THE EVALUATION OF SPECTRAL WAVE MODELS IN HAWAI‘I RELATIVE TO OBSERVATIONS

A THESIS SUBMITTED FOR PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE

IN

GLOBAL ENVIRONMENTAL SCIENCE

MAY 2025

By
JACKSON C. RICHARDS

Thesis Advisor

DR. JUSTIN STOPA

I certify that I have read this thesis and that, in my opinion, it is satisfactory in scope and quality as a thesis for the degree of Bachelor of Science in Global Environmental Science.

THESIS ADVISOR

Justin Stopa
Department of Ocean and Resources Engineering
School of Ocean and Earth Science and Technology
The University of Hawai'i at Mānoa
2540 Dole Street, Holmes 405
Honolulu, Hawaii 96822 USA
Phone: +1 808 956 6818

I would like to dedicate this project to my family who have supported and pushed me to chase my dreams, both in and out of the water. I would like to also dedicate this project to all my GES friends for all the great memories throughout school as well as all the feedback on my project. I would like to give a special dedication to all the surfers/lifeguards who I've spent time with in the ocean who rely on wave forecasts to save lives. I hope this research will help us score more perfect waves and also keep people in the ocean safe.

ACKNOWLEDGEMENTS

Thank you to my advisor, Dr. Justin Stopa, for guiding me through the research and introducing me to the science side of waves. Thank you, Dr. Li, for permission to use your model data as well as all the feedback you provided to my project. Thank you, Dr. Cheung, for permission to use your WW3 data. Thank you, Dr. Potemra, for access to the WW3 data. Thank you HSGC for the research internship opportunity which allowed me to expand my wave model evaluation using satellites rather than only buoys. I appreciate everyone's time dedicated, and I would not have been able to complete my GES thesis without any of you.

ABSTRACT

Waves impact the heavily populated coastal regions in Hawai‘i. Waves also affect commerce, maritime operations, renewable energy development, and recreational activities. In order to mitigate damages to offshore activities, coastal infrastructure, and life, reliable wave forecasts are needed. The Pacific Islands Ocean Observing System (PacIOOS) operates a nested spectral wave model that predicts waves in nearshore and offshore waters of Hawai‘i. This thesis analyzes the archived PacIOOS wave forecast from 2011 to 2021 using observations from 13 moored wave buoys and space-borne satellite altimeters from 7 different satellite missions to gain insight into the model errors. Using the observation datasets as reference, several error metrics are used to evaluate the forecast uncertainties, as well as the uncertainties that evolve with the increasing lead time. The wave forecast performs best for wave heights around 2 meters at the start of the forecast period. At the start of the forecast period during days 1 + 2 across the Hawai‘i regional model domain, NW wave events (281° - 348°) are overestimated by 6.0%, N-NE wave events (348° - 56°) are underestimated by -1.7%, SW wave events (191° - 236°) are underestimated by -2.7%, and S-SE wave events (146° - 191°) are underestimated by -1.5%. Most large wave events over 3 meters are underestimated at the start of the forecast period, and this occurs for most wave directions with the exception of some outlier wave events. The wave forecast skill decreases with increasing lead time, and the skill decreases at faster rates by days 3-4 of the forecast and occurs for all incoming wave directions. The approaches developed here can be used to evaluate other wave models worldwide and the results establish a baseline for future model improvements.

Keywords: Spectral Wave Model, Forecast, Significant Wave Height, Wave Buoys, Satellites, Nearshore, Offshore, Direction, Period

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 OCEAN SURFACE GRAVITY WAVES AND SPECTRAL WAVE MODELS

Ocean surface gravity waves play an important role in offshore regions for commerce, maritime operations, search and rescue, aquaculture, and renewable energy developments. Wave conditions often dictate the safety of recreational activities such as swimming, diving, surfing, sailing, or lifeguarding. In addition, large waves impact nearshore regions by driving hydrodynamic processes that can negatively impact coastal infrastructure via flooding, coastal erosion, or sediment transport (Barnard et al., 2015).

Wave forecasts have become powerful tools which are capable of forecasting wave conditions well in advance. Wave modeling is the standard approach to predicting wave conditions, which solve theoretical equations based on fluid dynamics and wave physics to predict ocean waves. Wave models are largely driven by surface winds predicted by atmospheric models which can forecast conditions 14 days into the future. Spectral wave models evolve the wave energy as a function of time, geographical space, and wavenumber space. Spectral wave models such as WAVEWATCH3 (WW3) (WAVEWATCH III Development Group (WW3DG) 2019) are the operational models of weather prediction agencies such as the National Oceanic and Atmospheric Administration (NOAA). Although useful, wave models contain errors and their performance oftentimes degrades as a function of forecast lead time (Janssen et al., 1997). Janssen et al., (1997) emphasized the importance of identifying regional biases in spectral wave models so that the models can be recalibrated or the physical parameterizations can be improved. Building upon this idea, the objective of this research is to rigorously compare the archived PacIOOS wave forecast (Cheung, 2011) from 2011 to 2021 to moored wave buoys and satellite observations. This will reveal any model errors and build upon similar work using wave hindcasts (Li et al., 2016).

1.2 THE PACIOOS WAVE FORECAST SYSTEM FOR HAWAIʻI

The PacIOOS wave forecast system contains 3 tiers of spectral wave models at a global, regional, and island scale. PacIOOS's global WW3 model is a 0.5 degree (~50km) resolution model with the purpose of capturing large-scale ocean waves in the open ocean and providing boundary conditions for higher resolution spectral wave models. PacIOOS's higher resolution wave models part of the nested wave forecast system include the WW3 Hawaiʻi regional model which has a resolution of 0.05 degrees (~5 km), and four Simulating WAves Nearshore (SWAN) (Booij et al., 1996) spectral wave models for the nearshore waters of Hawaiʻi which each have a resolution of 500 m. The nested wave model system is designed to capture the spatial scales of importance at global, regional, and island. The nearshore domains are designed to capture effects that may be due to island sheltering, refraction, and locally generated wind waves (Cheung 2010).

1.3 WAVE FORECAST PARAMETERS

This research investigates errors in the forecast in terms of their significant wave height (*Hs*), mean absolute wave period (*TM01*), peak period (*Tp*), and peak wave direction (*Dp*). *Hs* is traditionally defined as the average of the largest ⅓ waves or estimated as four times the standard deviation of the sea surface elevation, *TM01* is defined as the duration in seconds between the arrival of the mean sea surface waves, *Tp* is defined as the duration in seconds between the arrival of the dominant sea surface waves, and *Dp* is defined as the direction in which the dominant sea surface waves sea surface waves are traveling from (PacIOOS 2001, Cheung 2010). The standard parameters are typically measured over a 30-minute time period. It is important to note that these standard wave parameters represent average wave conditions and may not accurately represent individual swell events. The model and observation use the same method to calculate the wave parameters which allows us to compare the modeled outputs to observations.

2.0 METHODS

2.1 DATASETS

Two observed reference datasets are used to evaluate the wave forecast: 1) moored wave buoys and 2) space-borne satellite altimeters. Moored wave buoys are anchored to the ocean floor in strategic locations to measure the waves near Hawai‘i. Offshore wave buoys measure incoming waves in deep waters, and nearshore wave buoys measure waves as they begin to enter shallow waters. 5 offshore wave buoys, operated by NOAA, provide information of the offshore and deep water wave conditions. These data are compared to the Hawai‘i regional model. All offshore wave buoys except buoy 51003 exist just outside the domain of the Hawai‘i regional model (**Fig. 1.**), so we use the closest locations within the WW3 Hawai‘i domain to describe the wave conditions at the buoys. This approximation is justified because the wave conditions do not have a lot of variations in deep water. In addition to the 5 offshore wave buoy locations, 8 nearshore wave buoys operated by the Coastal Data Information Program (CDIP) provide nearshore wave information that is used to evaluate the island-scale SWAN models (**Fig. 1**). All wave buoy locations are listed in **Table 1**, along with the model grid comparison, location, depth, and temporal coverage used for the analysis. Wave buoys often experience outages due to maintenance and the temporal coverage does not reflect any missing data that may be within that period of time as well as any missing model data.

Table 1. Offshore and nearshore wave buoys

Buoy	**Computational Grid**	**Latitude (°N)**	**Longitude (°W)**	**Depth (m)**	**Temporal Coverage**
51000	Hawai‘i WW3	23.528	153.792	4762	2011.1 - 2023.12
51001	Hawai‘i WW3	24.475	162.030	4895	2015.8 - 2023.12
51002	Hawai‘i WW3	17.070	157.755	4979	2011.1 - 2023.8
51003	Hawai‘i WW3	19.143	160.645	5023	2011.1 - 2023.12
51004	Hawai‘i WW3	17.538	152.230	5278	2011.3 - 2023.12
51201	Oahu SWAN	21.671	158.118	200	2016.1 - 2023.6
51202	Oahu SWAN	21.417	157.680	86	2016.1 - 2021.7
51203	Maui SWAN	20.788	157.010	201	2016.3 - 2017.4
51205	Maui SWAN	21.018	156.421	200	2016.1 - 2022.10
51206	Hawaii Island SWAN	19.779	154.970	345	2016.3 - 2023.1
51207	Oahu SWAN	21.477	157.752	81	2016.1 - 2023.1
51208	Kauai SWAN	22.285	159.574	200	2016.1 - 2023.11
51212	Oahu SWAN	21.323	158.149	280	2018.5 - 2023.12

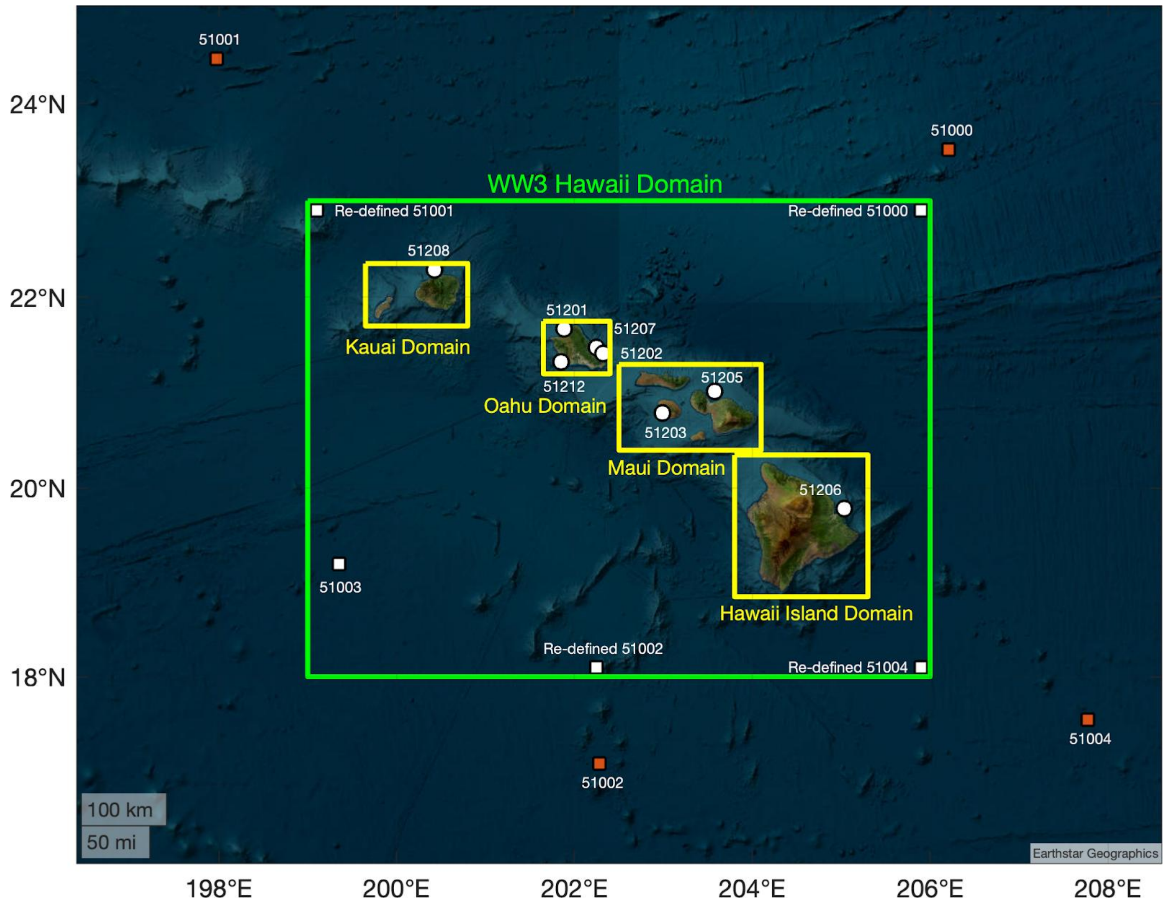

Fig. 1. The Hawaiian Island Chain's bathymetry, the location of buoys used, and the model grids used. The original locations of the offshore buoys are shown in orange and the re-defined offshore buoys along with the nearshore buoys are shown in white. The WW3 Hawaiʻi regional model is shown in green, and the four SWAN island-scale models are shown in yellow.

Satellite altimeters measure the *Hs* across large expanses of geographic space. Dodet et al., (2022) created a merged satellite-derived and calibrated *Hs* dataset across the global ocean. This dataset contains 7 different satellite missions (**Fig. 2b.**) from both the European Space Agency (ESA) and NASA and spans several decades. For this research, the satellite-derived *Hs* data in the WW3 Hawaiʻi regional model domain (**Fig. 2.**) from 2011 - 2021 is used. **Fig. 2a.** highlights the uneven distribution of altimeter observations across the domain for the given time period. Given the uneven distribution of observations, this might have implications when conducting spatial analysis, especially when analyzing extreme waves. Despite the uneven sampling size, the satellite *Hs* dataset provides spatial coverage to evaluate the Hawai‘i regional model complementing the fixed locations of the moored wave buoys.

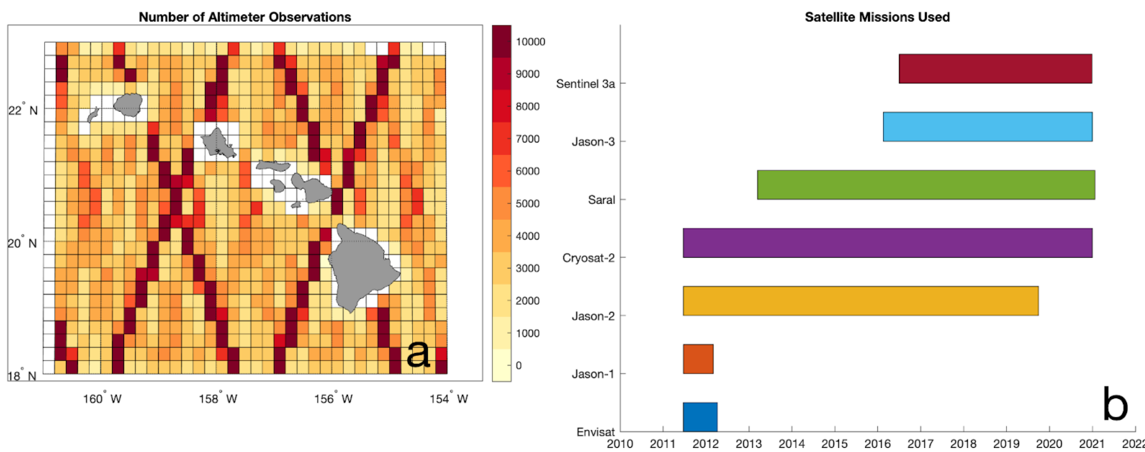

Fig. 2. (a) Altimeter sampling size per grid cell from 2011 to 2021 using the Dodet et al., (2022) satellite-derived and calibrated *Hs* dataset. The color bar represents the number of altimeter observations for each 0.2° bin. (b) The length in time of the satellite missions in the dataset.

We evaluate two model datasets from the PacIOOS nested forecast system: 1) the WW3 Hawai‘i regional model and 2) the four SWAN nearshore models (**Fig. 1.**). Both the regional and island-scale models produce an hourly forecast which output *Hs*, *TM01*, *Tp*, and *Dp* and extends out to 7 days. To directly compare the modeled outputs to the observations, we match the model to the observation in time and space. We find the closest forecast time to buoy time and set an absolute difference threshold of 60 minutes. We interpolate the model to the spatial coordinates of the observation.

To ensure the model and observation data are accurate, we conduct several assessments to quality assess/control (QA/QC) the datasets. We remove all ‘flagged’ data expected to be of low quality from model and observation datasets. We remove the modeled *Hs* values <= 0.3 m for all modeled outputs. In addition, we remove the first 2-

25 hours of the forecast depending on the location of the forecast-observation comparison due to the model's "spin up time." The nearshore SWAN domains do not use initial conditions and instead start with a flat ocean and propagate the waves from the boundaries throughout the domain. Consequently, the model needs time to propagate the waves from the boundaries across the domain.

2.2 WAVE FORECAST EVALUATION STRATEGY

Four error metrics are used to evaluate the wave forecast system. We use Root Mean Square Error (RMSE) to investigate the model performance. A RMSE value of 0 indicates a perfect model performance. We use a normalized RMSE (NRMSE) to account for the regional differences in average wave heights across Hawaiʻi and to intercompare the different wave parameters. We use BIAS to investigate where the model is underestimating and overestimating wave events. We use a normalized BIAS (NBIAS) for the same reason as normalizing RMSE while investigating the underestimation and overestimation. The four error metrics are as follows, where o = observation, p = prediction, and n = sampling size.

(1) $\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (p_i - o_i)^2}$

(2) $\mathrm{NRMSE} = \frac{RMSE}{mean(o)}$

(3) $\mathrm{BIAS} = \frac{1}{n} \sum_{i=1}^{n} (p_i - o_i)$

(4) $\mathrm{NBIAS} = \left(\frac{\frac{1}{n} \sum_{i=1}^{n} (p_i - o_i)}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} o_i^2}} \right) \times 100$

2.3 WAVE SEASONALITY IN HAWAI‘I

Unobstructed by any major land masses, Hawai‘i experiences waves from all directions. During December-January-February, Hawai‘i experiences its largest waves. Mid-latitude cyclones originating in the far North Pacific near the Aleutian Island Chain generate waves that typically take ~1-3 days to propagate to Hawai‘i. These waves travel away from its storm source and often arrive as organized ‘swells’ containing wave periods up to ~20 seconds and ranging from the WNW to NE. During the Austral winter (Hawai‘i’s summer), mid-latitude cyclones in the South Pacific generate waves that propagate from the south. Waves from the south have a travel time of ~6-9 days to Hawai‘i and are much smaller than waves from the north. These waves can also contain wave periods up to ~20 seconds and typically come from directions between the SW and SE. Due to the semi-permanent high-pressure system north of Hawai‘i, windward regions experience persistent year-round ‘wind waves’ from the ENE. These ENE ‘wind waves’ have very short wave periods, typically under 10 seconds. The waters across Hawai‘i contain a very steep continental slope, and the offshore waters quickly transition to depths of 1000s of meters. Due to this rapid transition of water depth, most of the wave transformations occur close to the coastlines (e.g. refraction and shoaling). When waves begin to feel the ocean floor at depths of half their wavelength, they undergo wave transformations, oftentimes altering the original direction of the waves.

We highlight the strong seasonal variability of Hawai‘i’s wave climate (**Fig. 3.**) using the satellite altimeter dataset of Dodet et al., (2022). The December-January-February (DJF) period experiences much larger waves on average than the June-July-August (JJA) period. During this time the largest wave heights occur north of the islands generated by mid-latitude cyclones located well north of Hawai‘i in the Pacific Ocean. During the DJF period, sheltering occurs in regions where waves cannot refract all the way into protected regions such as the south side of the Hawai‘i island. Sheltering also occurs during persistent east wind waves on the west side of the Hawai‘i island, where the JJA period shows the largest waves occurring offshore to the east, and the smallest waves occurring on the immediate western sides.

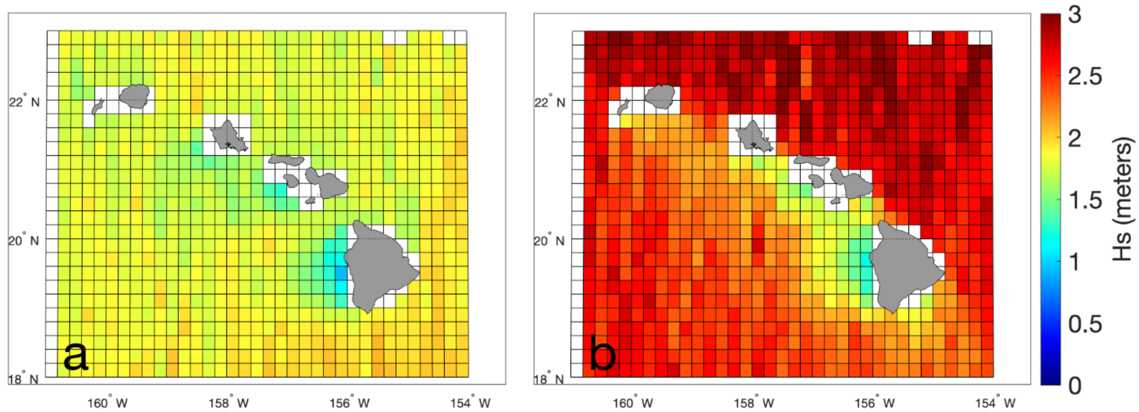

Fig. 3. Wave seasonality in Hawai‘i calculated using 10 years of satellite altimetry *Hs* for (a) June-July-August (JJA) and (b) December-January-February (DJF) . The colors represent the average *Hs* in meters for each 0.2° bin.

3.0 RESULTS

The results are organized into two main sections: 1) Baseline Uncertainty (days 1 + 2) and 2) Forecast Lead Time Uncertainty. The baseline uncertainty allows us to investigate the performance of the wave forecast system closest to the actual time of the observation, and this will establish the reference error. We use days 1 + 2 to establish the baseline uncertainty to have a sufficiently large sampling size. Janssen et al., (1997) show that days 1 and 2 perform quite similarly and consistent with our results in **Fig. 12.** and **Fig. 13.** Analysis as a function of lead time allows us to investigate the performance of the wave forecast system as the forecast lead time approaches day 7.

Throughout the analysis, 5 directional bins are used to reference dominant wave generation regions and defined as follows: 281° - 348° (NW waves), 348° - 56° (N-NE waves), 191° - 236° (SW waves), 146° - 191° (S-SE waves), 65° - 100° (E waves). In order to measure all angles in a circular system, we account for the wrap-around effect at 360° throughout the entire analysis and measure the smallest angular distance. Analyzing the significant wave heights into directional bins helps elucidate errors as a function of storm generation source. This directional bin analysis is conducted for both buoys and altimeters, at both the start of the forecast and as a function of forecast lead time.

3.1 BASELINE UNCERTAINTY

To intercompare different wave parameters, NRMSE is used rather than RMSE. The NRMSE for *Hs* is calculated at both offshore and nearshore buoy locations and NRMSE for *TM01* is only calculated at the nearshore buoys because *TM01* was not archived for the regional WW3 domain. The NRMSE for forecast-altimeter comparisons is computed within a 50 km radius of each offshore buoy and can be compared to the NRMSE forecast-offshore buoy comparisons. The largest NRMSE values of 22.7% and 22.3% for *Hs* and 21.2% and 27.8% for *TM01* from forecast-buoy comparisons exist at the 51212 and 51203 buoys (**Fig. 4.**). These two locations are both exposed to the smaller southerly waves, and mostly sheltered from the year-round east wind waves as well as large northerly waves. The other buoys contain lower NRMSE values mostly ranging between 10% and 20% and for *Hs* and *TM01*. Most of these locations are exposed to a variety of wave directions, especially the offshore buoys. When focusing on only the nearshore buoys, the NRMSE values for *TM01* are lower than that for *Hs*, with the exception of buoy 51203. It is also notable that 4 of 5 offshore locations recorded a lower NRMSE value for the forecast-altimeter comparisons vs the forecast-offshore buoy comparisons. This means the altimeter observations align more closely with the model than the buoy. The only instance the buoy recorded a lower NRMSE for forecast-offshore buoy vs. forecast-altimeter was for buoy 51002.

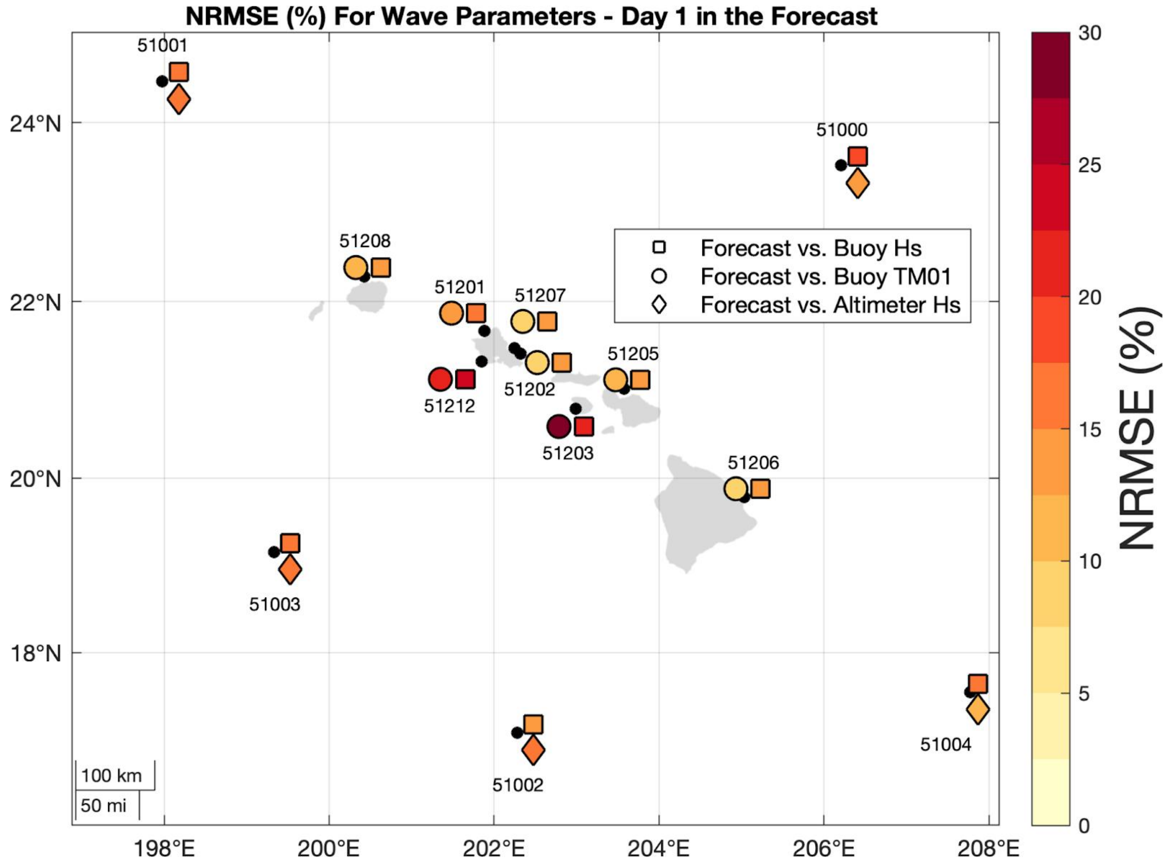

Fig. 4. NRMSE for forecast-buoy *Hs*, forecast-buoy *TM01*, and forecast-altimeter *Hs* at day 1 of the forecast. The forecast-altimeter NRMSE is computed within a 50 km radius of each offshore buoy. The color bar represents the NRMSE value in %.

In **Fig. 5.** the forecast-altimeter data is evaluated using the modeled *Dp* and the observed wave height as reference and the forecast-buoy data uses the observed *Dp* as reference. With the exception of the forecast-altimeter NW wave events, wave heights above 3 meters are underestimated. For all wave directions the smallest wave heights under 1.5 meters are overestimated. There are differences in performance between NW and N-NE waves. On average, N-NE wave events are underestimated more than NW wave events. For the forecast-altimeter comparison, the wave heights of NW waves are overestimated, with the exception of 1 wave height bin at 4 meters as shown in **Fig. 5a.** The increase in overestimation with larger NW wave events over 4 meters occurs over a relatively small sampling size as shown **Fig. 5c** and consists of ~105 wave events over 4 meters and only ~25 wave events over 5 meters. Due to the small sampling size of comparisons and limited number of wave events, it is difficult for us to interpret the forecast-altimeter comparisons for NW waves since the uncertainty is likely large. The forecast-altimeter comparison shows underestimation in wave forecast for N-NE wave events over 2 meters, and the result is consistent with the forecast-buoy comparison for N-NE wave events. For NW wave events, there is a noticeable difference between the two reference datasets: altimeter and buoy. Negative NBIAS referenced by buoy data for NW waves consistently increase as a function of wave height unlike the NBIAS referenced by altimeter, where the NBIAS remains positive for the large wave events. The sampling size for the large wave events are much smaller as shown in **Fig. 5c.** and any wave height bin with less than ~100 comparisons was removed. There is not a large discrepancy between the NBIAS values of SW waves vs. S-SE waves, but the waves generated in the South Pacific show a greater underestimation than the waves generated in the North Pacific, particularly for wave heights from 2 - 2.5 meters (**Fig. 5b.**). There is decent agreement between the two reference datasets for E waves. The smaller E waves under 2 meters are generally overestimated and the larger E waves above 2 meters are generally underestimated.

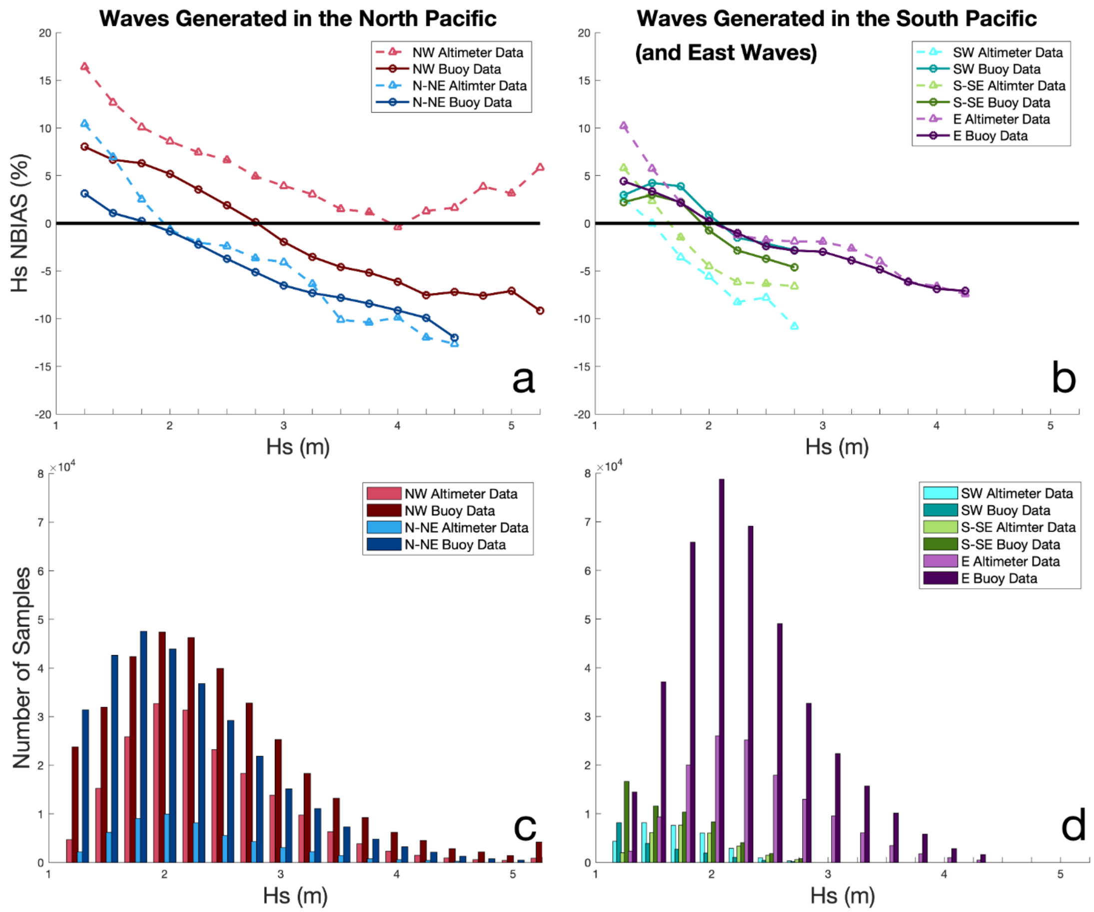

Fig. 5. (a & b) *Hs* NBIAS (%) as a function of *Hs* in meters using the 4 directional bins during the beginning of the forecast (1-2 days). The dotted lines represent the forecast-altimeter comparisons where the modeled *Dp* is used as reference and the solid lines represent the combined forecast-buoy comparisons where the observed *Dp* is used as reference. Above the zero line in black represents an overestimation and below the zero line represents an underestimation. (c & d) The sampling size from a and b also as a function of *Hs* in meters.

In **Fig. 6.** the forecast-nearshore buoy data uses the observed *Dp* as reference when binning. Here, rather than directly comparing wave heights, we directly compare *Tm01* values for forecast vs. nearshore buoy. Because *TM01* is the mean wave period rather than the peak wave period, there are no *TM01* values larger than 16s. For NW waves, N-NE waves, and E waves, the *TM01* NBIAS values are within +/- 5% despite the wave period value. Both NW and N-NE waves follow a similar trend which is a positive

NBIAS from *TM01* of 4-10s, then approaches a negative NBIAS after 12s (**Fig. 6a.**), but the sampling size rapidly decreases at *TM01* >= 12s (**Fig. 6c.**). E waves behave similarly showing a slight decrease in *TM01* NBIAS with larger *TM01* values, but do not reach negative NBIAS values. The *TM01* NBIAS for SW and S-SE waves are underestimated with the exception of *TM01* of 4s. Both SW and S-SE waves surpass an underestimation of 15% when *TM01* = 10 seconds and approach but not reach an underestimation of 20% when *TM01* = 12 seconds (**Fig. 6b.**). Similarly to the larger *TM01* values for waves generated in the North Pacific, the sampling size becomes quite small for the larger *TM01* values for waves generated in the South Pacific, with 96 comparisons for SW waves at 12 seconds and 178 comparisons for S-SE waves at 12 seconds (**Fig. 6d.**).

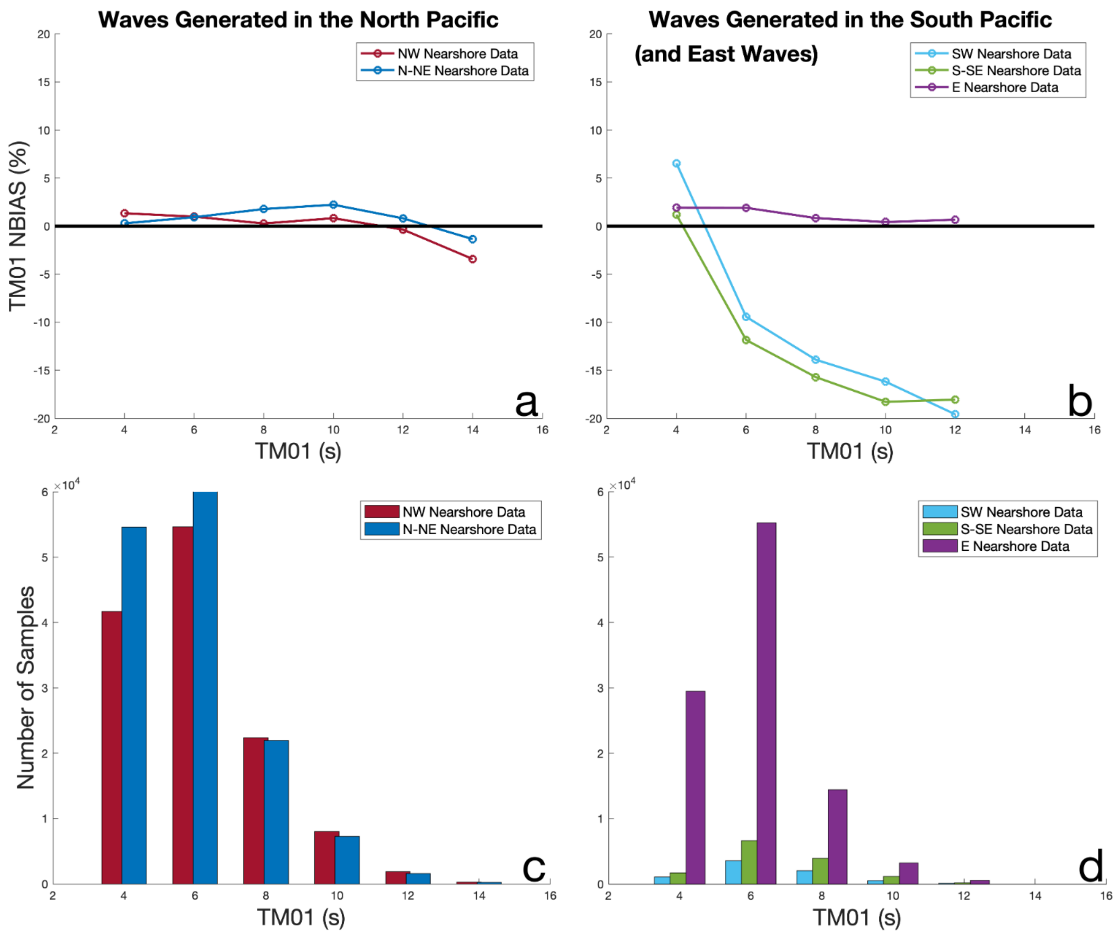

Fig. 6. (a & b) *TM01* NBIAS (%) as a function of *TM01* in seconds using the 4 directional bins during the beginning of the forecast (1-2 days) for nearshore forecast-buoy comparisons. The observed *Dp* is used as reference. Above the zero line in black represents an overestimation and below the zero line represents an underestimation. (c & d) The sampling size from a and b also as a function of *TM01* in seconds.

The density scatter plots in **Fig. 7.** provide insight into the prevailing wave conditions at nearshore locations as well as the intercomparison between the 3 wave parameters. The nearshore *Hs* scatter plot reveals an underestimation of the *Hs* when looking at the data density distribution relative to the perfect line and plotted percentiles below the perfect line, particularly for *Hs* > 2 meters. The underestimation of *Hs* is confirmed with the overall NBIAS calculation of -2.4% and BIAS of -0.05 meters. The nearshore *TM01* scatter plot reveals a larger spread in the data, but experiences a lower

NRMSE value of 14.2% compared to the nearshore *Hs* NRMSE value of 16.7%. The wave forecast also tends to overestimate the *TM01* very slightly at NBIAS of 0.6% and BIAS of 0.04 seconds, with the percentiles landing close to but slightly above the perfect line for *TM01* values above 8 seconds. Overall, the forecast tends to overestimate the nearshore *Dp* by 1.5% and 3.1 degrees, but it depends on the direction of the incoming waves as the data density and percentiles reveal positive biases for waves incoming from ~75 degrees and negative biases for waves incoming from 300 - 360 degrees. The denser data in the bottom left of **Fig. 7c.** represents the E 'wind waves' with more of the data above the perfect line, the not as dense but still common data in the middle represents the SW and S-SE waves generated in the South Pacific Ocean, and the dense data in the top right represents the NW and N-NE waves generated in the North Pacific Ocean, which is slightly below the perfect line.

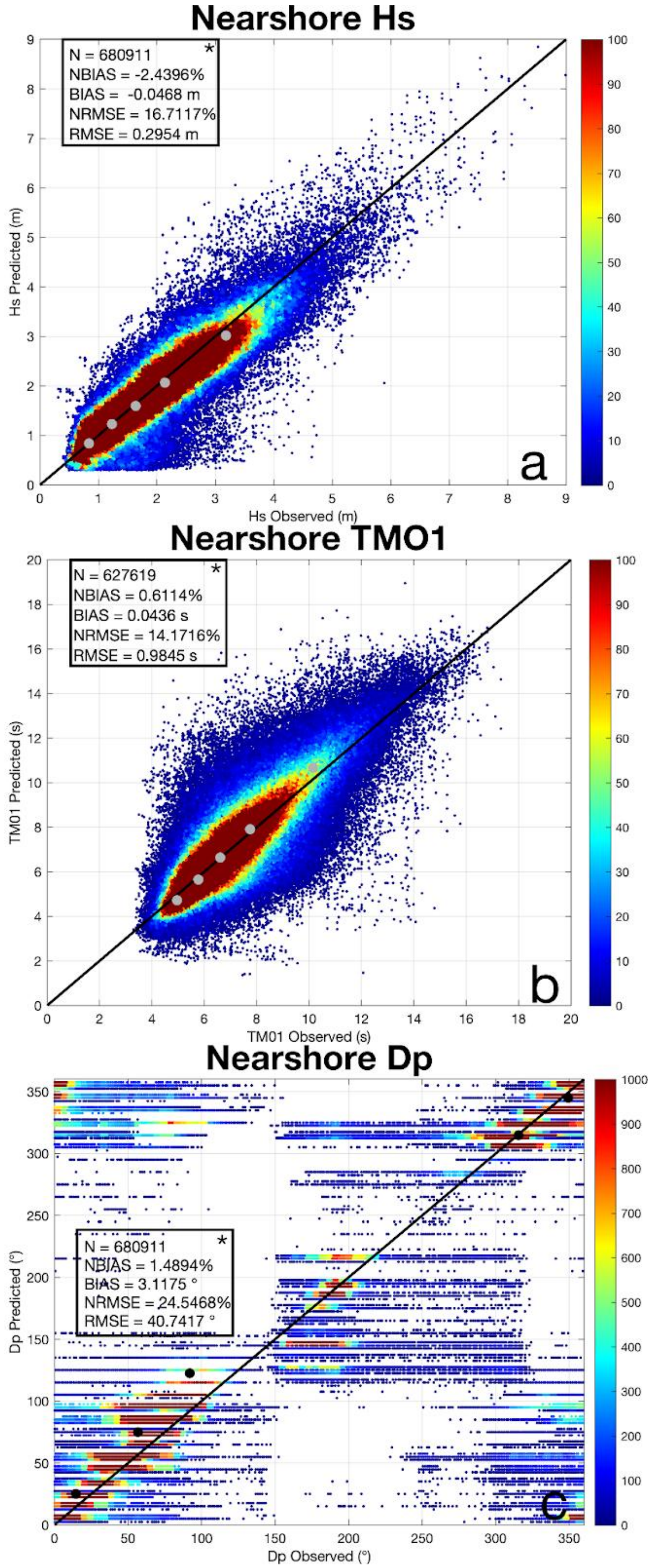
Nearshore Hs
N = 680911
NBIAS = -2.4396%
BIAS = -0.0468 m
NRMSE = 16.7117%
RMSE = 0.2954 m
Hs Predicted (m)
Hs Observed (m)
a
Nearshore TMO1
N = 627619
NBIAS = 0.6114%
BIAS = 0.0436 s
NRMSE = 14.1716%
RMSE = 0.9845 s
TM01 Predicted (s)
TM01 Observed (s)
b
Nearshore Dp
N = 680911
NBIAS = 1.4894%
BIAS = 3.1175 °
NRMSE = 24.5468%
RMSE = 40.7417 °
Dp Predicted (°)
Dp Observed (°)
c

Fig. 7. Nearshore buoy data density scatter plots for *Hs*, *TM01*, and *Dp* during the beginning of the forecast (1-2 days) with overall error metrics calculated in the top left of each plot. An asterisk denotes significant difference between the observed and predicted datasets at the 95th confidence limit. Colors denote data density where red/blue colors indicate large/small sample sizes. Points above the black perfect line represent the wave forecast overestimating, while points below the black perfect line represent the wave forecast underestimating. The percentiles are plotted at the 5, 25, 50, 75, and 95% level.

The density scatter plots in **Fig. 8.** give further insight into prevailing wave conditions with respect to storm generation source as well as outlier wave events, similar to the sampling size plots in **Fig. 5c,d.** NW waves are much more common than the other wave events with a sampling size of 165928 compared to the second most common N-NE waves with a sampling size of 47311. On average, NW waves get overestimated by 6.0% while N-NE waves get underestimated by -1.7%. The data density and percentiles reveal this, and the overestimation for NW waves and underestimation for N-NE waves becomes more pronounced with waves larger than 2 meters. This overestimation for NW waves and underestimation of N-NE waves is consistent with **Fig. 5.** The waves from N-NE, SW, and S-SE, all experience an underestimation on average based on the negative BIAS, which is also consistent with **Fig. 5.** The SW and S-SE wave events get underestimated the most on average based on the NBIAS of -2.7% and -1.5%, but are not as noticeable when looking at the data density scatter plots and percentiles. The NW wave events contain the highest RMSE and NRMSE values of 0.4 meters and 17.4%.

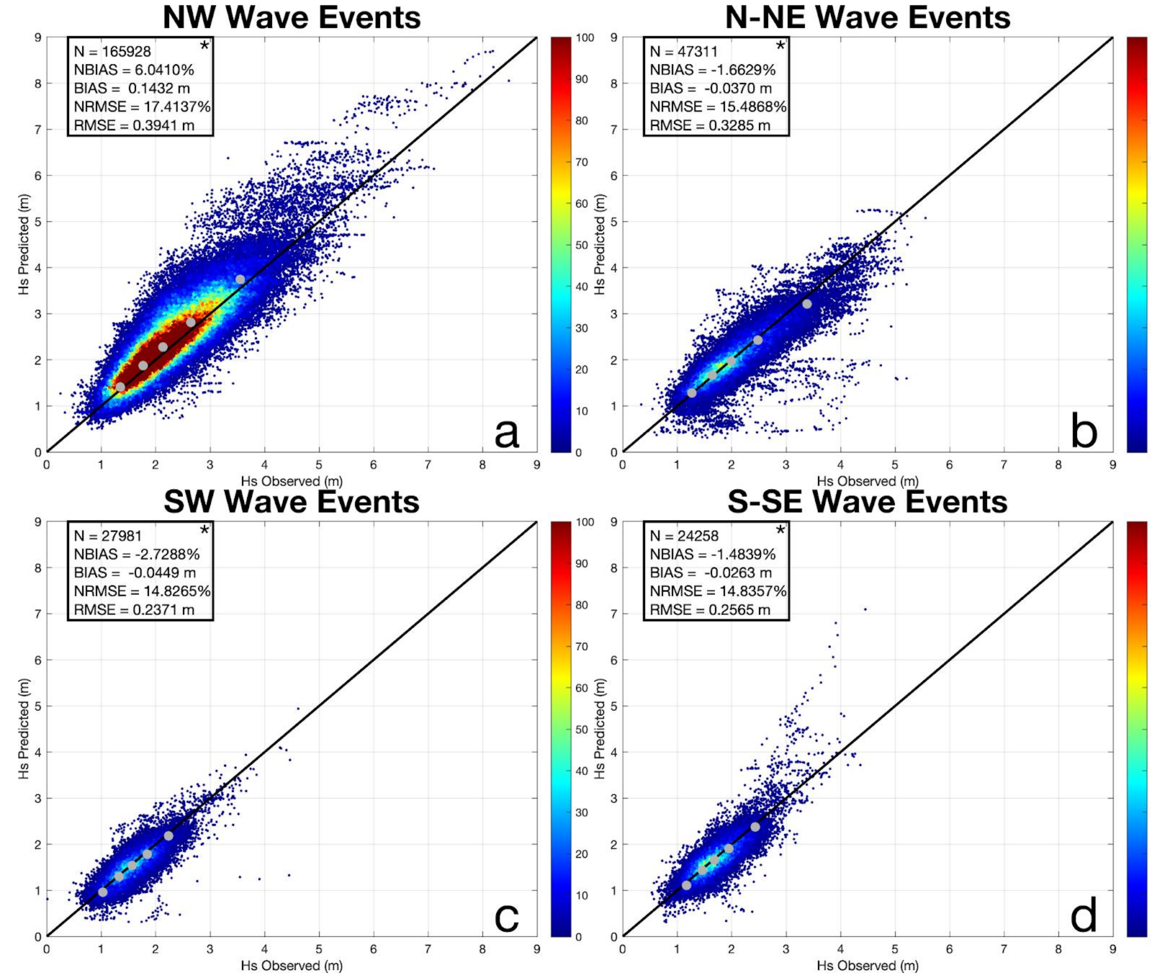

Fig. 8. Altimeter density scatter plots for the 4 directional bins using modeled *Dp* as the reference during the beginning of the forecast (1-2 days) with overall error metrics calculated in the top left of each plot. An asterisk denotes significant difference between the observed and predicted datasets at the 95th confidence limit. The colors denote data density in each 0.045 m by 0.045 m bin where red/blue colors indicate large/small sample sizes. Points above the black perfect line represent the wave forecast overestimating, while points below the black perfect line represent the wave forecast underestimating. The percentiles are plotted at the 5, 25, 50, 75, and 95% level.

Fig. 9. shows the overall spatial error metrics for *Hs* using the altimeter-forecast comparisons. The largest RMSE values, around 0.4 meters, exist offshore and the smallest RMSE values, around 0.2 meters exist in the sheltered regions (**Fig. 9a.**). The opposite occurs for NRMSE where the largest values occur nearshore, particularly in the wake of the Hawaiʻi Island (**Fig. 9b.**) with NRMSE values over 20%. Higher NRMSE

values nearshore are similar with the NRMSE results from **Fig. 4.** where nearshore buoy locations 51212 and 51203 exposed to south waves experienced the largest NRMSE values among the forecast-buoy comparisons at day 1 in the forecast. The lowest NRMSE values exist to the east and southeast of the Hawai‘i Island. The overall spatial analysis also reveals an overestimation of the *Hs* in the offshore regions well north and south of the island chain, and an underestimation of the *Hs* in the immediate sheltered regions of the island chain (**Fig. 9c,d**). The largest underestimation of *Hs* NBIAS, around 10-15%, occurs in the wake of the Hawaii Island, which is in the same location as the highest NRMSE values. The RMSE and BIAS/ME values are consistent with Li et al., (2016) Hawai‘i hindcast analysis where they also found the highest RMSE values offshore around 0.4 meters, and BIAS/ME values offshore around 0.2 meters.

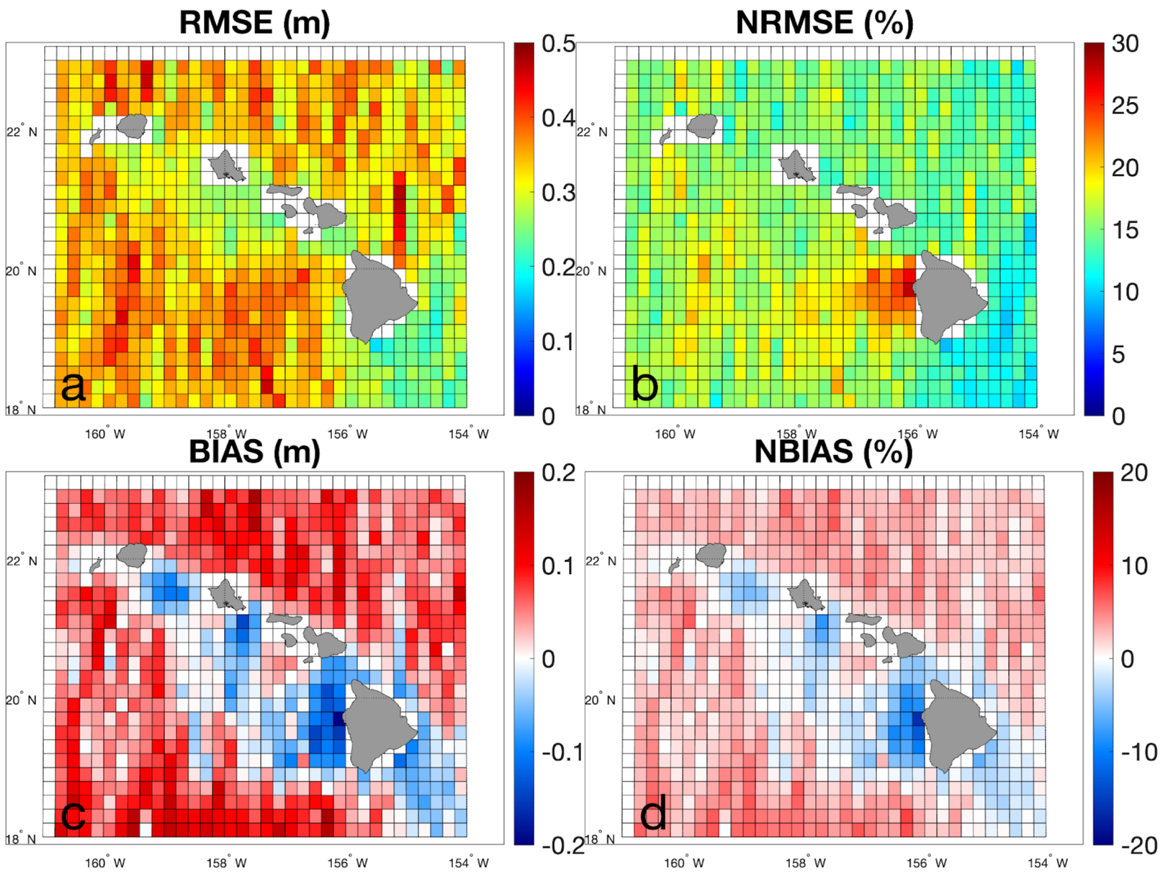

Fig. 9. Overall spatial error metrics for *Hs* at the beginning of the forecast (1-2 days). (a & b) The warm/orange colors indicate higher errors and the cool/green colors represent lower errors for each 0.2° bin. (c & d) The red represents an overestimation and the blue represents an underestimation for each 0.2° bin.

In **Fig. 10.** waves are divided into 2 wave period bins referenced by the modeled *Tp* and all wave heights are included in the error metric calculations. The NRMSE values for *Hs* during short period waves are lower overall, but high NRMSE values of 30% exist in the wake of the Hawaiʻi island (**Fig. 10a.**), similar to **Fig. 9c**. While the NRMSE values for *Hs* during long period waves are higher overall, they are more evenly spread throughout the domain, with more NRMSE values over 20% but no NRMSE values up to 30% (**Fig. 10b.**). Wave heights that are associated with short periods are underestimated much more than long period waves. The only underestimation for long period waves occurs in the sheltered regions and the rest of the offshore regions are generally overestimated (**Fig. 10d.**). The sampling size is likely dominated by the stronger long period NW waves as noted previously in **Fig. 8a.** and therefore likely more representative of long period NW waves rather than long period N-NE, SW, or S-SE waves. A larger underestimation in the wake of the Hawai‘i island is apparent in both NBIAS panels as shown in darker blue (**Fig. 10c,d**).

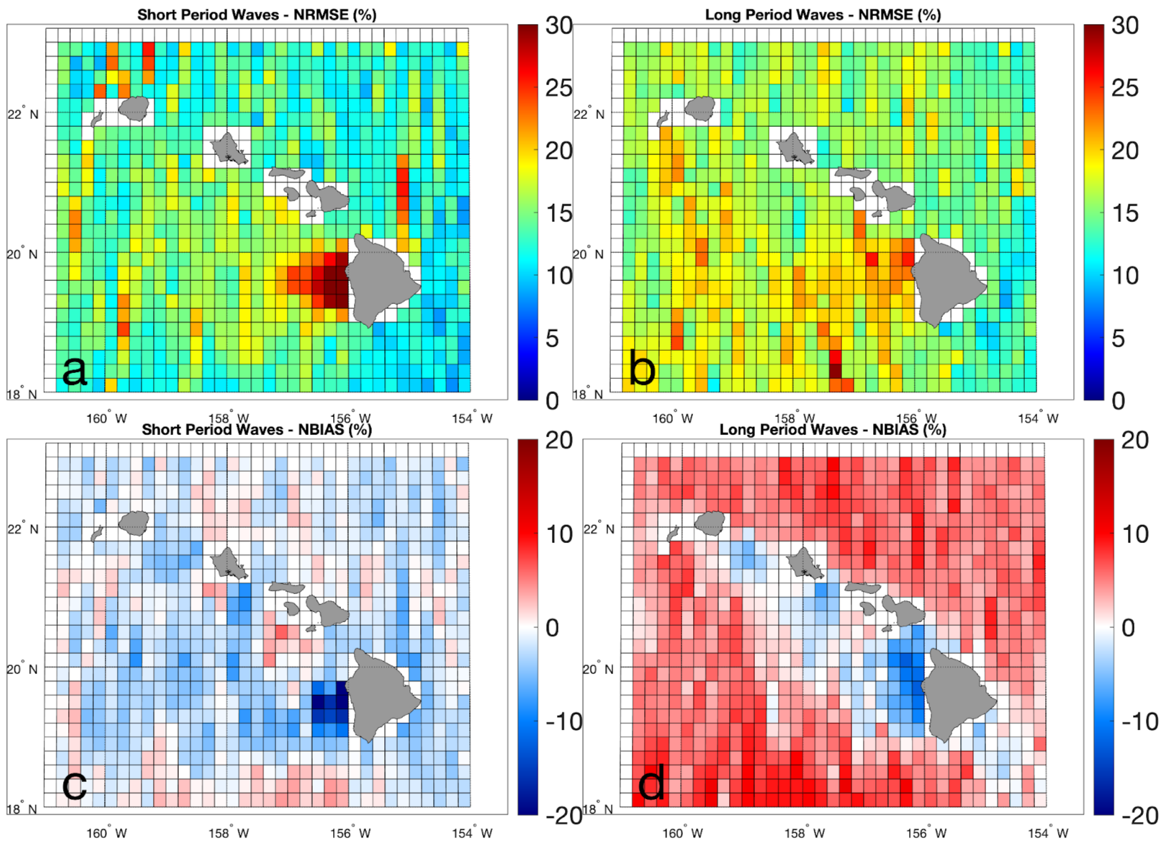

Fig. 10. Spatial NRMSE and NBIAS for *Hs* at the beginning of the forecast (1-2 days) using 2 peak period (*Tp*) bins which are referenced by the model. Short period waves are Tp < 10s and long period waves are Tp > 10s.

In **Fig. 11,** waves are divided into the same directional bins referenced by the modeled *Dp*. Here, all wave heights are included in the error metric calculations and we also set a minimal sampling size of 50 to avoid erroneous NBIAS values, especially for areas where a particular wave direction is not dominant and cannot refract into sheltered locations. Missing data or bins with less than 50 comparisons is shown in white and could be caused by lacking *Dp* in that area. The lack of data is particularly noticeable in the southerly waves plots (**Fig. 11c,d**) as a result of the islands blocking southerly waves. Because all wave height sizes are used, and because the smaller wave heights ($Hs < 3$) are much more common (**Fig. 5 & 8**), the results here are likely more representative of how smaller wave heights perform. **Fig. 11a** reveals that the NW wave events experience the highest overestimation, most notably in the southern portion of the domain, with some underestimation around sheltered areas. These results align with the results from Li et al., (2016) Hawai‘i hindcast analysis where they also found overestimation in NW waves and underestimation in sheltered areas. The N-NE wave events are underestimated much more than the NW wave events, consistent with the results from **Fig. 5.** and **Fig. 8.**. The less frequent coverage of N-NE waves in **Fig. 11b.** is likely a result of the intensity of N-NE wave events which are typically not as strong as NW wave events. Despite the smaller sample sizes in N-NE, SW, and S-SE waves, these comparisons reveal a noticeable underestimation compared to the NW wave events. The most notable underestimations occur for SW and S-SE waves in the wake of the islands, where east wind waves are less common due to sheltering from the islands. The blue colors in the wake of the islands may give a more accurate representation of south swells instead of a mix of south and east waves further offshore. This is likely why there are less areas offshore where south waves are dominant (**Fig. 11c,d.**).

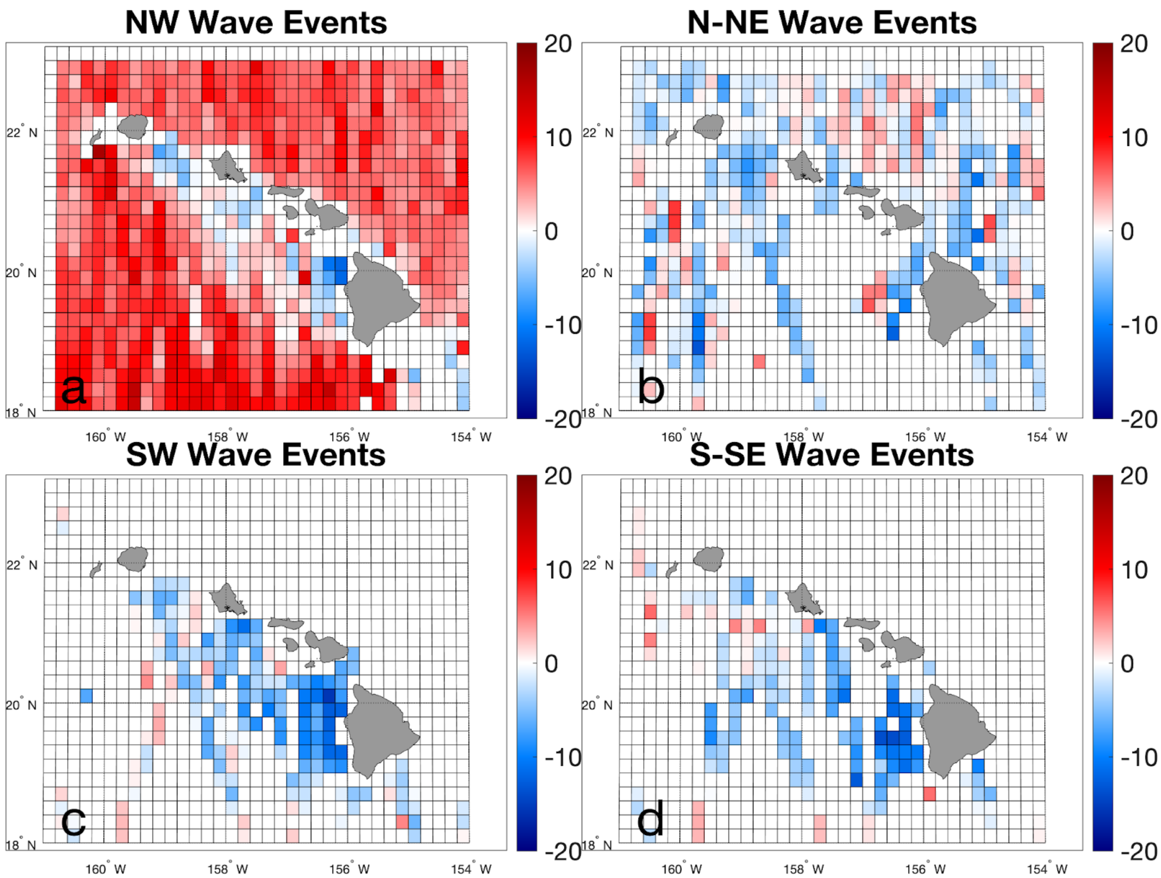

Fig. 11. *Hs* NBIAS (%) plotted spatially at the beginning of the forecast (1-2 days) using the 4 directional bins. Red colors represent overestimation and blue colors represent underestimation for each 0.2° bin.

3.2 FORECAST LEAD TIME UNCERTAINTY

In **Fig. 12.** RMSE, NRMSE, BIAS, and NBIAS are plotted as a function of forecast lead time using all *Hs* comparisons. We use all the forecast-altimeter data and all the combined forecast-buoy data. The first 24 hours of the forecast-altimeter comparisons are removed because of poor performance due to the model "spin up time." There is still some "spin up time" effect for the combined buoy data at day 1 in (**Fig. 12a,b.**), even with the first 12 hours of the forecast removed. We find that the RMSE increases as a function of forecast lead time for both forecast-altimeter and forecast-buoy comparisons. The slight difference between the forecast-altimeter and forecast-buoy lines are likely a result in how the forecast-altimeter data is more spread with data across the domain, and the same can be said for the other error metrics. For both datasets, the RMSE increases at a faster rate by days 3 and 4 of the forecast. The results here are relatively consistent with Janssen et al., (1997) who find increased errors as a function of forecast time when verifying the ECMWF Wave Forecasting System against wave buoy and altimeter data. The NRMSE values as a function forecast lead time follow the same increasing trend as the RMSE values (**Fig. 12a,b.**), and both comparisons reach a NRMSE value of around 20% by day 7 in the forecast compared to a NRMSE value of around 16% at days 1 and 2 of the forecast. We also find that further into the forecast period, the wave forecast tends to overestimate the *Hs*. Both the BIAS and NBIAS for the forecast-buoy lines begin close to 0 m or 0% at the beginning of the forecast, followed by a linear increase to about 0.05 m or 2% by day 7 in the forecast (**Fig. 12c,d.**). The forecast-altimeter comparison overestimates the *Hs* slightly more at the beginning of the forecast, with a BIAS value of about 0.04 meters at day 2 of the forecast and about 0.1 m by day 7. The forecast-altimeter comparison overestimates the *Hs* by about 2% at day 2 of the forecast and about 5% by day 7. Again, the difference in BIAS and NBIAS calculations between the forecast-altimeter comparisons and forecast-buoy comparisons are likely caused by the differences in data distribution across the domain.

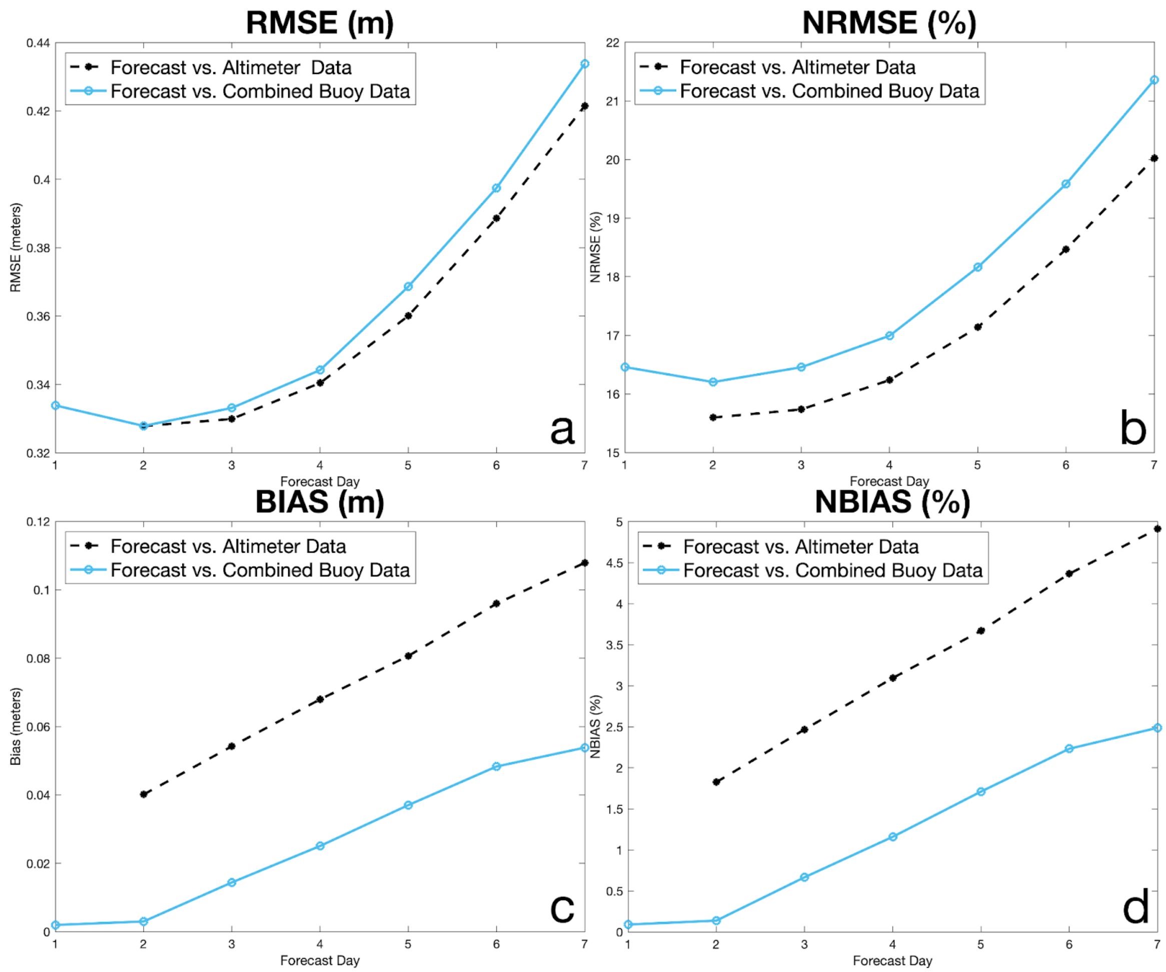

Fig. 12. (a) *Hs* RMSE in meters, (b) *Hs* NRMSE %, (c) *Hs* BIAS in meters, and (d) *Hs* NBIAS % as a function of forecast lead time. The black dotted lines represent all the forecast-altimeter comparisons and the solid blue lines represent the combined forecast-buoy comparisons.

In **Fig. 13.** NRMSE is plotted as a function of forecast lead time for the same directional bins used throughout the analysis. The results here are consistent with the results from **Fig. 12b.** which both reveal an increase in NRMSE as a function of forecast lead time, and that is true for waves incoming from all directions. For waves generated in the North Pacific, we see consistency between the two reference datasets. The forecast-altimeter comparison for NW waves follows a similar trend to the forecast-buoy comparison for NW waves. The same can be said for N-NE waves with slightly less consistency between the two reference datasets (**Fig. 13a.**). There is a clear difference in NRMSE values between the NW waves and N-NE waves, and on average according to both reference datasets, NW waves have higher NRMSE values to start and end the forecast. Consequently, there is no clear distinction between the slope values for NW vs. N-NE waves, which indicates that a particular direction for waves generated in the North Pacific does not necessarily have larger NRMSE values as the forecast lead time increases. For waves generated in the South Pacific, there is much less consistency between the two reference datasets. The NRMSE values for the forecast-altimeter comparisons are much lower than the NRMSE values for the forecast-buoy comparisons (**Fig. 13b.**). The NRMSE values for forecast-altimeter comparisons begin around 15%, while the forecast-buoy comparisons begin closer to 20%. This inconsistency between the reference datasets is again likely due to the differences in data distribution across the domain, where only so many wave buoys are exposed to south waves, while the forecast-altimeter comparisons cover the entire domain, and likely covering most of the Hawai‘i Island wake as shown in (**Fig. 11c,d.**). There is also no clear distinction between the slope values for waves generated in the South Pacific, so it is difficult to come to any conclusions if SW or S-SE waves have higher NRMSE values as the forecast lead time increases, as well as if a certain storm generation source (North Pacific vs. South Pacific) contains higher NRMSE values as the forecast lead time increases. The most consistent comparison between the two reference datasets occurs for E waves, where the reference datasets follow a similar trajectory and contain similar slope values (**Fig. 13b.**).

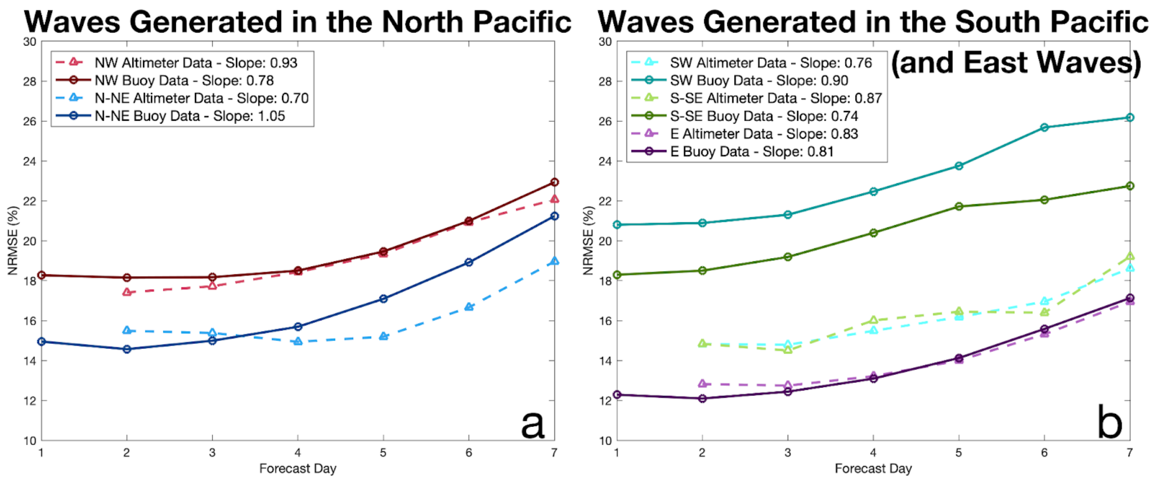

Fig. 13. NRMSE (%) as a function of forecast lead time using the 4 directional bins. The forecast-altimeter comparisons in dotted lines and forecast-buoy comparisons in solid solids are split up for (a) waves generated in the North Pacific and (b) waves generated in the South Pacific plus east waves.

The "forecast error" (Hs $NRMSE_{day6+7}$ - Hs $NRMSE_{day6+7}$) for the forecast-altimeter comparisons across the Hawai'i domain is shown in **Fig. 14**. The forecast error across the domain is positive with the exception of a few outliers likely due to small sampling size (**Fig. 2a.**). Forecast error values of 5-10% are most common across the domain, and there is no geographical location that experiences higher forecast errors than others. The results here are relatively consistent with **Fig. 12b.** and **Fig. 13.** when subtracting the start of the forecast NRMSE values from the end of the forecast NRMSE values. **Fig. 14.** provides a spatial view of the forecast error which the previous figures do not do. Positive forecast errors indicate that the NRMSE values are higher at the end of the forecast than the start of the forecast, and this is true for the entire Hawai'i domain.

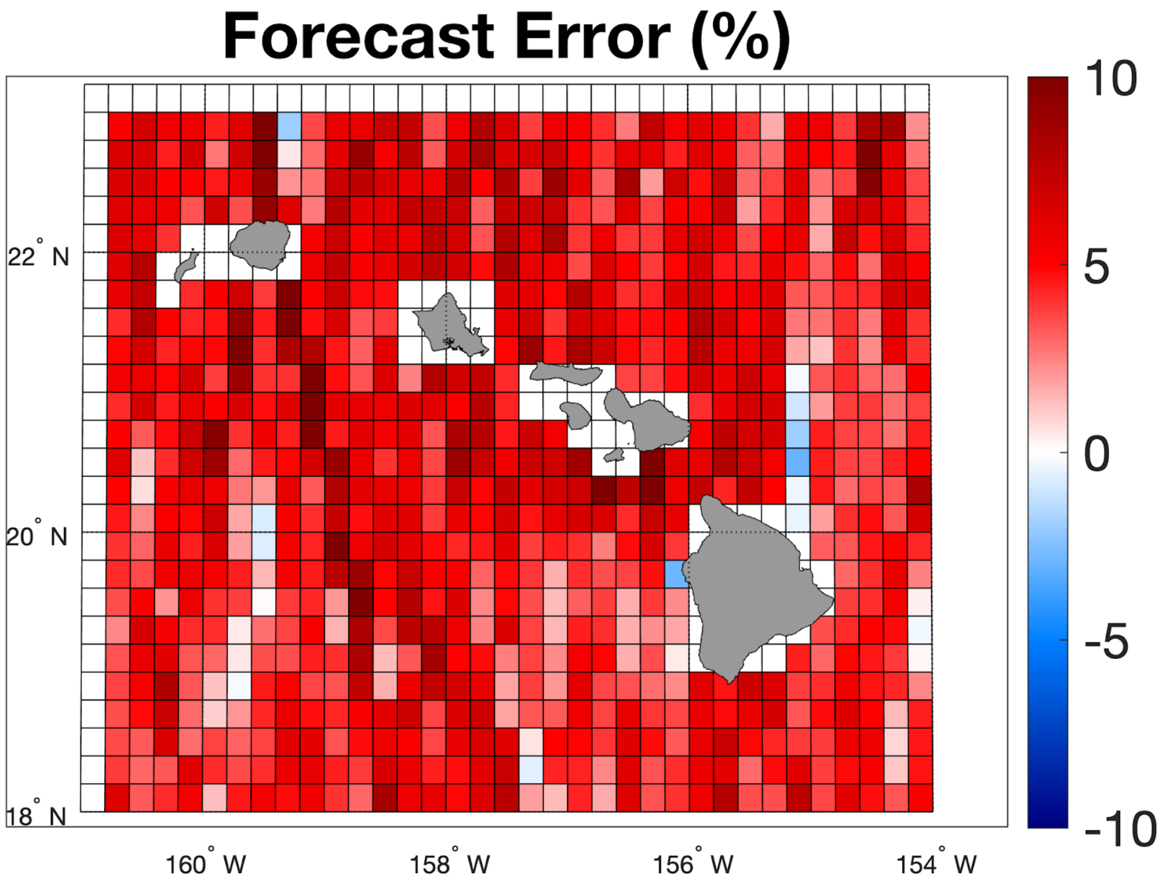

Fig. 14. The forecast error (Hs NRMSE$_{day6+7}$ - Hs NRMSE$_{day6+7}$) plotted spatially using the forecast-altimeter comparisons. Reds represent positive forecast error values and blues represent negative forecast error values for each 0.2° bin.

4.0 DISCUSSION

The PacIOOS nested wave forecast performs generally well relative to moored wave buoys and satellite altimeter *Hs* observations. Our results show that the wave forecast performs best at the start of the forecast, and in some aspects performs similarly to the Hawaiʻi hindcast from 1991 to 2011 (Li et al., 2016). Our results also show that there are multiple areas where the wave forecast system can be improved, and this includes both at the start of the forecast and as a function of forecast lead time.

When intercomparing *Hs* and *TM01*, the results show that these two wave parameters behave differently at day 1 of the forecast. *TM01* performed better than *Hs* overall, but the performance changed depending on the geographic location of the comparison. **Fig. 4.** highlights the larger NRMSE values for both *Hs* and *TM01* in the nearshore areas of 51212 and 51203, with respective depths of 280 and 201 meters. This suggests that relative errors may be larger nearshore in shallower water, even though wave heights at these locations are much smaller on average (**Fig. 3.**). Alday et al., (2022) discuss that wave models generally perform less accurately nearshore than offshore due to local bathymetry, local wind waves, and local currents. These higher NRMSE values may also suggest that waves from the south have the larger normalized errors than other directions because buoys 51212 and 51203 are the only buoys exposed to south waves and protected from east wind waves. This may not be true because the directional bin NRMSE values in **Fig. 8.** are higher for NW waves than SW and S-SE waves and results from **Fig. 4.** do not bin the wave heights for directions.

At the beginning of the forecast, the wave forecast performs best for wave heights around 2 to 3 meters and this is evident in **Fig. 5a,b.** as well as in **Fig. 8**. N-NE waves perform best around 2 meters, while NW waves perform best around 3 meters. SW, S-SE, and E waves all perform best around 2 meters. Waves that are near 2 meters are also the most frequent (**Fig. 5** and **Fig. 8.**), so the wave forecast seems to be well-calibrated for these common conditions. The model has the largest discrepancies with the smallest and the largest waves. The larger waves are a greater concern because they are most impactful to Hawaiʻi for reasons such as coastal erosion, infrastructure, marine usage, search and rescue, etc.

There is not a specific wave period that performs best at the beginning of the forecast unlike the ~2 meter wave heights that perform best. For most waves, the *TM01* NBIAS values in **Fig. 6.** do not fluctuate as the wave period increases. The most noticeable change in *TM01* NBIAS occurs for the waves that are generated in the South Pacific, which is an underestimation of *TM01* with an increase in wave period. This is consistent with wave heights from waves generated in the South Pacific which are underestimated with an increase in wave height. Both an underestimation in wave height and wave period for the largest and longest south waves could have important implications when predicting large south wave events because underestimating the period can also result in a wave event coming in larger than initially predicted with an increase in shoaling.

The nearshore data density scatter plots in **Fig. 7.** provide further insight into parameters other than *Hs*, and there is also consistency with results from **Fig. 4**. Both **Fig. 7.** and **Fig. 4.** reiterate that *TM01* forecast-buoy comparisons contain lower NRMSE values than *Hs* forecast-buoy comparisons. The overall *TM01* forecast-buoy comparison does not contain a large NBIAS value, which demonstrates that breaking the analysis into directional bins like in **Fig. 6.** can reveal systematic errors with respect to storm generation source. The nearshore error metrics calculated for *Dp* are much larger than *Hs* and *TM01*, particularly the NRMSE and RMSE, and this likely due to a combination of factors such as the calculation of *Dp*, especially when multiple wave events are in the water at that time, or the error metric calculation when normalizing by the mean. The *Dp* scatter plot also reveals biases in the model's ability to predict the *Dp* at the beginning of the forecast period, and this changes depending on the direction of the incoming waves. The majority of the data for E waves lies above the perfect line and confirmed by the plotted percentiles, meaning that the observed directions were more to the east than predicted by the model. The opposite occurs for waves generated in the North Pacific where the majority of the data lies slightly below the perfect line and confirmed by the percentiles, meaning that the observed directions were more to the north than predicted by the model. Despite the model's biases and errors for *Dp*, they are likely not as important as the biases for *TM01* because the wave period ultimately determines how and where waves refract into the coastline.

The binned data density scatter plots in **Fig. 8.** provide additional insight into the wave height errors in terms of their storm generations source which was not shown in **Fig. 7**. Similar to the nearshore scatter plots in **Fig. 7.**, both figures reveal that most of the data falls close to the perfect line at the beginning of the forecast, indicating good model performance. The binned data density scatter plots results are consistent with the results from **Fig. 5.**, reiterating that NW waves tend to get overestimated and N-NE, SW, and S-SE waves all get underestimated. Using a data density scatter plot also provides context into how outliers perform, and how these outliers, particularly larger wave events, may not be representative of the prevailing wave conditions.

Fig. 9. provides spatial insight of the model's performance at the beginning of the forecast and calculating the error metrics for all wave heights provides a baseline that can be compared to past and future studies. Here, we are able to compare our spatial forecast-altimeter results to Li et al., (2016) Hawai‘i hindcast who spatially compared the Hawai‘i WW3 hindcast to GlobWAVE altimetry measurements from 1991 to 2011. Our RMSE and ME/BIAS plots follow similar trends, which include the highest RMSE values offshore and lowest RMSE nearshore and to the east-southeast of the Hawai‘i island. The higher RMSE values offshore are likely because the largest wave heights exist offshore (**Fig. 4.**) which emphasizes the use of NRMSE to analyze relative errors. Both our results and Li et al., (2016) find that the highest ME/BIAS values occur offshore to the north and south of the islands, and the lowest ME/BIAS values occur nearshore, especially in the wake of the Hawai‘i island.

Splitting the analysis based on *Tp* spatially reveals additional errors and biases that would likely be masked if only conducting the analysis for the overall data or based on directions. But the underestimation of waves with *Tp* < 10s reveals inconsistencies with results from **Fig. 5b.** because typically short period waves are associated with smaller wave heights, especially east wind waves that are locally generated, and we find that smaller wave heights, and smaller east wind waves, are generally overestimated (**Fig. 5a,b.**). The downside to this analysis is that we do not know what directions these waves are originating from, and our results may be more representative of a particular storm generation source than another.

In **Fig. 11.** we are able to get insight into the directional analysis spatially. Using the modeled *Dp* as reference is a powerful way to bin the direction since altimeters do not measure *Dp*. Even if the modeled *Dp* is off by a few degrees, like we show in **Fig. 7c.**, it still approximates the general direction and therefore provides a reasonable estimate of the storm generation source. Our results in **Fig. 11.** reveal many consistencies with results from **Fig. 5.** and **Fig. 8**. The spatial analysis is a good way for us to confirm that NW waves are overestimated and N-NE, SW, and S-SE waves are underestimated, and additionally provides where these biases are occurring. Underestimation of wave heights in sheltered areas is a notable finding, and is most notable for NW waves. These sheltered areas are not consistently exposed to NW waves, and in some cases, it takes a long-period wave event for NW waves to reach these areas. In order for NW waves to reach these areas, they undergo refraction and bend around the islands. The sharp difference between the overestimation offshore and the underestimation immediately nearshore may indicate that the wave forecast may have difficulty predicting waves that refract behind the island chain. The WW3 Hawai‘i regional model's resolution of 0.05 degrees (~5 km) may have too coarse of resolution which does not capture the full effects of refraction and shoaling. This could be a reason why the large NW waves are overestimated for the forecast-altimeter comparisons, while large NW waves were underestimated in forecast-buoy comparisons (**Fig. 5a**). However, confidence with large NW wave events remains low because of the small sampling size for both altimeters and buoys (**Fig. 5c.**). Analysis over a longer data period could be useful, and additional research will need to be conducted to answer this question, possibly by including the utilization of swell partitions or evaluating errors at the storm source. If errors are occurring at the storm source and then occurring near Hawai‘i once the wave event reaches the vicinity of the island chain, this would likely indicate an issue with the WW3 global wave model or GFS global atmospheric model rather than a local wave model error. Gibbs et al., (2018) discuss this same issue with the global WW3 model and how it may be providing boundary condition biases which affect the modeled outputs of the regional Hawai‘i wave models. Gibbs et al., (2018) propose that data assimilation techniques with wave buoys and altimeter *Hs* observations could be a way to improve model performance.

Our results indicate that the wave forecast performs best at the start of the forecast, as expected (Janssen et al., 1997), and our results suggest there is room for model improvement later in the forecast period. RMSE and NRMSE values begin to increase quicker at days 3 and 4 of the forecast (**Figs. 12 & 13.**) which displays more of an exponential-like curve than the BIAS. The reason for this difference is unclear, but may be related to how the calculations are performed in the RMSE vs. BIAS equations. When calculating the NRMSE as a function of lead time for the directional bins, our results do not reveal any systematic differences in the forecast error in terms of storm generation source. The results from **Fig. 13.** are still consistent with **Fig. 12.** because the NRMSE values increase as a function of forecast lead time. The spatial forecast error analysis in **Fig. 14.** also confirms this, but does not reveal any spatial patterns.

4.1 OUTLOOK

The methods used here are bulk statistics that do not indicate the timing of the wave event. For example, large wave events may only be overestimated during the peak of the event, followed by a more accurate prediction later in the event. The peak of a wave event typically occurs for a few hours before slowly declining, so our results do not reflect the performance of peak wave events. We note that it could be important to isolate the peak of the wave event, because oftentimes the peak of the event is most consequential which ultimately leads to the strongest coastal impacts. Another area for further research would be investigating the arrival time of wave events. For example, are wave events arriving later or earlier than forecasted and what does this mean in terms of how their significant wave heights are predicted? Predicting the arrival time of a wave event can be just as important as predicting the height, direction, or period, because a wave event coming in earlier or later than forecasted can lead to society unprepared. It is also likely that the forecast's arrival time is affecting our analysis. Gibbs et al., (2018) showed that the model predicting wave events too early is a systematic pattern when analyzing storm tracks for NW wave events in the winter of 2016-17. When a wave event comes in later than forecasted, the forecast would be overestimating the waves because the waves are not there yet. If NW waves in our analysis consistency arrived later than the forecast, it is likely that this influenced our results and may be why we found that NW waves are overestimated by the forecast, when in reality they may be underestimated because they're peaking sometimes 6-18 hours later than forecasted (Gibbs et al., 2018).

The analysis conducted for this research only uses standard wave parameters and does not use swell partitions. The results give us a general insight into how different wave events perform, but conducting research with only the standard wave parameters is likely misrepresenting what is actually occurring during a 'swell' event. For example, in the bottom panels of **Fig. 11.** during south waves, the underestimation values in blue near the immediate coastal areas, especially near the Hawaiʻi Island, are likely more representative of waves from the south compared to the values offshore. This is because farther offshore, consistent east wind waves are exposed there, and mixing in with the south waves. Using swell partitions isolates only on swell and eliminates the effect of other waves that may be in the water at the same time. We conclude that it would be beneficial to confirm our results with swell partitions, such as with Synthetic Aperture Radar or partitions from buoy spectra. Using these datasets would likely give us a more accurate representation of how the different wave events are performing.

5.0 CONCLUSION

The PacIOOS wave forecast system provides important spatial-temporal information to the nearshore and offshore waters of Hawai‘i which global spectral wave models like WW3 cannot do. For this reason, it is crucial that the PacIOOS wave forecast system provides consistent and reliable wave forecasts in Hawai‘i. To get an insight into potential systematic model errors, we evaluated the WW3 Hawai‘i regional model and 4 nearshore SWAN models from 2011 to 2021 relative to measurements from 5 offshore wave buoys, 8 nearshore wave buoys, and a decade of space-borne *Hs* observations from 7 different satellite missions. We found that PacIOOS wave forecast system performs well relative to *Hs* record from moored wave buoys and space-borne *Hs* observations, and the error metrics follow similar trends to previous Hawai‘i hindcast studies. We also found that the wave forecast contains various errors at the start of the forecast and as a function of forecast lead time. Based on the results, we established a baseline error, and they can provide a baseline for future model improvement and insight for maritime users, forecasters, and other model developers.

At the beginning of the forecast, forecast-buoy comparisons revealed that the highest normalized errors exist nearshore, particularly at buoys 51212 and 51203 which are exposed to southerly waves and sheltered from NW and E waves. On average *Hs* comparisons contained higher normalized errors than *TM01*. At the start of the forecast, the wave forecast system performed best for ~2 meter wave heights, but began to underestimate the less frequent larger wave events. The start of the forecast analysis also revealed that the wave forecast performs differently as a function of storm generation source, with waves propagating from the NW (281° - 348°) containing the largest errors compared to other directional bins. The NW waves were also overestimated by the wave forecast the most at 6.0% compared to the other directional bins, which all underestimate the wave forecast on average. Waves generated in the South Pacific were generally underestimated, sometimes by as much as 10-20% near the wake of the Hawaiʻi island and during wave events larger than 2.5 meters. Spatial analysis revealed that these errors do not occur uniformly, and were sometimes altered around the island chain in nearshore areas. The NW wave events experienced underestimation in immediate sheltered regions while overestimation dominated the north-facing shores. The overall spatial analysis revealed that the highest RMSE and BIAS values exist offshore and follows similar results to a Hawaiʻi hindcast study by Li et al., (2016). Our analysis revealed that the wave forecast's performance degrades as a function of forecast lead time, and both RMSE and NRMSE begin to increase at faster rates by days 3 and 4 in the forecast. Later in the forecast, the wave forecast typically overestimated the wave heights, with agreement from both the forecast-altimeter comparisons and forecast-buoy comparisons. Overall, the results presented here establish a baseline for future model improvements.

LITERATURE CITED

Alday, M., Ardhuin, F., Dodet, G., and Accensi, M.: Accuracy of numerical wave model results: application to the Atlantic coasts of Europe, Ocean Sci., 18, 1665–1689, https://doi.org/10.5194/os-18-1665-2022, 2022.

Alex Gibbs, Andre van der Westhuysen, Stylianos Flampouris, Patrick Caldwell, Pablo Santos, Roberto Padilla-Hernandez. (2018). Evaluation and challenges using the Nearshore Wave Prediction System on Oahu, Hawaii through the 2016-2017 winter season. American Meteorological Society, 98th Annual Meeting.

Booij, N., Holthuijsen, L.H. and R.C. Ris, 1996, The SWAN wave model for shallow water, Proc. 25th Int. Conf. Coastal Engng., Orlando, USA, Vol. 1, pp. 668-676.

Cheung, K.F. (2011). *WaveWatch III (WW3) Hawaii Regional Wave Model.* [2011 - 2021] Distributed by the Pacific Islands Ocean Observing System (PacIOOS). Accessed [February 7, 2024].

Cheung, K.F. (2010). *Simulating WAves Nearshore (SWAN) Regional Wave Model: Oahu.* [2011 - 2021] Distributed by the Pacific Islands Ocean Observing System (PacIOOS). Accessed [February 7, 2024].

Coastal Data Information Program (CDIP), M.A. McManus, M.A. Merrifield, and Pacific Islands Ocean Observing System (PacIOOS). (2001) *PacIOOS Wave Buoy 106: Waimea Bay, Oahu, Hawaii.* Distributed by the Coastal Data Information Program

Janssen, P. A. E. M., B. Hansen, and J. Bidlot. (1997) Verification of the ECMWF Wave Forecasting System against Buoy and Altimeter Data. *Wea. Forecasting*. **12**, 763–784.

Ning Li, Kwok Fai Cheung, Justin E. Stopa, Feng Hsiao, Yi-Leng Chen, Luis Vega, Patrick Cross. (2016) Thirty-four years of Hawaii wave hindcast from downscaling of climate forecast system reanalysis. Ocean Modeling. **100**,78-95

Patrick L. Barnard, Andrew D. Short, Mitchell D. Harley, Kristen D. Splinter, Sean Vitousek, Ian L. Turner, Jonathan Allan, Masayuki Banno, Karin R. Bryan, André Doria, Jeff E. Hansen, Shigeru Kato, Yoshiaki Kuriyama, Evan Randall-Goodwin, Peter Ruggiero, Ian J. Walker & Derek K. Heathfield. (2015) Coastal Vulnerability across the Pacific Dominated by El Niño/Southern Oscillation. *Nature Geosci*. **8**, 801–807.

Piollé, J.-F.; Dodet, G.; Quilfen, Y.; Schwatke, C.; Passaro, M.; Quartly, G.; Thibaut, P. (2022) ESA Sea State Climate Change Initiative (Sea_State_cci): Global remote sensing multi-mission along-track significant wave height from altimetry, L2P product, version 3. NERC EDS Centre for Environmental Data Analysis.

The WAVEWATCH IIIR Development Group (WW3DG), (2019) User manual and system documentation of WAVEWATCH IIIR version 6.07. Tech. Note 333, NOAA/NWS/NCEP/MMAB, College Park, MD, USA, 465 pp. + Appendices.