Tilt Error in NDBC Ocean Wave Height Records

C. O. COLLINS III[a] AND R. E. JENSEN[b]

[a] Field Research Facility, Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, Duck, North Carolina
[b] Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi

(Manuscript received 28 June 2021, in final form 3 February 2022)

ABSTRACT: We identify and characterize an error in the National Data Buoy Center (NDBC) wave records due to the sustained tilt of a buoy under high winds. We use a standard, operational 3-m aluminum discus buoy from NDBC with two wave systems, one gimballed, and the other strapped down but uncorrected. By comparing the two, we find that the most extreme significant wave heights are systematically overestimated. The overestimation is shown to be confined to a region around the peak frequency in the spectra: 0.05–0.15 Hz. Wave direction and directional spread are unaffected. A bias due to tilt error can be observed starting at winds of 10 m s^{-1} or wave heights of 4 m. The bias increases as a function of wind speed and wave height, i.e., the bias is +10% when winds are 20 m s^{-1}. Very high waves and winds are relatively rare, so while the tilt error does not affect overall statistics and basic analyses it could potentially affect analysis sensitive to the extremes. A correction is derived for significant wave height, which is a quadratic function of wind speed. The correction is shown to reduce wave heights in uncorrected records, but is found inadequate for general use. There is evidence of tilt error at other NDBC stations, but the full extent of prevalence in the record is not known at this time.

KEYWORDS: Ocean; North Pacific Ocean; Pacific Ocean; Wave properties; Oceanic waves; Wind waves; Sea state; Wind; Climatology; Automatic weather stations; Buoy observations; Climate records; Data processing/distribution; In situ oceanic observations; Instrumentation/sensors; Bias; Error analysis; Quality assurance/control; Wind effects

1. Introduction

The National Atmospheric and Oceanic Administration maintains the National Data Buoy Center (NDBC) as an operational headquarters for ocean weather data. NDBC data are mission critical to the National Weather Service; NDBC data are ingested and used to produce weather and marine forecasts. Therefore, providing current and reliable oceanographic and meteorological information in support of forecast centers is NDBC's primary purpose, and the data collected typically serve this purpose within about 6 h of being recorded. NDBC buoy design is linked to this mission and thus is based on robustness, reliability, and, in an environment of ever increasing demands and shrinking resources, cost effectiveness. They have a number of standards in place that assure that, on average, their wave measurement products of significant wave height, peak period, and peak wave direction are withing a certain tolerance (NDBC 2009). Most new systems are evaluated against older systems before becoming operational.

NDBC's purpose is important to keep in mind when using their data for purposes outside the scope of their mission. An illustrative example comes from wave climate studies that require long records—record lengths only available from NDBC. Gemmrich et al. (2011) questioned previous studies of trends in the northeast Pacific wave climate, specifically Allan and Komar (2000), Gower (2002), Menéndez et al. (2008), and Ruggiero et al. (2010). Gemmrich et al. (2011) showed that step changes in the mean wave height, an artifact of operational changes to the buoys overtime (or poor data), exacerbated ostensible trends. Moral of the story: small artifacts can go undetected because NDBC's standards and checks are limited by their mission scope, and these artifacts could wreak havoc on your analysis.

The value of a long wave record is obvious in a changing climate, so the more we understand about these artifacts the better. In this spirit, we document another data artifact in the NDBC wave height record. A particular version of NDBC wave measurement system overestimates significant wave heights compared to standard measurements. The overestimation is only significant for the high wave heights ($H_S > 4$ m) records, and thus affects only a very small percentage of the total time series, depending of course on the local wave climate. Although the population of wave heights affected is small, some analyses would be very sensitive to this error including, e.g., extreme value analysis, calculating trends and variability of the 99th+ percentiles.

The artifact is due to heel, or sustained tilt, of the buoy. To understand why this is important, recall the time series of sea surface elevation at a single point can be measured, in principle, by a vertically stabilized accelerometer on a platform that follows the surface. This is the physical basis of a wave measurement from a buoy: they carry an accelerometer and bob with the waves. Within the buoy, an accelerometer can be suspended in a viscous fluid on a gimbal so that as the buoy pitches and rolls the accelerometer is always measuring the true vertical. This is the design of the iconic HIPPY 40 sensor (abbreviated as HIPPY) produced by Datawell BV. If not gimballed, the sensor is mounted in a static position within the buoy. This is referred to as "strapped down," and the accelerometer collects data in the frame of reference of a buoy

Denotes content that is immediately available upon publication as open access.

Corresponding author: Tripp Collins, Clarence.O.Collins@usace.army.mil

DOI: 10.1175/JTECH-D-21-0079.1

that has the usual 6 degrees of freedom: heave, pitch, roll, surge, sway, and yaw.

A strapped-down accelerometer could introduce a bias into significant wave height given a sustained heel. There can be over- or underestimation depending on the relative orientation of the heel to the axis of the buoy's heave accelerometer. This is demonstrated in Bender et al. (2010), hereafter B10) and Riley et al. (2011), hereafter R11). Given the 6 degrees of freedom are measured (e.g., from an inertial motion unit) and the orientation on Earth (compass bearings), one can account for heel using an Euler angle transformation (e.g., Anctil et al. 1994). Assuming small angles, the transformation effectively takes the measurement from a buoy reference frame to an Earth reference frame giving true vertical displacement (Drennan et al. 2014; Collins et al. 2014).[1]

The buoy feels a drag force due to wind on the air side of the buoy. Wind pushes the buoy in space, once the mooring nears its maximum extent it limits the movement of the buoy with force in the opposing direction applied to bottom of the buoy. The balance of wind and mooring forces results in heel. The heel projects a component of gravity out of the vertical, which is measured as an acceleration. Here we show that the overestimation is shown to be confined to frequencies 0.05–0.15 Hz; this band corresponds with the spectral peak when wave heights are large. Wave direction and directional spread are relatively unaffected. We observe the effect starting at winds of 10 m s^{-1}, and it increases as a function of wind speed. Because high winds and high waves tend to cooccur, the effect is shown to be most prevalent in high waves. A correction is derived for significant wave height, which is a quadratic function of wind speed. The correction is shown to reduce wave heights in uncorrected records, but it is not appropriate for general use. There is evidence of this effect at other NDBC stations, but the full extent of prevalence in the record is not known at this time.

2. Methods

In this study we analyze data from a strapped-down sensor prior to the implementation of tilt correction. The data come from a standard operational NDBC 3-m aluminum discus buoy deployed with two wave sensor packages—one gimballed, one strapped down—at station 46029. NDBC station 46029, known as Columbia River Bar, is located on the Pacific coast continental shelf approximately 37 km west of Columbia River Mouth in depth of 134 m (see Fig. 1). The Pacific Northwest is known for large waves produced by strong winter lows (e.g., Tillotson and Komar 1997). The strapped-down system used a MicroStrain 3DM-GX1 accelerometer (3DMG), processed with NDBC's Digital Directional Wave Module (DDWM) v2.0 (Teng et al. 2009). The other system was a HIPPY 40 sensor processed by NDBC's Wave Processing Module (WPM) (Chaffin et al. 1994). The dual-sensor time series is about 5 years long, from 2011 to 2016. Some analysis on these data has been presented by NDBC at conferences and are generally consistent with our findings (Bouchard et al. 2013, 2014, 2016). The total number of paired observations for the period of record was 30 893. In addition to buoy 46029, Coastal Data Information Program (CDIP) station 179 (NDBC 46248) was moored about 15 km to the west in a depth of 175 m. CDIP station 179 was a Datawell Directional Waverider using a HIPPY 40 sensor. The total number of tripleted observations was 23 207. Of special relevance is the comparisons by O'Reilly et al. (1996) and Jensen et al. (2021). O'Reilly et al. (1996) showed good agreement between a NDBC 3-m discus buoy with a HIPPY sensor and a Datawell Waverider on wave energy and mean wave direction in the spectral band 0.06–0.14 Hz. Jensen et al. (2021) showed excellent agreement between a tilt-corrected 3DMG and a HIPPY 40 that were mounted on the same 3-m discus buoy, their data serve as a benchmark from which we present comparable figures in the appendix.

The 3-m discus buoy has a weather vane on the frame that is a standard part of its design, an example of which is shown in Fig. 2. Thus in high winds, the buoy is always oriented within 45° of the wind (not shown).

Frequency spectra and the four directional moments are available from each sensor package on the 3-m discus and the Waverider. The NDBC wave modules produced similar spectra for both 3DMG and HIPPY sensors. For comparison, the Waverider spectra were interpolated to the NDBC spectral bands, and then for all we used the range 0.035–0.485 Hz. Spectral parameters were recalculated from the frequency spectra $S(f)$ as follows:

$$m_n = \int_{0.485}^{0.035} f^n S(f) df, \qquad (1)$$

$$H_S = H_{m_0} = 4\sqrt{m_0}, \qquad (2)$$

$$T_{p4} = \frac{\int_{0.485}^{0.035} S^4(f) df}{\int_{0.485}^{0.035} f S^4(f) df}. \qquad (3)$$

The spectral moments m_n are frequency weighted integrals over the spectrum; H_S is the significant wave height, which is the mean of the highest 1/3 wave heights; and T_{p4} is a metric for the frequency of the spectral peak (Young 1995). The directional parameters of mean wave direction $\theta(f)$, directional spread $\sigma(f)$, and mean direction at the spectral peak $\theta(f_p)$ were calculated from the first two directional coefficients, $a_1(f)$ and $b_1(f)$, which relate the co- and quadrature spectra to the directional–frequency spectrum (Longuet-Higgins et al. 1963):

$$\theta(f) = \arctan\left[\frac{b_1(f)}{a_1(f)}\right], \qquad (4)$$

$$\sigma(f) = \sqrt{1 - \sqrt{(a_1^2 + b_1^2)}}, \qquad (5)$$

$$\theta_p = \theta(f_p) \text{ and } f_p = \max_f[S(f)]. \qquad (6)$$

[1] This procedure is also referred to as a "tilt correction" or "numerical vertical stabilization," and the error caused by not accounting for it is sometimes referred to as "the Bender effect."

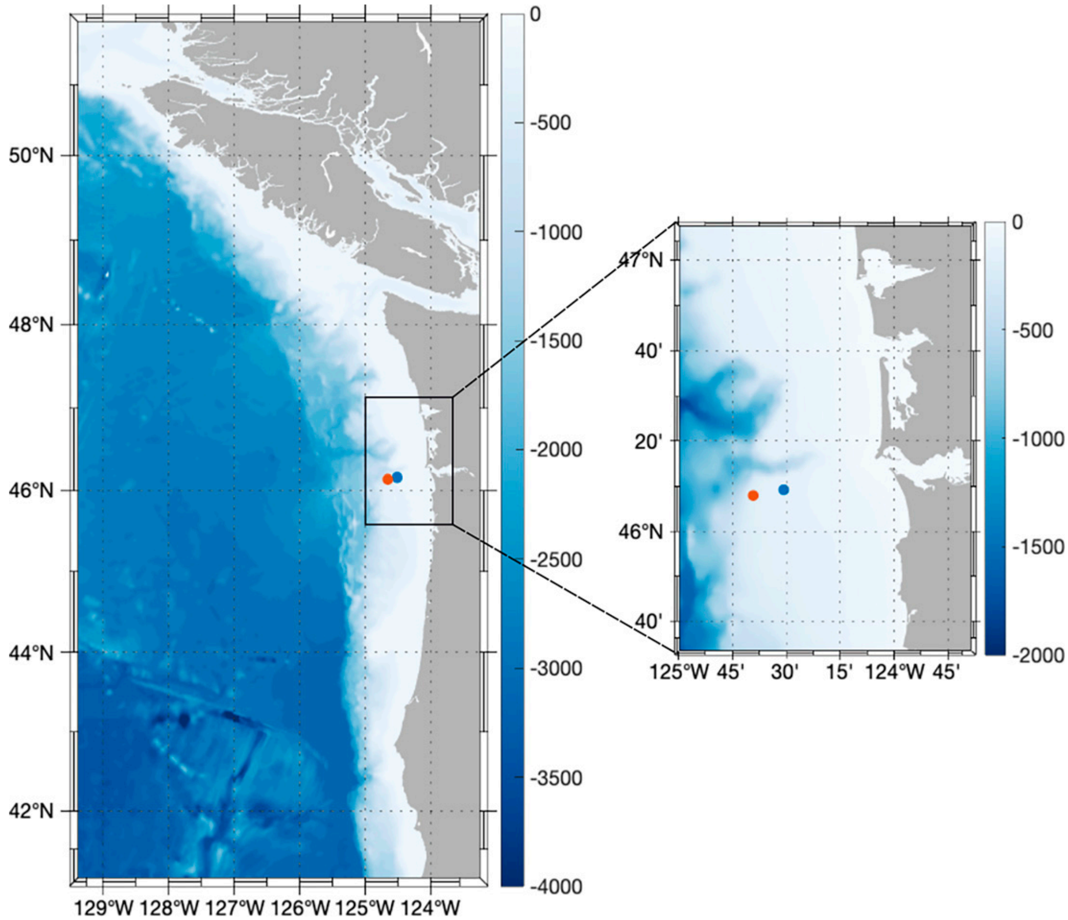

FIG. 1. Map of the study area with land shown in gray. (left) A regional map showing western Washington State. (right) A zoom of the coastal area near the buoys. The 3-m discus buoy is shown in blue and the CDIP buoy is shown in red. Water depth is shown on a color scale with units of meters.

Figure 3 shows the time series of wind and wave parameters averaged daily. The time series considered contains a wide variety of conditions with typical seasonal variation including wind speeds up to 24 m s^{-1}, with H_S up to 10 m and peak periods that range from 5 to 25 s. The wave direction is primarily from the west and northwest. The time series qualitatively shows good agreement for H_S and T_{p4}. There is episodic disagreement between the sensors on the θ_p. This is typical when multiple seas of similar energy are present, and either direction could be chosen as the peak direction given innate sampling variability. Next we examine spectral parameters in detail to identify and characterize the tilt error.

3. Results

a. Identifying tilt error

1) TIME SERIES

Figure 4 shows a detailed look at the H_S time series that spans about 6 weeks. The time series of significant wave height shows very close correspondence between the three systems, with the exception that the 3DMG has an elevated peak compared to either the HIPPY or the Waverider during the highest wave event.

2) COMPARISON

Figures 5a–c show the scatterplot with 50%, 90%, 99%, and 99.9% quantiles are also indicated. Table 1 gives the correlation coefficient (R^2), bias, and root-mean-square errors (rmse). The scatterplots and error statistics show nearly identical comparison between either 3-m discus system and the Waverider. The differences are small considering sampling variability and 15-km distance between buoys (e.g., Collins 2012). The scatterplot between the two systems on the 3-m discus buoy shows very little scatter (0.07- versus 0.23-m rmse). This is because mounting the two systems on the same platform essentially removes the sampling variability. However, there is a noticeable shift in the bias and rmse as a function of increasing wave height. As wave height increases, the 3DMG overestimates wave height. Indeed the q–q plot shows that there is a noticeable divergence from the 1:1 line around 99% level (~6 m) and continues to diverge as wave height increases.

Apparently, the tilt error is undetectable in the statistics until an appreciable percentage of the data are affected. Indeed, p values from a Kolmogorov–Smirnov test are shown in Fig. 6

FIG. 2. An example of a 3-m aluminum discus buoy (image from https://www.ndbc.noaa.gov/images/buoys/3moly.jpeg).

for data populations subject to various thresholds. Depending on your chosen significance value, the population distributions are not statistically different, i.e., the difference was unlikely to arise randomly, until a threshold of about 4 m. Whereas when distributions from both 3-m discus systems are compared to the that from the Waverider, they fail the null hypothesis without thresholds.

We check the error statistics for a particular threshold, 6-m H_S, which is approximately the 99th percentile. Table 1 gives error statistics for high waves (≥6-m H_S). The scatter for very high waves is expected to increase because the sampling variability increases as a function of wave height [e.g., Eq. (11) of Anctil et al. 1993]. The dual system still compares relatively well with high correlation reasonable rmse, but the 3DMG is biased nearly 0.3 m higher than the HIPPY about 4.3%. There is a 0.15-m difference in bias between the Waverider and the two systems, with the HIPPY biased low by 0.26 m and the 3DMG biased low about 0.11 m. Without the wave height threshold, the tilt error is buried. Even with a 6-m threshold, the 0.30-m bias between the systems is reduced to a 0.10-m difference compared to a buoy only 15 km away where waves are apparently slightly larger farther out to sea.

Next, we perform a similar analysis but instead of thresholding by wave height, the data are binned by wave height. Figure 7 shows the percent bias and normalized rmse (nrmse) between wave heights measured by the 3DMG, HIPPY, and Waverider. For each pair, error is calculated for each wave height bin; bin widths are 1 m. From 1- to 4-m H_S, the Waverider measured 5% higher wave heights than either system on the 3-m discus buoy. This is because the buoy is farther out to sea, and the waves are dissipating over the shelf as they approach the coast. The difference between the Waverider and the HIPPY remains constant near −5% over all wave heights, whereas the bias increases for the 3DMG as a function of wave height. For the 3DMG relative to the HIPPY, there is slight negative bias until about 3 m. There is an almost linear increase in the relative bias from 0% at 4 m to 6% at 8 m. The comparison between the Waverider and the 3DMG is ostensibly better for high wave heights than that of Waverider and HIPPY, but only because an error is compensating for what should be real differences. Notice that the error bars (nrmse) are much larger between 3-m discus and Waverider (15%–20%) compared to the two collocated systems (5%). Even within 15-km distance, there is a systematic difference between the two 3-m discus systems relative to the Waverider, but the difference in bias is small compared to the random error. Only for the highest wave height bins is the difference in systematic error a comparable magnitude to the random error. It appears that wave measurements are accurate, such that with enough measurements we can identify small systematic differences (<5%), but the precision is limited by sampling variability.

3) DISTRIBUTIONS

A slightly different perspective comes from the empirical distributions, which are similar to *q*–*q* plots in that time information is discarded. Figure 8 shows the long term empirical distribution from each sensor (for the same times). The distributions are nearly identical below 7 m. At 7 m, the HIPPY sensor and the Waverider diverge from the 3DMG, which maintains a higher probability into the most extreme wave heights. The divergence grows as a function of wave height (or probability level). At the 0.1% level, HIPPY, Waverider, and 3DMG measured 7.9, 8.0, and 8.3 m, respectively, and at the 0.01% level, 8.8, 9.0, and 9.8 m, respectively. Although there is uncertainty due to the reduced number of samples in the tail of the distribution, there is clearly a divergence in 3DMG from the other two sensors. The difference is about ~1 m (13%) at the 0.01% level.

b. Characterizing the error

1) SPECTRAL CHARACTERISTICS OF TILT ERROR

In this and the following subsections, we characterize the tilt error of the 3DMG assuming the HIPPY system on the same 3-m discus to be ground truth. In Fig. 9, frequency spectra are binned by wave height and the percent error is shown as a function of frequency. The errors below 0.06 are likely related to low signal-to-noise and the application of different low-frequency cutoffs (Earle 1996). Outside of the very low frequencies, the largest error (20%) occurs in the vicinity of

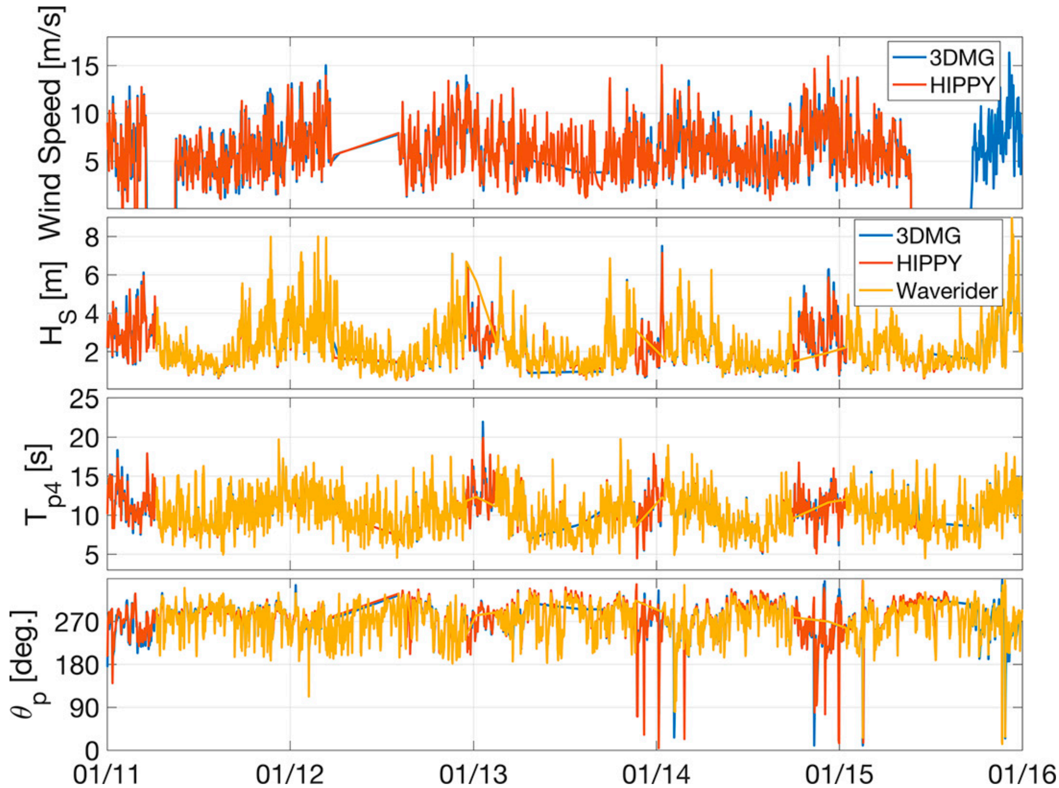

FIG. 3. Time series of (from top to bottom) wind speed, significant wave height, peak period, and peak direction.

the peak frequency for the highest wave heights, and falls off sharply toward the high frequencies. The errors also drop as the frequency increases until the energy is underestimated at all frequencies > 0.25 Hz. This underestimation is nearly the same across wave height bins indicating that it is not related to tilt. (Although 20% underestimation may be of concern for certain applications, the difference in energy is small because rear face of frequency spectra tend to decay with f^{-4}.) This shows that the overestimation related to tilt is confined to the energy around the spectra peak.

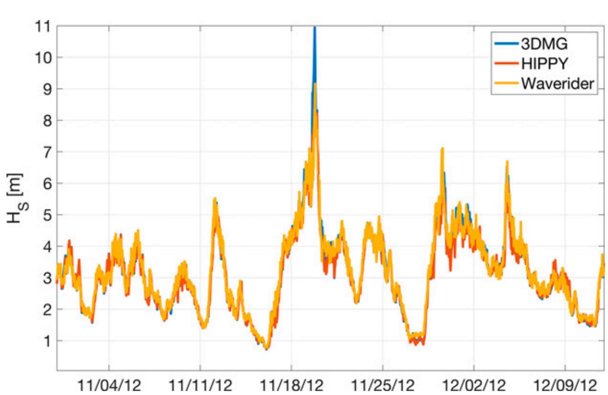

FIG. 4. Short time series extracted from Fig. 3b of significant wave height.

Similarly, Fig. 10 shows the spectra as binned percent error in H_S, and where the error is distributed across the spectra. The subsets are selected from all cases of HIPPY $H_S > 6$ m and H_S percent error threshold, where all data with H_S percent errors greater than 1%–20% are shown. Top plot shows the mean spectra from HIPPY and the bottom plot shows the mean percent difference between 3DMG and HIPPY. The spectral shape can be thought of as a weighting function for the error that is distributed across frequency space. For most spectral bands at or above the peak, the mean spectra are ordered by error where the highest error corresponds to the spectra with the highest energy, but there is also overlap. As the percent error in H_S increases, it appears as an increase in the error around 0.05–0.10 Hz, which corresponds to the location of the peak frequency. Although there appears to be significant underestimation in the high frequencies, this does not affect H_S because the weight is low in the tail of the spectrum.

2) DIRECTIONAL CHARACTERISTICS OF THE TILT ERROR

Figure 11 shows the difference between HIPPY and 3DMG measurement for mean direction and directional spread binned by wave height. There are no meaningful differences between the two systems for mean direction, on average the bias falls within quality standards of ±10° defined by NDBC (https://ndbc.noaa.gov/rsa.shtml). There are systematic

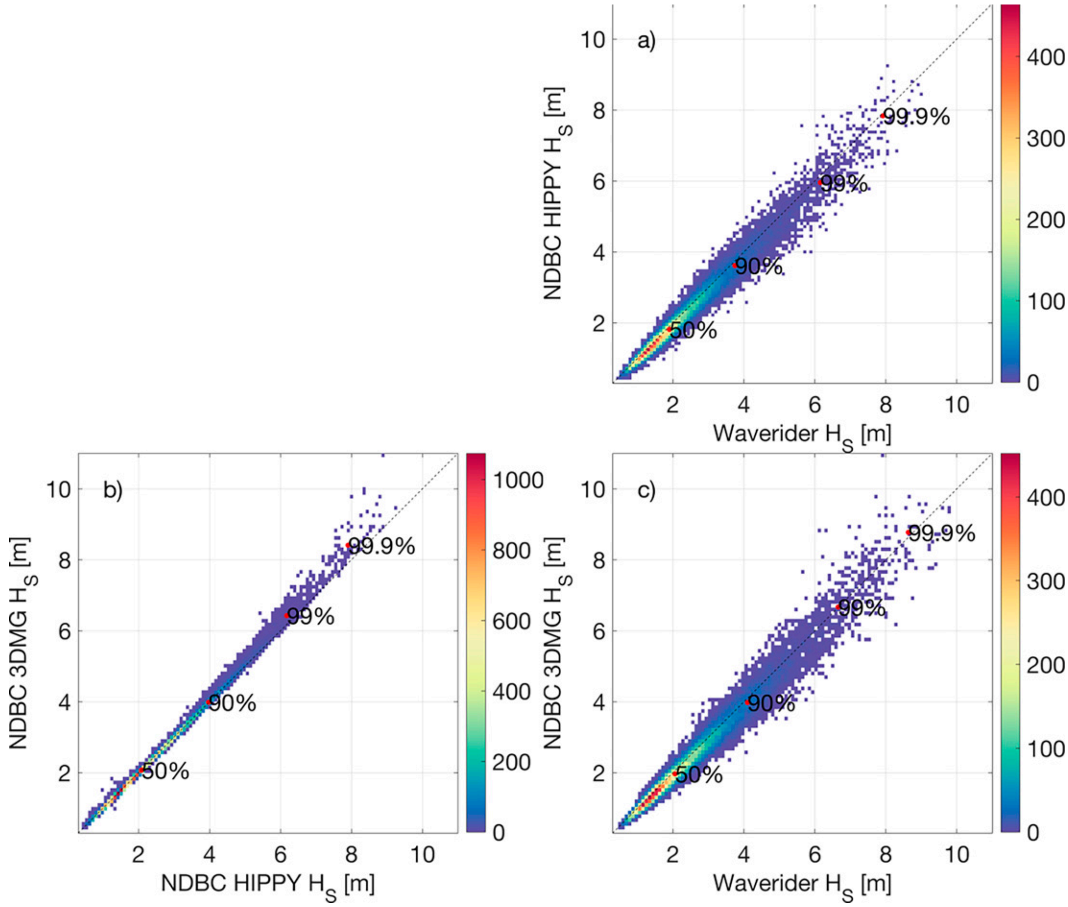

FIG. 5. Scatterplots with density shown in color. Specific quantiles are indicated at the 50%, 90%, 99%, and 99.9% levels with a red dot against a black dashed 1–1 line. (a) 3-m discus HIPPY vs 3-m discus 3DMG. (b) Datawell Waverider vs 3-m discus 3DMG. (c) Datawell Waverider vs 3-m discus HIPPY.

differences that appear in the very low frequencies below the typical peak frequencies, with no apparent delineation as a function of wave height. The information in this region is likely dominated by noise or artifacts. The directional spread shows systematic differences where 3DMG measures higher directional spread through the energy containing frequencies. The error, both systematic and random, appears to improve as wave height increases. We can conclude that tilt error does not have a meaningful impact on the measuring wave direction and directional spread.

TABLE 1. Error statistics including correlation coefficient (R^2), bias, and root-mean-square error (rmse).

	HIPPY			Waverider		
Sensor	R^2	Bias (m)	rmse (m)	R^2	Bias (m)	rmse (m)
3DMG	1.00	0.003	0.070	0.97	−0.088	0.241
HIPPY				0.97	−0.087	0.222
$H_S \geq 6$ m						
3DMG	0.88	0.29	0.42	0.57	−0.11	0.68
HIPPY				0.52	−0.26	0.65
$H_S \geq 6$ m and two-part correction						
3DMG	0.93	−0.03	0.18			
$H_S \geq 6$ m and H_S-only correction						
3DMG	0.88	−0.06	0.25			

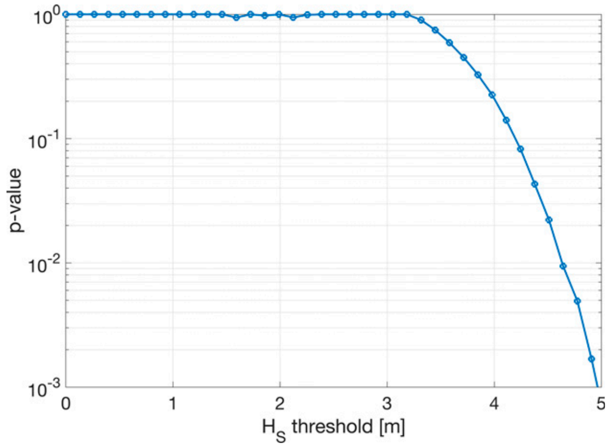

FIG. 6. The p value from two-sample Kolmogorov–Smirnov test as a function of threshold wave height.

c. A parametric correction for H_S

1) FIT TO ERROR

Using HIPPY as ground truth 30 893 paired observations, we can visualize 3DMG error as a function of both significant wave height and wind speed. According to B10 and R11, the major source of sustained tilt is due to the action of wind on the buoy. If this is true, then this error should increase as a function of wind speed. And indeed it does as shown on the right-hand side of Fig. 12. Figure 12 shows that both the absolute and relative (normalized by the wave height) differences increase as a function of wave height and wind speed for the period of record when both sensors are reporting. There is a slight negative bias at low wave heights and wind speeds. Starting at $H_S = 4$ m or $U = 10$ m s^{-1}, this shifts to positive bias that continues to increase in magnitude as a function of wave height and wind speed. The increase of relative error appears to be linear with wave height and quadratic with wind speed.

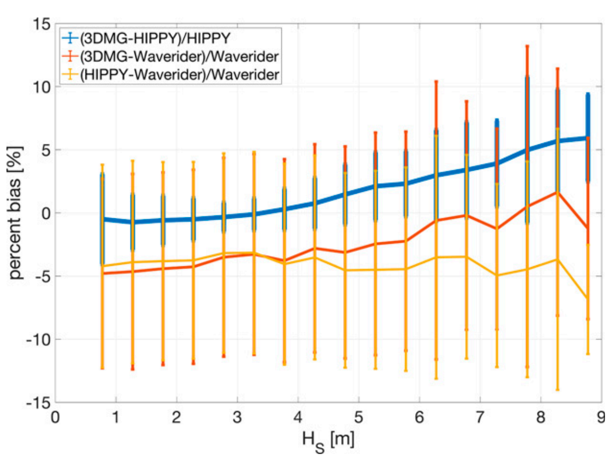

FIG. 7. Relative bias as binned with wave height. The nrmse is shown with vertical bars. HIPPY vs 3DMG in blue, Waverider vs 3DMG in red, and Waverider vs HIPPY in yellow.

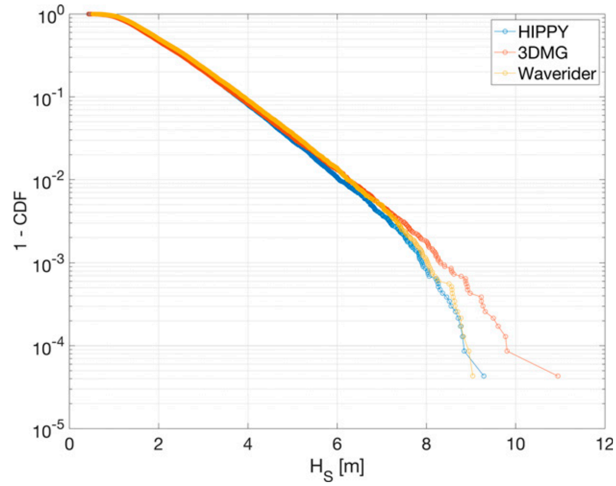

FIG. 8. The empirical cumulative distribution function from HIPPY (blue), 3DMG (red), and Waverider (yellow).

Since NDBC does not transmit tilt data, the best proxy is wind speed. Therefore, it is desirable to develop a correction based on wind speed. We start with a quadratic fit to the data as a function of wind speed. As evident in the fit to the data in Fig. 13, the correction is useful when the winds are over about 10 m s^{-1}. The correction did not fully eliminate the linear trend as a function of wave height, so an additional step based on a linear fit to the difference as a function of wave height was applied. This two-part correction could be used when wind speed is available. If wind speed is not available, then one could use just a linear wave height correction.

2) HOW TO APPLY A CORRECTION

The corrections were derived from relative error space, so it is applied with the following equations:

$$c_1 = a_1U^2 + a_2U + a_3, \quad (7)$$

$$H_{S_{c_1}} = H_S(1 - c_1)/100. \quad (8)$$

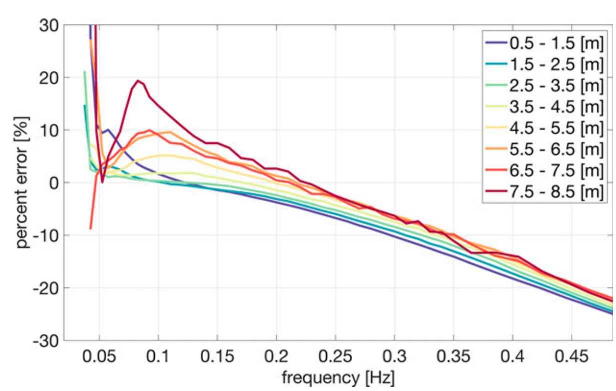

FIG. 9. Percent bias as a function of frequency. 3DMG spectra evaluated against HIPPY spectra. Colors are wave height bins as indicated in the legend.

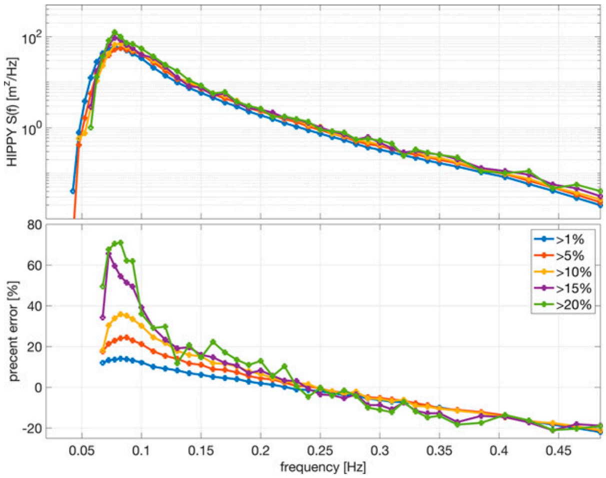

FIG. 10. As in Fig. 9, but binned by H_S percent bias. (top) The spectral shapes as measured by HIPPY are shown for the different error bins indicated by color. (bottom) How the error, (HIPPY − 3DMG)/HIPPY, is distributed across frequency space.

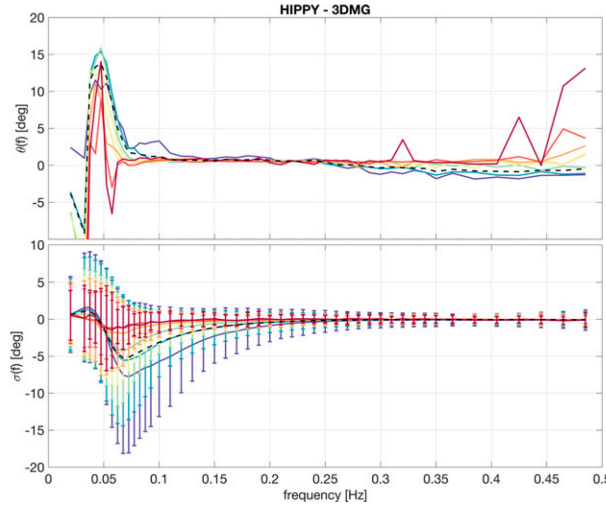

FIG. 11. The difference between HIPPY and 3DMG for (top) wave direction and (bottom) directional spread as a function of frequency. The color indicates binning by wave height with bin centers from 1 to 8 m (blue to red) as defined in Fig. 9. The mean of all bins is a dashed black line. Error bars are given for directional spread representing plus and minus one standard deviation.

The coefficients were determined by the quadratic fit above: $a_1 = 0.057$, $a_2 = 0.6217$, and $a_3 = 0.3145$. Now, the second part of the procedure (or the only part if wind speed is unavailable) applies to wave heights greater than 4 m:

$$c_2 = b_1 H_S + b_2, \tag{9}$$

$$H_{S_{c_2}} = H_{S_{c_1}}(1 - c_2)/100. \tag{10}$$

For the two-part correction, $b_1 = 0.06455$ and $b_2 = -2.576$. If wind speed is unavailable, the coefficients are then $b_1 = 1.4804$ and $b_2 = -5.6255$. We applied both the two-step and the H_S-only correction and recalculated the statistics in Table 1. The H_S-only correction provides a marked improvement and the two-step is better still. Figure 14 shows the wave height distribution plot with the corrections applied. Although some small differences remain for return periods of wave heights 5–7 m, beyond 7 m the corrections lead to a distribution more consistent with the HIPPY distribution. There is little difference between the two correction methods.

4. Discussion

We compared data from a strapped-down wave system on an NDBC 3-m discus buoy to a gimballed system on the same buoy and to a nearby Waverider. We uncovered evidence of an error due to not correcting for sustained tilt of the buoy. It was difficult to identify because it progressively affects high wave heights and wind speeds, and thus it is not significant for the majority of the data. Identifying the error is only possible after thresholding for high wave heights (Fig. 6). Using a buoy only 15 km away, we could no longer identify the error with precision due to sampling variability and slight inhomogeneity in the wave field. Even with good metadata identifying the NDBC stations affected by tilt error, it may be difficult to establish a baseline for comparison. The error is found to be most prevalent in high waves and high winds. This is because the heel is controlled mostly by wind force on the buoy and high winds and large waves tend to cooccur and be more or less aligned,[2] particularly in large, slowly developing extratropical and polar lows that frequent the eastern Pacific.

Since wind is the primary mechanism inducing heel, and wind also controls the orientation of the buoy through drag on the weather vane, the bias (magnitude and direction) is determined by the relative angle between the weather vane the accelerometer. R11 showed positive bias when tilt was oriented along the axis with the weather vane (following winds or head on winds), negative when at 90°, and slightly positive when at 45°. B10 showed an overestimation of 60% around the peak of Hurricane Katrina, but what is often overlooked directly following the overestimation, there was underestimation of 43%. This was due to a unique situation where extreme surge was relaxing and dragging the buoy along, mooring anchor and all. The mooring line likely reached its maximum extent, exerting some force on the buoy through the bridle that, in turn, controlled the buoy heel. As demonstrated in Fig. 16 of B10, the orientation of the heel was determined by the wind until the point where the buoy was dragged, then orientation of the heel was somehow determined by relaxing surge. The result was overestimation then underestimation. The surge relaxation is likely extremely rare and we speculate that wind is the primary mechanism for buoy heel in the NDBC database at large. Furthermore, a buoy with a weather vane, like most 3-m discus buoys, always

[2] Tropical cyclones offer a counterexample (Collins et al. 2018, 2021).

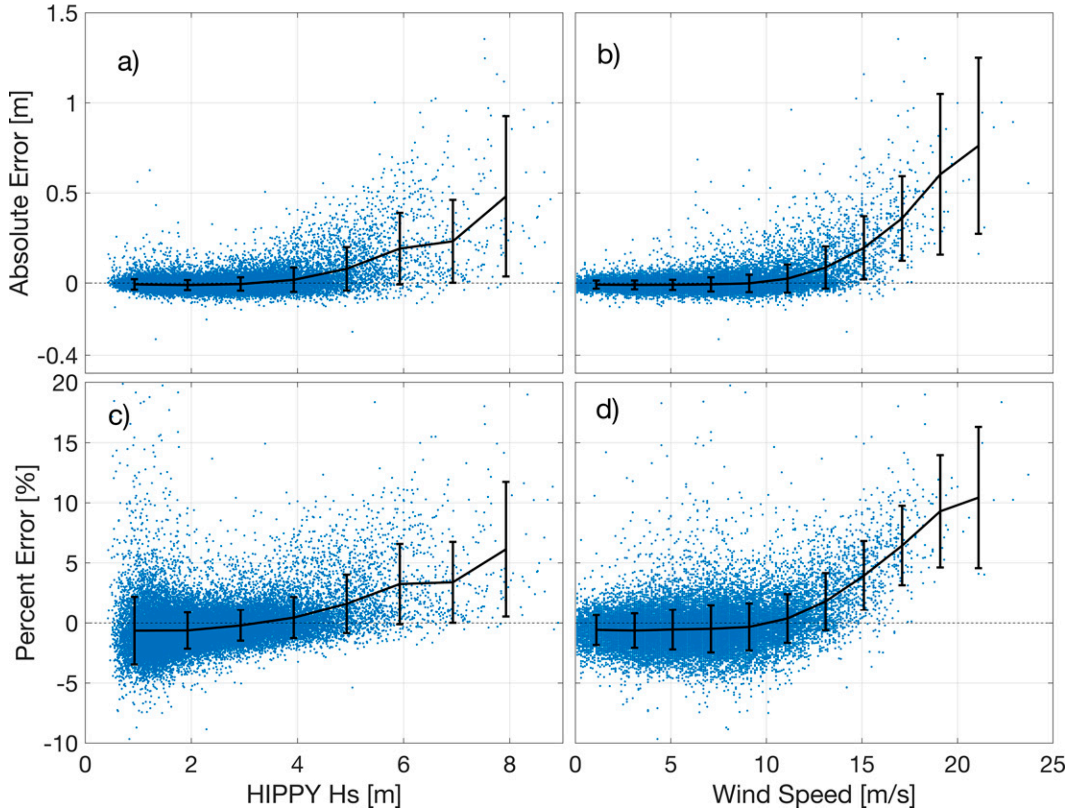

FIG. 12. H_S bias and percent bias as a function of wave height and wind speed. The blue dots are each individual measurement; the black line with error bars is the bin average (bias) and standard deviation (rmse).

faces within 45° of the wind direction. This is why, in contrast to B10, we observe on average only positive bias.

Along with bias, variability also increases with wave height and wind speed. This is expected when accounting for sampling variability, but these two systems are sampling the same buoy motion. We do not know the cause of the variability, so here we venture some speculation. As the winds and waves increase, a buoy is forced to the edge of its mooring scope through wind drag on the structure. A counterforce is only felt once a buoy reaches the edge of its watch circle when the mooring is taut. So the heel is actually due to the counterforce being applied through the bridle on a taut mooring below the surface in response to a wind force above the surface. Forces of opposite sign applied to the top and bottom of a solid form create a moment, and the object rotates in response to the moment. Perhaps the reason that the error is higher in B10 is the water depth of 19 m, and thus a much shorter mooring. The hypothesis is that for a short mooring it would require less work (and time) for a buoy to reach its maximum extent and thus tilt would be more prevalent. Varying winds would push the buoy around in a circle before it had a chance to reach its maximum extent. Whether or not the mooring is taut controls the tilt and introduces variability. Logically, a longer mooring would have more variability and less total bias than a short mooring.[3]

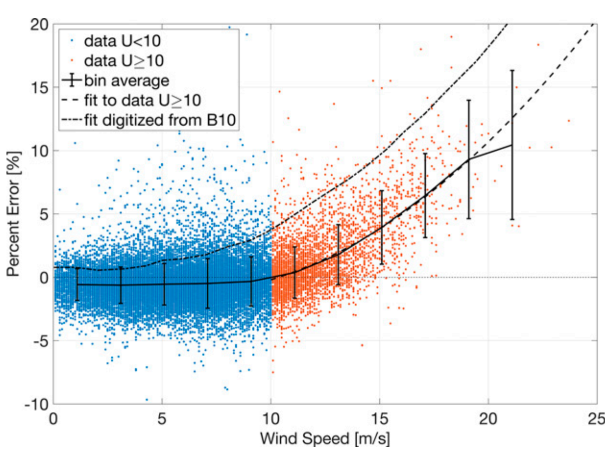

FIG. 13. As in Fig. 12d. Now red dots indicate data with wind speed > 10 m s^{-1}. A dashed line shows a quadratic fit to the red data. A dash–dotted line is a digitized fit to the data from Fig. 17 of B10.

Figure 13 shows the relative error as a function of wind speed. Also shown is a digitization of the fit in the bottom panel of Fig. 17 in B10. One can draw some similarities from

[3] If the data exist to test this hypothesis, please let us know.

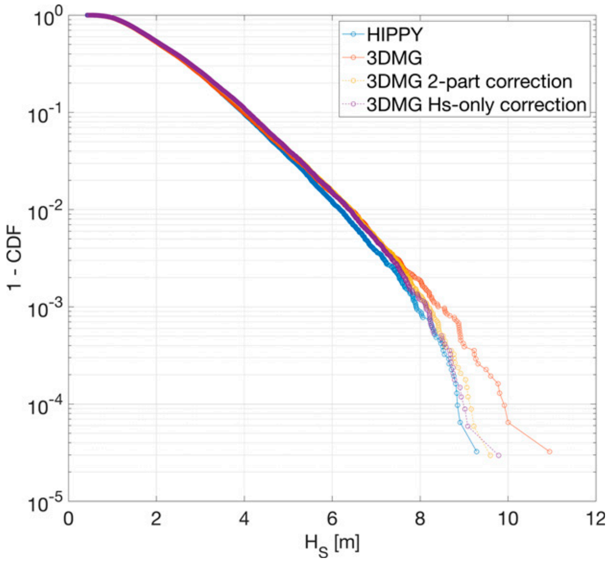

FIG. 14. Wave height distributions from HIPPY (blue), 3DMG (red), 3DMG with two-part correction (yellow), and 3DMG with H_S-only correction (purple).

the curve of the fit from B10 to the curve of fit derived here, but there is a significant offset. One reason is that B10 had access to raw signals of uncorrected and corrected to compare. Here we treat the gimballed system as ground truth. Apparently there is a small offset between the two systems, about 0.5% before tilt error becomes apparent. In addition their buoy was in 19-m water depth, and their data are limited to one very extreme case study. Regardless, the parametric error correction derived from our dual-system comparison would not have performed well for the case B10.

The parametric correction was derived from a fit to the error. It is likely that the error, and thus the coefficients of the fit depend on several factors. The windage would be altered by different hull types and differences in superstructure, e.g., the use of a seal cage. The mooring force would be affected by depth and mooring design. Both forces on the buoy depend on the local wind and wave climate. The platform examined here is a standard 3-m aluminum discus buoy in relatively deep water during the DDWM era, and we would caution against applying it to anything else. Other buoy hulls will almost certainly require a different correction, or perhaps not require a correction at all. For example, Collins et al. (2014) found that a NOMAD-type hull buoy required no tilt correction for significant wave height (though it did matter for the very highest individual waves). This is because the heel of that buoy was not controlled by windage but by the balance between the buoyancy and the ballast attached to a keel on the bottom of the buoy. Other factors, say, depth and the potential for the mooring to be under tension may also have an effect on the heel. Last, since orientation matters, we do not know how consistent NDBC was when mounting the IMUs relative to the weather vane. All of these factors prevent general application.

a. *Specific application*

Short of an exhaustive list of NDBC stations affected, for this publication we have identified several stations and time periods in Table 2.

Table 2 details deployments of three other stations with 3-m discus buoys with a strapped-down accelerometers using DDWM firmware < v2.03. Stations 41002 and 42056 were aluminum and 41013 was foam. After this period the firmware was updated, so data after this period would have been tilt corrected. Stations 41002 and 42056 used different hulls prior to the 3-m discus, and station 41013 used the same form factor but with a foam composition. First, let us take a look at the data from before and after the implementation of tilt correction (via firmware upgrade).

Figure 15 shows the long-term distribution of wave heights at each location. The colors indicate status of the buoy either before the firmware update (uncorrected) or after (firmware update). Also shown are the application of the parametric corrections applied to the uncorrected data. To deal with the variable size of the records, we took a equal number of random subsamples from each record, 12 000 for 41002 and 42056 and 8000 for 41013. We repeated the process 1000 times and took the mean. All records show significant differences between the corrected and uncorrected records. The uncorrected and firmware updated distributions are similar until they deviate near 4-m H_S. We interpret this as evidence of tilt error in the buoy record during that uncorrected time period.

There are data at each station prior to the uncorrected data, but for most stations the hull types were significantly different: 6-m NOMAD or 10- or 12-m discus. Since the hull type likely effects the heel, these were not suitable for comparison. For station 41013, the metadata indicated that prior to the use of a strapped-down accelerometer, there was 3-m discus with a HIPPY 40 module, so we include it in the center panel of Fig. 15). The material of the hull is different, although with a similar shape the wind force over hull structure should be similar. This being said, how buoy material effects heel is not known. The distribution from HIPPY sensor matches nicely with the distribution after the firmware update. This suggests that the implementation of tilt correction in the firmware update was effective.

Although the parametric correction developed nudges the distribution in the right direction, it does not match the post firmware update distributions. Even accounting for hull and depth (station 41002 and 42056), evidently there are still other factors. At all three stations, the tilt errors were apparently larger than those found in the dual-system buoy. At station 41013, the distributions are similar until $H_S = 4$ m, and then distributions drastically diverge with differences greater than 3 m. This location has relatively shallow water, perhaps wave shoaling and breaking, and strong currents, all these factors may work to increase the buoy heel.

b. *Prevalence*

Even if we cannot confidently correct the data at present, one may like to know whether or not the data they are analyzing contain this type of error. The first NDBC wave buoys

TABLE 2. Sample table caption and table layout.

Station	Name	Nominal depth (m)	Nominal location (lat, lon)	Service dates	Hull notes
41002	South Hatteras	3950	31.989°N, 74.962°W	18 May 2012–12 Apr 2016	Prior: 6-m NOMAD 12-m discus
41013	Frying Pan Shoals, North Carolina	33	33.441°N, 77.764°W	23 Aug 2011–11 Sep 2015	Before 2005: 3 m with HIPPY
42056	Yucatan Basin	4554	19.820°N, 84.945°W	17 May 2011–3 Apr 2014	Scoop after 2015 Prior: 12-m discus

deployed in the late 1970s used gimballed systems. Sometime in the 1990s they began testing and implementing systems that were strapped down. Internal investigations by NDBC did not show a significant difference between the strapped-down and gimballed systems [Teng et al. (2009) and an unpublished study referred to in R11]. The comparisons were limited to wave heights < 3.2-m H_S. Our work shows that the tilt error would not have appeared within this wave height range. By the 2000s, most NDBC buoys were using a strapped-down sensor and the DDWM payload. B10 revealed the importance of tilt correcting in large waves generated by Hurricane Katrina, and, ostensibly in response, NDBC implemented a tilt correction in the DDWM later 2011 (R11). This update came in version 2.03 of the DDWM firmware, which had to be updated manually during regular buoy maintenance. Therefore, NDBC buoys with DDWM with version < 2.03 were not tilt corrected until sometime (possibly years) after the 2011 update. Using this information it is possible to catalog stations potentially affected by this error if detailed metadata are available. Indeed, this was how the three stations examined above were identified. Obviously, there are no guarantees that tilt error is the only issue affecting the data. The bigger obstacle to knowing whether or not data are affected is that, at the time writing, there are no publicly available metadata with firmware and module information. However, there is work underway at NDBC and elsewhere to collect, curate, and make more detailed metadata available. There is also work to combine disparate versions of the data into a single, quality controlled database (e.g., Hall and Jensen 2021).

Although tilt error is shown to affect only a very small portion of the data starting at wave heights of ~4 m, it is conceivable that tilt error could have exacerbated cases of model underestimation in the literature, e.g., Cavaleri (2009) and Cardone et al. (1996). Models and other sensors, most notably satellite altimeters, rely on NDBC buoy data for evaluation and calibration (Queffeulou 2004; Zieger et al. 2009; Ribal and Young 2019; Dodet et al. 2020). Different altimeter databases use different buoys over different time periods. This could have important implications for studies of climate and extremes (e.g., Stopa et al. 2019; Collins et al. 2021). Again, the portion of data affected is small so that the calibrations

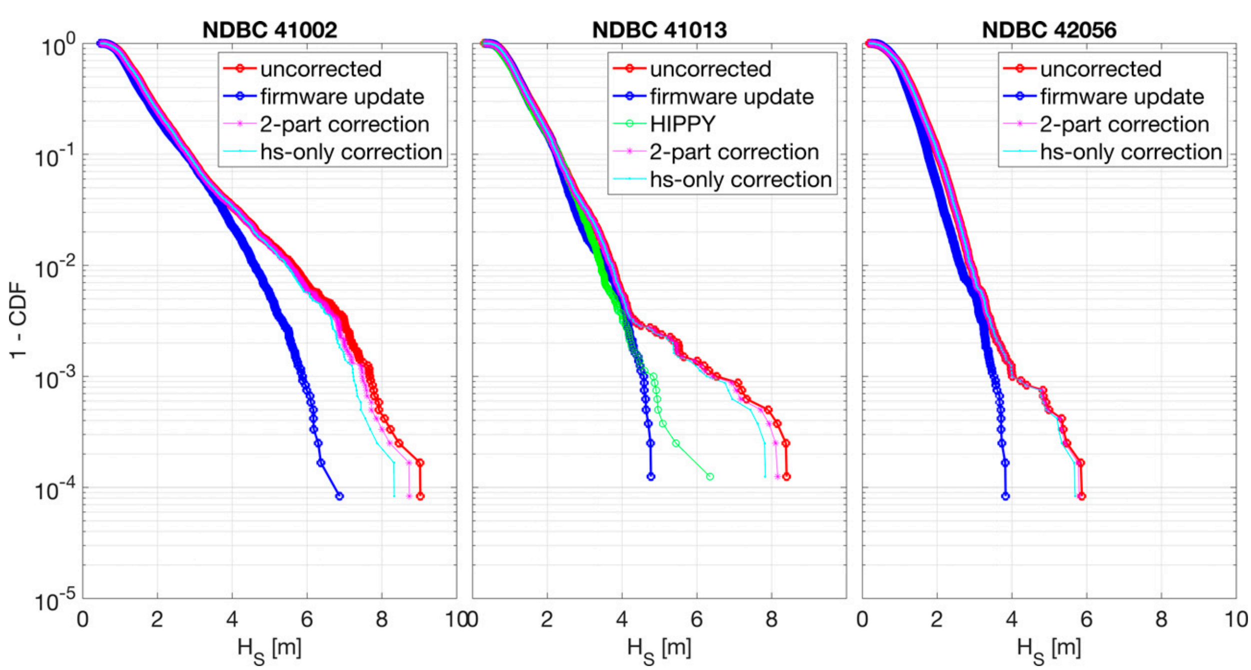

FIG. 15. Wave height distributions for three NDBC stations. Blue data are likely affected by tilt error; red data are after the implementation of tilt correction in the NDBC software. Also shown are the parametric corrections developed here as indicated in the legends.

based on the bulk of data remain unaffected, but conceivably, tilt errors could have an impact wave heights greater than 4 m. Since tilt error increases in magnitude as a function of wind speed (and wave height), one should be careful that analysis based on the extreme values is not affected by this issue. All of this suggests a need for better understanding the measurement of very large significant wave heights.

Although we focused on error caused by tilt of the buoy, the analysis also revealed systematic spectral differences in the high frequencies. As evident in Figs. 9, 10, and A2, the 3DMG underestimates spectral density which increases as a function of frequency. It appears that tilt has a small effect outside the most energetic range (0.05–0.15 Hz), where the differences in spectra a stratified by wave height in Fig. 9, but not Fig. A2. Although not important for H_S, differences in high frequencies have implications for air–sea interaction, breaking and transport, and remote sensing, and deserve further investigation.

5. Summary

Tilt error can occur if a buoy experiences heel, or sustained tilt, and the wave sensor is a strapped-down accelerometer. If a strapped-down sensor collects all 6 degrees of freedom, then a numerical correction can be performed (e.g., Anctil et al. 1994; B10; R11; Collins et al. 2014). The NDBC DDWM did not implement the tilt correction until version 2.03 of the firmware. We set out to identify and characterize the error with wave data from two different wave measurement systems on a single buoy (NDBC 46029), one strapped down and another a gimballed. We also compared both systems to a nearby Datawell Waverider. It appears that uncorrected strapped-down system introduces a positive bias into the measured wave heights starting at about 4-m H_S. Buoy heel is induced by buoy windage and the error can be described in terms of wind speed. The error was about +5% in H_S at wind speeds of 15 m s^{-1} and +10% at wind speeds of 20 m s^{-1}. The error was shown to be relatively small compared to the sampling variability, and is more difficult to detect when comparing the two systems to a nearby Waverider (15-km separation).

The bias in H_S is related to overestimation of spectral energy around the peak frequencies (0.05–0.15 Hz). Tilt error outside the spectral peak is small but evident across all frequencies; differences in directional quantities are negligible. Spectral differences unrelated to tilt were also revealed: systematic underestimation in the high frequencies of the 3DMG relative to the HIPPY. A parametric correction was derived using wind speed and wave height. Using metadata, several NDBC stations were found that were potentially affected by tilt error, and their wave height distributions before and after the firmware update shows evidence of tilt error. However, the correction derived from the dual-system buoy was not particularly effective at correcting distributions at the other sites where the magnitude of the error was apparently higher. It is doubtful that a general tilt correction could be derived from the dual-system buoy given the disparities in buoy hulls, depths, currents, mooring systems, and other factors that may contribute to buoy heel.

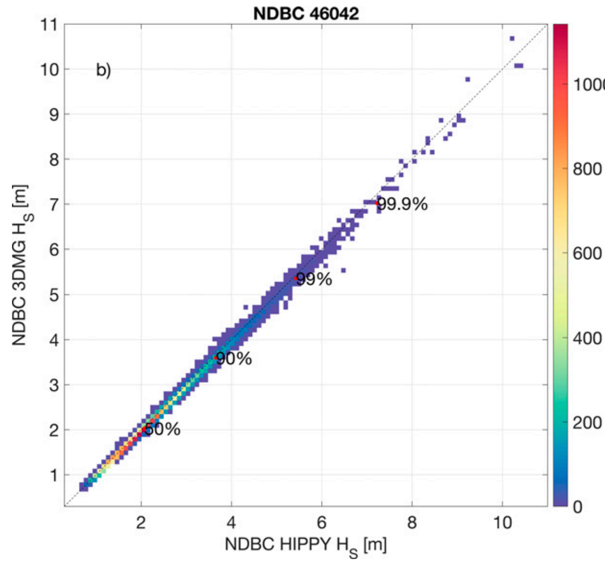

FIG. A1. As in Fig. 5c, but for NDBC buoy 46042. Scatterplots with density shown in color. Specific quantiles are indicated at the 50%, 90%, 99%, and 99.9% levels with a red dot against a black dashed 1–1 line

Luckily, the error uncovered here affects only the highest wind speeds and wave heights (>4-m H_S), and therefore a rarefied population of data at most locations. The extent to which the fleet of NDBC buoys affected is not well known at this time. On the whole, this error does not detract from NDBC's history of providing reliable data, well suited for informing weather prediction, model evaluation, and remote sensor calibration under the vast majority of conditions. However, before performing analysis sensitive to the extremes, users should try to ensure the buoy data are unaffected by tilt error.

Acknowledgments. Rodney Riley (NOAA/NDBC) retrieved, documented, and shared buoy data and metadata and Steven N. DiNapoli (NOAA/NDBC/PAE) collected and organized metadata. This work also benefited from conversations with Richard Bouchard, Dr. David Wang, and Candice Hall. COC and REJ were funded under the USACE civil works managed by CODS program managed by Dr. Spicer Bak and Dr. Julie Rosati.

Data availability statement. The data used here are publicly available through NCEI THREDDS server: https://www.ncei.noaa.gov/thredds-ocean/catalog/ndbc/cmanwx/catalog.html. Author COC will accommodate any reasonable request for data and code to recreate plots.

APPENDIX

Benchmark Comparison

NDBC buoy 46042 was another 3-m discus buoy with both 3DMG and HIPPY 40 sensor systems. It was deployed during the FLOSSIE experiment (Jensen et al. 2021). This

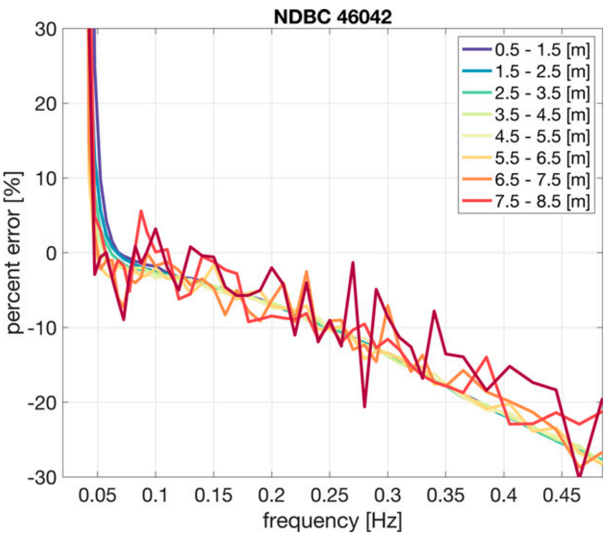

FIG. A2. As in Fig. 9, but for NDBC buoy 46042. Percent bias as a function of frequency. 3DMG spectra evaluated against HIPPY spectra. Colors are wave height bins as indicated in the legend.

3DMG system was a version after the application of tilt correction, so it may serve as a benchmark for our dual-sensor comparison. Figure A1 shows scatterplot built from about 4 years' worth of data. Comparing 32 945 paired data points, the bias was −0.03, rmse was 0.06, and the correlation was 1.00. Figure A2 shows the spectral difference binned by wave height. Besides the increased scatter, there is no difference between the wave height bins. There is a similar disagreement below the spectral peak, and a pattern of underestimation from low to high frequencies, but over estimation around the peak frequencies.

REFERENCES

Allan, J., and P. Komar, 2000: Are ocean wave heights increasing in the eastern North Pacific? *Eos, Trans. Amer. Geophys. Union*, **81**, 561–567, https://doi.org/10.1029/EO081i047p00561-01.

Anctil, F., M. A. Donelan, G. Z. Forristall, K. E. Steele, and Y. Ouellet, 1993: Deep-water field evaluation of the NDBC-SWADE 3-m discus directional buoy. *J. Atmos. Oceanic Technol.*, **10**, 97–112, https://doi.org/10.1175/1520-0426(1993)010<0097:DWFEOT>2.0.CO;2.

——, ——, W. M. Drennan, and H. C. Graber, 1994: Eddy-correlation measurements of air–sea fluxes from a discus buoy. *J. Atmos. Oceanic Technol.*, **11**, 1144–1150, https://doi.org/10.1175/1520-0426(1994)011<1144:ECMOAS>2.0.CO;2.

Bender, L. C., III, N. Guinasso Jr., J. N. Walpert, and S. D. Howden, 2010: A comparison of methods for determining significant wave heights—Applied to a 3-m discus buoy during Hurricane Katrina. *J. Atmos. Oceanic Technol.*, **27**, 1012–1028, https://doi.org/10.1175/2010JTECHO724.1.

Bouchard, R., R. Riley, R. Dinoso, K. Grissom, W. McCall, G. Mungov, and T. Ryan, 2013: Recent advances in the NDBC wave and tsunami measurement programs. *13th Int. Workshop on Wave Hindcasting/Fourth Coastal Hazards Symp.*, Banff, AB, Canada, Environment Canada.

——, R. Jensen, D. Wang, G. Forristall, O. Francis, and R. Riley, 2014: Extreme wave height reports from NDBC buoys during Hurricane Sandy. *Tropical Cyclone Research Forum/68th Interdepartmental Hurricane Conf.*, College Park, MD, and Miami, FL, NOAA.

——, K. Steele, C.-C. Teng, L. Fiorentino, and T. Rutledge, 2016: Some theory and application of calibration techniques for NDBC wave measurement buoys. *RMIC WMO Region IV Workshop*, Gulfport, MS, NDBC.

Cardone, V. J., R. E. Jensen, D. T. Resio, V. R. Swail, and A. T. Cox, 1996: Evaluation of contemporary ocean wave models in rare extreme events: The "Halloween storm" of October 1991 and the "Storm of the Century" of March 1993. *J. Atmos. Oceanic Technol.*, **13**, 198–230, https://doi.org/10.1175/1520-0426(1996)013<0198:EOCOWM>2.0.CO;2.

Cavaleri, L., 2009: Wave modeling—Missing the peaks. *J. Phys. Oceanogr.*, **39**, 2757–2778, https://doi.org/10.1175/2009JPO4067.1.

Chaffin, J. N., W. Bell, K. O'Neil, and C.-C. Teng, 1994: Design and testing of the NDBC wave processing module. *Ocean Wave Measurement and Analysis*, ASCE, 277–286.

Collins, C. O., III, 2012: In situ wave measurements: Sensor comparison and data analysis. M.S. thesis, Applied Marine Physics, University of Miami, 208 pp., https://scholarship.miami.edu/permalink/01UOML_INST/1grrnr5/alma991031447696402976.

——, B. Lund, T. Waseda, and H. C. Graber, 2014: On recording sea surface elevation with accelerometer buoys: Lessons from ITOP (2010). *Ocean Dyn.*, **64**, 895–904, https://doi.org/10.1007/s10236-014-0732-7.

——, H. Potter, B. Lund, H. Tamura, and H. C. Graber, 2018: Directional wave spectra observed during intense tropical cyclones. *J. Geophys. Res. Oceans*, **123**, 773–793, https://doi.org/10.1002/2017JC012943.

——, T. Hesser, P. Rogowski, and S. Merrifield, 2021: Altimeter observations of tropical cyclone-generated sea states: Spatial analysis and operational hindcast evaluation. *J. Mar. Sci. Eng.*, **9**, 216, https://doi.org/10.3390/jmse9020216.

Dodet, G., and Coauthors, 2020: The Sea State CCI dataset v1: Towards a sea state climate data record based on satellite observations. *Earth Syst. Sci. Data*, **12**, 1929–1951, https://doi.org/10.5194/essd-12-1929-2020.

Drennan, W. M., H. C. Graber, C. Collins III, A. Herrera, H. Potter, R. Ramos, and N. Williams, 2014: EASI: An air–sea interaction buoy for high winds. *J. Atmos. Oceanic Technol.*, **31**, 1397–1409, https://doi.org/10.1175/JTECH-D-13-00201.1.

Earle, M. D., 1996: Nondirectional and directional wave data analysis procedures. NDBC Tech. Rep. 96-01, 43 pp.

Gemmrich, J., B. Thomas, and R. Bouchard, 2011: Observational changes and trends in northeast Pacific wave records. *Geophys. Res. Lett.*, **38**, L22601, https://doi.org/10.1029/2011GL049518.

Gower, J., 2002: Temperature, wind and wave climatologies, and trends from marine meteorological buoys in the northeast Pacific. *J. Climate*, **15**, 3709–3718, https://doi.org/10.1175/1520-0442(2002)015<3709:TWAWCA>2.0.CO;2.

Hall, C., and R. E. Jensen, 2021: Utilizing data from the NOAA National Data Buoy Center. USACE Rep. CHETN I-100, 20 pp.

Jensen, R. E., V. Swail, and R. H. Bouchard, 2021: Quantifying wave measurement differences in historical and present wave buoy systems. *Ocean Dyn.*, **71**, 731–755, https://doi.org/10.1007/s10236-021-01461-0.

Longuet-Higgins, M. S., D. Cartwright, and N. Smith, 1963: Observations of the directional spectrum of sea waves using the motions of a floating buoy. *Ocean Wave Spectra*, Prentice-Hall, 111–132.

Menéndez, M., F. J. Méndez, I. J. Losada, and N. E. Graham, 2008: Variability of extreme wave heights in the northeast Pacific Ocean based on buoy measurements. *Geophys. Res. Lett.*, **35**, L22607, https://doi.org/10.1029/2008GL035394.

NDBC, 2009: Handbook of automated data quality control checks and procedures. NDBC Tech. Doc. 09-02, 78 pp.

O'Reilly, W., T. Herbers, R. Seymour, and R. Guza, 1996: A comparison of directional buoy and fixed platform measurements of Pacific swell. *J. Atmos. Oceanic Technol.*, **13**, 231–238, https://doi.org/10.1175/1520-0426(1996)013<0231:ACODBA>2.0.CO;2.

Queffeulou, P., 2004: Long-term validation of wave height measurements from altimeters. *Mar. Geod.*, **27**, 495–510, https://doi.org/10.1080/01490410490883478.

Ribal, A., and I. R. Young, 2019: 33 years of globally calibrated wave height and wind speed data based on altimeter observations. *Sci. Data*, **6**, 77, https://doi.org/10.1038/s41597-019-0083-9.

Riley, R., C.-C. Teng, R. Bouchard, R. Dinoso, and T. Mettlach, 2011: Enhancements to NDBC's digital directional wave module. *OCEANS'11 MTS/IEEE*, Kona, HI, IEEE, https://doi.org/10.23919/OCEANS.2011.6107025.

Ruggiero, P., P. D. Komar, and J. C. Allan, 2010: Increasing wave heights and extreme value projections: The wave climate of the US Pacific Northwest. *Coastal Eng.*, **57**, 539–552, https://doi.org/10.1016/j.coastaleng.2009.12.005.

Stopa, J. E., F. Ardhuin, E. Stutzmann, and T. Lecocq, 2019: Sea state trends and variability: Consistency between models, altimeters, buoys, and seismic data (1979–2016). *J. Geophys. Res. Oceans*, **124**, 3923–3940, https://doi.org/10.1029/2018JC014607.

Teng, C.-C., R. Bouchard, R. Riley, T. Mettlach, R. Dinoso, and J. Chaffin, 2009: NDBC's digital directional wave module. *OCEANS 2009*, Biloxi, MS, IEEE, https://doi.org/10.23919/OCEANS.2009.5422386.

Tillotson, K., and P. D. Komar, 1997: The wave climate of the Pacific Northwest (Oregon and Washington): A comparison of data sources. *J. Coastal Res.*, **13**, 440–452.

Young, I., 1995: The determination of confidence limits associated with estimates of the spectral peak frequency. *Ocean Eng.*, **22**, 669–686, https://doi.org/10.1016/0029-8018(95)00002-3.

Zieger, S., J. Vinoth, and I. Young, 2009: Joint calibration of multiplatform altimeter measurements of wind speed and wave height over the past 20 years. *J. Atmos. Oceanic Technol.*, **26**, 2549–2564, https://doi.org/10.1175/2009JTECHA1303.1.