR) (Figure 2).

(Figure 2 goes here)

In the first step, following the CAWCR data, a full directional wave spectrum was decomposed into 696 monochromatic bins, representing 29 frequencies and 24 directions (Figure 3a). For computational efficiency, only bins with landward wave directions relative to the NSW coastline (−15° to 195°) and occurrence rate of wave energy above 1% were retained (Figure 3b). The Simulating WAves Nearshore (SWAN) model [60] was used to simulate wave propagation from offshore to nearshore. The NSW coastline was divided into seven model domains, each covering one offshore buoy (Figure 1). The main grids were aligned parallel to the coastline and extended seaward beyond the 300-m depth contour. Each domain included one offshore buoy for validation and covered the entire shoreline in the landward end. Five of seven domains contain nearshore wave buoys with BYR and COH as exceptions (Figure 1). Each main grid was paired with a higher-resolution nested grid focused on the nearshore zone (depth < 50 m). Grid resolutions were 250 m for main grids and 100 m for nested grids. Positioning and dimension of grids are provided in Table 1. High-resolution nearshore bathymetry (e.g. LiDAR surveys) and coarse resolution offshore bathymetry data were stitched and resampled to 50-m resolution [61]. This merged dataset was used as the input bathymetric data to the SWAN wave models.

(Figure 3 goes here)

Table 1 Parameters for SWAN computational grids used. Xpc and Ypc are the x and y coordinates of the origin of the grid. Alpc is the rotation of the computational grid in Cartesian convention. Xlenc and Ylenc are the length of the grid in x and y directions. Mxc and Myc are the number of domains in the x and y directions.

Domain		Xpc	Ypc	Alpc (°)	Xlenc (°)	Ylenc (°)	Mxc	Myc
Byron Bay (BYR)	Main	153.83	-29.75	83	1.8	0.7	718	278
	Nested	153.56	-29.72	83	1.8	0.43	1798	428
	Main	153.37	-31.14	75	1.6	0.5	638	198

Coffs Harbour (COH)	Nested	153.13	-31.08	75	1.6	0.23	1598	248
Crowdy Head (CRO)	Main	152.73	-32.96	67	2.1	0.6	838	238
	Nested	152.41	-32.82	67	2.1	0.25	2098	248
Sydney (SYD)	Main	151.2	-34.7	50	2.3	0.72	917	286
	Nested	150.95	-34.49	50	2.3	0.4	2299	399
Port Kembla (KEM)	Main	151.00	-35.26	66	1.2	0.45	478	178
	Nested	150.82	-35.18	66	1.2	0.25	1199	249
Batemans Bay (BAT)	Main	150.39	-36.34	61	1.4	0.48	558	190
	Nested	150.23	-36.25	61	1.4	0.3	1398	299
Eden (EDE)	Main	150.29	-37.7	83	1.6	0.5	638	198
	Nested	150.07	-37.67	83	1.6	0.28	1598	278

For each bin representing a unique combination of frequency f_i and direction θ_j, a unitary wave boundary condition was created with H_s of 1 m to run the SWAN model in stationary mode. To simulate monochromatic wave boundary conditions, a unimodal theoretical JONSWAP spectrum was defined with a peak enhancement factor (γ) of 50 and a directional spreading coefficient of 3 degrees following Cagigal et al. [42]. For each grid point P of the SWAN model, given the unitary boundary spectra $U(f_i, \theta_j)$ and the propagated spectra $U_p(f_i, \theta_j)$, a propagation coefficient $K_p(f_i, \theta_j)$ was obtained using the quadratic relationship between the variance density spectrum and its associated bulk H_s [62] (Equation 2). To avoid unrealistic wave steepness, for peak periods smaller than 5 s (f_i>0.2 Hz), the unitary boundary condition used a H_s of 0.1 m and the corresponding K_p was rescaled by multiplying by 10 afterwards.

$$K_p(f_i, \theta_j) = \sqrt{\frac{U_p(f_i, \theta_j)}{\int\int U(f_i, \theta_j)\, df\, d\theta}} \quad (2)$$

In the second step, full spectral reconstruction $S_p(f_i, \theta_j)$ can be implemented at a grid point P using the propagation coefficient K_p and boundary condition spectrum $S(f_i, \theta_j)$ by linearly superposing the propagation of each monochromatic bin (equation 2).

$$S_p(f_i, \theta_j) = \sum_i \sum_j S(f_i, \theta_j) * K_p^2(f_i, \theta_j) \quad (3)$$

To reduce data storage, the wave propagation coefficient K_p and directional wave spectra S_p were saved only at grid points within 9–11 m water depth, as well as at grid points closest to the nearshore and offshore buoys for validation purposes (Figure 1). For the NSW nearshore wave data herein, spectral timeseries from seven CAWCR stations are used as the offshore boundary condition $S(f_i, \theta_j)$ to reconstruct the nearshore grid points $S_p(f_i, \theta_j)$ for the seven domains, respectively (Figure 1). All these CAWCR stations are adjacent to the seaward boundary of model domains and located in deep waters (depth >300 m) where the non-linear interactions can be assumed negligible. Hourly wave parameters including H_s, T_p, T_{m01}, T_{m02}, D_p, and D_m were derived from the reconstructed nearshore spectra $S_p(f_i, \theta_j)$. Only the parametric outputs are included in the published dataset.

To associate the nearshore wave conditions with shoreline transects, each transect was extended seaward to identify its intersection with the 10 m depth contour, where the assumption of linear wave theory remains valid. SWAN grid points within 500 m of this intersection and between 9–11 m depth were used to interpolate wave parameters at the intersection point using the Inverse Distance Weighting (IDW) method. In overlapping zones between adjacent domains, results were merged based on the wave direction as follows. Since constant boundary conditions were applied on all grid sides, near-boundary areas can exhibit directional biases. Specifically, wave parameters may be less accurate for northern and southern boundaries under North or South wave conditions, respectively. To address this, if the incident wave direction was between −90° and 90° (North waves), results from the northern domain were used; otherwise, results from the southern domain were retained.

Data records

The dataset is publicly available via the Zenodo data repository [63] (https://doi.org/10.5281/zenodo.17693915). It is organized into two primary components: one for

shoreline data and another for nearshore wave data. In addition to these core datasets, four supplementary geospatial layers are provided in GeoJSON format: the region of interest (ROI_NSW.geojson), a reference shoreline for each beach (shorelines_NSW.geojson), the location of transects (transects_NSW.geojson), and the intersection of extended transects at 10 m contour with quality assurance (intersections_10m_QA.geojson). These files can be easily viewed and processed using common GIS software such as ArcGIS, QGIS, or GDAL.

The CoastSat_shorelines folder contains shoreline data from 1984 to 2024, organized into subfolders for each beach site (e.g., aus0045, aus0046, etc.), with folder names matching the id field in ROI_NSW.geojson. Each site folder includes CSV files with raw shoreline time series (time_series_raw.csv), tidally corrected shoreline time series (time_series_tidally_corrected.csv), tidal levels from the FES2022 model (tide_levels_fes2022.csv), and beach-face slopes (transect_beach_slopes.csv) used for tidal correction. Shorelines were mapped in January 2025 using CoastSat (v3.2).

The wave component contains hourly nearshore wave data from 1979 to 2024, stored in 10 NetCDF files with size ranging from 5 to 10 GB. File names (e.g., aus0045_aus0066.nc) indicate the range of beach site IDs included. All NetCDF files follow a consistent structure with two dimensions: time and tran_id (Transect ID, matching transects_NSW.geojson). Each file includes standard reconstructed wave parameters from the spectral output of BinWaves: significant wave height (H_s), peak period (T_p), mean wave periods (T_{m01}, T_{m02}), and mean and peak wave directions (D_m, D_p). Other properties for each transect include the longitude and latitude of the intersection point between the extended transect and the 10 m depth contour (where waves are extracted), the distance from the seaward end of the transect to the intersection, and the angle between the transect and the 10m contour at the intersection. Time for both shoreline and wave datasets is in Coordinated Universal Time (UTC).

In addition to the examples presented in the Technical Validation section, the shoreline and wave validation figures for all sites in NSW are also provided in the comparison results folder.

Technical validation

Shoreline comparison

The SDS shoreline data is validated against in-situ field survey data from 12 beaches in New South Wales (Figure 1a). These include: Narrabeen Beach, with multi-decadal monthly surveys [8]; three beaches – Bilgola, Dee Why, and Mona Vale – with 3.5 years of weekly to bi-weekly surveys (Dec 2013 to Jun 2017); six beaches—Dixon Park, Manly, Shoalhaven, Thirroul, Wamberal, and Wanda—with monthly surveys over two years (2011–2012) [64]; and two beaches—Moruya and Pedro—with 10 years of monthly surveys (Dec 2007 to Jun 2017) [28].

For each transect and site, error metrics including mean bias, root-mean-square error (RMSE), and coefficient of determination (R^2) are computed. With Moruya Beach as an example (Figure 4) and results for other sites provided in the Zenodo data repository [63] (https://doi.org/10.5281/zenodo.17693915), the SDS shoreline time series closely matches the in-situ observations. For Moruya Beach, transect-wise RMSEs range from 11.21 m at the northern end (Profile 1) to 15.01 m at the southern end (Profile 3). The bias is positive (seaward) for Profile 1, negative (landward) for Profile 2 and negligible for Profile 3. The site-level statistics at Moruya are bias = -0.26 m, RMSE = 12.71 m, and $R^2 = 0.67$. Across all 12 sites (Figure 5), Moruya has the largest RMSE. Most beaches show a seaward (positive) bias, except Shoalhaven, a more dissipative beach, which exhibited a landward (negative) bias. Bias values range from −4 m to +6 m, RMSE values are consistently within 7–13 m, and R^2 values exceed 0.4 for all sites.

(Figure 4 goes here)

(Figure 5 goes here)

Wave comparison

Wave parameters derived from the BinWaves approach are validated against measurements from 7 offshore [65] and 19 nearshore wave buoys [66] located along the NSW coast (Figure 1) with details provided in the supplementary Table S1. The comparison results for the Sydney (SYD) offshore buoy are shown in Figure 6, with results for the other offshore buoys provided in the Zenodo data repository [63] (https://doi.org/10.5281/zenodo.17693915). The comparison shows strong agreement (RMSE=0.34 m and $R^2 = 0.84$) and negligible bias (Bias = 0.13 m) for H_s. Among other parameters, T_p shows the largest discrepancies, with an RMSE of 2.31 s, R^2 of 0.26,

and a negative bias of -0.73 s, likely due to differences in spectral binning between the buoy and CAWCR data. In comparison, T_{m02} has a lower RMSE of 0.94 s and larger R^2 of 0.42. The model also effectively reconstructs wave directions with RMSE of 29° and negligible bias. These metrics at SYD are broadly representative of other offshore buoys (Table 2), with slightly larger errors observed at COH and EDE. Across all buoys, H_s and T_{m02} are generally overestimated, while T_p is consistently underestimated.

(Figure 6 goes here)

Table 2 Summary of offshore wave comparison metrics. Blank cells are for buoys without direction information.

	H_s (m)			T_p (s)			T_{m02} (s)			D_m (°)		
	RMSE	Bias	R^2	RMSE	Bias	R^2	RMSE	Bias	R^2	RMSE	Bias	R^2
BYR	0.38	0.16	0.76	2.00	-0.49	0.31	0.92	0.09	0.43	29.61	-7.42	0.53
COH	0.41	0.23	0.77	2.02	-0.43	0.28	1.21	0.77	0.45	29.44	1.05	0.52
CRO	0.36	0.18	0.81	2.14	-0.53	0.29	1.08	0.56	0.45			
SYD	0.34	0.13	0.86	2.31	-0.73	0.24	0.94	0.25	0.42	29.10	-3.38	0.52
KEM	0.36	0.11	0.77	2.37	-0.59	0.23	1.25	0.60	0.29	31.87	-6.51	0.48
BAT	0.36	0.18	0.78	2.18	-0.50	0.27	1.10	0.50	0.37			
EDE	0.42	0.20	0.77	2.38	-0.54	0.28	0.85	0.14	0.49	38.72	-0.42	0.55

The BinWaves outputs are also compared against nearshore wave buoy data [66]. The Woonona buoy is used as an example (Figure 7) while results for other buoys are provided in the Zenodo data repository [63] (https://doi.org/10.5281/zenodo.17693915) and summary statistics across all sites are shown in Table 3. The results demonstrate that BinWaves can effectively reproduce nearshore wave parameters. Compared to the offshore validation, nearshore H_s is even better reconstructed, with RMSE consistently below 0.35 m. However, the accuracy of wave period estimates declines in the nearshore. T_p is generally underestimated, showing a similar magnitude of negative bias as in the offshore case, while T_{m02} is overestimated with a larger bias. Both parameters have increased RMSE. For directional accuracy, D_m generally improves for nearshore buoys with RMSE reducing to below 25°. In contrast, D_p exhibits greater bias and RMSE due to

the coarse (15°) directional binning of the CAWCR input spectra, which results in discretised outputs and limits the resolution of directional changes.

(Figure 7 goes here)

Table 3 Summary of nearshore wave comparison metrics. Blank cells are for buoys without direction information.

	Hs (m)			T_p (s)			T_{m02} (s)			D_p (°)			D_m (°)		
	RMSE	Bias	R^2	RMSE	Bias	R^2	RMSE	Bias	R^2	RMSE	Bias	R^2	RMSE	Bias	R^2
Farquhar	0.29	0.15	0.78	3.92	-0.65	0.21	1.12	0.63	0.43	26.35	-2.06	0.27	13.23	0.67	0.64
OldBar A	0.32	0.20	0.76	3.78	-0.65	0.10	1.23	0.86	0.40	22.25	1.02	0.38	12.97	1.16	0.64
OldBar B	0.34	0.26	0.80	2.52	-0.19	0.14	1.28	1.04	0.37	23.51	-6.33	0.32	11.07	-1.26	0.70
Boomerang A	0.24	0.13	0.72	2.20	-0.73	0.26	0.96	0.52	0.44	20.88	-2.31	0.45	10.64	-4.92	0.79
Boomerang B	0.29	0.18	0.78	4.36	-1.34	0.17	1.05	0.61	0.29	39.06	-0.37	0.22	22.63	-2.48	0.50
Worimi	0.33	0.16	0.80	2.72	0.23	0.15	1.50	1.27	0.47	24.67	5.60	0.40	13.00	5.97	0.70
Stockton	0.30	0.09	0.78	3.18	-0.60	0.17	1.68	1.19	0.39	20.27	0.22	0.32	10.21	0.95	0.61
Narrabeen	0.27	0.16	0.61	3.41	-1.49	0.07	1.34	0.87	0.45	22.26	-5.28	0.21	10.18	-3.63	0.58
Collaroy	0.25	0.10	0.79	2.41	-0.56	0.32	1.85	1.49	0.49	15.46	-1.84	0.36	9.20	-3.46	0.62
Bronte	0.26	0.08	0.70	3.94	-0.97	0.05	1.21	0.77	0.22	33.18	3.65	0.26	16.41	2.43	0.65
Maroubra	0.25	0.02	0.73	4.37	-1.28	0.04	1.17	0.57	0.36	32.28	-5.62	0.32	17.45	-4.25	0.67
Woonona	0.21	0.05	0.77	2.88	-0.78	0.15	1.39	0.86	0.39	19.11	3.96	0.30	10.39	5.50	0.48

Fairy-Meadow	0.22	0.00	0.76	3.10	-0.69	0.15	1.18	0.62	0.42	24.86	4.32	0.21	12.02	1.81	0.59
Broulee	0.19	0.02	0.86	2.45	-0.59	0.22	1.51	1.02	0.37	15.53	0.30	0.40	7.46	0.55	0.65
Bengello	0.29	0.16	0.84	2.34	-0.97	0.25	1.12	0.62	0.46	18.86	0.96	0.38	9.81	0.84	0.65
Merimbula	0.20	0.03	0.82	2.73	0.31	0.25	1.96	1.56	0.21	15.98	-5.54	0.41	9.10	-4.20	0.60

Usage Notes

Unlike most existing datasets that focus on individual beaches, this dataset supports regional-scale analysis by providing spatially consistent wave and shoreline data across 307 beaches. This facilitates deeper understanding of regionally consistent or disparate patterns of shoreline change related to subtle differences in wave energy along this coastline. It may also help to serve in the development of shoreline models that can generalize across various beach settings.

The nearshore wave and shoreline datasets can be used independently to investigate historical patterns of wave climate and shoreline change. When linked via the common transect IDs, these datasets enable deeper insights into the relationship between nearshore wave dynamics and shoreline response. This integration supports both qualitative and quantitative analyses of beach change over the past ~ 40 years. Additionally, it can be used for calibration and training of shoreline evolution models at seasonal and interannual scales. Shoreline data were extracted from the Landsat archive. Prior to 2000, when only Landsat 5 was operational, data availability was limited to approximately fortnightly to monthly intervals. With the addition of Landsat 7, 8, and 9, the temporal resolution improved to around weekly by 2024. Despite this improvement, the data may still lack the temporal resolution needed to capture event-scale shoreline changes.

Although the default output shoreline data (time_series_tidally_corrected.csv) is tidally corrected to MSL, given the tide levels (tide_levels_fes2022.csv), the instantaneous waterline (time_series_raw.csv), and the slope (transect_beach_slopes.csv), the shoreline data can be adjusted to other reference water levels following equation 1. The correction for wave runup and setup was not applied because these processes have mixed and site-dependent effects on the accuracy of shoreline position estimates [41]. As most validation sites exhibit a seaward bias

(Figure 5), applying runup and setup corrections would further amplify this existing bias [41]. Consequently, uncertainties associated with wave runup and setup are not explicitly resolved in the provided shoreline positions. These processes may contribute to short-term shoreline variability and residual bias, particularly under energetic wave conditions, and should be considered when interpreting the dataset or applying it to site-specific analyses. We provide (i) the raw shoreline positions and (ii) wave parameters for each transect. This allows users to compute runup and setup appropriate for their specific site and apply customised corrections where required. Spatio-temporal smoothing can also be applied to reduce the noise of shoreline data related to water level variations for shoreline modelling [46].

It is important to note that the nearshore wave data are extracted at the intersections between extended transects and the 10 m depth contour (Figure 8), and that wave transformation landward of this contour is not included. As waves propagate from offshore toward the break point, their direction tends to refract toward the shoreline normal. Because many physics-informed shoreline models (e.g. CERC equation [68]) require wave conditions at the breaking point, using wave parameters sampled at the 10 m contour may introduce additional uncertainty, particularly directional bias, to which models such as CERC are highly sensitive [19, 46]. This issue becomes more pronounced for transects located near headlands or strongly curved coastlines (e.g. Transect 00 in Figure 8), where the extended transect can intersect the contour at a large angle. In such cases, the extracted wave angle of incidence may be larger and less representative of the local shoreline dynamics. To support transparency and informed use of the dataset, NetCDF files for wave data includes the distance from the transect origin to the intersection point and the angle between the transect and the 10 m contour. These metadata can be used to assess the representativeness of the wave data and inform the selection or weighting of transects in subsequent analyses. Nearshore wave conditions are generally more representative for transects with shorter distances to the intersection point and intersection angles closer to 90 degrees, indicating a more perpendicular alignment between the transect and the depth contour. Nearshore reefs landwards of the 10m depth contour can also pose challenges in transforming wave data to the shoreline, resulting in less representative nearshore wave conditions [69].

(Figure 8 goes here)

The metrics for both offshore and nearshore buoys (Table 2&Table 3) are comparable to the existing wave transformation products in NSW [8, 67]. The dominant source of uncertainty is associated with the offshore CAWCR wave hindcasts used as forcing, which exhibit average RMSE values of approximately 0.37 m for significant wave height (H_s) and 3 s for peak period (T_p) when evaluated against buoy observations across the South Pacific [55]. Additional uncertainty arises from non-linear nearshore wave processes that are not fully resolved by the BinWaves framework, contributing RMSE values of approximately 0.11 m for H_s and 1.93 s for T_p when compared against full nearshore wave simulations [67]. As evidenced by the validation results (Table 2&Table 3), the peak wave period (T_p) and peak direction (D_p) are notably less accurate than their mean counterparts (T_{m02} and D_m). This reduced accuracy is largely attributable to the directional and period binning resolution of the offshore wave hindcasts, compared to the resolution of the inshore wave buoys, which is particularly pronounced for D_p. Therefore, for shoreline modelling applications, $D_m$is a more reliable parameter for wave direction. Although T_{m02} also shows better metrics than T_p, the latter is a more direct indicator of shoreline dynamics [26]. In addition, the reconstructed T_p can sufficiently capture the observed temporal trends with metrics comparable to the state of the art [37, 67]. For these reasons, it is still recommended to use T_p as the primary wave-period parameter. This strategy has been adopted in previous shoreline model benchmarking exercises based on this dataset [26, 46].

The current version of the dataset is based on the CAWCR wave hindcasts [55, 56]. However, the archive of the propagation coefficient K_p is independent of the offshore/source wave spectra and enables the reconstruction of nearshore waves from any offshore wave conditions. With a growing number of downscaled wave hindcasts for Southeast Australia now available, which have improved accuracy, higher spatial resolution, and finer directional and period binning [70-72], future iterations of this dataset will evaluate the performance of multiple offshore wave products. The updated dataset will incorporate the best-performing hindcast or an ensemble of wave hindcasts to enhance the robustness and reliability of the nearshore wave estimates.

Data availability

The NSW wave and shoreline datasets are publicly available in the Zenodo repository [63] under the following link (https://doi.org/10.5281/zenodo.17693915).

Code availability

The source code of CoastSat for satellite-derived shoreline and beach slope estimation is available at https://github.com/kvos/CoastSat and archived at a Zenodo repository [73] (https://doi.org/10.5281/zenodo.2779293). The source code for BinWaves is available at https://github.com/GeoOcean/BlueMath/tree/main. The source code for NSW nearshore wave reconstruction and interpolation is available at https://github.com/yongjingmao/BinWaves_NSW/tree/main. In addition, the offshore CAWCR wave data and the propagation coefficients (K_p) used to transform offshore wave conditions to nearshore grid points are archived at a Zenodo repository [74] (https://doi.org/10.5281/zenodo.15678828).

References

1. Nicholls, R.J., J.R. French, and B. van Maanen, *Simulating decadal coastal morphodynamics.* Geomorphology, 2016. **256**: p. 1-2. https://doi.org/10.1016/j.geomorph.2015.10.015
2. Ranasinghe, R., *Assessing climate change impacts on open sandy coasts: A review.* Earth-Science Reviews, 2016. **160**: p. 320-332. https://doi.org/10.1016/j.earscirev.2016.07.011
3. Yates, M.L., R.T. Guza, and W.C. O'Reilly, *Equilibrium shoreline response: Observations and modeling.* Journal of Geophysical Research: Oceans, 2009. **114**(C9). https://doi.org/10.1029/2009JC005359
4. Hans, H., *Genesis: A Generalized Shoreline Change Numerical Model.* Journal of Coastal Research, 1989. **5**(1): p. 1-27.
5. Van Dongeren, A., N. Plant, A. Cohen, D. Roelvink, M.C. Haller, and P.J.C.e. Catalán, *Beach Wizard: Nearshore bathymetry estimation through assimilation of model computations and remote observations.* 2008. **55**(12): p. 1016-1027.
6. Woodroffe, C., D. Callaghan, P. Cowell, D. Wainwright, K. Rogers, and R. Ranasinghe, *A framework for modelling the risks of climate-change impacts on Australian coasts*, in *Applied Studies in Climate Adaptation*. 2014. p. 181-189.
7. Ranasinghe, R., *On the need for a new generation of coastal change models for the 21st century.* Scientific Reports, 2020. **10**(1): p. 2010. 10.1038/s41598-020-58376-x
8. Turner, I.L., M.D. Harley, A.D. Short, J.A. Simmons, M.A. Bracs, M.S. Phillips, and K.D. Splinter, *A multi-decade dataset of monthly beach profile surveys and inshore wave forcing at Narrabeen, Australia.* Scientific Data, 2016. **3**: p. 160024. 10.1038/sdata.2016.24
9. Banno, M., S. Nakamura, T. Kosako, Y. Nakagawa, S.-i. Yanagishima, and Y. Kuriyama, *Long-Term Observations of Beach Variability at Hasaki, Japan.* Journal of Marine Science and Engineering, 2020. **8**(11): p. 871.

10. Castelle, B., S. Bujan, V. Marieu, and S. Ferreira, *16 years of topographic surveys of rip-channelled high-energy meso-macrotidal sandy beach.* Scientific Data, 2020. **7**(1): p. 410. 10.1038/s41597-020-00750-5
11. Larson, M. and N.C. Kraus, *Temporal and spatial scales of beach profile change, Duck, North Carolina.* Marine Geology, 1994. **117**(1-4): p. 75-94.
12. Hoover, D.J., Snyder, A.G., Barnard, P.L., Hansen, J.E. and J.A. Warrick, *Shoreline data for Ocean Beach, San Francisco, California, 2004 to 2021*. 2024. https://doi.org/10.5066/P13CAWLM
13. Ludka, B.C., R.T. Guza, W.C. O'Reilly, M.A. Merrifield, R.E. Flick, A.S. Bak, T. Hesser, R. Bucciarelli, C. Olfe, B. Woodward, W. Boyd, K. Smith, M. Okihiro, R. Grenzeback, L. Parry, and G. Boyd, *Sixteen years of bathymetry and waves at San Diego beaches.* Scientific Data, 2019. **6**(1): p. 161. 10.1038/s41597-019-0167-6
14. Harley, M.D., G. Masselink, A. Ruiz de Alegría-Arzaburu, N.G. Valiente, and T. Scott, *Single extreme storm sequence can offset decades of shoreline retreat projected to result from sea-level rise.* Communications Earth & Environment, 2022. **3**(1): p. 112. 10.1038/s43247-022-00437-2
15. Harley, M.D., I.L. Turner, M.A. Kinsela, J.H. Middleton, P.J. Mumford, K.D. Splinter, M.S. Phillips, J.A. Simmons, D.J. Hanslow, and A.D. Short, *Extreme coastal erosion enhanced by anomalous extratropical storm wave direction.* Scientific Reports, 2017. **7**(1): p. 6033. 10.1038/s41598-017-05792-1
16. Harley, M.D., I.L. Turner, and A.D. Short, *New insights into embayed beach rotation: The importance of wave exposure and cross-shore processes.* Journal of Geophysical Research: Earth Surface, 2015. **120**(8): p. 1470-1484. https://doi.org/10.1002/2014JF003390
17. Boak, E.H. and I.L. Turner, *Shoreline Definition and Detection: A Review.* Journal of Coastal Research, 2005. **2005**(214): p. 688-703, 16.
18. Robinet, A., B. Castelle, D. Idier, G. Le Cozannet, M. Déqué, and E. Charles, *Statistical modeling of interannual shoreline change driven by North Atlantic climate variability spanning 2000–2014 in the Bay of Biscay.* Geo-Marine Letters, 2016. **36**(6): p. 479-490. 10.1007/s00367-016-0460-8
19. Repina, O., R.C. Carvalho, G. Coco, J.A.Á. Antolínez, I. de Santiago, M.D. Harley, C. Jaramillo, K.D. Splinter, S. Vitousek, and C.D. Woodroffe, *Evaluating five shoreline change models against 40 years of field survey data at an embayed sandy beach.* Coastal Engineering, 2025. **199**: p. 104738. https://doi.org/10.1016/j.coastaleng.2025.104738
20. Jaramillo, C., M.S. Jara, M. González, and R. Medina, *A shoreline evolution model for embayed beaches based on cross-shore, planform and rotation equilibrium models.* Coastal Engineering, 2021. **169**: p. 103983. https://doi.org/10.1016/j.coastaleng.2021.103983
21. Calcraft, K., K.D. Splinter, J.A. Simmons, and L.A. Marshall, *Do LSTM memory states reflect the relationships in reduced-complexity sandy shoreline models.* Environmental Modelling & Software, 2025. **183**: p. 106236. https://doi.org/10.1016/j.envsoft.2024.106236
22. Montaño, J., G. Coco, L. Cagigal, F. Mendez, A. Rueda, K.R. Bryan, and M.D. Harley, *A Multiscale Approach to Shoreline Prediction.* Geophysical Research Letters, 2021. **48**(1): p. e2020GL090587. https://doi.org/10.1029/2020GL090587
23. Davidson, M.A., K.D. Splinter, and I.L. Turner, *A simple equilibrium model for predicting shoreline change.* Coastal Engineering, 2013. **73**: p. 191-202. https://doi.org/10.1016/j.coastaleng.2012.11.002
24. Miller, J.K. and R.G. Dean, *A simple new shoreline change model.* Coastal Engineering, 2004. **51**(7): p. 531-556. https://doi.org/10.1016/j.coastaleng.2004.05.006
25. Yates, M.L., R.T. Guza, W.C. O'Reilly, J.E. Hansen, and P.L. Barnard, *Equilibrium shoreline response of a high wave energy beach.* Journal of Geophysical Research: Oceans, 2011. **116**(C4). https://doi.org/10.1029/2010JC006681
26. Splinter, K.D., I.L. Turner, M.A. Davidson, P. Barnard, B. Castelle, and J. Oltman-Shay, *A generalized equilibrium model for predicting daily to interannual shoreline response.* Journal of

Geophysical Research: Earth Surface, 2014. **119**(9): p. 1936-1958. https://doi.org/10.1002/2014JF003106
27. Simmons, J.A. and K.D. Splinter, *Data-driven shoreline modelling at timescales of days to years.* Coastal Engineering, 2025. **197**: p. 104685. https://doi.org/10.1016/j.coastaleng.2024.104685
28. Short, A.D., M.A. Bracs, and I.L. Turner, *Beach oscillation and rotation: local and regional response at three beaches in southeast Australia.* Journal of Coastal Research, 2014(70 (10070)): p. 712-717. 10.2112/SI-120.1
29. Barnard, P.L., A.D. Short, M.D. Harley, K.D. Splinter, S. Vitousek, I.L. Turner, J. Allan, M. Banno, K.R. Bryan, A. Doria, J.E. Hansen, S. Kato, Y. Kuriyama, E. Randall-Goodwin, P. Ruggiero, I.J. Walker, and D.K. Heathfield, *Coastal vulnerability across the Pacific dominated by El Niño/Southern Oscillation.* Nature Geoscience, 2015. **8**(10): p. 801-807. 10.1038/ngeo2539
30. Vos, K., M.D. Harley, I.L. Turner, and K.D. Splinter, *Pacific shoreline erosion and accretion patterns controlled by El Niño/Southern Oscillation.* Nature Geoscience, 2023. **16**(2): p. 140-146. 10.1038/s41561-022-01117-8
31. Mao, Y., D.L. Harris, D.P. Callaghan, and S. Phinn, *Determining the Shoreline Retreat Rate of Australia Using Discrete and Hybrid Bayesian Networks.* Journal of Geophysical Research: Earth Surface, 2021. **126**(6): p. e2021JF006112. https://doi.org/10.1029/2021JF006112
32. Al Najar, M., R. Almar, G. Thoumyre, E.W. Bergsma, J.-M. Delvit, and D.G. Wilson, *GLOBAL SHORELINE FORECASTING USING SATELLITE-DERIVED DATA AND INTERPRETABLE MACHINE LEARNING.* Coastal Engineering Proceedings, 2024(38): p. 219-219.
33. Castelle, B., E. Kras, G. Masselink, T. Scott, A. Konstantinou, and A. Luijendijk, *Satellite-derived sandy shoreline trends and interannual variability along the Atlantic coast of Europe.* Scientific Reports, 2024. **14**(1): p. 13002. 10.1038/s41598-024-63849-4
34. Vitousek, S., P.L. Barnard, P. Limber, L. Erikson, and B. Cole, *A model integrating longshore and cross-shore processes for predicting long-term shoreline response to climate change.* Journal of Geophysical Research: Earth Surface, 2017. **122**(4): p. 782-806. 10.1002/2016jf004065
35. Vitousek, S., K. Vos, K.D. Splinter, L. Erikson, and P.L. Barnard, *A Model Integrating Satellite-Derived Shoreline Observations for Predicting Fine-Scale Shoreline Response to Waves and Sea-Level Rise Across Large Coastal Regions.* Journal of Geophysical Research: Earth Surface, 2023. **128**(7): p. e2022JF006936. https://doi.org/10.1029/2022JF006936
36. Vitousek, S., K. Vos, K.D. Splinter, K. Parker, A. O'Neill, A.C. Foxgrover, M.K. Hayden, J.A. Thomas, L. Erikson, and P.L. Barnard, *Scalable, data-assimilated models predict large-scale shoreline response to waves and sea-level rise.* Scientific Reports, 2024. **14**(1): p. 28029. 10.1038/s41598-024-77030-4
37. Turner, I.L., C.K. Leaman, M.D. Harley, M.C. Thran, D.R. David, K.D. Splinter, N. Matheen, J.E. Hansen, M.V.W. Cuttler, D.J.M. Greenslade, S. Zieger, and R.J. Lowe, *A framework for national-scale coastal storm hazards early warning.* Coastal Engineering, 2024. **192**: p. 104571. https://doi.org/10.1016/j.coastaleng.2024.104571
38. Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, and I. Polosukhin, *Attention is all you need.* Advances in neural information processing systems, 2017. **30**.
39. Mao, Y., R.D.R. Turner, J.M. McMahon, D.F. Correa, D.A. Chamberlain, and M.S.J. Warne, *Predicting Ground Cover with Deep Learning Models—An Application of Spatio-Temporal Prediction Methods to Satellite-Derived Ground Cover Maps in the Great Barrier Reef Catchments.* Remote Sensing, 2024. **16**(17): p. 3193.
40. Xu, J., Y. Zhu, R. Zhong, Z. Lin, J. Xu, H. Jiang, J. Huang, H. Li, and T. Lin, *DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping.* Remote Sensing of Environment, 2020. **247**: p. 111946. https://doi.org/10.1016/j.rse.2020.111946
41. Vos, K., K.D. Splinter, J. Palomar-Vázquez, J.E. Pardo-Pascual, J. Almonacid-Caballer, C. Cabezas-Rabadán, E.C. Kras, A.P. Luijendijk, F. Calkoen, L.P. Almeida, D. Pais, A.H.F. Klein, Y.

Mao, D. Harris, B. Castelle, D. Buscombe, and S. Vitousek, *Benchmarking satellite-derived shoreline mapping algorithms.* Communications Earth & Environment, 2023. **4**(1): p. 345. 10.1038/s43247-023-01001-2

42. Cagigal, L., F.J. Méndez, A. Ricondo, D. Gutiérrez-Barceló, C. Bosserelle, and R. Hoeke, *BinWaves: An additive hybrid method to downscale directional wave spectra to nearshore areas.* Ocean Modelling, 2024. **189**: p. 102346. https://doi.org/10.1016/j.ocemod.2024.102346
43. Ricondo, A., L. Cagigal, A. Rueda, R. Hoeke, C.D. Storlazzi, and F.J. Méndez, *HyWaves: Hybrid downscaling of multimodal wave spectra to nearshore areas.* Ocean Modelling, 2023. **184**: p. 102210. https://doi.org/10.1016/j.ocemod.2023.102210
44. Short, A.D., *Beaches of the New South Wales Coast*. 2007: Sydney University Press.
45. Short, A.D., *Australian Beach Systems—Nature and Distribution.* 2006: p. 11-27. 10.2112/05a-0002.1
46. Mao, Y., G. Coco, S. Vitousek, J.A.A. Antolinez, G. Azorakos, M. Banno, C. Bouvier, K.R. Bryan, L. Cagigal, K. Calcraft, B. Castelle, X. Chen, M. D'Anna, L. de Freitas Pereira, I. de Santiago, A.N. Deshmukh, B. Dong, A. Elghandour, A. Gohari, E. Gomez-de la Peña, M.D. Harley, M. Ibrahim, D. Idier, C.J. Cardona, C. Lim, I. Mingo, J. O'Grady, D. Pais, O. Repina, A. Robinet, D. Roelvink, J. Simmons, E. Sogut, K. Wilson, and K.D. Splinter, *Benchmarking shoreline prediction models over multi-decadal timescales.* Communications Earth & Environment, 2025. **6**(1): p. 581. 10.1038/s43247-025-02550-4
47. Leaman, C.K., M.D. Harley, K.D. Splinter, M.C. Thran, M.A. Kinsela, and I.L. Turner, *A storm hazard matrix combining coastal flooding and beach erosion.* Coastal Engineering, 2021. **170**: p. 104001. https://doi.org/10.1016/j.coastaleng.2021.104001
48. Vos, K., K.D. Splinter, M.D. Harley, J.A. Simmons, and I.L. Turner, *CoastSat: A Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery.* Environmental Modelling & Software, 2019. **122**: p. 104528. https://doi.org/10.1016/j.envsoft.2019.104528
49. Vos, K., *CoastSat*. 2025: GitHub.
50. Xu, H., *Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery.* International Journal of Remote Sensing, 2006. **27**(14): p. 3025-3033. 10.1080/01431160600589179
51. Otsu, N., *A threshold selection method from gray-level histograms.* IEEE transactions on systems, man, and cybernetics, 1979. **9**(1): p. 62-66.
52. Cipolletti, M.P., C.A. Delrieux, G.M.E. Perillo, and M. Cintia Piccolo, *Superresolution border segmentation and measurement in remote sensing images.* Computers & Geosciences, 2012. **40**: p. 87-96. https://doi.org/10.1016/j.cageo.2011.07.015
53. CNES, *FES2022 (Finite Element Solution) Ocean Tide (Version 2022)* 2022. https://doi.org/10.24400/527896/A01-2024.004
54. Vos, K., M.D. Harley, K.D. Splinter, A. Walker, and I.L. Turner, *Beach Slopes From Satellite-Derived Shorelines.* Geophysical Research Letters, 2020. **47**(14): p. e2020GL088365. https://doi.org/10.1029/2020GL088365
55. Durrant, T., D. Greenslade, M. Hemer, and C. Trenham, *A global wave hindcast focussed on the Central and South Pacific*. Vol. 40. 2014: Citeseer.
56. Smith, G.A., M. Hemer, D. Greenslade, C. Trenham, S. Zieger, and T. Durrant, *Global wave hindcast with Australian and Pacific Island Focus: From past to present.* Geoscience Data Journal, 2021. **8**(1): p. 24-33. https://doi.org/10.1002/gdj3.104
57. Ribal, A. and I.R. Young, *33 years of globally calibrated wave height and wind speed data based on altimeter observations.* Scientific Data, 2019. **6**(1): p. 77. 10.1038/s41597-019-0083-9
58. Albuquerque, J., J.A.A. Antolínez, A. Rueda, F.J. Méndez, and G. Coco, *Directional correction of modeled sea and swell wave heights using satellite altimeter data.* Ocean Modelling, 2018. **131**: p. 103-114. https://doi.org/10.1016/j.ocemod.2018.09.001

59. Zornoza-Aguado, M., B. Pérez-Díaz, L. Cagigal, S. Castanedo, and F.J. Méndez, *HyWaThy: Hybrid modeling of nearshore Waves with different baThymetric states.* Coastal Engineering, 2025. **202**: p. 104837. https://doi.org/10.1016/j.coastaleng.2025.104837
60. Booij, N., R.C. Ris, and L.H. Holthuijsen, *A third-generation wave model for coastal regions: 1. Model description and validation.* Journal of Geophysical Research: Oceans, 1999. **104**(C4): p. 7649-7666. https://doi.org/10.1029/98JC02622
61. O'Grady, J.T., Claire; Hoeke, Ron, *Updated Australian bathymetry: merged 250m bathyTopo. v2.* . 2021: CSIRO. https://doi.org/10.25919/cm17-xc81
62. Holthuijsen, L.H., *Waves in oceanic and coastal waters*. 2010: Cambridge university press.
63. Mao, Y., Vos, K., Cagigal, L., Harley, M., & Splinter, K. D., *Shoreline and Nearshore Wave Data for New South Wales, Australia*. 2025, Zenodo: Zenodo. https://doi.org/10.5281/zenodo.17693915
64. Bracs, M., *Efficient monitoring of sandy shoreline variability at the regional scale*. 2016, UNSW Sydney.
65. Manly Hydraulics Laboratory, N.G., *NSW Ocean Wave Data Collection Program*. 2025.
66. Kinsela, M.A., B.D. Morris, T.C. Ingleton, T.B. Doyle, M.D. Sutherland, N.E. Doszpot, J.J. Miller, S.F. Holtznagel, M.D. Harley, and D.J. Hanslow, *Nearshore wave buoy data from southeastern Australia for coastal research and management.* Scientific Data, 2024. **11**(1): p. 190. 10.1038/s41597-023-02865-x
67. Doyle, T.B., A. Bradford, S. Garber, R. Ibaceta, B.D. Morris, M.A. Kinsela, T.C. Ingleton, I. Jizan, D. Taylor, D.J. Hanslow, and K. Bilham, *The New South Wales nearshore wave tool – an interactive platform integrating high-resolution wave data for enhanced coastal science and management.* Environmental Modelling & Software, 2025. **194**: p. 106686. https://doi.org/10.1016/j.envsoft.2025.106686
68. CERC, *Shore protection manual*. Vol. I, 597, II, 603. 1984: US Army Corps of Engineers, Washington DC. 37-53.
69. State Government of NSW and NSW Department of Climate Change, E., the Environment and Water, *NSW seabed landforms derived from marine lidar data 2022*. 2025.
70. Smith, G.M., Alberto; Zieger, Stefan; Trenham, Claire; Hally, Bryan; Seers, Blake; Spillman, Claire; & Hoeke, Ron, *WHACS: Wave Hindcast for the Australian Climate Service.* , CSIRO, Editor. 2024. https://doi.org/10.25919/83kj-kp95
71. Hernaman, V.H., Ron; Hally, Bryan; Trenham, Claire; Seers, Blake; Leighton, Ben; Meucci, Alberto; & Echevarria, Emilio, *Coupled Coastal Hazard Prediction System (CCHaPS) Hindcast for Australia*, CSIRO, Editor. 2025. https://doi.org/10.25919/6tbn-px91
72. Dong, X., Q. Liu, S. Zieger, I.R. Young, R. Li, A. Meucci, J. Sun, K. Wu, and A.V. Babanin, *A 45-year high-resolution unstructured wave hindcast for the Australian coast: Validation and climatological insights.* Coastal Engineering, 2026. **204**: p. 104892. https://doi.org/10.1016/j.coastaleng.2025.104892
73. Kilian Vos, A.T., Kristen Splinter, Sharon Fitzpatrick, Chris Leaman, Spicer Bak, DanieTheron, Darío Hereñú, Jim Paulo Bautista, borgstad, hcastrol, lan Turner., *kvos/CoastSat: CoastSat v3.2 minor bugfixes*. 2025: Zenodo. https://doi.org/10.5281/zenodo.14606615
74. Mao, Y., Splinter, K. D., Cagigal, L., *BinWaves for NSW nearshore wave reconstruction*. 2025: Zenodo. https://doi.org/10.5281/zenodo.15678828

Author Contributions

Y.M. conducted the data validation, wrote the manuscript, and prepared the repository data files.

K.S. supervised the data analysis and secured funding. L.C. developed the code for the BinWaves

approach. K.V. generated the shoreline dataset. M.H. provided the ground-truth data for shoreline validation. K.S., L.C., K.V., and M.H. contributed to writing the manuscript. V.B. contributed to the wave data validation.

Competing Interests

The authors declare no competing interests.

Acknowledgements

Y.M. and K.S. are supported by ARC Future Fellowship FT220100009 and the US Geological Survey Research Co-op (G21AC10672). L.C. acknowledges support from the Government of Cantabria, and the European Union NextGenerationEU/PRTR under projects Perfect-Storm (2023/TCN/003) and CE4Wind (CPP2022-010118). We thank Brendan Crisp for contributing to the wave data validation. We thank Andy Short for providing the Moruya/Pedro beach profile datasets, and Melissa Bracs, Joshua Simmons, Matthew Philips, Tom Beuzen, and Ed Kearney for contributions of additional profile data. We acknowledge Manly Hydraulics Laboratory (MHL), on behalf of NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW), for access to the offshore buoy records available at https://mhl.nsw.gov.au/Data and NSW DCCEEW for the nearshore buoy records available at https://datasets.seed.nsw.gov.au/dataset/nsw-nearshore-wave-buoy-parameter-time-series-data-completed-deployments. Landsat data were provided by USGS and NASA, and accessed via Google Earth Engine. We also acknowledge CSIRO for access to the CAWCR dataset available at https://data.csiro.au/collection/csiro:39819.

Figure legends

Figure 1 Site maps. (a) Map of shoreline-normal transects used to define shoreline positions across NSW. Red points mark the locations of field surveys, with the inset map highlighting transects used by SDS and field surveys at Moruya. (b) Locations of wave buoys in NSW: green points indicate nearshore buoys and red points denote offshore buoys. (c) Bathymetry map of NSW showing the model domains used in wave downscaling. Dashed and solid polygons represent the extent of main and nested domains for the SWAN models with details described in Table 1. Green stars mark offshore CAWCR stations used as boundary conditions.

Figure 2 Workflow of BinWaves. This figure is modified from Cagigal et al. [42].

Figure 3 Binning of CAWCR wave spectra and distribution of wave energy occurrence rates at CAWCR station 657–658. (a) Occurrence rate of wave energy across all spectral bins, plotted in polar coordinates with frequency as the radial axis and direction as the angular axis. Each cell represents a bin defined by a specific frequency and direction; colour shading indicates the frequency of occurrence of wave energy in each bin. (b) Bins retained after filtering, with blank cells indicating bins excluded due to low energy occurrence (<1%) and seaward wave directions.

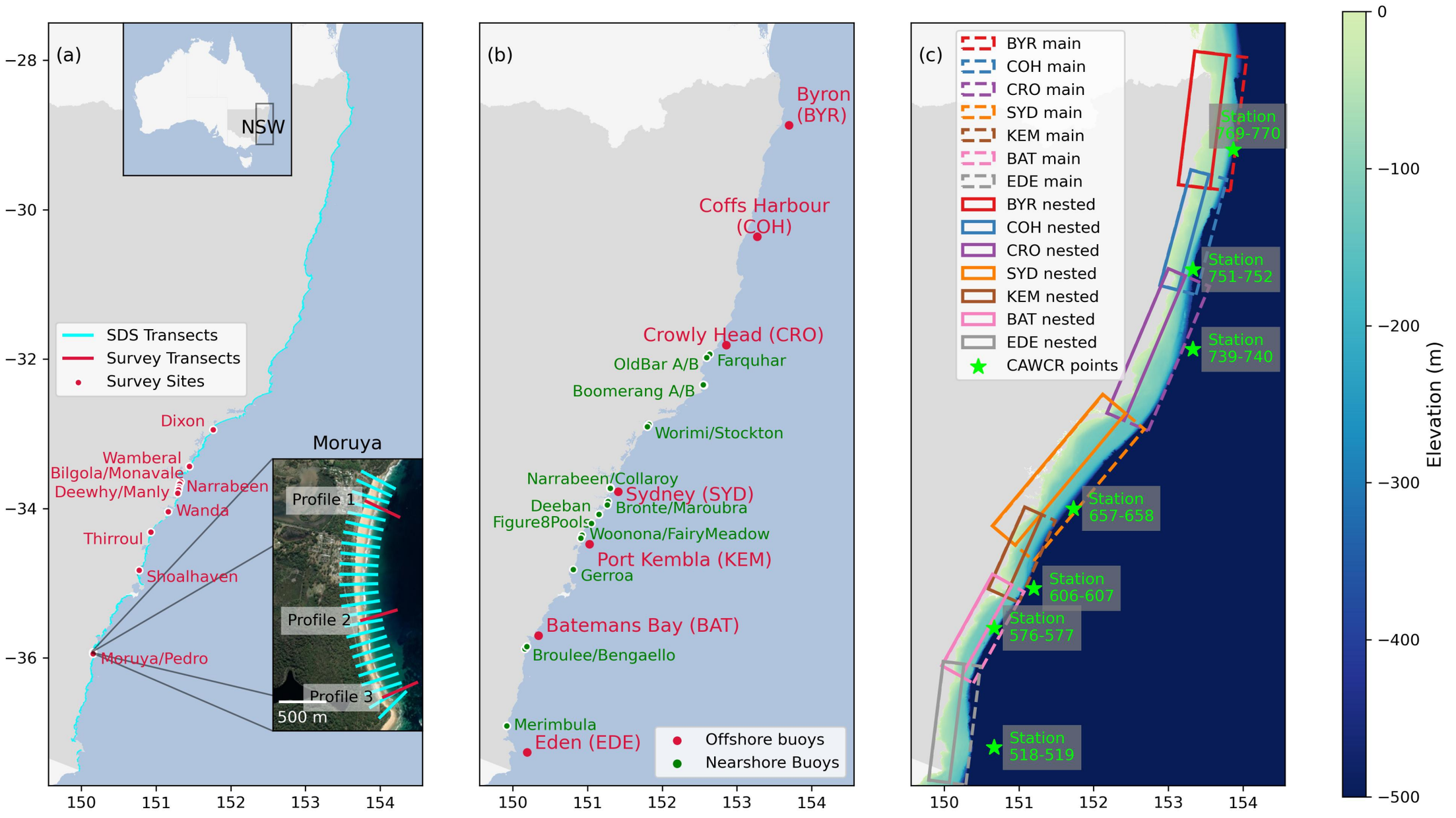
Figure 4 Shoreline validation results for Moruya. (a–c) Time series of shoreline positions from in-situ surveys and SDS for three surveyed transects shown in Figure 1a. (d) Histogram of shoreline position errors, defined as SDS minus in-situ survey values. (e) Scatter plot showing the correlation between SDS and in-situ shoreline positions.

Figure 5 Validation metrics of SDS against field surveys across 12 NSW beaches (locations shown in Figure 1). The colour gradient indicates the average number of in-situ survey observations across all transects per site.

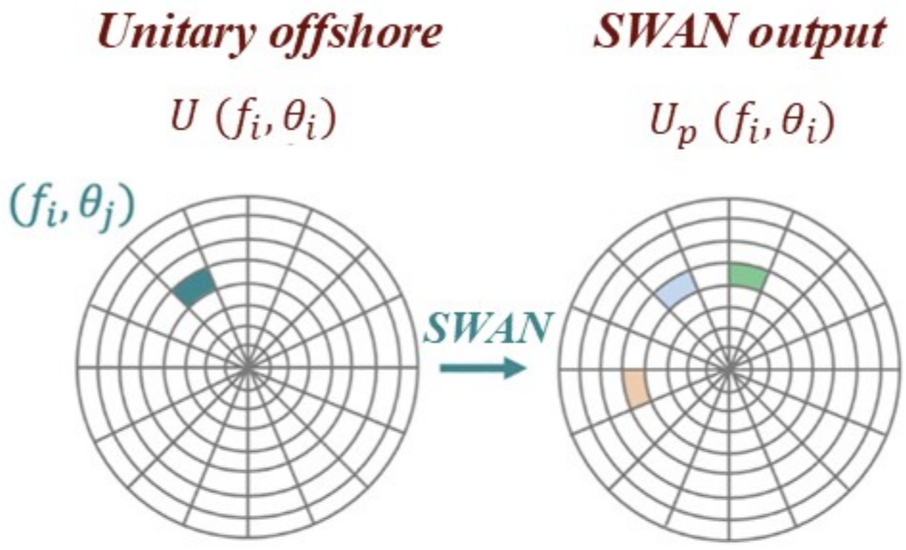
Figure 6 Wave comparison for Sydney offshore wave buoy. (a1)~(d1) Timeseries of buoy and BinWaves reconstructed H_s, T_p, T_{m02} and D_m. Only the data in the last 100 days of buoy record is included in the plot for better visualization but the metrics were calculated based on the whole record. (a2)~(d2) Correlation between buoy and BinWaves modelled wave parameters based on the whole record.

Figure 7 Wave comparison for Woonona nearshore wave buoy. (a1)~(e1) Timeseries of buoy and BinWaves modelled H_s, T_p, T_{m02}, D_p and D_m. (a2)~(e2) Correlation between buoy and BinWaves modelled wave parameters.

Figure 8 Example of wave data extraction for transects at Bondi Beach. Solid black lines show the original transects, while dashed black lines indicate their seaward extensions. Red crosses mark the intersection points with the 10 m depth contour. Arrows illustrate the corresponding average wave directions.

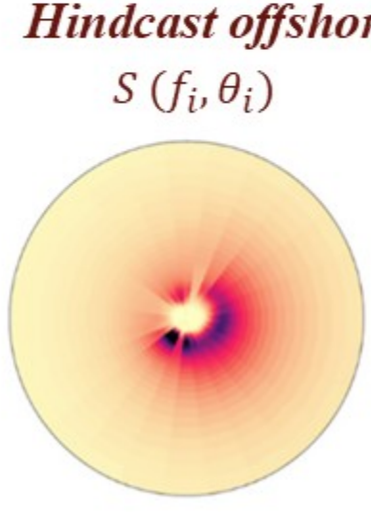

(a)
NSW
SDS Transects
Survey Transects
Survey Sites
Dixon
Wamberal
Bilgola/Monavale
Deewhy/Manly
Narrabeen
Wanda
Thirroul
Shoalhaven
Moruya/Pedro
Moruya
Profile 1
Profile 2
Profile 3
500 m
−28
−30
−32
−34
−36
150
151
152
153
154
(b)
Byron (BYR)
Coffs Harbour (COH)
Crowly Head (CRO)
OldBar A/B
Farquhar
Boomerang A/B
Worimi/Stockton
Narrabeen/Collaroy
Sydney (SYD)
Deeban
Bronte/Maroubra
Figure8Pools
Woonona/FairyMeadow
Port Kembla (KEM)
Gerroa
Batemans Bay (BAT)
Broulee/Bengaello
Merimbula
Eden (EDE)
Offshore buoys
Nearshore Buoys
(c)
BYR main
COH main
CRO main
SYD main
KEM main
BAT main
EDE main
BYR nested
COH nested
CRO nested
SYD nested
KEM nested
BAT nested
EDE nested
CAWCR points
Station 769-770
Station 751-752
Station 739-740
Station 657-658
Station 606-607
Station 576-577
Station 518-519
Elevation (m)
0
−100
−200
−300
−400
−500

Step1: Create library of wave propagation coefficients



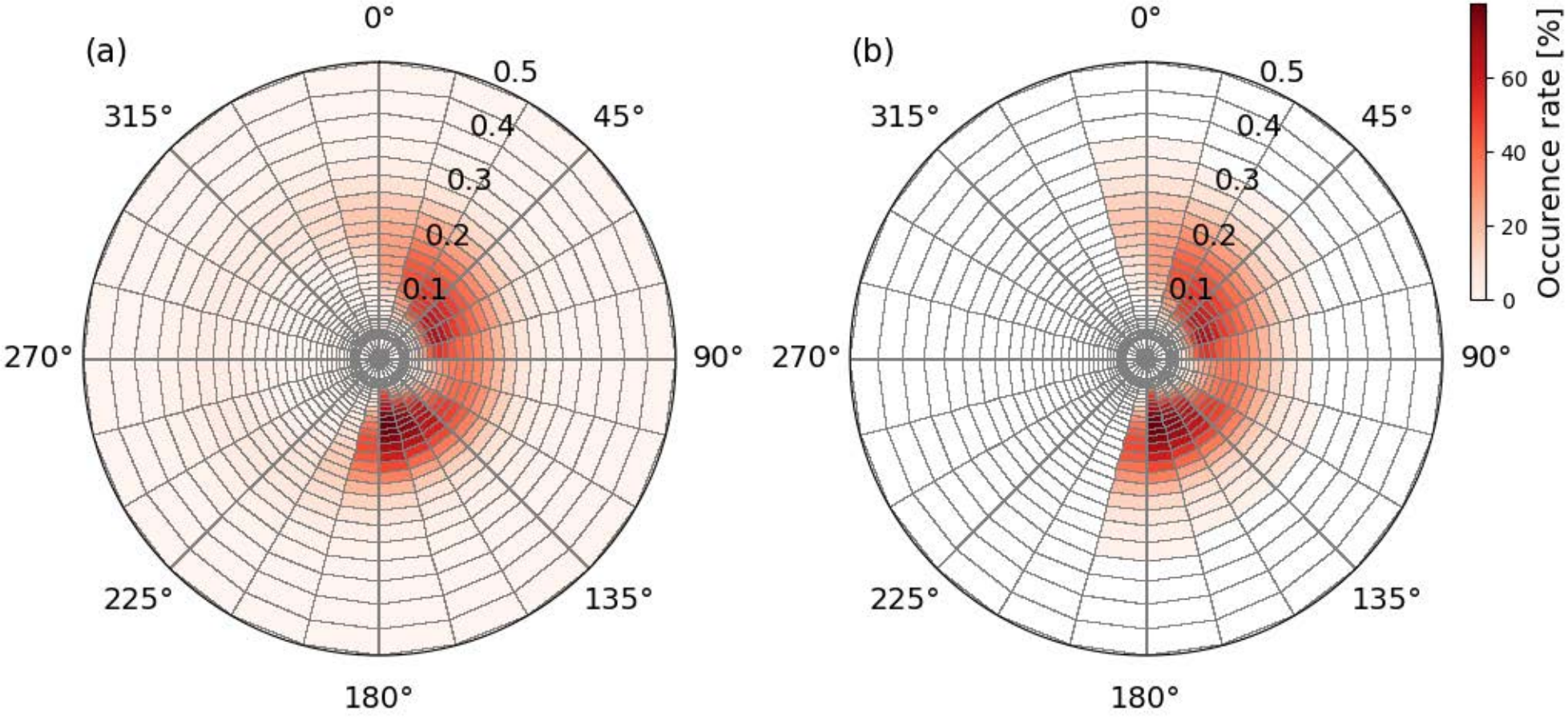
$$Kp\ (f_i, \theta_j) = \frac{U_p(f_i, \theta_j)}{\iint U\ (f_i, \theta_j)\, df\, d\theta}$$

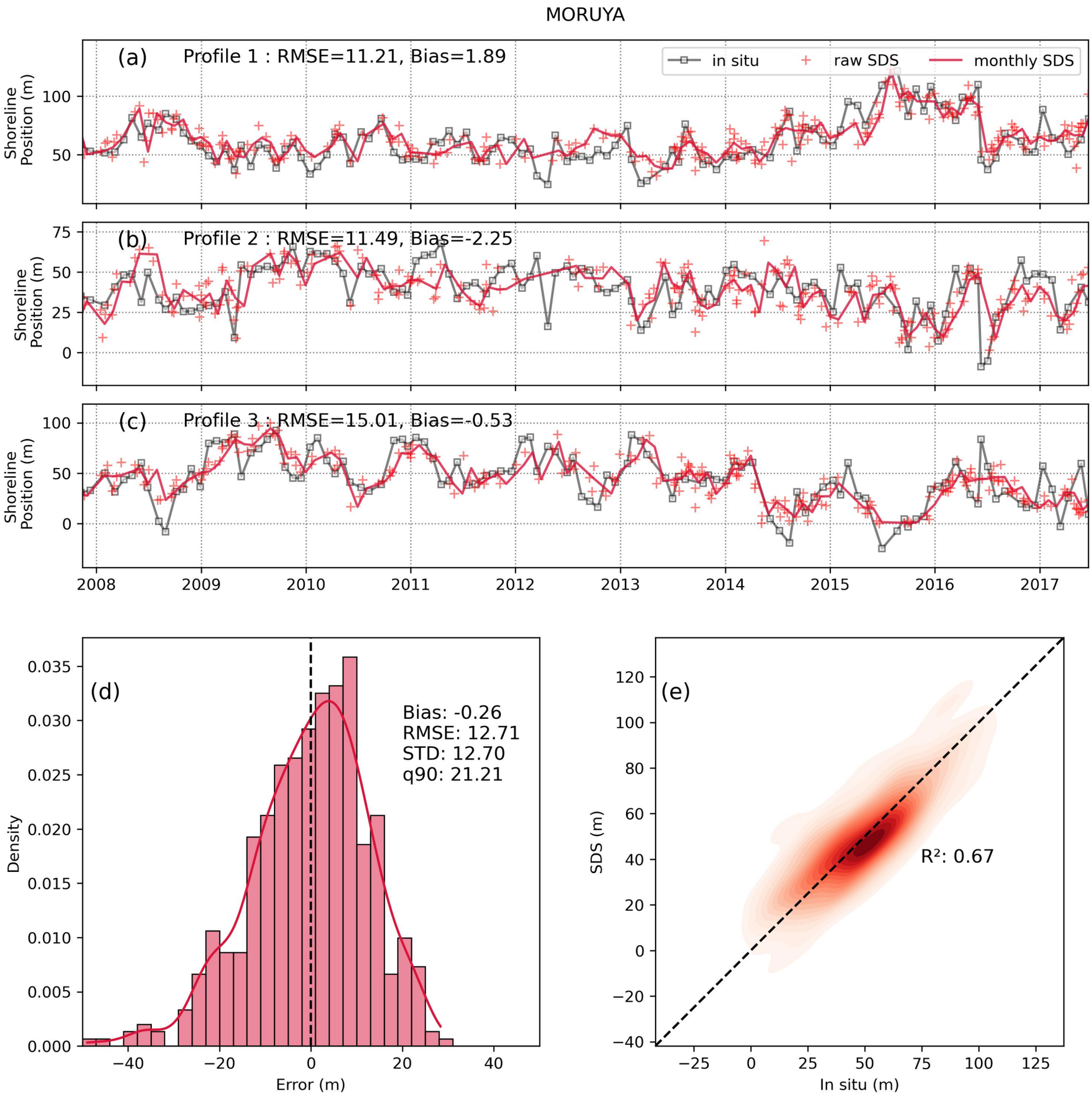
Step2: BinWaves Reconstruction

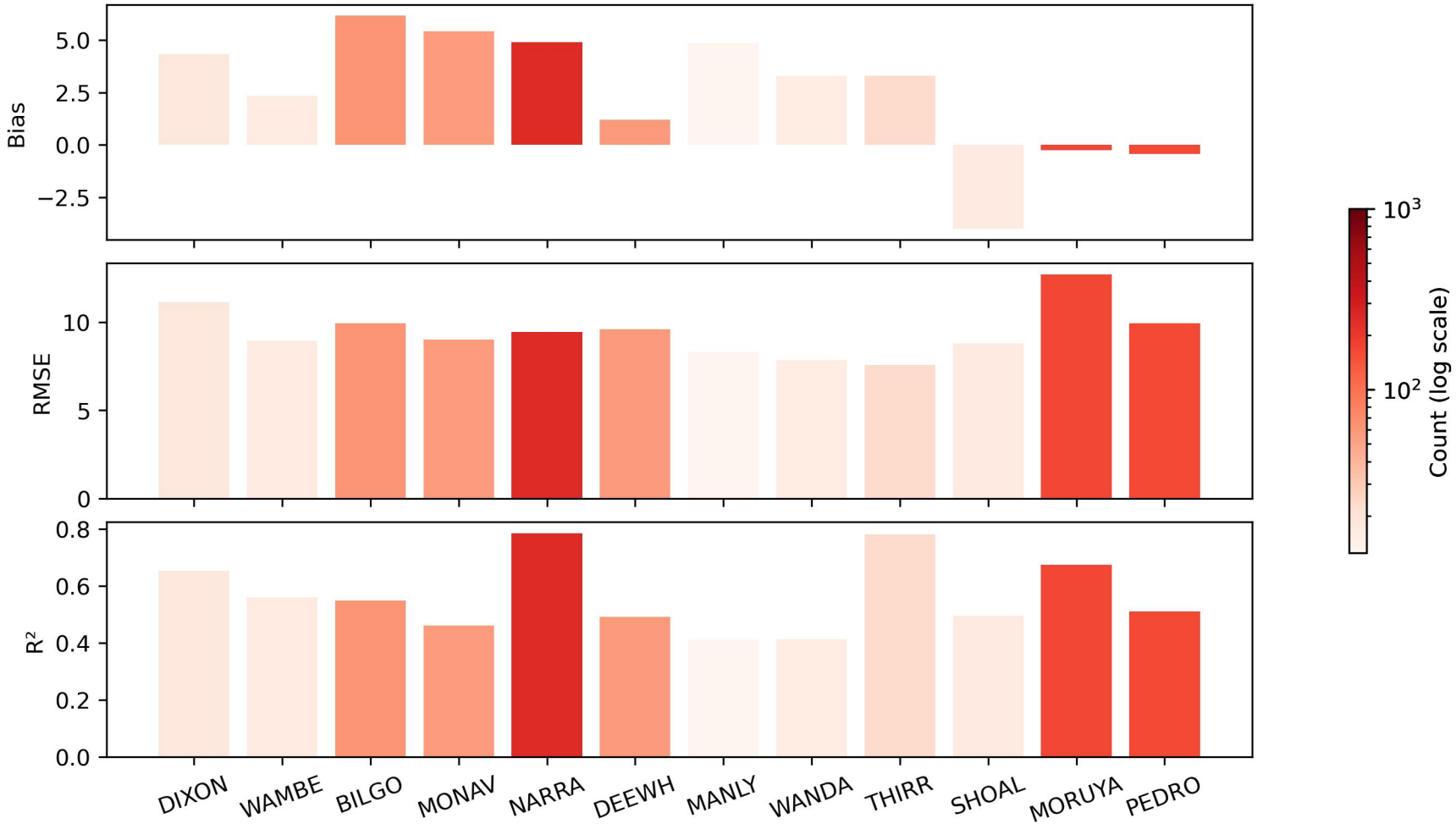


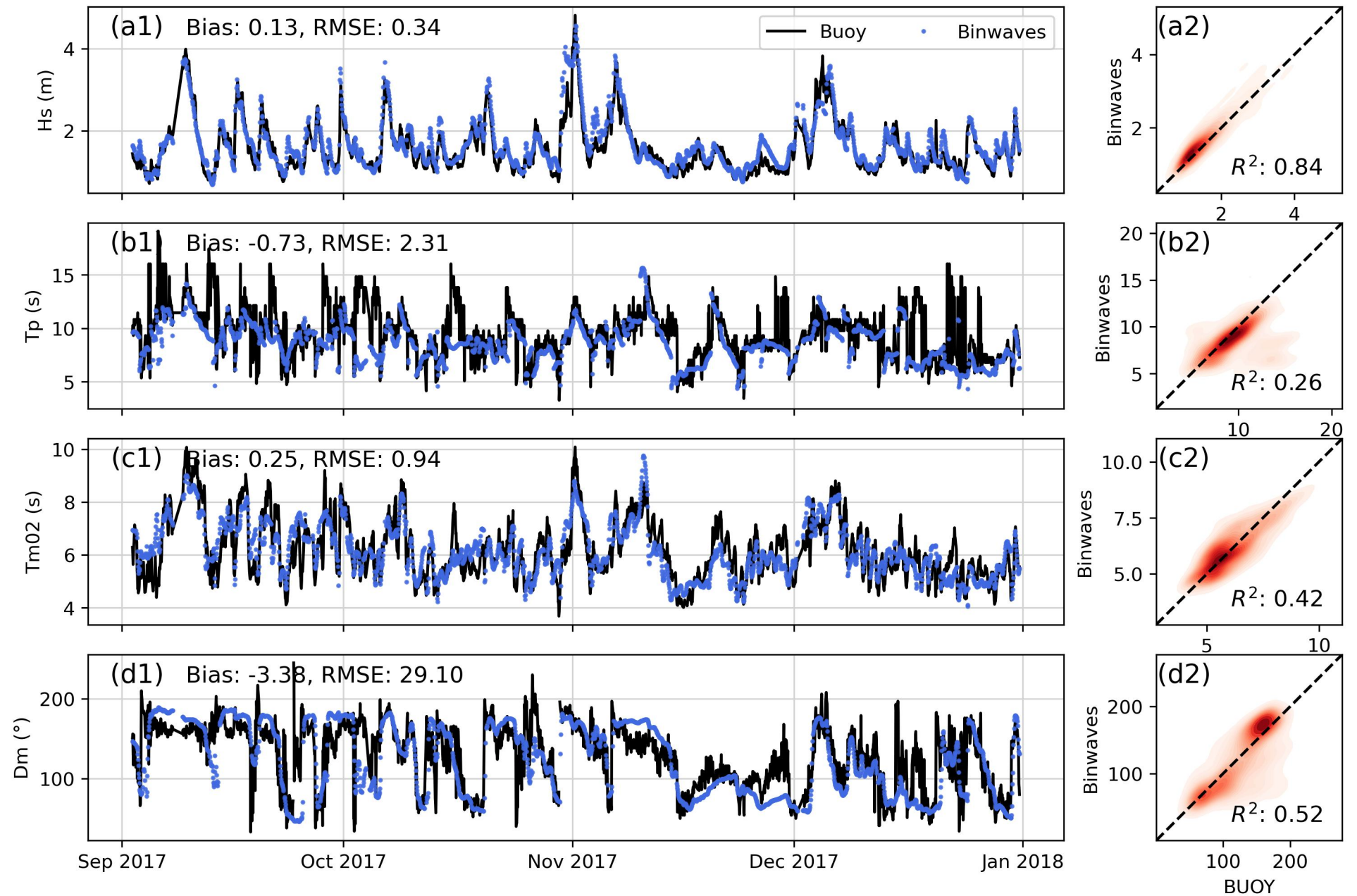
Each point

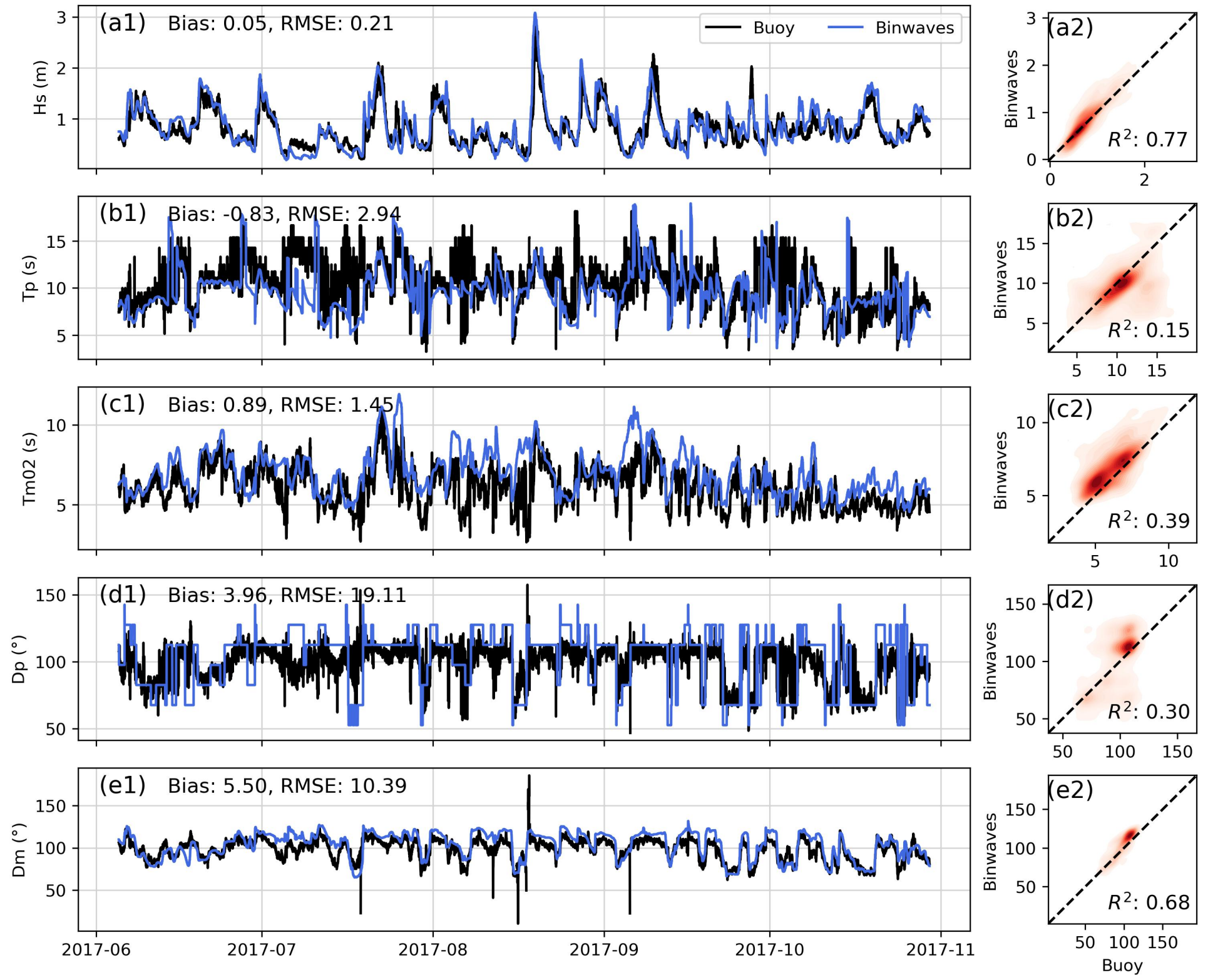
$$Sp\ (f_i, \theta_j) = \sum_i \sum_j S\ (f_i, \theta_j)\ *\ {K_p}^2 (f_i, \theta_j)$$

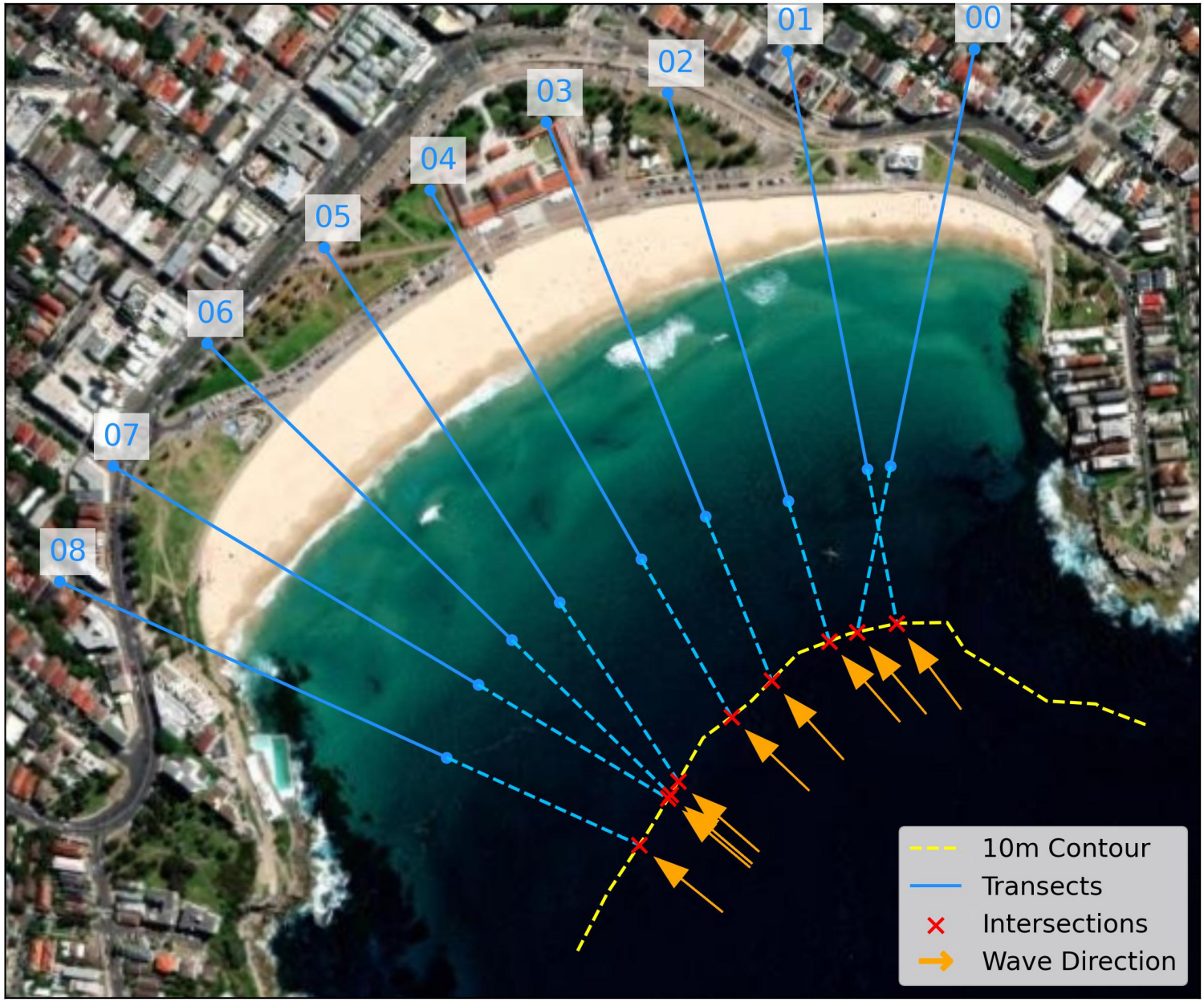

(a)
(b)
0°
45°
90°
135°
180°
225°
270°
315°
0.1
0.2
0.3
0.4
0.5
Occurence rate [%]
0
20
40
60


MORUYA
(a)
Profile 1 : RMSE=11.21, Bias=1.89
in situ
raw SDS
monthly SDS
Shoreline Position (m)
(b)
Profile 2 : RMSE=11.49, Bias=-2.25
(c)
Profile 3 : RMSE=15.01, Bias=-0.53
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
(d)
Bias: -0.26
RMSE: 12.71
STD: 12.70
q90: 21.21
Density
Error (m)
(e)
R²: 0.67
SDS (m)
In situ (m)


Bias
5.0
2.5
0.0
−2.5
RMSE
10
5
0
R²
0.8
0.6
0.4
0.2
0.0
DIXON
WAMBE
BILGO
MONAV
NARRA
DEEWH
MANLY
WANDA
THIRR
SHOAL
MORUYA
PEDRO
10³
10²
Count (log scale)


(a1) Bias: 0.13, RMSE: 0.34
Buoy
Binwaves
Hs (m)
(a2)
Binwaves
R^2: 0.84
(b1) Bias: -0.73, RMSE: 2.31
Tp (s)
(b2)
Binwaves
R^2: 0.26
(c1) Bias: 0.25, RMSE: 0.94
Tm02 (s)
(c2)
Binwaves
R^2: 0.42
(d1) Bias: -3.38, RMSE: 29.10
Dm (°)
(d2)
Binwaves
R^2: 0.52
Sep 2017
Oct 2017
Nov 2017
Dec 2017
Jan 2018
BUOY


(a1) Bias: 0.05, RMSE: 0.21
Buoy
Binwaves
Hs (m)
(a2)
Binwaves
R^2: 0.77
(b1) Bias: -0.83, RMSE: 2.94
Tp (s)
(b2)
Binwaves
R^2: 0.15
(c1) Bias: 0.89, RMSE: 1.45
Tm02 (s)
(c2)
Binwaves
R^2: 0.39
(d1) Bias: 3.96, RMSE: 19.11
Dp (°)
(d2)
Binwaves
R^2: 0.30
(e1) Bias: 5.50, RMSE: 10.39
Dm (°)
(e2)
Binwaves
R^2: 0.68
2017-06
2017-07
2017-08
2017-09
2017-10
2017-11
Buoy


01
00
02
03
04
05
06
07
08
10m Contour
Transects
Intersections
Wave Direction