Proxy observations of surface wind from a globally distributed network of wave buoys

Ciara Dorsay, Galen Egan, Isabel Houghton, Christie Hegermiller, and Pieter B. Smit

Sofar Ocean, San Francisco, USA

Corresponding author: Ciara Dorsay, ciara.dorsay@sofarocean.com

ABSTRACT: In the equilibrium range of the wave spectrum's high frequency tail, energy levels are proportional to the wind friction velocity. As a consequence of this intrinsic coupling, spectral tail energy levels can be used as proxy observations of surface stress and wind speed when direct observations are unavailable. Proxy observations from drifting wave-buoy networks can therefore augment existing remote sensing capabilities by providing long dwell observations of surface winds. Here we consider the skill of proxy wind estimates obtained from observations recorded by the globally distributed Sofar Spotter network (observations from 2021–2022) when compared with collocated observations derived from satellites (yielding over 20000 collocations) and reanalysis data. We consider physics motivated parameterizations (based on frequency^{-4} universal tail assumption), inverse modelling (estimate wind speed from spectral energy balance), and a data driven approach (artificial neural network) as potential methods. Evaluation of trained/calibrated models on unseen test-data reveals comparable performance across methods with generally order 1 m/s root-mean-square-difference with satellite observations.

1. Introduction

In situ observations of ocean surface winds are sparse. The cost of deploying and maintaining deep sea moorings restricts the use of moored buoys to select locations (e.g., the TOA array, Hayes et al. 1991). Ship based anemometer readings (e.g., the voluntary observing ship scheme, Kent et al. 2010) are restricted to major trading routes, often have sub-optimal placement of instruments, are biased towards low wind conditions (due to storm avoidance), and have seen a decline in participation rates from the merchant marine fleet over the last decade (Smith et al. 2019).

While less complex than moored systems, the cost of deploying and maintaining drifting instrument arrays (e.g., buoys with anemometers), still prohibits efforts to establish large oceanic observational networks that directly observe surface wind or stress. As a consequence, remote sensing observations from satellites, rather than in situ observations, are the predominant source of (near) real-time global oceanic surface wind observations (Ribal and Young 2019).

Programs to observe sea surface temperature, surface barometric pressure or surface waves are actively growing through efforts such as the Global Drifter Program (Niiler 2001; Maximenko et al. 2013) or the Sofar Spotter network (Houghton et al. 2021). Given the strong relation between waves and wind – in particular in the high frequency wave spectral tail – wave observations from these efforts could provide valuable proxy observations of surface stress and wind at 10 meter elevation (U_{10}), greatly expanding available surface wind observations.

Proxy observations of wind from spectral wave observations are presently based on assumed proportionality between tail spectral energy levels and wind friction velocities. Observations show that for mature wind seas, energy levels of the frequency spectrum $e(f)$ for frequencies f above the peak scale as $e(f) \propto g u_* f^{-4}$, with u_* denoting the wind friction velocity and a constant of proportionality $2\pi\alpha_{\text{Toba}}$ (Toba 1973; Thomson et al. 2013). Physical motivation for this dependency on u_*, and the f^{-4} shape is based on the assumption of local equilibrium between generation, dissipation and nonlinear interactions in the tail of the spectrum (Phillips 1985). On the open ocean ocean, the evolution of the wave variance density spectrum $E(f, \theta, \boldsymbol{x}, t)$ as a function of frequency, direction θ, space $\boldsymbol{x}$, and time t, is described by a wave energy transport equation of the form

$$\frac{\mathrm{D}E}{\mathrm{D}t} = S_{\text{gen}} + S_{\text{diss}} + S_{\text{nl}} \tag{1}$$

which describes the evolution of wave-energy in (spectral) space and time (left-hand-side) under the action of generation by wind (S_{gen}), dissipation due to white-capping (S_{diss}), and nonlinear interactions (S_{nl}). Assuming statistical stationarity and homogeneity of the wave field in the tail ($DE/Dt \approx 0$), and using approximate forms for the source terms, Phillips (1985) found theoretical direct proportionality between spectral tail levels and friction velocity, as established in earlier experimentations.

Efforts to estimate wind-stress and U_{10} from buoy observed $e(f)$ are generally based on these findings, and reported accuracy of wave-derived U_{10} estimates have been found to be $O(1\text{ m/s})$ when evaluated in coastal regions and moderate ($U_{10} < 10\,\text{m/s}$) wind conditions (Thomson et al. 2013; Voermans et al. 2019; Shimura et al. 2022; Beckman and Long 2022). However, comparison of proxy estimates from Spotter buoys with altimeter-derived wind observations appeared to indicate that buoy-derived wind estimates saturate at 10 m/s and sometimes (severely) underestimate U_{10} under strong wind conditions (Houghton et al. 2021, their Fig. 3).

Here, we revisit the potential of proxy U_{10} estimates for several reasons. Foremost, the saturation reported in Houghton et al. (2021) is likely not due to fundamental saturation of spectral levels, but is rather attributable to algorithmic issues in determining the spectral region actively interacting with the wind (Shimura et al. 2022, e.g., their Fig. 7). Second, initially only bulk parameters were reported by Spotter buoys, but from 2021 onward, wave spectra from all operational Spotter buoys (as of March 2023, 570 buoys) are available hourly. Hence, there is now a 2 year long spectral dataset that can be used to calibrate/train and test different methods to obtain proxy wind observations from the Sofar Spotter network.

We first consider the physics-motivated parameterizations by Voermans et al. (2019) and Shimura et al. (2022) which both relate wind stress to observed representative energy levels, but differ in how they are defined. Second, motivated by advances in understanding of wind-wave interaction (Janssen 1989, 1991; Ardhuin et al. 2009) we consider explicit stationary solutions of the source term balance to determine U_{10}. Lastly, given the size of the observation dataset, and the rapid developments in data driven (or machine learning) methods over the last decade, we train a shallow, artificial neural network to infer U_{10} from raw observational data. To calibrate and test these methods we use collocated observations between wave buoys and satellite altimeter derived wind speeds. In this work we focus on wind speed estimation (because altimeters do not directly report direction), but we will comment on directional estimates in passing.

2. Proxy estimates

a. Physics motivated parameterizations

Observations in mature wind seas show that energy levels in the tail of the wave spectrum are proportional to the wind friction velocity (Toba 1973; Thomson et al. 2013) through Toba's relation

$$\tilde{\epsilon} = \frac{g\epsilon'}{u_*}, \tag{2}$$

where the dimensionless slope spectrum $\tilde{\epsilon}$ is an $O(1)$ empirical constant in the equilibrium range and ϵ' is a representative value of squared slope density $\epsilon(f) = k^2 e(f)$ in the equilibrium range. The wavenumber $k(f)$ is defined through the deep-water dispersion relation $k = (2\pi f)^2/g$, and $\epsilon(f)$ may be interpreted as the spectral squared-slope density since $\int_0^\infty \epsilon(f)df$ represents the mean squared slope. Further, $\tilde{\epsilon}$ may be expressed as (Phillips 1985; Thomson et al. 2013),

$$\tilde{\epsilon} = 2\pi\alpha_{\mathrm{Toba}} = 8\pi\beta I, \tag{3}$$

where β is the proportionality constant between the saturation spectrum and inverse wave age ($0.006 \leq \beta \leq 0.024$, Juszko et al. 1995), and I (≈ 2.5, Thomson et al. 2013) accounts for wind-wave directional misalignment. In this work, $\tilde{\epsilon}$, rather than β and I, is the calibrated parameter with literature values of β and I corresponding to $\tilde{\epsilon}$ as an $O(1)$ parameter ($0.4 \leq \tilde{\epsilon} \leq 1.4$). When calibrated, $\tilde{\epsilon}$ also compensates for errors relating U_{10} to u_* and estimating ϵ'. Therefore, we consider $\tilde{\epsilon}$ a model parameter (absorbing 2π) and not directly representative of α_{Toba}.

Given a u_* estimate from Equation 3, a proxy estimate of U_{10} may be obtained from a constant-stress boundary layer approximation. For neutrally stable conditions, the sustained wind profile is well-represented by a logarithmic profile (Janssen 1989), with U_{10} approximated by

$$U_{10} = \frac{u_*}{\kappa}\log\left(1 + \frac{10}{z_r}\right). \tag{4}$$

Here $\kappa \approx 0.41$ is the von Karman constant and z_r the sea-surface roughness in the presence of waves. Here we approximate z_r through Charnock's relation as $z_r = z_c = \alpha u_*^2/g$ (Charnock 1955), with $\alpha = O(0.01)$ as the Charnock parameter.

Estimation of wind direction θ_{10} depends on the assumption that the mean wave direction θ_w in the equilibrium range is generally aligned with the wind stress direction θ_*, so that $\theta_* = \theta_w$. In the atmospheric boundary layer near the ocean surface, stress and sustained winds are typically aligned; therefore we have, to a good approximation, $\theta_* = \theta_{10} = \theta_w$. To define θ_w, we assume representative values of directional moments in the equilibrium range are available, and (following Kuik et al. 1988) define the wind direction as

$$\theta_{10} = \mathrm{atan2}(b_1', a_1') \tag{5}$$

where a_1', b_1' are the representative moments, and atan2 is the two argument inverse tangent.

Thus, to arrive at an estimate of wind speed from the wave spectrum, a choice of representative values of the spectrum and directional moments, ϵ', a_1', b_1', is necessary. In this work, we explore two different approaches for calculating these representative values.

1) V2019: Best fit approximation for ϵ', a_1', b_1' (Voermans et al. 2019)

Due to sampling, instrument noise, and the idealized assumptions underlying equilibrium range theory, observed spectra will only approximately follow a f^{-4} power law. To account for limitations in real data, Voermans et al. (2019) define the representative value of the compensated spectra ϵ' and the directional moments a_1', b_1' as the mean over a spectral region with size Δf and bounds $f_0, f_0 + \Delta f$ where squared steepness ϵ is approximately constant, i.e.

$$\begin{bmatrix} \epsilon' \\ a_1' \\ b_1' \end{bmatrix} = \frac{1}{\Delta f} \int_{f_0}^{f_0+\Delta f} \begin{bmatrix} \epsilon(f) \\ a_1(f) \\ b_1(f) \end{bmatrix} \mathrm{d}f. \tag{6}$$

The lower bound, f_0, of the best fit frequency interval is found through minimization of the relative difference of a constant relative slope with observed slope spectra over the averaging window,

$$f_0 = \operatorname{argmin} \int_{f_0}^{f_0+\Delta f} \frac{[\epsilon - \epsilon']^2}{(\epsilon')^2} \mathrm{d}f. \tag{7}$$

In practice, integrals are substituted with approximate discrete sums, and the fitting range Δf is effectively a model parameter. In this work, we do not further attempt optimizing Δf, but instead use $\Delta f = 0.2\,\mathrm{Hz}$ as used by Voermans et al. (2019) and onboard the Spotter buoy currently. In the rest of the text, we will refer to this method as V2019.

2) S2022: Max ϵ' approximation for ϵ', a_1', b_1' (Shimura et al. 2022)

As an alternative to the best fit approximation of V2019, Shimura et al. (2022, S2022 hereafter) proposed defining ϵ' as the maximum of ϵ, i.e.,

$$\epsilon' = \epsilon(f_0), \text{ where } f_0 \equiv \operatorname{argmax} \epsilon(f). \tag{8}$$

The representative moments are analogously defined as $b_1' = b_1(f_0), a_1' = a_1(f_0)$. The novel estimation method for ϵ' was principally motivated based on observed algorithm performance, with the resulting proxy estimates of wind speeds by S2022 being superior to V2019 (Shimura et al. 2022). Physically, increased wave steepness is strongly correlated to wind forcing - therefore, it is plausible that the frequencies being actively energized through interaction with the wind contribute most to the mean-squared-slope. Moreover, the peak value is likely closest to the saturated maximum ϵ for a given u_*, and assuming relaxation times are short, may be a good estimate even under changing conditions (e.g., rotating or reduction of winds), and potentially superior to a fitted approach if interaction timescales vary significantly across the tail. In the present context, we will evaluate the algorithm developed by Shimura et al. (2022) solely on performance relative to other methods.

b. Inverse Modelling (IM)

Beyond estimations via the parameterized solutions described above, we consider solving for wind speed and direction directly. Assuming (quasi-)homogeneous and (quasi-)stationary conditions, the source term balance (approximately) closes at all frequency/direction components, so that

$$S_{\text{gen}}(f,\theta;E,U_{10},\theta_{10})+S_{\text{diss}}(f,\theta;E)+S_{\text{dist}}(f,\theta;E)\approx 0. \tag{9}$$

We only consider processes directly associated with a wind-driven sea (wind generation, white-capping, quadruplet wave-wave interaction), so that only S_{gen} has an explicit dependence on (unknown) wind speed and direction. The dependency of source terms on frequency, direction, and known wave spectral densities is implied. Given $E(f,\theta)$ approximated from observations (more on this below), wind speed and direction may in principle be inferred from the above balance, though numerical approximation is required given complex expressions for the different source terms in the balance.

In practice, this is difficult. The spectral distribution of generation and dissipation are not well understood, and modern approximations have been tailored to produce correct results in bulk parameters (specifically significant wave height) when operating on model spectral shapes. However, because quadruplet interactions are conservative (and vanish in the bulk; Hasselmann 1962), wind generation is strictly positive (neglecting transfer from waves to wind), and white-capping strictly negative, the source term balance may be simplified through integration over all frequencies and direction,

$$\int_0^\infty\int_0^{2\pi}\left[S_{\text{gen}}(U_{10},\theta_{10};E)+S_{\text{diss}}(E)\right]\mathrm{d}\theta\mathrm{d}f=S_{\text{gen}}^{\text{bulk}}(U_{10},\theta_{10})+S_{\text{diss}}^{\text{bulk}}=0, \tag{10}$$

This bulk source term balance is expected to be more robust as it does not rely on the intricacies of spectral distribution. To estimate wind direction, we assume that bulk kinematic stress $\vec{\tau}$ is aligned with the mean wind direction such that

$$\vec{\tau}=\underbrace{\iint\frac{gS_{\text{gen}}}{\rho_{\text{a/w}}c}\frac{\vec{k}}{k}\mathrm{d}\theta\mathrm{d}f}_{\vec{\tau}_{\text{wave}}}+\vec{\tau}_{\text{viscous}}+\underbrace{\frac{z_c^2}{z_r^2}\vec{\tau}}_{\vec{\tau}_{\text{background}}}\quad, \tag{11}$$

with $\rho_{\text{a/w}}$ the air/water density ratio, $c=\omega/k$ the wave celerity, $\vec{\tau}_{\text{wave}}$ the contribution to the stress of sea waves, $\vec{\tau}_{\text{viscous}}$ the wind-aligned viscous-stress contribution which is only significant at very low wind speeds, and $\tau_{\text{background}}$ the contribution of unresolved background gravity-capillary waves which is also assumed to be aligned with the wind. The wave stress exerted on the atmosphere is estimated from the rate of change of wave momentum due to energy transfer from the atmosphere to the waves (Janssen 1989). The background stress is parameterized (following Janssen 1989) through a Charnock-like relation, with z_c the roughness length following from Charnock's relation, and z_r the surface roughness length in Equation (4) that relates u_* to U_{10}.

Given expressions for generation and dissipation and an estimate of the directional wave spectrum, Equations (4), (10), and (11) form a system of three coupled nonlinear equations for wind speed, direction (U_{10}, θ_{10}) and surface roughness z_r, which may be solved in an iterative fashion.

1) Source term approximations

To estimate energy transfer from wind to waves (S_{gen}), we use the quasi-linear approximation (Janssen 1991) to model energy transfer due the resonant shear-instability mechanism (Miles 1957),

$$\frac{S_{\mathrm{gen}}}{\omega E} = \rho_{\mathrm{a/w}} \beta \chi^2 \cos^2(\Delta\theta) \tag{12}$$

with $\chi = u_*/c$ the inverse wave age, $\Delta\theta$ the smallest mutual angle between waves and wind ($S_{\mathrm{gen}} = 0$ if the absolute angle exceeds $\pi/2$), and where the Miles parameter β is expressed in terms of the relative critical height μ as

$$\beta = \frac{\beta_{\max}}{\kappa^2} \mu \ln^4 \mu \qquad\qquad \mu = k z_r \exp\left(\frac{\kappa}{(\chi + \chi_0)\cos(\Delta\theta)}\right),$$

with $\beta = 0$ for $\mu > 1$. $\beta_{\max}$ was set to 1.2 in Janssen (1991), but has since essentially been regarded as a model tuning parameter. The wave age tuning parameter χ_0 is typically set to values of 0.006–0.008. Here, both parameters are considered model parameters to be calibrated. For frequencies beyond what is observed ($f > 0.5$ Hz here), we extrapolate the spectrum until $\mu = 1$ using a f^{-5} tail based on the last resolved frequency $f_{\max}$, since a large proportion of the stress is carried by the tail.

To estimate dissipative effects (S_{diss}), we adopt the direction-dependent saturation from Ardhuin et al. (2010), which may be expressed as

$$\frac{S_{\mathrm{diss}}}{\omega E} = C_{\mathrm{sat}} \left(\frac{B(k,\theta) - B_{\mathrm{threshold}}}{B_{\mathrm{threshold}}}\right)^2 \tag{13}$$

where $B_{\text{threshold}}$ is a saturation-based threshold, B is a representative spectral saturation for the given direction, C_{sat} is a tuning coefficient, and $S_{\text{diss}} = 0$ if the saturation is below the threshold (i.e., if $B(f,\theta) - B_{\text{threshold}} < 0$). In terms of the frequency spectrum B is expressed as

$$B(f,\theta) = \frac{c_g}{2\pi} \int_0^{2\pi} F(\theta - \theta') k^3 E(f,\theta')\, \mathrm{d}\theta' \tag{14}$$

where the integration kernel is $F = \cos^2(\alpha)$ if the mutual angle $|\alpha| \leq \theta_0$ and 0 elsewhere (with θ_0 as a calibration parameter). Here we set θ_0 to 80° and no further calibration is attempted.

The justification for this simplified form of the source term balance is our focus on bulk estimates, for which simplicity is preferred since the balances were tuned for use within a wave model and require re-calibration when applied to observational spectra. Specifically, we will calibrate for: the Miles scale parameter β_{max}, the wave age tuning parameter χ_0, the Charnock parameter α, the saturation threshold parameter $B_{\text{threshold}}$ and the breaking strength parameter C_{sat}.

2) Directional spectrum reconstruction

Direct observations of $E(f,\theta) = E(f)D(f,\theta)$ are not available from directional wave buoys, and instead the directional distribution $D(f,\theta)$ (with $\int D \mathrm{d}\theta = 1$) has to be reconstructed based on knowledge of the frequency spectrum and the lowest two Fourier coefficients of the the directional distribution. Here we will use a maximum entropy method (MEM) to define the directional distribution (Kobune and Hashimoto 1986, referred to as 'MEM2') that generally produces spectra which compare favorably to target spectra in controlled settings (Benoit and Teisson 1995) – though field performance is unknown. Preliminary investigation shows that different methods can produce similar skill in terms of wind inference (not shown), though optimum calibration coefficients differ slightly.

c. Data Driven (DD)

In addition to the physics-based estimates, we explored the potential for a purely data-driven algorithm to infer wind speed from observations of directional wave spectra. Recent studies have shown promise in applying data-driven methods to explore the coupling between wind and waves. For example, Peres et al. (2015) were able to extend an observational significant wave height record back by multiple decades by training an artificial neural network on reanalysis wind data. More recently, Shamshirband et al. (2020) compared significant wave height predictions from a numerical wave model with those estimated from a neural network trained on wind data, finding comparable accuracy between the two methods. Tackling the inverse problem, as we are in the present study, Zeng et al. (2016) trained a neural network to predict wind speed based on the echo spectra of high-frequency radar data, which are traditionally used to measure wave height and direction. Relative to ground-truth buoy data, the neural network achieved an root-mean-square-error (RMSE) of 1.7 m/s.

In order to expand on these studies with a global, multi-year dataset, we trained a neural network to learn a mapping from buoy-observed $e(f)$, $a_1(f)$, and $b_1(f)$ to satellite altimeter measurements of U_{10}. Input data were detrended and normalized by their standard deviation (at each frequency) across the training set (see Section 3d for details regarding the training/evaluation/test split). In order to set the network architecture, we conducted a parameter sweep over: the number of hidden rectified linear unit (ReLU) layers (ranging from 1–16), the size of each hidden layer (ranging from 2–128 neurons), and the strength of an L2 regularization term applied to each layer's kernel (ranging from a proportionality factor of 10^{-4}–10^{-1}). The neural networks were constructed using `keras`, and optimized using the Adam scheme (Kingma and Ba 2014) with a Huber loss function. The accuracy of each network was evaluated through the root-mean-square-difference (RMSD) on a 20% evaluation set.

The network that achieved the lowest RMSD on the evaluation set consisted of 2 densely connected ReLU layers with 64 neurons each, followed by 2 densely connected ReLU layers with 32 neurons each, and 2 densely connected ReLU layers with 16 neurons each. The optimal L2 regularization strength was 0.005 at each layer. This architecture and the structure of the input to the neural network are depicted in Fig. 1.

3. Data

a. Buoy observations

Wave spectrum observations used to calculate U_{10} and θ_{10} come from a global, distributed sensor network of several hundred Sofar Spotter buoys (Fig. 3). The Spotter buoy is a surface-following drifter that is approximately spherical in shape with a pentagonal horizontal profile, a mass of 5.5 kg, and a diameter of 38 cm. In the free-drifting configuration, half of the Spotter is submerged

$$X = \begin{pmatrix} e^{(1)}(f) & a_1^{(1)}(f) & b_1^{(1)}(f) \\ \vdots & \vdots & \vdots \\ e^{(m)}(f) & a_1^{(m)}(f) & b_1^{(m)}(f) \end{pmatrix}$$

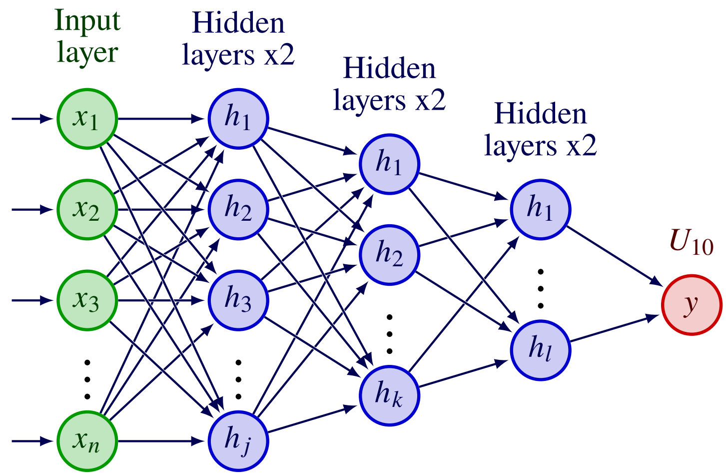

FIG. 1. Neural network architecture for wind speed prediction. The input $X \in \mathbb{R}^{m\times n}$ contains m training examples. Each example (row) has length $n = 294$, and consists of the frequency-dependent variance density and first two Fourier coefficients of the directional distribution. The input layer is followed by three sets of two ReLU layers of sizes $j = 64$, $k = 32$, and $l = 16$, respectively.

beneath the ocean surface (Fig. 2). The top half is exposed, allowing an array of hull-mounted solar panels to continuously power and charge the unit.

FIG. 2. (a) Top-view of Spotter showing array of solar panels which provide power to the unit, allowing it to continuously transmit information. (b) Spotter deployed in Half Moon Bay, CA.

As of September 2021, all Spotter buoys deployed include sensors for barometric pressure and sea surface temperature along with GPS to observe surface waves. The wave spectra are derived using the horizontal and vertical displacements of the unit which are recorded at 2.5 Hz for a period of 30 minutes in the default setting. From horizontal and vertical (co-)spectra the wave energy density $e(f)$ and four directional moments (canonically referred to as $a_1(f), b_1(f), a_2(f), b_2(f)$ Kuik et al. 1988) are calculated. These form the primary directional spectral observations.

For efficient data transmission, a variable spectral resolution is used of approximately 0.01 Hz between 0.03 and 0.35 Hz and a resolution of 0.03 Hz from 0.35 Hz to 0.5 Hz. In this study, spectra are interpolated onto a regular 0.01 Hz grid, and above 0.5 Hz an extrapolated tail (f^{-4} or f^{-5} depending on local best fit on last 10 resolved bins) is appended up to 1.0 Hz such that the integrated energy matched the reported lumped contribution.

Following onboard processing of sensor inputs, Spotter transmits oceanic and atmospheric measurements once every hour through Iridium. Given the current size of the global Spotter network, approximately 14,880 unique information transmissions are available daily. In January 2023, there were 619 actively reporting buoys, a marked increase from early 2019 when the deployment of free-drifting Spotters as part of the Sofar Ocean-owned global drifter network first began (Fig. 3). Transmission was increased beyond the bulk parameters to include the directional spectra in December 2020, which led us to select the subsequent 2 year time period (January 2021 to December 2022) for our wind comparison.

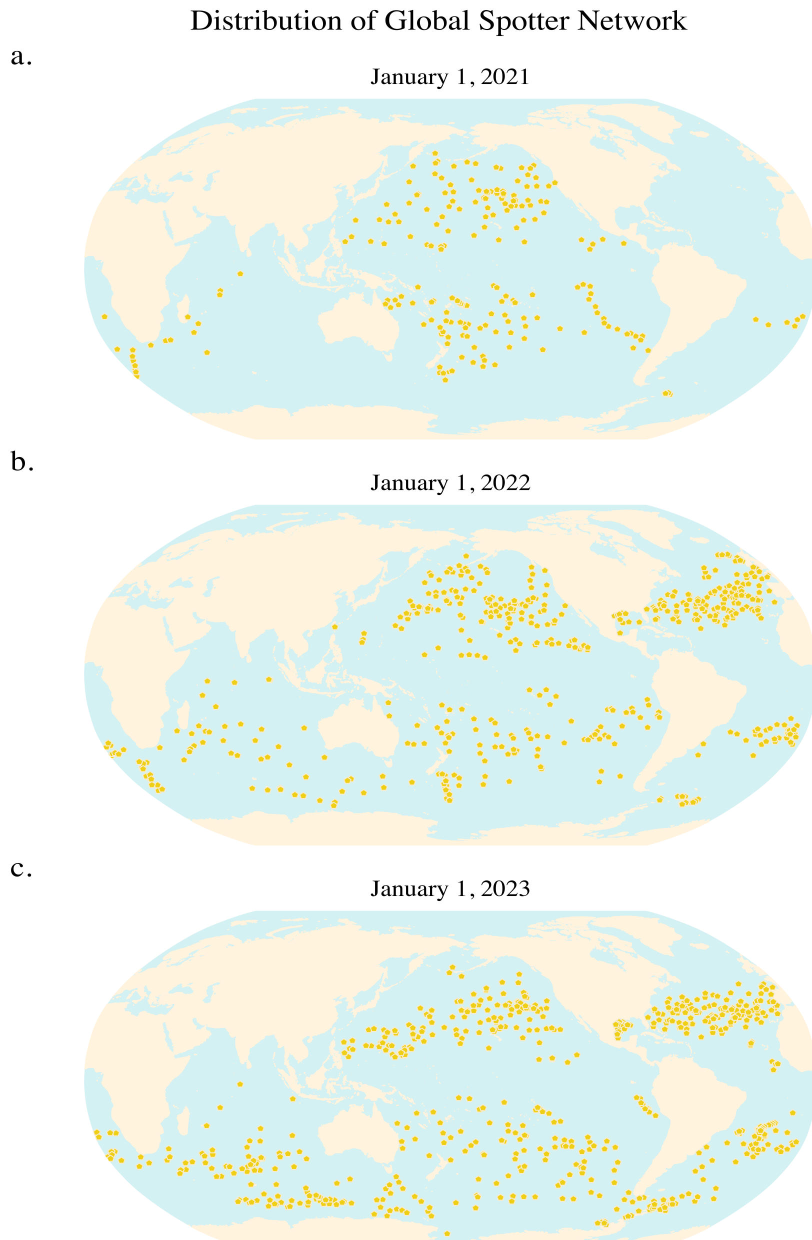

Fig. 3. Distribution of global Spotter network at the beginning of years (a) 2021, (b) 2022, and (c) 2023. On January 1 of each of those years there were 213, 527, and 619 Spotters actively reporting, respectively.

b. Comparison data: Satellite observations and reanalysis data

For comparison data with global coverage, satellite altimeter measurements of wind speed were chosen to assess the skill of the Spotter U_{10} estimation methods. We choose altimeters because they produce estimates of both wind speed and wave height, allowing us to quality-control the satellite observations via the Spotter/altimeter significant wave height mismatch (a large mismatch presumably implies an altimeter error, or that the instruments were not sampling the same sea-state). Data from multiple altimeter platforms were included in the collocation with Spotter data: Jason-3, Satellite with ARgos and ALtiKA (SARAL), and Sentinel-6 Michael Freilich (Sentinel-6). Observations corresponding to non-physical satellite values for U_{10} were excluded from the Spotter comparison.

Due to orbit characteristics and sampling footprints, a large portion of the collocated measurements are associated with an observation made by Jason-3 (43%) and SARAL (45%). Only 12% of the collocated measurements were associated with an observation made by Sentinel-6 due to its later launch date. Reported maximum RMS errors in wind speed observations from altimeters are 1.43 m/s for Jason-3 (Yang et al. 2020), 1.83 m/s for SARAL (Li et al. 2020) and 1.2 m/s for Sentinel-6 (Jiang et al. 2022). Some portion of the error values reported in the Yang et al. (2020), Li et al. (2020), and Jiang et al. (2022) studies can be attributed to the fact that satellite altimeters provide proxy measurements of U_{10} and are therefore subject to their own errors.

In lieu of additional, in situ data sources we used the global ERA5 reanalysis dataset (Hersbach et al. 2020) as an additional point of comparison. For this analysis, we only considered the eastward and northward components of U_{10} from ERA5 (1/4° resolution). For every collocated satellite altimeter/Spotter observation pair, the corresponding ERA5 data was obtained, interpolated in space and time to the altimeter/Spotter observation pair, and converted to magnitude and direction for comparison. Because altimeters do not provide direction estimates, directional information is only available from the model.

c. Triple collocation

In order to obtain estimates of error between the three collocated datasets, we follow an approach outlined in Janssen et al. (2007), which assumes no correlation between the errors associated with each of the wind speed measurement instruments/methods, and a linear relationship between the measurements and the ground truth. The method is only applied to wind speed magnitude as satellite altimeters do not provide directional information. Values for the wind speed linear calibration constants β_{Spotter}, $\beta_{\text{satellite}}$, and β_{ERA5} can be found in Table 1.

Estimation method	β_{Spotter}	$\beta_{\text{satellite}}$	β_{ERA5}
V2019	1.0	0.919	0.912
S2022	1.0	0.982	0.972
IM	1.0	0.949	0.954
DD	1.0	1.026	1.016

TABLE 1. Values of the linear calibration constant β as defined in Janssen et al. (2007) for the four Spotter U_{10} estimation methods.

To note, ERA5 does assimilate satellite altimeter data, specifically from SARAL's AltiKa instrument and other generations of the Jason satellite. However, the primary objective of validating Spotter's estimation of U_{10}, rather than conclusions on independent altimeter or model accuracy, makes this error assessment approach sufficient for the current analysis.

d. Training/calibration and evaluation datasets

To collocate Spotter and satellite data, any observed pair within 25 km and 30 minutes was considered a match. Matching was performed using a kd-tree data structure, in which the latitude, longitude and time triplet were converted to a four dimensional spatial vector $\boldsymbol{x} = [x, y, z, tv]$, with x, y, z the 3D-representation of the latitude/longitude pair (using mean radius of the WGS84 ellipsoid), and tv the time coordinate t expressed as a spatial coordinate using the velocity v defined by the time and space limits, i.e. $v = 25\,\text{km}/30\,\text{min} \approx 14\text{m/s}$ (time deltas of 30 minute are converted to 25 km differences). Any two points A and B for which the Euclidean norm $|\boldsymbol{x}_A - \boldsymbol{x}_B|$ was ≤ 25 km were identified as a match.

If multiple consecutive satellite observations all mapped to the same Spotter observation the observed mean was used as a representative best estimate. To ensure both instruments are sampling a sea state representing the same weather conditions, and to filter for potential outliers, we further restricted matches to data points where observed wave height from satellite and buoy agreed to within 0.25 m. With these restrictions in place the total dataset yielded 21,843 pairs over the two year period, excluding any erroneous observations that were discarded for this analysis.

To train the various models we apply a 50/50 split to the data to form training-evaluation and testing datasets. To avoid biases due to unequal distribution of the Spotter network between 2021 to 2022 the split is performed randomly across the dataset (versus splitting by year). Observed satellite derived wind speeds in both training and testing-evaluation data sets were similarly distributed (Fig.

4). The meta-parameters for the neural network were optimized using a further 80/20 split of the training-evaluation data. To calibrate the physics based estimates all training-evaluation data was used (no further split).

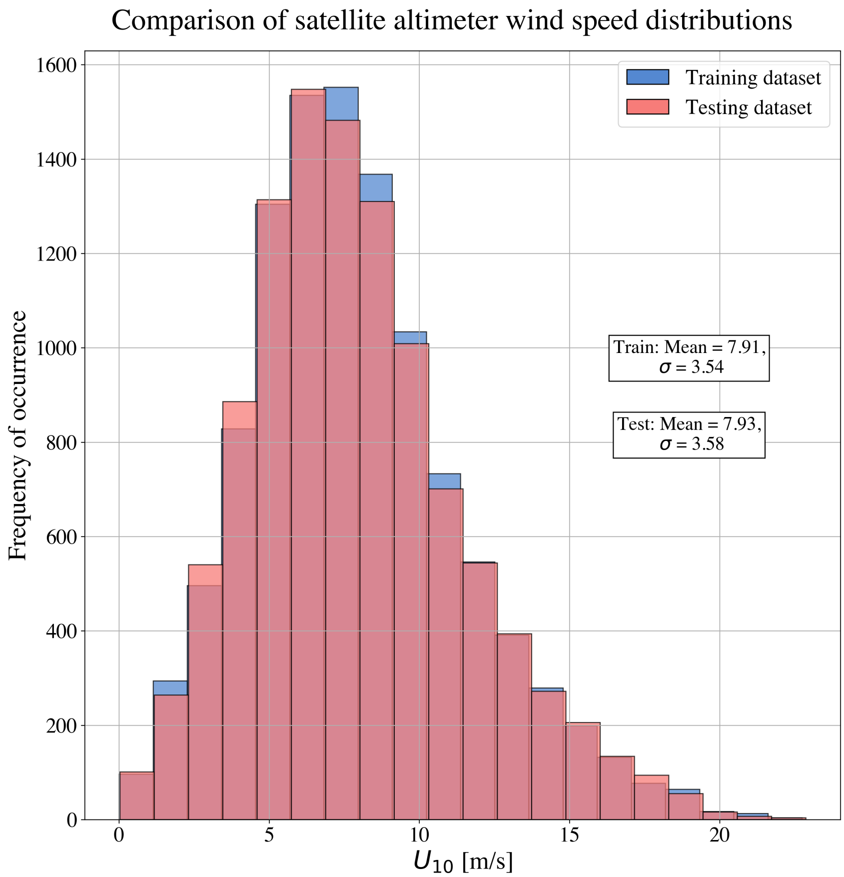

FIG. 4. Distribution of satellite U_{10} observations for the training (red) and testing (blue) datasets used for the evaluation and analysis of the different wind speed estimation methods. Mean and standard deviation values for the distributions indicate that the splitting of the full 2021/2022 dataset did not produce significant biases.

e. Calibration/Training

Calibration of the three physics-based methods was performed using a nonlinear, gradient-descent based optimization algorithm (SLSQP algorithm as implemented in SciPy Virtanen et al. 2020). Optimization was loosely constrained (bounds of 0.01 and 100 times initial value) with initial values given by literature values. We opted for a weighted calibration target to avoid over-fitting on intermediate wind speeds (Fig. 4) because training and test observations are heavily concentrated

Calibration Parameter	Best-fit	Max ϵ'	Voermans et al. (2019)
Dimensionless slope spectrum $\tilde{\epsilon}$	0.55	0.83	0.75
Charnock parameter α_{ch}	0.02	0.018	0.012

TABLE 2. Calibrated parameter values for the physics-based parameterizations (V2019, S2022) compared with literature values.

Calibration Parameter	Inverse model	ST4 value (TEST405 Ardhuin et al. 2009)
Miles scale parameter $\beta_{\max}$	1.57	1.55
Wave age tuning parameter χ_0	0.004	0.006
Charnock parameter α_{ch}	0.012	0.0095
Saturation threshold $B_{\mathrm{threshold}}$	5×10^{-4}	9×10^{-4}
Breaking strength parameter C_{sat}	2.6×10^{-5}	2.2×10^{-5}

TABLE 3. Calibrated values for the inverse model parameters compared to representative ST4 values.

in the 5–10 m/s range. Specifically, calibration/training cost function $\overline{\mathrm{RMSD}}$ was defined as a weighted error,

$$\overline{\mathrm{RMSD}} = \frac{1}{20}\sum_{j=1}^{20} RMSD_j. \tag{15}$$

Here, $RMSD_j$ was defined as the RMSD of all satellite/proxy estimate pairs for which the satellite observation of U_{10} fell within $j-1 <= U_{10} < j$, with values exceeding 20 m/s all collected in the last bin. Calibration on this target reduces overall skill, but significantly improves performance at intermediate wind speeds.

4. Results

a. Calibration/Training

For the physics-based parameterizations (V2019 and S2022 models), RMSD values with the training data set were 1.84 m/s (V2019 model) and 1.43 m/s (S2022 model) when compared to corresponding satellite altimeter observations. Model optimum parameters are of similar order of magnitude to typical literature values, though Charnock values are generally higher (Table 3). Satellite comparison RMSD for the IM method with the training set was 1.23 m/s, and model optimum parameters (Table 3) are generally comparable to values used within operational wave models (e.g. ST4, (Ardhuin et al. 2009)). Lastly, training of the data-driven approach typically converged to RMSD ≈ 1.25 m/s after 30-40 epochs.

b. Wind Speed

Comparison of V2019 with satellite and ERA5 data (Fig. 5) clearly exhibits the saturation earlier observed in Houghton et al. (2021). RMSD (1.84 m/s) and bias (-0.92 m/s) values across the dataset are the highest for V2019. Spread at values for $U_{10} > 15$ is high, with estimates biased low. The weighted calibration does diminish severity of errors (compared with default V2019 parameters, not shown), but at the expense of bias in the mid-range, evident from the curve in the quantile-quantile line.

Performance of the other methods is generally better, with DD obtaining the lowest RMSD (1.16 m/s) value, followed by IM (1.20 m/s) and S2022 (1.42 m/s). Bias is lowest for IM (-0.05 m/s), followed by the DD method (0.12 m/s) and S2022 (-0.27 m/s). All three methods capture data distribution well (quantile-quantile lines close to one-to-one), though the DD approach starts to bias low at high winds, potentially inhibiting its ability to extrapolate beyond 20 m/s wind speeds. All methods perform poorly at the low wind speed values, with generally large scatter compared to satellite observations likely due to buoy limitations. This is addressed further in the discussion.

Errors for all methods tend to increase with increasing wind speed (Fig. 6). The random error for the DD and IM method demonstrate very similar characteristics for $U_{10} > 3$ m/s, with RMSD around 10% of U_{10}. S2022 performs slightly worse across intermediate winds, whereas V2019 generally performs the worst, with particularly high errors of 5 m/s at the upper range and > 1.5 m/s RMSD values at lower wind speeds. Better performance may be gained by non-weighted calibration (comparable to other methods), but at the expense of even larger errors elsewhere (not shown).

High bias for the DD method at higher wind speeds is noteworthy, and indicative of over-fitting on the training data. The sample size at high wind speeds is low and the current approach of weighted calibration likely amplifies over-fitting in this range. Both S2022 and (more-so) V2019 exhibit a bias trade-off from calibration: compensating negative bias in the mid-range with positive bias at the upper range of wind speeds. The observed bias compensation in the physics motivated methods may indicate that the physics are not fully parameterized, which contrasts near zero bias of the IM method output above 3 m/s.

Comparisons with ERA5 data show broadly similar trends, though RMSD values (distributed or bulk) are higher, which is expected if satellite data is closest to truth at the Spotter observation location. Triple collocation results indicate that this is likely the case (Table 4). Regardless of the proxy method, ERA5 error is estimated at $\sim$ 1 m/s, whereas (with more variation) satellite errors are limited to $\sim$ 0.5 m/s. Of the proxy methods, the DD approach has the lowest bulk error magnitudes. Errors associated with the ERA5 and satellite observations are likely underestimated due to the assimilation of satellite observations into ERA5.

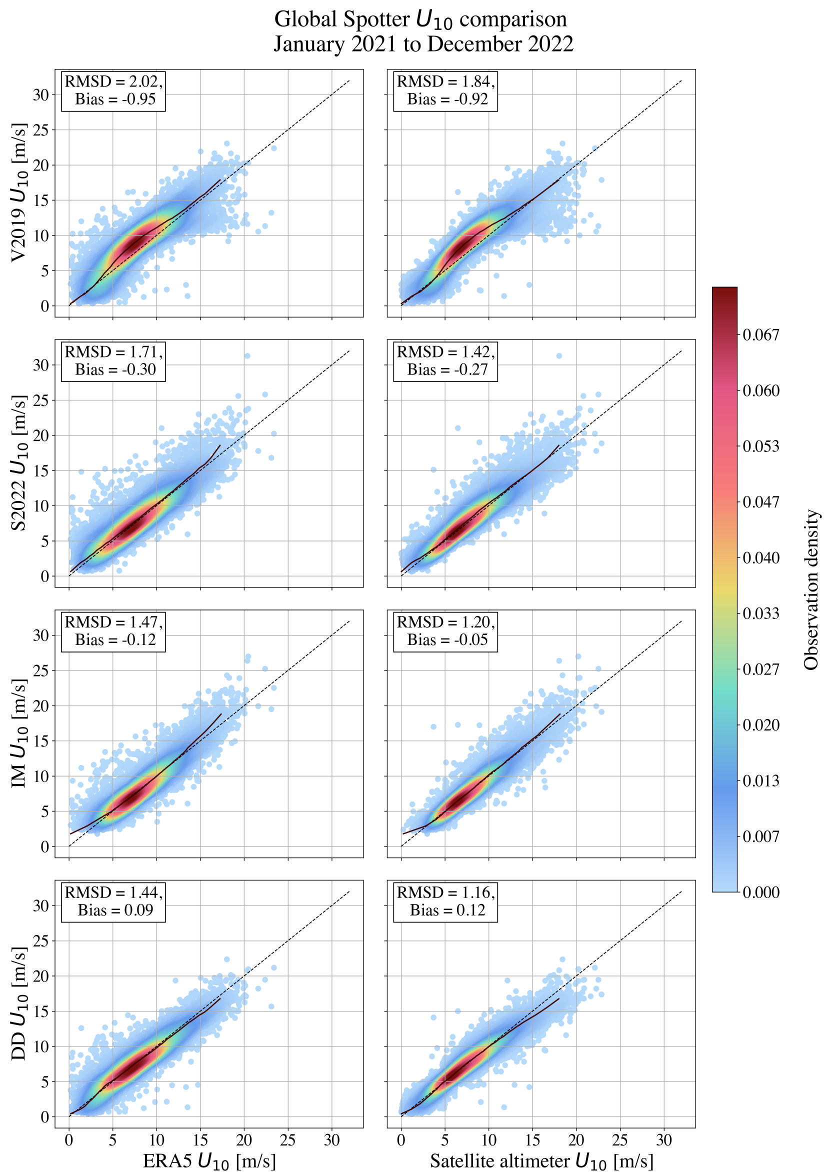

FIG. 5. From top to bottom in the left column, ERA5 U_{10} values are compared to the estimation methods used to produce estimates of U_{10} from Spotter spectra in the following order: (1) V2019, (2) S2022, (3) IM, and (4) DD. From top to bottom in the right column, satellite altimeter U_{10} values are compared to the estimation methods used to produce estimates of U_{10} from Spotter spectra in the following order: (1) V2019, (2) S2022, (3) IM, and (4) DD. Dashed line indicates one-to-one correspondence. The dark, maroon line is the quantile-quantile line.

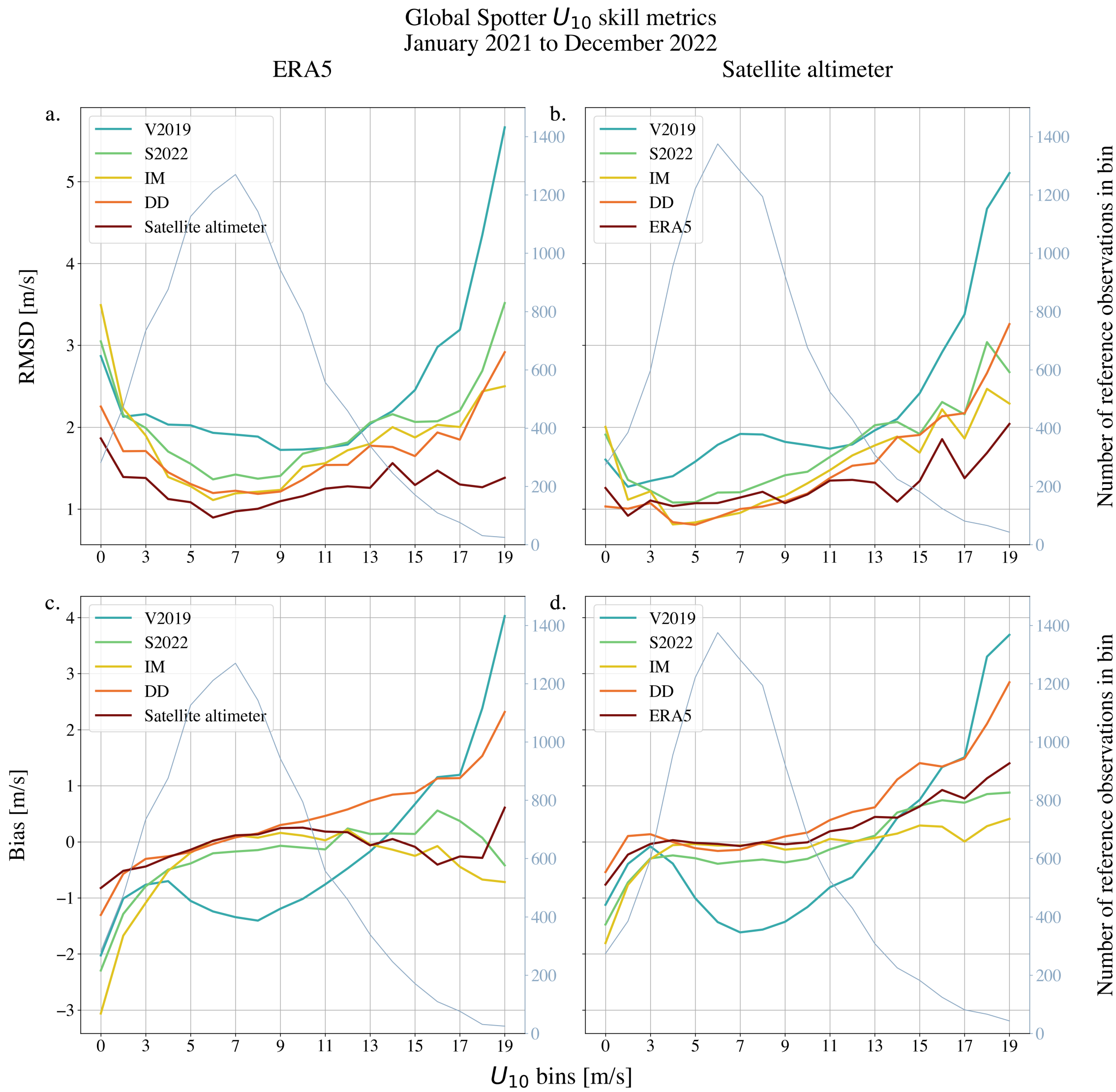

FIG. 6. RMSD (bias) values for U_{10} comparisons between the output from Spotter estimation methods and (a, c) ERA5 as a reference/(b, d) satellite altimeter data as a reference are shown. Spotter/reference observation pairs are binned by the reference U_{10} value (bin edges from 0 m/s to 19 m/s, bin width of 1 m/s for the range 2 m/s to 19 m/s, values above 19 m/s are collected in the last bin). The light blue, gray line indicates the number reference observation in each bin. Bias is defined as the Spotter estimation subtracted from the reference.

c. Wind Direction

Directional estimates from V2019, S2022 and IM perform similarly compared with ERA5 (Fig. 7). The mean difference (smallest mutual angle) is small $O(1°)$, indicating virtually unbiased esti-

Estimation method	e_{Spotter}	$e_{\text{satellite}}$	e_{ERA5}
V2019	1.557 m/s	0.616 m/s	0.986 m/s
S2022	1.327 m/s	0.485 m/s	1.055 m/s
IM	1.167 m/s	0.609 m/s	1.015 m/s
DD	1.002 m/s	0.544 m/s	1.026 m/s

TABLE 4. Values of the residual measurement errors e as defined in Janssen et al. (2007) for the four Spotter U_{10} estimation methods.

mators for direction. Differences are distributed quasi-normally, with smallest standard deviations for the stress-based IM estimate (RMSD 23°) , whereas V2019 and S2022 (both based on tail direction) perform comparably (RMSD of 33.6° and 37.8°, respectively). Directional moments from buoys are generally noisy, therefore both the V2019 and IM methods would likely benefit from a more integrated nature of the estimation, as opposed to the point-like estimate of S2022 implemented here.

No comparison to the data driven method nor satellite altimeters is shown because the satellite instruments included in our analysis do not produce estimates of wind direction, preventing training or direct comparison. This naturally prohibits further definitive conclusions regarding reliability of estimates. That said, the high frequency tail generally reliably follows the wind direction, and errors of $O(20°–30°)$ are in line with previous reported values (Voermans et al. 2019) at coastal sites. We suspect actual error may be lower given that sheltering and fetch limitations (influencing estimates at coastal sites Voermans et al. 2019; Shimura et al. 2022) do not apply.

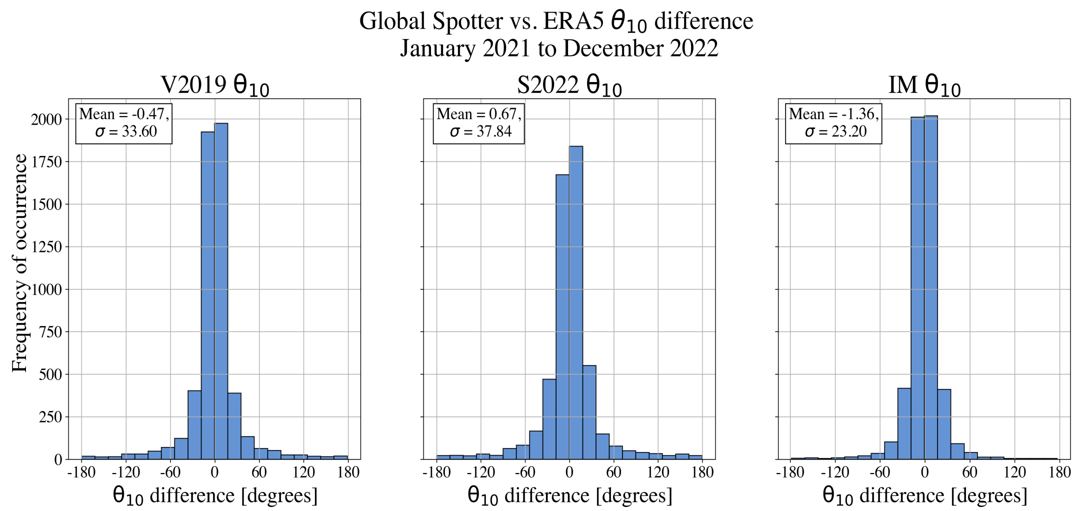

Fig. 7. Differences between ERA5 θ_{10} and θ_{10} as estimated by three of the Spotter wind estimation methods: V2019, S2022, and IM.

5. Discussion and Conclusion

For all but V2019 (best fit) the reported errors of U_{10} are comparable to those obtained from altimeters, and differences among methods are small. Thus, within the $5 \leq U_{10} \leq 25$ m/s range, wave-derived wind observations can augment satellite derived wind products to provide additional long-dwell coverage in deep-water environments. Assessments of the skill of the Spotter wind estimation methods in shallow water environments will require further work.

The more advanced methods (IM and DD) do reduce errors, but judged by this data alone, not by a sufficient margin to justify their more complex implementation. Therefore, we plan to pursue an embedded implementation of S2022 for the Spotter platform using the calibration coefficients derived here.

There are reasons to believe that inference using the IM or DD methods could be further improved. At $O(1\text{ m/s})$, observed differences ("errors") are comparable to those of altimeters when compared to fixed platforms. Assuming altimeter errors are random, $O(1\text{ m/s})$ differences are therefore likely a lower skill limit when calibrating/evaluating against altimeter data. The similarity in error characteristics between the inverse model and the data-driven approach confirms that the remaining error is likely effectively random, but what fraction is attributable to altimeter errors, wave observation noise, or unobserved features (e.g., atmospheric stability, heterogeneity in space or time of wind and waves, etc.) is unknown and requires higher accuracy reference data to investigate further. Paucity of Spotter data collocated with other *in situ* observations currently prevents us from pursuing this calibration further, though the addition of calibrated scatterometer data in future analyses would address the data deficiencies faced in this work.

a. Performance at low wind speeds ($U_{10} < 5\,m/s$)

At very low wind speeds, performance is poor. At O(1 m/s), errors approach 100% and other than qualitative information (e.g., winds are mild, which can have operational use), quantitative utility is low. Reduced skill is in part explained by a change in exchange processes at very low wind speeds, where skin-drag dominates and momentum is directly transferred to currents rather than waves (Kudryavtsev and Makin 2001).

Poor performance can also be linked to the frequency cut-off at 0.5 Hz presently used on Spotter when sending information through Iridium. At 0.5 Hz the wave speed is ~ 3 m/s. Consequently for O(1 m/s) winds, waves and winds are only weakly (or not) interacting in the resolved frequencies ($f \leq 0.5$ Hz) since wave age $\gg 1$. When using full spectra (up to 1.0 Hz) errors in inference may potentially be reduced (e.g. 0.5 m/s error for Spotter at 2 m/s winds were reported by Voermans et al. 2019). In practice, given device dimensions and GPS accuracy this may be a practical lower limit. At 1 Hz the device diameter (~ 0.4 m) is an appreciable fraction of the wavelength (~ 1.5 m) and will display a damped response. Further, heave motions will approach the centimeter scale which is at the limit of what is resolvable from the motion package.

b. Performance at high wind speeds ($U_{10} > 25m/s$)

The collocated altimeter dataset is restricted to wind speeds under 25 m/s, and performance of wind inference from wave measurements at higher wind speeds is unclear. There is reason to doubt the presented methods will extrapolate well to high wind speeds (e.g., in tropical storms). The drag coefficient estimated from Charnock-like relations calibrated on <25m/s winds is known to overestimate drag at wind speeds in excess of 30 m/s (Holthuijsen et al. 2012), even if wave effects on the drag are taken into account (as is done in the inverse model, (Janssen 1991)). This overestimation of roughness will lead to reduced shear in the profile, and consequently an underestimation of wind speed at 10 m height if extrapolated from friction velocity estimates alone.

Further, under strong forcing conditions the f^{-4} equilibrium range vanishes (dissipative range starts at the peak), and assumptions of equilibrium are suspect.

Anecdotally, from samples where Spotters encountered hurricanes (e.g., Hurricane Ian, 2022), we do find (not shown) that neither the IM nor S2022 saturates, and in fact often report comparable wind speeds (up to 50 m/s), while V2022 and the DD method saturate to 30 m/s. Further, output of the IM method and S2022 during Hurricane Ian did not exhibit marked lags (compared with ERA5) in capturing higher wind speeds, indicating that exploring the performance of these methods in high wind regimes is worthwhile. Resolving this is out of scope for the current work, and care should be taken for reported winds well above 25 m/s.

Acknowledgments. The authors acknowledge support from the Office of Naval Research through grants N00014-21-1-2185, N00014-22-1-2394 and N00014-22-1-2405.

Data availability statement. All satellite altimeter data used in this study are openly available from the NASA Physical Oceanography Distributed Active Archive Center (NASA/JPL 2013; Desai 2016; SENTINEL-6 2021). All wind model data used in this study are made openly available via the Amazon Web Services' ECMWF ERA5 Reanalysis bucket (ERA5). Historical data from Spotter buoys, including those used in this study, are available for academic use through Sofar Ocean Technologies by contacting the authors or requesting them online from `https://content.sofarocean.com/free-academic-license`.

To access an example dataset and the tooling necessary to implement the physics based methods outlined in this work, visit the wind-proxy-observations repository hosted at `https://github.com/sofarocean/wind-proxy-observations`.

References

Ardhuin, F., B. Chapron, and F. Collard, 2009: Observation of swell dissipation across oceans. *Geophysical Research Letters*, **36 (6)**.

Ardhuin, F., and Coauthors, 2010: Semiempirical dissipation source functions for ocean waves. part I: Definition, calibration, and validation. *Journal of Physical Oceanography*, **40 (9)**, 1917–1941.

Beckman, J. N., and J. W. Long, 2022: Quantifying errors in wind and wave measurements from a compact, low-cost wave buoy. *Frontiers in Marine Science*, **9**.

Benoit, M., and C. Teisson, 1995: Laboratory comparison of directional wave measurement systems and analysis techniques. *Coastal Engineering 1994*, 42–56.

Charnock, H., 1955: Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society*, **81 (350)**, 639–640.

Desai, S., 2016: Jason-3 GPS based orbit and SSHA OGDR. NASA Physical Oceanography DAAC, URL https://podaac.jpl.nasa.gov/dataset/JASON_3_L2_OST_OGDR_GPS, https://doi.org/10.5067/J3L2G-OGDRF.

ERA5, 2023: ECMWF ERA5 Reanalysis. https://registry.opendata.aws/ecmwf-era5/, Accessed 03/23/2023.

Hasselmann, K., 1962: On the non-linear energy transfer in a gravity-wave spectrum part I. General theory. *Journal of Fluid Mechanics*, **12 (4)**, 481–500.

Hayes, S., L. Mangum, J. Picaut, A. Sumi, and K. Takeuchi, 1991: TOGA-TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bulletin of the American Meteorological Society*, **72 (3)**, 339–347.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049.

Holthuijsen, L. H., M. D. Powell, and J. D. Pietrzak, 2012: Wind and waves in extreme hurricanes. *Journal of Geophysical Research: Oceans*, **117 (C9)**.

Houghton, I., P. Smit, D. Clark, C. Dunning, A. Fisher, N. Nidzieko, P. Chamberlain, and T. Janssen, 2021: Performance statistics of a real-time Pacific Ocean weather sensor network. *Journal of Atmospheric and Oceanic Technology*, **38 (5)**, 1047–1058.

Janssen, P. A., 1989: Wave-induced stress and the drag of air flow over sea waves. *Journal of Physical Oceanography*, **19 (6)**, 745–754.

Janssen, P. A., 1991: Quasi-linear theory of wind-wave generation applied to wave forecasting. *Journal of Physical Oceanography*, **21 (11)**, 1631–1642.

Janssen, P. A. E. M., S. Abdalla, H. Hersbach, and J.-R. Bidlot, 2007: Error estimation of buoy, satellite and model wave height data. *Journal of Atmospheric and Oceanic Technology*, **24**, 1665–1677.

Jiang, M., K. Xu, and J. Wang, 2022: Evaluation of Sentinel-6 altimetry data over ocean. *Remote Sensing*, **15**, 12–38.

Juszko, B.-A., R. F. Marsden, and S. R. Waddell, 1995: Wind stress from wave slopes using Phillips equilibrium theory. *Journal of Physical Oceanography*, **25 (2)**, 185–203.

Kent, E., A. Hall, and V. T. T. Leader, 2010: The voluntary observing ship (VOS) scheme. *Proceedings from the 2010 AGU Ocean Sciences Meeting*, American Geophysical Union, 2000 Florida Ave., N. W. Washington DC 20009 USA,.

Kingma, D. P., and J. Ba, 2014: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kobune, K., and N. Hashimoto, 1986: Estimation of directional spectra from the maximum entropy principle. *International Offshore Mechanics and Arctic Engineering. Symposium. 5*, 80–85.

Kudryavtsev, V., and V. Makin, 2001: The impact of air-flow separation on the drag of the sea surface. *Boundary-layer meteorology*, **98**, 155–171.

Kuik, A., G. P. Van Vledder, and L. Holthuijsen, 1988: A method for the routine analysis of pitch-and-roll buoy wave data. *Journal of Physical Oceanography*, **18 (7)**, 1020–1034.

Li, X., P. Mitsopoulos, Y. Yin, and M. Peñas, 2020: SARAL-Altika wind and significant wave height for offshore wind energy applications in the New England region. *Remote Sensing*, **13**, 57–81.

Maximenko, N., R. Lumpkin, and L. Centurioni, 2013: Ocean surface circulation. *International Geophysics*, Vol. 103, Elsevier, 283–304.

Miles, J. W., 1957: On the generation of surface waves by shear flows. *Journal of Fluid Mechanics*, **3 (2)**, 185–204.

NASA/JPL, 2013: SARAL near-real-time value-added operational geophysical data record sea surface height anomaly. NASA Physical Oceanography DAAC, URL http://podaac.jpl.nasa.gov/dataset/ALTIKA_SARAL_L2_OST_XOGDR, https://doi.org/10.5067/AKASA-XOGD1.

Niiler, P., 2001: The world ocean surface circulation. *International geophysics*, Vol. 77, Elsevier, 193–204.

Peres, D., C. Iuppa, L. Cavallaro, A. Cancelliere, and E. Foti, 2015: Significant wave height record extension by neural networks and reanalysis wind data. *Ocean Modelling*, **94**, 128–140.

Phillips, O., 1985: Spectral and statistical properties of the equilibrium range in wind-generated gravity waves. *Journal of Fluid Mechanics*, **156**, 505–531.

Ribal, A., and I. R. Young, 2019: 33 years of globally calibrated wave height and wind speed data based on altimeter observations. *Scientific data*, **6 (1)**, 77.

SENTINEL-6, 2021: Sentinel-6A MF/Jason-CS L2 P4 Altimeter Low Resolution (LR) NRT Ocean Surface NASA Physical Oceanography DAAC, URL https://podaac.jpl.nasa.gov/dataset/JASON_CS_S6A_L2_ALT_LR_STD_OST_NRT_F, https://doi.org/10.5067/S6AP4-2LSNR.

Shamshirband, S., A. Mosavi, T. Rabczuk, N. Nabipour, and K.-w. Chau, 2020: Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines. *Engineering Applications of Computational Fluid Mechanics*, **14 (1)**, 805–817.

Shimura, T., N. Mori, Y. Baba, and T. Miyashita, 2022: Ocean surface wind estimation from waves based on small GPS buoy observations in a bay and the open ocean. *Journal of Geophysical Research: Oceans*, **127 (9)**, e2022JC018 786.

Smith, S. R., and Coauthors, 2019: Ship-based contributions to global ocean, weather, and climate observing systems. *Frontiers in Marine Science*, 434.

Thomson, J., E. D'Asaro, M. Cronin, W. Rogers, R. Harcourt, and A. Shcherbina, 2013: Waves and the equilibrium range at ocean weather station P. *Journal of Geophysical Research: Oceans*, **118 (11)**, 5951–5962.

Toba, Y., 1973: Local balance in the air-sea boundary processes: III. On the spectrum of wind waves. *Journal of the Oceanographical Society of Japan*, **29**, 209–220.

Virtanen, P., and Coauthors, 2020: SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, **17 (3)**, 261–272.

Voermans, J., A. Babanin, J. Thomson, M. Smith, and H. Shen, 2019: Wave attenuation by sea ice turbulence. *Geophysical Research Letters*, **46 (12)**, 6796–6803.

Yang, J., J. Zhang, Y. Jia, C. Fan, and W. Cui, 2020: Validation of Sentinel-3A/3B and Jason-3 altimeter wind speeds and significant wave heights using buoy and ASCAT data. *Remote Sensing*, **12**, 2079–2096.

Zeng, Y., H. Zhou, H. Roarty, and B. Wen, 2016: Wind speed inversion in high frequency radar based on neural network. *International Journal of Antennas and Propagation*, **2016**.