Quantifying Tropical Cyclone-Generated Waves in Extreme-Value-Derived Design for Offshore Wind

Sarah McElman[1,2], Amrit Shankar Verma[1], and Andrew Goupee[1]

[1]University of Maine, Orono, ME, USA
[2]Metocean Expert Americas, Concord, NH, USA

Correspondence: Sarah McElman (sarah.mcelman@maine.edu)

Abstract. Wave extreme values, such as significant wave height, peak period, and crest height, are central to design and operation practices for offshore wind structures. However, the most suitable methods for deriving these extremes, both statistically and from numerical models, is not straightforward. This is especially acute in mixed-type climates, as in the Atlantic coast of the US, where tropical cyclones (hurricanes) and extra-tropical cyclones (winter storms) occur at the same locations with varying frequency and intensity. Limited guidance is provided in major offshore wind energy standards for the minimum requirements of these ocean models and methods used for determining accurate design and operational metocean conditions for regions with tropical cyclones and mixed-type environments. This study investigates the representation of extreme significant wave heights on the US Atlantic coast generated by mixed storm types, as represented in numerical simulations and univariate extreme value analysis. Notable differences between N-year design values are found, as projected by the two different modeled conditions with both block maxima and peaks-over-threshold methods. Attributing factors include hindcast duration, proximity of design location to historical track storm centers, and single analysis of mixed-type distributions. This paper is the first of its kind to propose a methodology for defining extreme significant wave heights due to tropical cyclones for offshore wind design and operation in Mid- and North-Atlantic waters. Recommendations for achieving accurate and representative extreme values for offshore design on the US Atlantic coast are provided.

1 Introduction

As offshore wind development expands to locations with mixed storm types, such as a combination of tropical and extra-tropical cyclones, new meteorological and oceanographic features influence infrastructure planning, design, and operation. The successful design and operation of offshore projects require long-term metocean data, traditionally in the form of a locally-validated, high-fidelity multi-decade hindcast of coupled winds, waves, currents, and water levels. These models are typically forced with or derive boundary conditions from global or downscaled reanalysis data sets, such as Climate Forecast System Reanalysis (CFSR) or European Climate Mid-Range Weather Forecast Reanalysis v5 (~~ERA-5~~ERA5) (e.g., as in Groll and Weisse (2017)). Data from these models form the basis for Extreme Value Analysis (EVA), the statistical determination of an N-year parameter. The results of these methods can be sensitive to method choice, parameterization, or data fit (for example, Haselsteiner and Thoben (2020)).

25 The representation of coastal extreme events by numerical models are also sensitive to parameterization choices and model design decisions. Investigations in the literature have quantified tropical cyclone features and their influence on offshore wind design on turbine-scale dynamics. Kim and Manuel (2019) simulated the local wind, waves, and hydrodynamic features of Hurricane Sandy at a number of offshore wind development areas on the Atlantic coast with the Miami Coupled Model. They recommend coupled wind-ocean modeling for the best representation of the features and evolution of tropical cyclones as a

30 prerequisite to intensity assessment and wind-wave probability distribution. Gomez et al. (2023) quantified the convective momentum transfer of Category 1, 2, and 3 tropical cyclones, developed from a catalog of synthetic events on the US East Coast, and found that the resulting turbulence and gust characteristics of these scenarios at times exceeded current IEC standards.

Fewer studies have focused on tropical cyclone-generated waves and wave growth and how they are represented in offshore wind engineering decisions. Additionally, limited guidance is provided in major offshore standards (API RP 2MET, 2019; DNV,

35 2018; IEC–614000–3, 2019) for the minimum requirements of ocean models and methods for capturing tropical cyclone-generated N-year wave heights and periods. As a result, many different approaches have been taken to model and quantify these ocean design values. Few studies have investigated how current methods treat mixed climates with annual tropical and extra-tropical cyclone events. In the absence of abundant buoy observations in the path of a tropical cyclone at the points of interest, the regionally-validated GROW-Fine East Coast model is referenced as the ground truth to assess the results of current

40 methods–modeling and statistical analysis–in capturing ocean design parameters.

In this study, the selection of proper design values depends on accuracy in three tiers: statistical methodology, model capacity, and model design. Statistical characterization of the meteorological and oceanographic extremes on the US Atlantic Coast with General Extreme Value and Generalized Pareto methods are frequently selected in the literature for similar applications, for example in Jonathan and Ewans (2013), Haselsteiner and Thoben (2020), Northrop et al. (2017), Barthelmie et al. (2021),

45 and Bhaskaran et al. (2023). In contrast, **?** proposes the STM-E (space-time maximum and exposure) method for determining wave extremes from synthetic tropical cyclone data. In a broader study, Kresning et al. (2024) investigated uncertainty in loads assessment for offshore wind on the US Atlantic coast resulting from a variety of univariate and bivariate extreme value methods, finding differences of up to 6 % between calculated return values. Neary et al. (2020) found that discrepancies between univariate and bivariate methods based on the same data were random and indicated "reasonable agreement on average".

50 The capacity of spectral wave models to resolve tropical storm features has been investigated by MacAfee and Wong (2007), O'Grady et al. (2022), and Padilla-Hernández et al. (2007), and have shown that Simulating Waves Nearshore~~(SWAN~~, SWAN (**?**) and the Wave Modeling Project~~(WAM~~, WAM (**?**), the basis for wave models in the GROW-Fine East Coast hindcast, are capable of capturing peak wave values and trapped-fetch swell generated by tropical cyclones. ~~Additionally~~Recently, a comparison of the performance of SWAN and ~~MIKE-21~~ MIKE21 (**?**) in coastal Portugal by Fonseca et al. (2016) has shown

55 similar overall behavior of wave growth and propagation.

In addition to tool capacity, modeling choices such as boundary condition quality must be investigated for sufficient ability to resolve tropical cyclone features. As in Campos et al. (2022) and Gandoin and Garza (2024), metocean models frequently employ the ~~ERA-5~~ ERA5 reanalysis dataset, for direct analysis, or as a boundary condition to high-resolution modeling. Campos et al. (2022) and Gandoin and Garza (2024) found that ~~ERA-5~~ ERA5 under-captured peak winds during storm events,

60 while Caires and Sterl (2005) and Stephens and Gorman (2006) found that ERA-based wave models under-captured significant wave heights (Hs). Similarly, Neary et al. (2020) investigated 50-year significant wave height return values on the Atlantic coast and found a "systematic underbias for extreme significant wave height derived from model hindcasts as compared to those derived from buoy measurements." Each author suggests additional calibration for improved performance.

This paper investigates the performance of two ~~high-resolution~~ metocean models forced by global ocean reanalysis datasets

65 ~~to~~ (hereafter referred to as "high-resolution" for their high-resolution wave component) to (1) represent significant wave height due to tropical and extra-tropical events, and (2) quantify differences in return period values between these high-resolution models and datasets generated from direct modeling of tropical and extra-tropical storm events. Two locations on the US east coast are considered, which experience tropical and extra-tropical events with varying frequencies and magnitudes: in the North Atlantic ("NA", New England Wind) and in the Mid-Atlantic Bight ("MAB", Kitty Hawk Wind), to assess sensitivity

70 of the analysis to different mixed-type climates (see map, Figure 1). A schematic of the study method is provided in Figure 2. Wind and ocean parameterization of all three models are documented in Table 1 and Table 2, respectively. To ensure that the statistical findings from this study are not significantly sensitive to the selected extreme value method employed, a subset of model ~~results~~ data are additionally analyzed with an alternative extreme value method~~and evaluated~~.

2 Methods

75 Two ~~validated and calibrated~~ calibrated and validated high-resolution models of wind, waves, and hydrodynamics are used to simulate a multi-decade hindcast of hourly conditions around the two wind project areas~~(hereafter referred to as the North Atlantic and Mid-Atlantic "high-resolution" models)~~. Post-calibrated model results are compared to buoy observations and the reconstructed tropical and extra-tropical storm models (collectively referred to as the "GROW-Fine East Coast" model). Results are presented as a collection of absolute and normalized values to protect intellectual property of the GROW-Fine data, where

80 necessary.

Return period ~~results~~ estimates from all four models (high-resolution models in the North Atlantic and in the Mid-Atlantic Bight, the GROW-Fine tropical-cyclone-only model, and the GROW-Fine extra-tropical-cyclone-only model) are calculated by the Block Maxima (BM) method with a Gumbel distribution~~fit~~.

To ensure that the results were not significantly influenced by the chosen statistical method, a sensitivity analysis of a

85 subset of model results was carried out with Peaks-Over-Threshold (POT)~~, using both 2-parameter Weibull and Gumbel distribution fits~~. A sample of results from this analysis is available in Appendix A. ~~For both the North Atlantic (Figure A2a) and Mid-Atlantic Bight (Figure A2b) cases, fits by Weibull or Gumbel resulted in 4 % or less variation between return values in all model subsets, except for the tropical 10,000-year MAB significant wave height (+8 %, Weibull). Larger variation was observed between return values by POT and BM analyses; however, the trends between models and storm types were~~

90 ~~consistent. Following Barthelmie et al. (2021) and Bhaskaran et al. (2023), the block maxima method was considered suitable for this study.~~ To enable long hindcast periods for the Tropical and Extra-Tropical GROW-Fine models, only storm events are represented in these model datasets~~. As the dataset does not include normal sea states, it is not possible to determine extreme~~

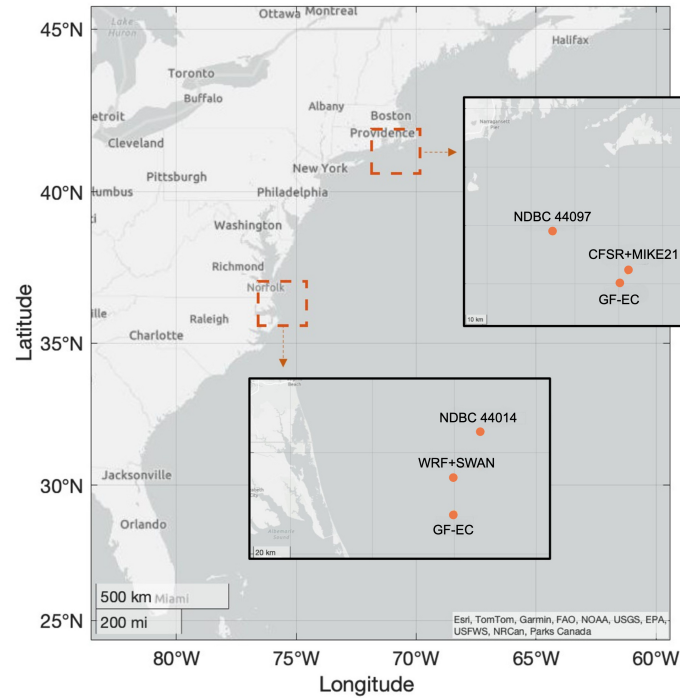



Figure 1. Location of study data: buoy observations and associated turbine analysis locations in the North Atlantic (top insert) and in the Mid-Atlantic Bight (bottom insert).

~~values for this model~~, with modeled periods of up to 1 week surrounding a peak event. Such a limited dataset poses challenges in establishing the correct threshold and clustering time for analysis by the Peaks-Over-Threshold method. Following results

95   of a Kolmogorov-Smirnov one-sample test (see section 2.2.1), the Annual Maxima method adequately characterized all sites to proceed.

2.1 Model descriptions

Different spectral wave and hydrodynamics coupled models were chosen for the two investigated sites (Figure 1) in this study. These combinations, described in Tables 1 and 2, were chosen as representative modeling choices for quantifying the waves and

100   hydrodynamics of the coastal United States. ~~Both North Atlantic and Mid-Atlantic coupled models are locally validated~~The domains and extents are shown in Figure 3.

A list of significant storms during the hindcast period, available observations, and storm events used for model calibration is provided in Appendix ~~B~~C, beginning with the first available buoy-based observations.

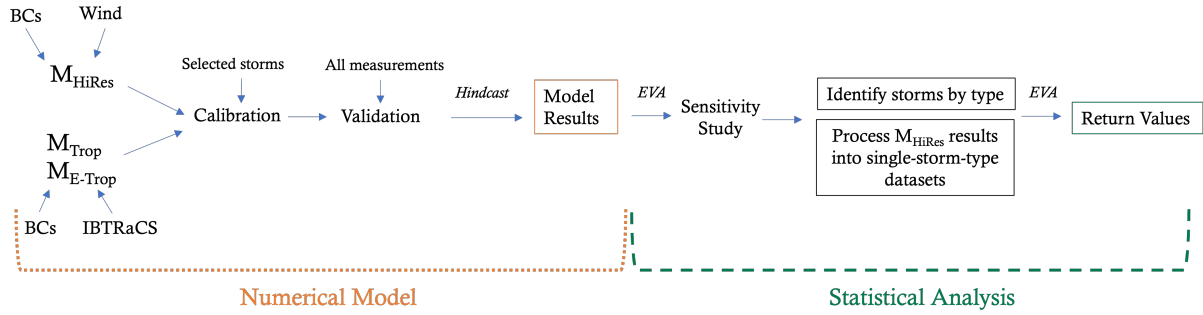



Figure 2. Study overview: Model ~~calibration is applied linearly~~ parameters are calibrated based on ~~a~~ subset of measurements ~~,~~ prior to ~~statistical analysis by~~ model validation and extreme value ~~methods~~analysis. Additional data sets are post-processed from the high-resolution models into time series of normal and single-type events. This process is applied in the North Atlantic and the Mid-Atlantic Bight analyses locations. Boundary conditions (BCs) and model parameterization can be found in Tables 1 and 2.

Model	Source	Resolution	Boundary Conditions	Duration
NA HiRes	CFSR	22km spatial (2D) 2-hour	CFSR	1979 - 2021
MAB HiRes	WRF	9km spatial (3D) 1-hour	CFSR	1989 - 2019
GROW-Fine East Coast	Tropical Boundary Layer Model	5.5km spatial (3D) 15-minute	*Wind*: Satellite reconstruction *Pressure*: Far-field tropical reconstruction	*Trop*: 1924 - 2021 *Extra-Trop*: 1954 - 2021

Table 1. Wind parameterization for the three investigated models in the North Atlantic, Mid-Atlantic Bight, and along the US Atlantic coast (GROW-Fine East Coast).

2.1.1 North Atlantic high-resolution model

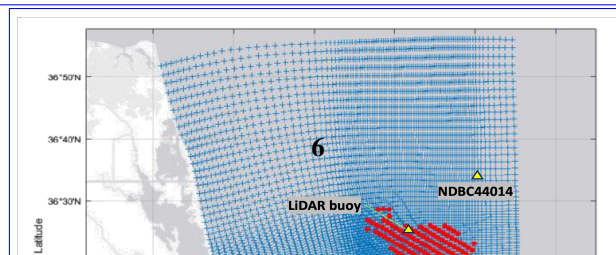
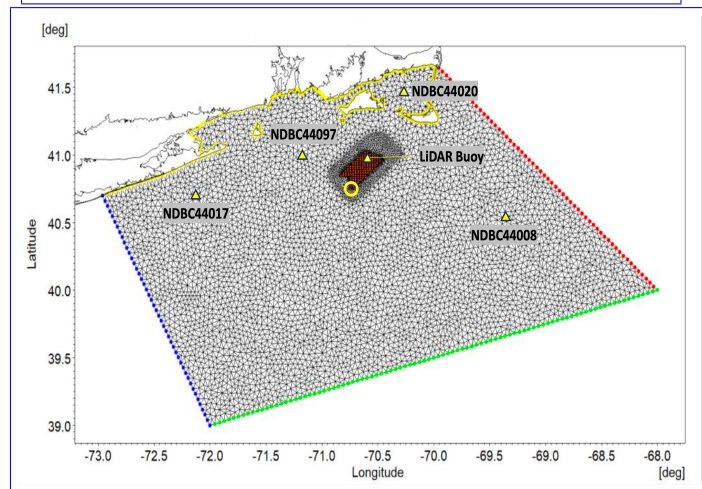
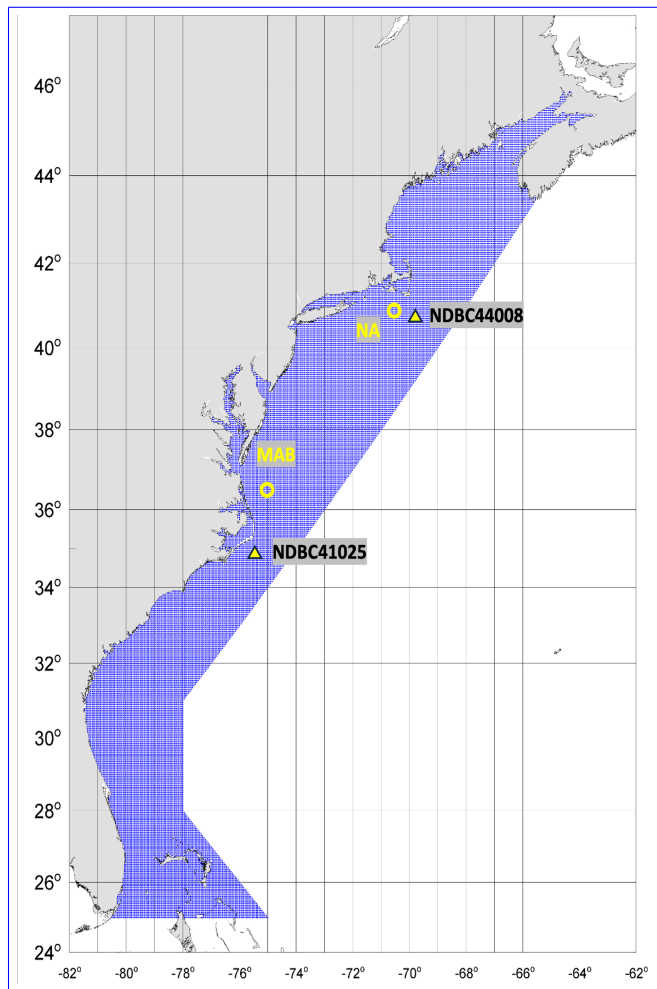
The North Atlantic model comprises 42 years of CFSR winds forcing a MIKE21 spectral wave and MIKE21 hydrodynamics hindcast, with one-way coupling from hydrodynamics to waves. The CFSR dataset is reanalyzed from the National Centers for Environmental Prediction (NCEP) Climate Forecast System model. In this study, 2-hour averaged winds are interpolated to the hydrodynamic and wave domain resolutions, with a wind-wave coupling time step of 1 hour. The MIKE21 wave model is based on the wave action conservation formulation and is run in fully-spectral mode for this study. Depth-induced wave breaking (breaking parameter $\gamma = 0.8$ and time step factor $\alpha = 1$), spatially-varying whitecapping (~~$C_{ds} = 2.1$ to 2.6~~coefficient of surface drag $C_{ds} = 0.021$ to 0.026 from the Charnock formulation), and nonlinear growth (coefficient $= 1.35$) are modeled. Wave calibration is based on a mixed-type set of 16 storms at four buoys, and is applied ~~spatially~~ uniformly throughout the domain. ~~A list of tropical storms during the hindcast period and nearby observations used for validation are provided in Appendix C.~~

Model	Tool	Resolution	Boundary Conditions	Coupling	Spec Parameti
NA HiRes	MIKE21	600m wave (2D) 600m hydro (2D) 1-hour	DHI ~~Global~~ East Coast Waves (waves) DHI East Coast (hydro)	1-way, ~~1-way, hydro to waves~~hydro to waves	36 dire 32 free 0.033 Hz 0.667 F
MAB HiRes	SWAN+DELFT3D	400m wave (2D) 400m hydro (3D) 1-hour	ERA5 (waves) HYCOM (hydro)	2-way, ~~2-way, waves and hydro~~waves and hydro	36 dire 24 free ~~0.005~~0.05 F 0.448 F
GROW-Fine East Coast	OWI3G+ADCIRC	5.5km wave (2D) 5.5km hydro (2D) 15-minute	GROW2012 (waves) Prevost '08 (hydro)	No dynamic coupling. Reanalysis of each modeled storm.	48 dire 26 free ~~0.0029~~0.029 0.322 F

Table 2. Wave and hydrodynamic parameterization for the three investigated models in the North Atlantic, Mid-Atlantic Bight, and along the US Atlantic coast (GROW-Fine East Coast). All model bathymetries are derived from GEBCO. The NA and MAB model bathymetries are supplemented with 1m geophysical survey measurements within the project area.

Tides are modeled by the DTU10 Global Tide Model to capture any tidally-induced hydrodynamic or wave effects. ~~More details on validation and the project background can be found in the Commonwealth Wind metocean report, ?.~~

This study investigates extremes at NA model location 40.75°N, 70.74°W~~, which~~. This locations has a depth of ~~62 m~~62m and is 45 km from buoy 44097 (measured depth: 49.4m), the closest observation to the ~~most "at-risk" turbine location~~ turbine location of interest. GROW-Fine East Coast ("GF-EC") model results are presented for 40.8°N, 70.7°W, which has a depth of 59m.

2.1.2 Mid-Atlantic model

The Mid-Atlantic model comprises 30 years of Weather Research and Forecasting (WRF) winds forcing a SWAN spectral wave and Delft3D hydrodynamics hindcast (?), with two-way coupling between the hydrodynamic and wave models. The WRF model is a mesoscale atmospheric model developed by the National Center for Atmospheric Research (NCAR) and partners. In this study, ~~vertically-nested~~ horizontally-nested domains with real lateral boundary conditions ~~are applied and model the boundary layer~~ (CFSR) are applied. The planetary boundary layer is modeled with the YSU ~~scheme. SWAN is a phase-averaged spectral wave model.~~ ("Yonsei University") scheme. Depth-induced wave breaking, whitecapping (Westhuysen formulation, ?), and nonlinear growth are similarly modeled as for the North Atlantic case.

Tides are modeled by the Oregon State University TPXO dataset to capture any tidally-induced hydrodynamic or wave effects. Wave calibration is based on observations of the 2020 "Hurricane Isaias" at 36.41°N, 75.23°W, and is applied uniformly throughout the domain. ~~More details on validation and the project background can be found in the Kitty Hawk Wind metocean report, ?.~~

This study investigates extremes at MAB model location 36.38°N, 75.00°W~~, which~~. This locations has a depth of ~~40~~ 38 m

and is 53 km from buoy 44014 (measured depth: 49.1m), the closest observation to the ~~most "at-risk" turbine~~ turbine location

of interest. GROW-Fine EC model results are presented for 36.2°N, 75.0°W, which has a depth of 40m.

2.1.3 GROW-Fine East Coast model

The GROW-Fine East Coast Tropical and Extra-Tropical models reconstruct storm winds and pressure fields based on multiple types of historical observations, including from satellites, aircraft flights, and SF microwave radiometetry. Waves are modeled by OWI3G, ~~a~~ the Oceanweather 3rd generation wave model based on WAM~~that~~. This model does not require a prescribed wave spectrum for initialization, and are modified by assimilation of ocean observations during storm events. The GF-EC domain spans from southern Florida through the Bay of Fundy.

To ~~feasibly~~ reconstruct 100 years of tropical storms and 75 years of extra-tropical storms, normal sea states are omitted from the dataset, and only major storm events are represented. The duration (associated data of storm development and decay) of individual events therefore varies for each storm throughout the dataset. The trajectory of tropical storms in the GF-EC Tropical model is based on International Best Track Archive for Climate Stewardship (IBTRaCS) records. While validation of individual events is difficult prior to large-scale observing systems like the National Data Buoy Center (NOAA), select events with limited observations are recreated from the verified model, balancing model fidelity with overall dataset duration.

To assess the influence of time (number of storms) on the statistical assessment of extremes, Extreme Value Analysis (EVA) is conducted for both the full duration of the GF-EC records and for the shorter duration of the high-resolution models: 1979–2020 (42 years), for the North Atlantic, and 1989–2019 (30 years) for the Mid-Atlantic Bight.

Model validation with ~~NDBC~~ NOAA buoy observations spanning the Atlantic coast from 1979, the beginning of data availability, has been carried out by Oceanweather. For more details on the storm wind reconstruction method, ~~boundary conditions, and overall model validation, refer to the~~ see **?**.

2.1.4 Model skill

The GROW-Fine ~~East Coast project description (Oceanweather (2022)).~~ East Coast winds and waves are validated against 20 buoy measurements, spanning from NDBC 41009, in coastal Florida, to NDBC 44027, in the northern Gulf of Maine, in addition to satellite altimetry. A selection of validation statistics are provided in Table 3.

2.1.5 ~~Model skill~~

~~The performance of the~~ The North Atlantic model ~~is assessed at 40.86N, 70.808W, which is 29 km from the nearest observation, NDBC 44097 from 2009 - 2020 (refer to Appendix C for Figure **??**, significant wave height, and Figure **??**, peak period). Nine significant tropical and twenty significant extra-tropical events occurred during this period. There is good agreement in low-to-mid energy waves, however, modelrepresentation of the largest waves is higher than those recorded at NDBC 44097, resulting in an over-representation of significant wave height in this region, post-calibration, compared to buoy observations~~

Location and Duration	Hs [m]			T02 [s]			Tp [s]		
	Correlation Coefficient [%]	Scatter Index	Bias [m]	Correlation Coefficient [%]	SI	Bias [m]	Correlation Coefficient [%]	SI	Bias [m]
NDBC44097 40.97 °N, 71.12 °W 2009–2020	91	0.25	0.03	82	0.13	-0.52	73	0.23	-0.01
NDBC44008 40.50 °N, 69.25 °W 1982–2020	93	0.24	-0.03	82	0.12	-0.39	75	0.17	0.06
NDBC44017 40.69 °N, 72.05 °W 2002–2020	92	0.25	0.10	83	0.11	-0.38	77	0.20	0.13
NDBC44020 41.50 °N, 70.28 °W 2009–2020	81	0.35	0.10	61	0.07	-0.65	44	0.11	-0.15
LiDAR Buoy 41.07 °N, 70.48 °W 2018–2020	91	0.26	0.08	84	0.11	-0.21	71	0.22	0.09

Table 3. Validation statistics throughout the North Atlantic model domain against continuous measurements, 2010 - 2020.

Location	Hs [m]			Tp [s]		
	Correlation Coefficient [%]	Scatter Index	Bias [m]	Correlation Coefficient [%]	SI	Bias [m]
Combined 20 Buoys	0.93	0.24	0.13	0.59	0.28	-0.07
NDBC44008 40.50 °N, 69.25 °W	0.90	0.21	0.04	0.47	0.32	-0.95
NDBC41025 35.01 °N, 75.45 °W	0.90	0.23	0.20	0.59	0.24	-0.27

Table 4. Selected validation statistics for the GROW-Fine East Coast model, overall and in the North and Mid-Atlantic regions.

during the same time, suggesting that modeled sea states over the hindcast period were larger than observed for the strongest

165 events . However, when model and observation quantiles are compared for tropical-cyclone-only events , calibration was determined from a range of cap to friction velocity values and nonlinear growth coefficients for overall performance during 1) a mixed set of storms, and 2) during the entire year of 2012. After this calibration, validation was conducted over a 10-year period at five observation locations throughout the model domain. Validation statistics are provided in Table 4.

Very few measurement data is available within the project area: in the case of the North Atlantic model, no tropical cyclones

170 were captured during the two-year deployment of the LiDAR buoy, and in the case of the model Mid-Atlantic model, there is only a 3-month overlap between observations and the hindcast period. The performance of the calibrated model during multiple tropical cyclone events is therefore assessed at a farther location: in the North Atlantic, the closest observation to the project

area is 29km away, at NDBC buoy 44097. Nine significant tropical and twenty significant extra-tropical events occurred during this ten-year period. For all waves during this period, the high-resolution root mean square error (RMSE) ~~increases from~~ is
0.338m. When only considering tropical cyclone-generated waves, the RMSE at this location increases to 0.474m, suggesting poorer performance during these events.

~~Model skill in~~ For the Mid-Atlantic model~~is assessed at 36.38N, 75.00W, which is 29 km from~~, the closest observation ~~with the longest overlapping record, NDBC 44014 from 1990–2019 (refer to Appendix C for Figure ??, significant wave height, and Figure ??, peak period).~~ to the project area is again 29km away, at NDBC buoy 44014. Twenty-seven significant tropical and thirty-five significant extra-tropical events occurred during this ~~period. The model shows good agreement in low-to-mid energy waves, however, under-representation is observed for large-amplitude and large-peak-period swell, suggesting that the model may not suitably capture sea states during~~ overlapping thirty-year period. Similar to the North Atlantic high-resolution model, when only tropical cyclone events ~~. When model and observation quantiles are compared for tropical-cyclone-only events, the RMSE is slightly raised from 0.455 to 0.468, suggesting weaker model performance during these events, as well.~~ are considered, the significant wave height RMSE remains elevated, at 0.468m.

2.2 Univariate extreme value methods

Return values are determined according to the general linear model,

$$Y = -\beta' \underline{\mathbf{x}} x + \epsilon \tag{1}$$

which associates covariates x to the return values Y in terms of regression coefficient, β' and error, ϵ. The return values are calculated based on the log-likelihood of the density function, described here as either block maxima (~~the Gumbel form of the Generalised Extreme Value distribution, GEV~~ fit with a Gumbel distribution) or as probability-based peaks-over-threshold (~~Generalized Pareto, GP), when assessing~~ fit with a Generalized Pareto distribution). Two methods were chosen to assess the sensitivity of results to ~~extreme value~~ method choice.

2.2.1 Block Maxima

The Generalized Extreme Value distribution describes a set of data in terms of ξ, shape, μ, location, and σ, scale. When $\xi = 0$, this distribution is equivalent to the 2-parameter Gumbel cumulative distribution function used in this study:

$$\underline{G} F(x) = \underline{exp-exp} \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right) \right), \quad for \quad -\infty < x < \infty \tag{2}$$

from Coles (2001). The location and scale variables were determined by ~~empirical estimation , and~~ maximum likelihood estimation (parameter values for each distribution are provided in Appendix ~~C~~ B). The associated return value, x_p, for ~~probability period~~ annual exceedence probability p are calculated as:

$$x_p = \mu + \sigma \left(\underline{-ln-ln} -\ln(-\ln(1-p)) \right), \quad for \quad -\infty < x_p < \infty \tag{3}$$

The statistical basis for the distribution fit is composed of the annual largest ~~values~~ significant wave heights of the model ~~data.~~ dataset. A one-sample K-S test was conducted to ensure that the data follows a Gumbel cumulative distribution function: the null hypothesis, H_0: the data follows a Gumbel distribution, was not rejected, indicating that at a 95% confidence level, the Gumbel distribution adequately fits the data. Figures of the fitted and empirical Gumbel CDF of the data is provided in Appendix A.

Based on an assessment of the Atlantic hurricane season between 1991 and 2020, an average of 14 named storms (depressions, storms, and hurricanes) and 7 hurricanes occur each year (see "Tropical Cyclone Climatology", **?**, for an overview of this assessment). The GF-EC MAB site data includes 81 tropical cyclone events. The GF-EC NA site data includes 80 tropical cyclone events. In the event that a tropical cyclone did not pass nearby the analysis location during the season, no maxima are recorded for that year in the GF-EC dataset. In the "tropical cyclone only, high-resolution" NA and MAB datasets, a smaller, non-extra-tropical cyclone is picked as a maxima for a year without nearby tropical cyclones activity. Annual maxima analysis for the "high-resolution" model datasets account for 42 storms (tropical or extra-tropical cyclones) at the NA location and 30 storms (tropical or extra-tropical cyclones) at the MAB location. The "high-resolution" hindcast duration, and therefore storm sample size, is set by the available duration of reanalysis data applied to the model boundaries.

2.2.2 Separating storm types

The high-resolution hindcasts comprise all normal and mixed-type storm periods in one dataset. In order to calculate extreme return values according to storm type, as established by Gomes and Vickery (1978), the high-resolution model results were post-processed for their entire hindcast period into two datasets: one, with tropical storm events removed, and another with extra-tropical storm events removed, aligned with storms represented in the overlapping period with the GROW-Fine datasets. ~~A list of tropical cyclones included in these datasets is available in Appendix C.~~

Extreme Value Theory assumes that ~~extremes~~ extreme values are independent variables (Mackay and Johanning (2018)). However, multiple peaks may be attributed to the same event during storm growth and dissipation. To preserve the independence criterion in this study, only the peak significant wave height is retained in a period of 98 hours ~~(storm length also assumed by Kresning et al. (2024), Oceanweather (2022), and others)~~ during an identified storm. For simplicity, the same storm duration is assumed for both tropical and extra-tropical events, however in a number of instances, the storm duration is shorter than this period. In general, tropical cyclone forward speed is highly dependent on local climatic conditions, which influences storm duration in a given location. For storm removal from a dataset, values 49 hours prior to and 49 hours after the peak significant wave height are removed.

3 Results and Discussion

As return periods extend to 50 years and beyond, tropical cyclones are observed to be the dominating storm type for extremes in both the North Atlantic (Figure 5a), and for all return periods in the Mid-Atlantic (Figure 5b). While extra-tropical cyclones are a more frequent occurrence at the North Atlantic site (there are 48 significant extra-tropical cyclone events recorded during

(a) NA high-resolution model   (b) NA GF-EC Tropical model

(c) MAB high-resolution model   (d) MAB GF-EC Tropical model

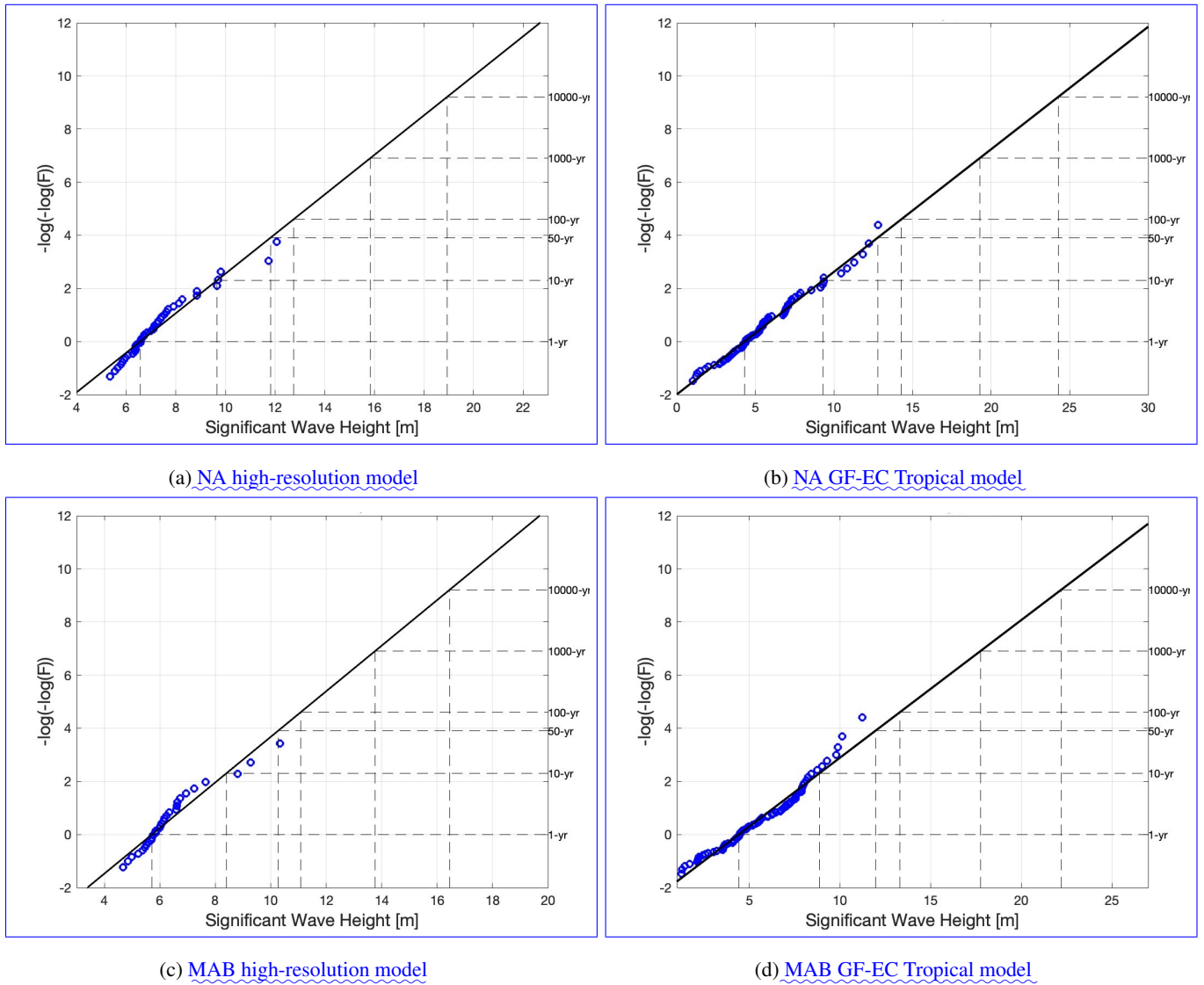
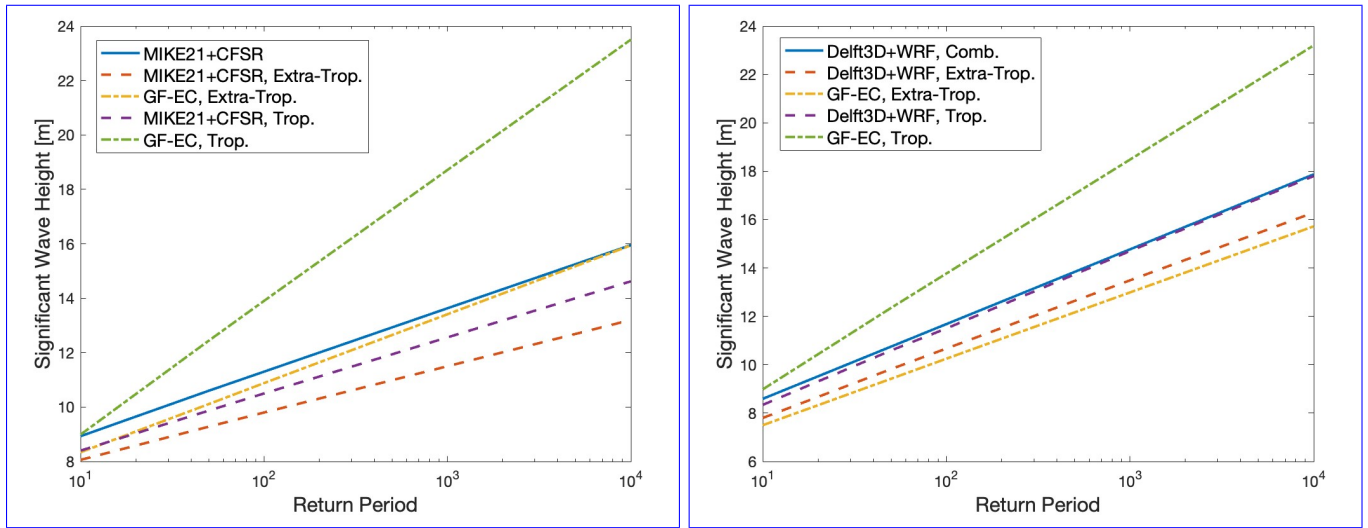



Figure 4. Model data with annual maxima fit to Gumbel distribution

the 42-year hindcast period, whereas there are 15 significant tropical cyclone events), tropical cyclones were able to reach
235   greater intensities in the studied record.

The return values calculated from the GF-EC Tropical model show diverging values, in trend and magnitude, from the
high-resolution hindcast for both the North Atlantic (Figure 5a, dash-dotted green line) and for the Mid-Atlantic (Figure 5b,
dash-dotted green line) scenarios, despite different calibrations, boundary conditions, and latitudes. The influence of statistical
and modeling choices on these diverging trends for tropical cyclone events are investigated further in the following sections.

(a) Significant wave height return values for the North Atlantic site at 40.8N, 70.7W.

(b) Significant wave height return values in the Mid-Atlantic Bight site at 36.2N, 75.0W.

Figure 5. Study return values for all models by annual maxima with Gumbel distribution fit.

In contrast, there is good agreement between the GF-EC Extra-Tropical model and the high-resolution model data, in both its original and post-processed (normal + extra-tropical events) data set forms (for the North Atlantic, see Figure 5a, dash-dotted orange line, and for the ~~the~~ Mid-Atlantic Bight, see Figure 5b, dash-dotted orange line). This suggests that both the storm ~~physics and storm~~ type and sample set of extra-tropical events at these latitudes are sufficiently represented by the high-resolution models in the 30–42 year periods investigated.

~~Significant wave height return values for the North Atlantic site at 40.8N, 70.7W. Significant wave height return values in the Mid-Atlantic Bight site at 36.2N, 75.0W. Study return values for all model results by annual maxima with Gumbel distribution.~~

3.1 Influence of Model Characteristics

3.1.1 Overall model performance

Representative model performance is presented here for similar-intensity storms: the January 2015 extra-tropical storm (Class 3 on the Dolan-Davis scale) and "Hurricane Dorian" (Category 2, Saffir-Simpson scale). For more information on these scaling systems, see Dolan and Davis (1992) and Simpson (1974). The time and duration of the extra-tropical wave height successive peaks (Figure 6a) agree between models and measurements, although the GF-EC Extra-Tropical model ~~over-represents the largest peak as compared to measurements . Evolution~~ represents a larger peak than measurements (45 km distance). The evolution of total wave period was captured very well for the extra-tropical event as exhibited by overlapping ~~timeseries evolution of~~ significant wave height and peak period data from both model results and buoy observations (Figure 6b). While

there is overall agreement between models and measurements on peak timing and storm duration during "Hurricane Dorian", the ~~the~~ Mid-Atlantic Bight high-resolution model ~~shows under-representation~~ returns a lower magnitude of the storm peak significant wave height (Figure 6c) by 10 % as compared to observations, and by 15 % as compared to the GF-EC tropical model. While the observed storm peak period (Figure 6d, green crosses) spans from 4 to 13 seconds, neither model captures the higher-frequency, wind-driven waves below 7 seconds in this event. The MAB high-resolution model follows the largest periods, while the ~~the~~ GF-EC Tropical model captures frequencies in the middle of the range.

Throughout the 30-year period of the MAB high-resolution hindcast, all identified tropical and extra-tropical events are represented. There is a notable trend of under-representation of the largest storm peaks, suggesting that calibration based on a single tropical cyclone observation was insufficient in terms of numerical modeling. However, calibration is not the only (or largest) influence on the return values calculated in this study, as discussed in Section 3.2.

Similarly, the North Atlantic model represents all significant tropical and extra-tropical events on record during the 42-year period, in a number of cases with higher significant wave heights than in the GF-EC Tropical dataset. Minor differences between the models are investigated in the following sections. ~~There~~ However, there is one case of substantial under-representation of significant wave height ~~, however:~~ during "Hurricane Bob", when the analysis point was within the storm fetch. Compared to the GF-EC Tropical model for this event, this was a consequential error for a dataset containing 15 tropical cyclone events. The maximum significant wave height in the GF-EC Tropical dataset during "Hurricane Bob" is on the order of the two largest wave heights in the high-resolution data set: "Hurricane Gloria" of 1985, and the "Storm of the Century" Blizzard of 1993. Given the rarity of this wave magnitude for this location, this numerical modeling error is expected to influence the overall statistical assessment of extreme values for the site.

3.1.2 Tropical Cyclone Sea State Representation

Model performance is first assessed for the representation of wind-sea- and swell-waves generated during storm evolution. For brevity, only the North Atlantic case is presented here. The two-dimensional wave spectrum is partitioned ~~(separated into wind-sea and swell wave systems)~~ by the watershed algorithm. ~~(See Portilla-Yandun et al. (2009) for more information on spectral partitioning schemes.)~~ For both the high-resolution and GF-EC data sets, the wave components are similarly captured during "Hurricane Bob" and during "Hurricane Dorian". Both models capture the physical progression of a tropical cyclone as signaled by the evolution of wave peak period: elevated swell frequencies (approximately 10 seconds, Bob, GF-EC Tropical in Figure 7a and NA high-resolution in Figure 7b; up to 16 seconds, Dorian, GF-EC Tropical in Figure 7c and NA high-resolution in Figure 7d) precede maximum winds. In the case of "Hurricane Bob", where the analysis point is within the storm fetch, peak significant wave height (solid vertical line) corresponds to the wind-sea and overall storm maximum peak periods (approximately 15 seconds). Both swell and wind-sea frequencies are characterized similarly by both the high-resolution and GF-EC Tropical models.

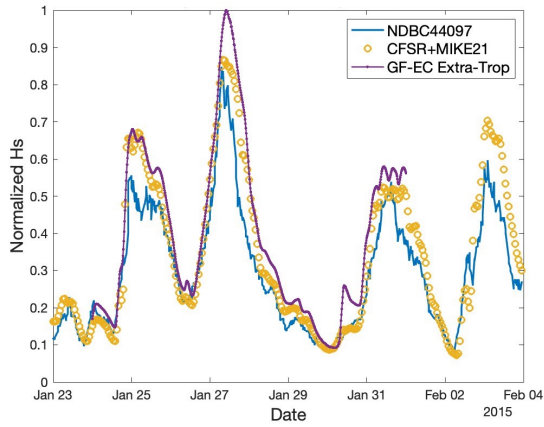
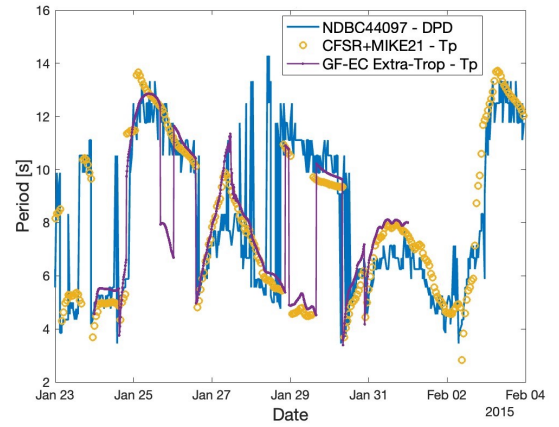
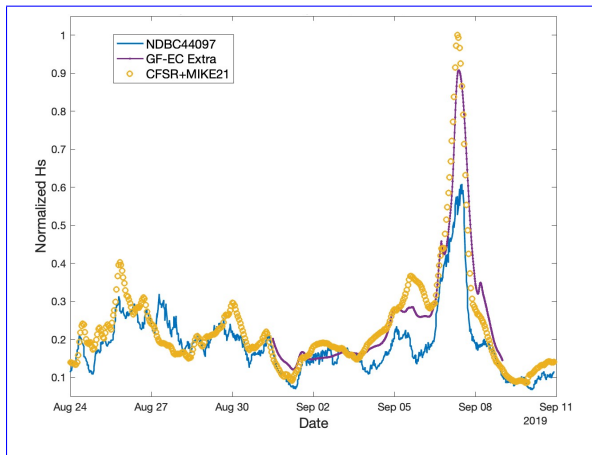
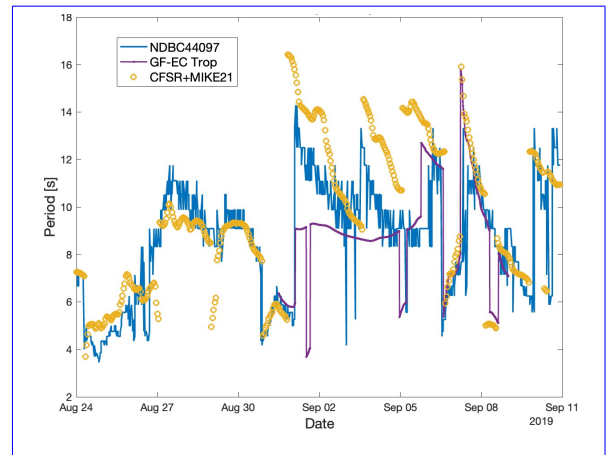
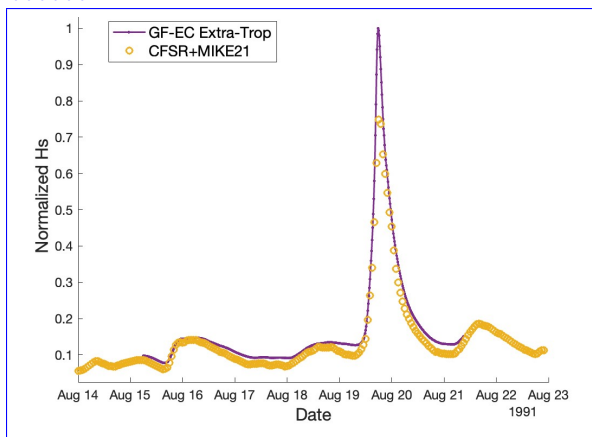
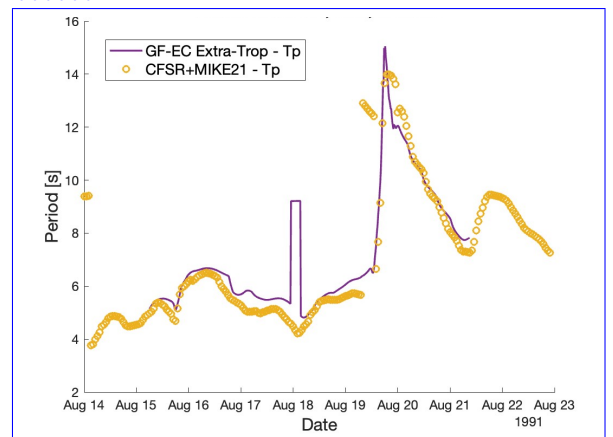
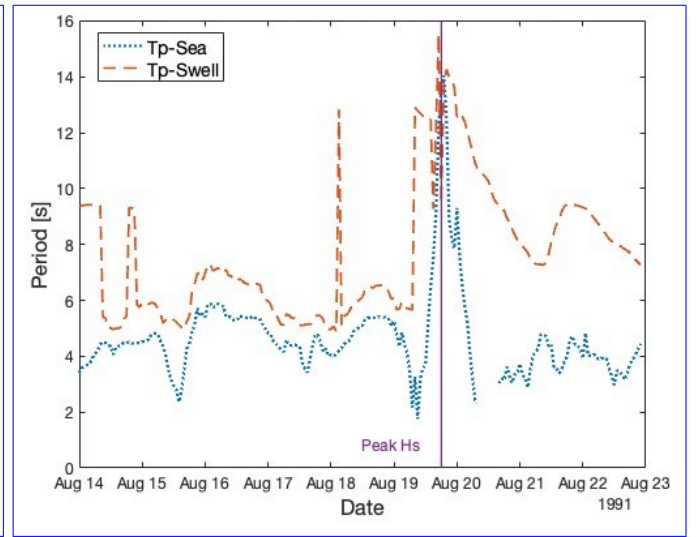
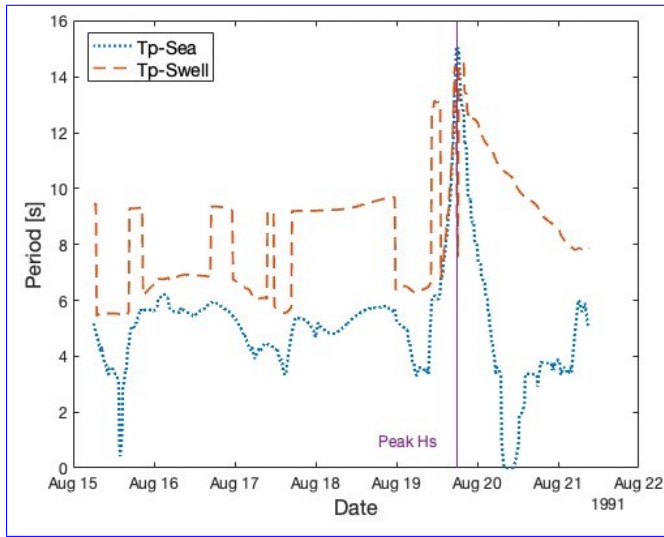
In the case of "Hurricane Dorian", where the analysis point is outside of the storm fetch, the swell maximum peak period (16 seconds) and wind-sea maximum peak period in both models arrive prior to the time of peak significant wave height (vertical solid line). During and after this time, both models represent "Hurricane Dorian" with similar magnitudes of wind sea and

(a) Hs during the 2015 Blizzard, North Atlantic

(b) Tp during the 2015 Blizzard, North Atlantic

(c) Hs during "Hurricane Dorian", ~~Mid-Atlantic Bight~~North Atlantic

(d) Tp during "Hurricane Dorian", ~~Mid-Atlantic Bight~~North Atlantic

(e) Hs during "Hurricane Bob", North Atlantic

(f) Tp during "Hurricane Bob", North Atlantic

Figure 6. Representative performance of the high-resolution and GF-EC model ocean wave results during (a–b) extra-tropical and (c–d) tropical cyclone storm events, with buoy measurements. To protect intellectual property and to preserve scale, significant wave height values are normalized to the single largest peak value.

(a) Elevated sea state during "Hurricane Bob", GF-EC Tropical model

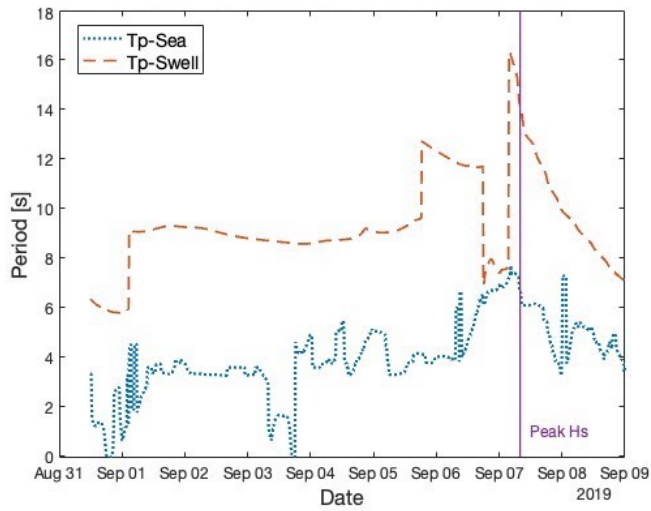
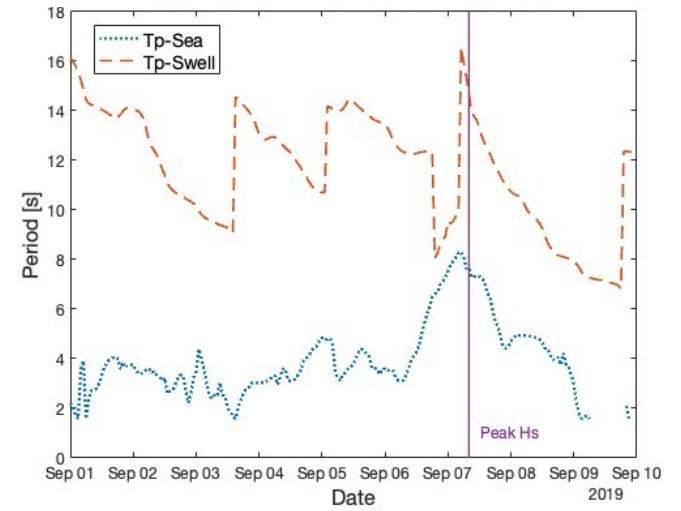
(b) Elevated sea state during "Hurricane Bob", NA high-resolution model

(c) Elevated sea state during "Hurricane Dorian", GF-EC Tropical model

(d) Elevated sea state during "Hurricane Dorian", NA high-resolution model

Figure 7. Partitioned wave systems during the arrival and passage of two tropical cyclone events by the high-resolution and GF-EC Tropical models for the North Atlantic site.

290 swell, however, there is a notable difference in swell frequencies prior to the storm peak between the two models. Section 3.1.4 investigates this discrepancy further.

(a) GF-EC Tropical model   (b) NDBC buoy 44097 observations   (c) NA high-resolution model

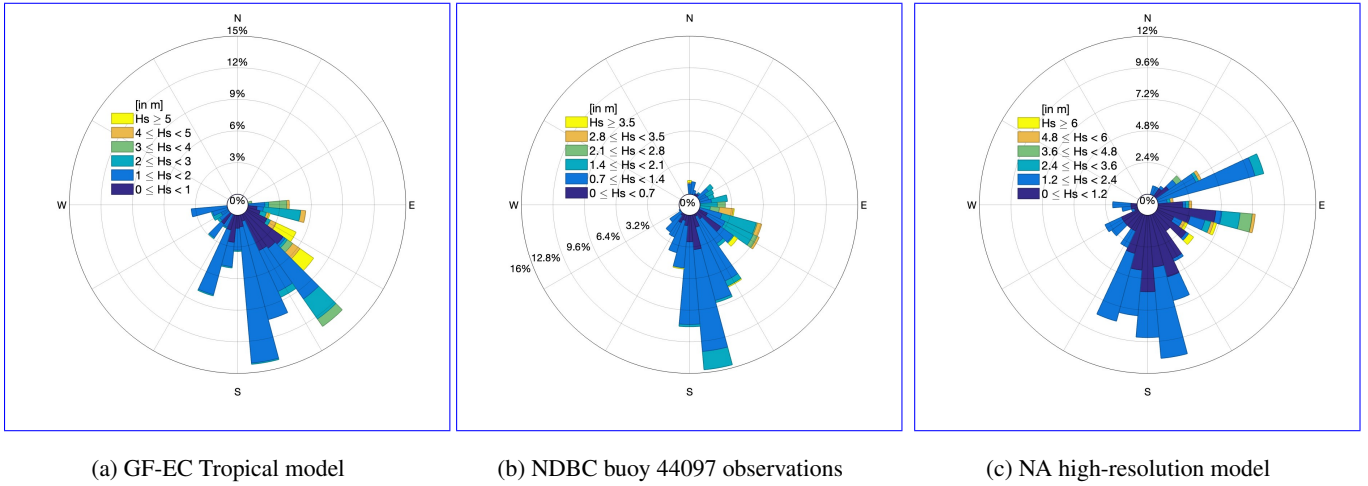



Figure 8. Mean wave direction during passage of "Hurricane Dorian" at the North Atlantic site.

(a) GF-EC Tropical model   (b) NA high-resolution model

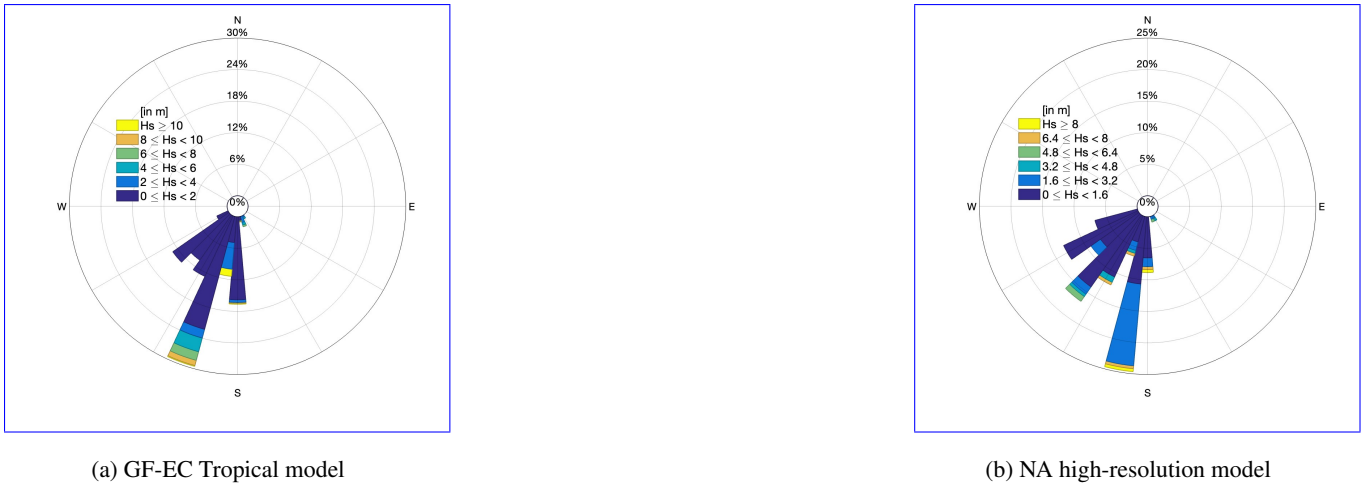



Figure 9. Mean wave direction during passage of "Hurricane Bob" at the North Atlantic site.

3.1.3 Wave directional spreading

In the case of "Hurricane Dorian", the high-resolution and GF-EC models show global similarities in mean wave direction, however, the NA high-resolution model (which represents 36 directional sectors and 32 frequency bins), Figure 8c, presents a narrower range of the largest waves from the North-North West as compared to GF-EC Tropical (which simulates 48 directional sectors and 26 frequency bins), Figure 8a, and buoy observations, Figure 10b.

In the case of "Hurricane Bob" in the North Atlantic, where analysis occurs inside the storm fetch, both models represent a concentrated range of waves similarly for the average and largest waves (GF-EC, Figure 9a, and high-resolution, Figure 9b), while the magnitude of the wave height differs. This narrower range ~~agrees with Forristall and Ewans (1998), which confirms~~

(a) GF-EC tropical model   (b) NDBC buoy 44097 observations   (c) NA high-resolution model

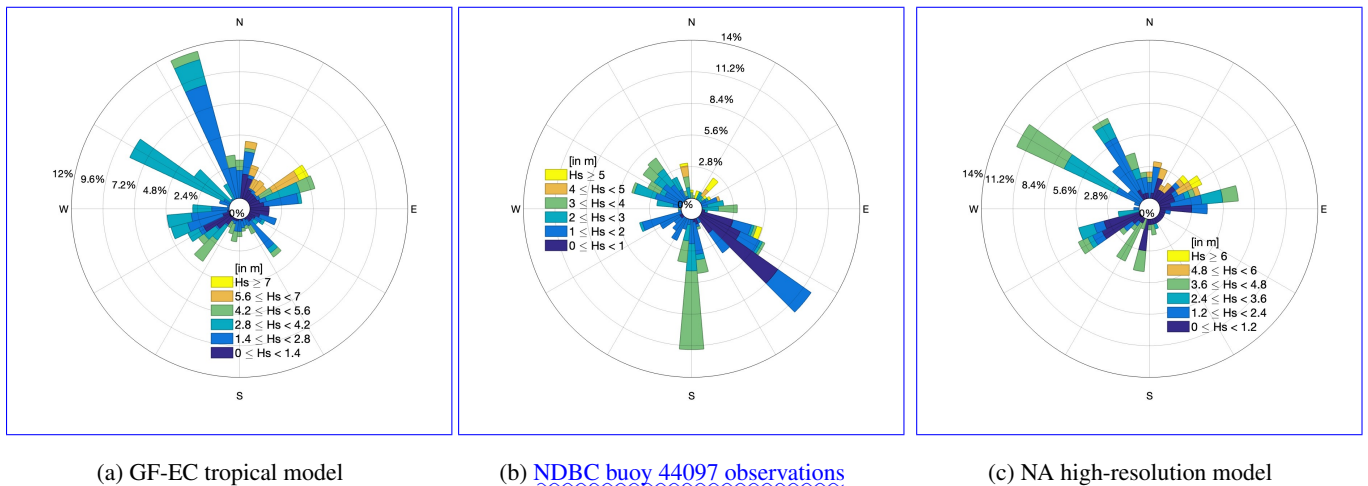



Figure 10. Mean wave direction during passage of the 2015 Blizzard at the North Atlantic site.

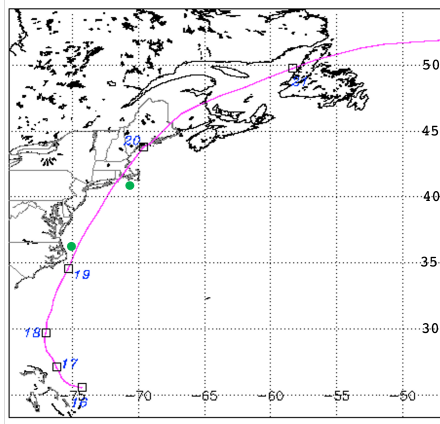
300 ~~that spreading factors are generally reduced the closer they are to the storm center. Storm~~ is an important feature of the storm type (Forristall and Ewans (1998)), and is a key model attribute to investigate within the storm fetch and approaching the storm center. One-minute-sustained storm surface winds (90 knots, or 46 m/s, at a 10m height) were recorded in this location and passage of the storm eye over the area took one hour. Considering wave directions represented, the GF-EC Tropical model captured seas on a 15-minute interval. Wave buoy measurements occurred on a 30-minute cycle, and the NA high-resolution

305 model ~~modeled~~ captured seas on an hourly time step. Given the high-intensity, fast-moving event through the region, an hourly time step likely contributed to the under-representation of sea states in the high-resolution model during "Hurricane Bob".

    In contrast, a large directional spread is observed in both models during the 2015 Blizzard (Figure 10a, GF-EC Extra-Tropical, and Figure 10c, NA high-resolution).~~ This again agrees with the findings of Forristall and Ewans (1998) that, while~~, which Forristall and Ewans (1998) identifies as an attribute of extra-tropical cyclones. While ocean response to tropical cy-

310 clones may result in a variety of wind sea and swell wave system interactions, there is a generally narrower range of directional spreading within the storm fetch for tropical cyclones than for extra-tropical cyclones. Given the larger spread of wave directional propagation expected during extra-tropical storms, combined with excellent alignment in magnitude and trend with buoy observations of significant wave heights, the high-resolution models are considered highly capable of resolving the physical properties of wave development by extra-tropical events on the Atlantic coast.

315 ### 3.1.4 Proximity to Tropical Cyclone Eye

The physical evolution of tropical cyclones varies from to storm to storm. The radius of a tropical cyclones can span from 20 to 250 km, while the radius of an extra-tropical cyclone can span 100 km to 2000 km. The representation of waves and wave growth due to a tropical cyclone in design values may therefore be sensitive to storm track, fetch, and the location of derived site conditions. The two tropical cyclone events investigated in detail here represent a close proximity to the storm path: From

(a) Path of "Hurricane Bob", 1991   (b) Path of "Hurricane Dorian", 2019

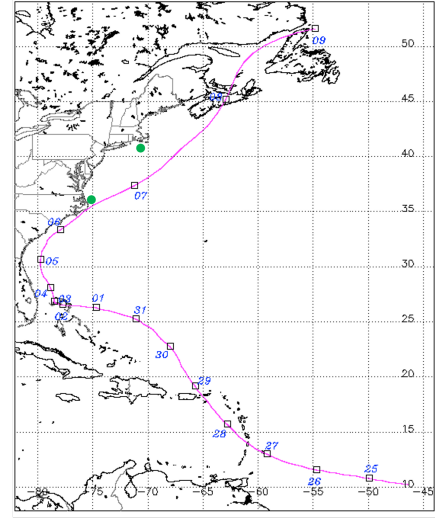



Figure 11. Historical best tracks of the studied tropical cyclones, from www.csics.org/ibtracs

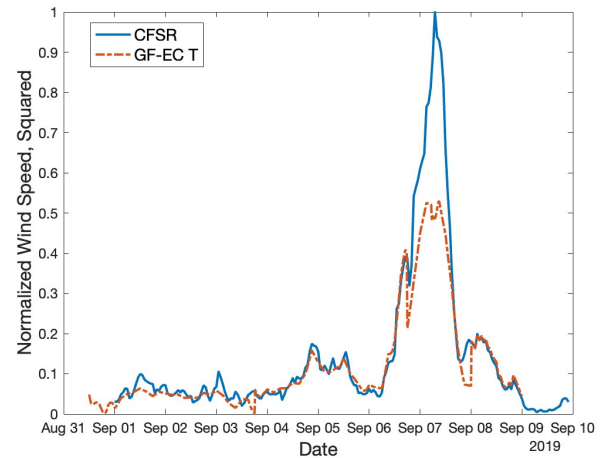
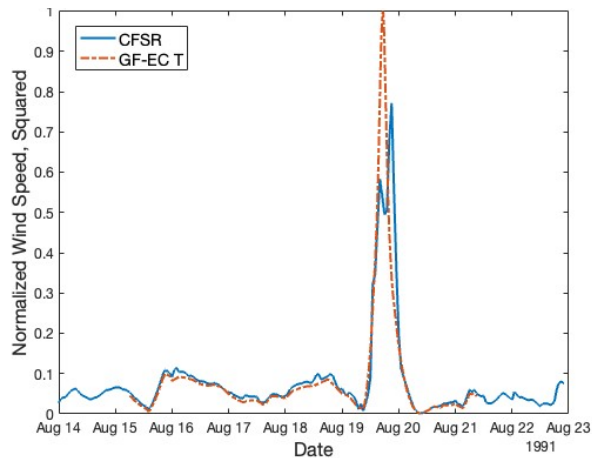
320   IBTRaCS records, the eye of "Hurricane Bob" reached its peak low pressure (950 mb, Category 3) 60km from the Mid-Atlantic analysis point, and came within 100km of the North Atlantic analysis point on August 19, 1991 as a Category 2 storm. The eye of "Hurricane Dorian" also came within 60km of the Mid-Atlantic analysis point as a Category 2 storm, and persisted as Category 2 within 260km from the North Atlantic analysis point.

As Hwang and Walsh (2018) show based on satellite observations of multiple tropical cyclone events, wave growth inside
325   tropical cyclones follows the same formulation as wave growth in other, non-tropical-cyclone conditions. As a result, Hwang and Walsh (2018) describe fetch and duration as linearly relating to storm radius. Differences in wave development in this study are therefore considered primarily to be a function of fetch or duration representation, and not model performance.

To further quantify the influence of storm fetch on peak significant wave heights, the momentum flux between the atmosphere and the ocean surface is investigated in terms of wind stress (from Jones (2011)):

330   $$\tau_i = \rho C_D U_{10m,i}^2 \qquad (4)$$

where ρ is air density and C_D is the coefficient of drag. Setting aside constants, effective stress is a function of the square of the wind speed, normalized here to the largest magnitude recorded from either the GF-EC Tropical and NA high-resolution models during the event. The square of normalized wind speed inside of the storm radius for "Hurricane Bob" is presented in Figure 12a, and outside of the storm radius for "Hurricane Dorian" in Figure 12b. There is a notable difference in these cases
335   in applied wind forcing to the sea surface, despite both storms persisting as Category 2 at the time and location of analysis.

During the modeled events of "Hurricane Bob", the NA high-resolution model applies 25 % less wind stress on the ocean surface than the GF-EC tropical simulation; peak significant wave height from the coupled MIKE21 wave model is 25 % lower than in the assimilated GF-EC Tropical wave dataset. In contrast, during "Hurricane Dorian", the NA high-resolution

(a) GF-EC Tropical and NA high-resolution normalized wind, squared, within the storm radius (fetch) of "Hurricane Bob".

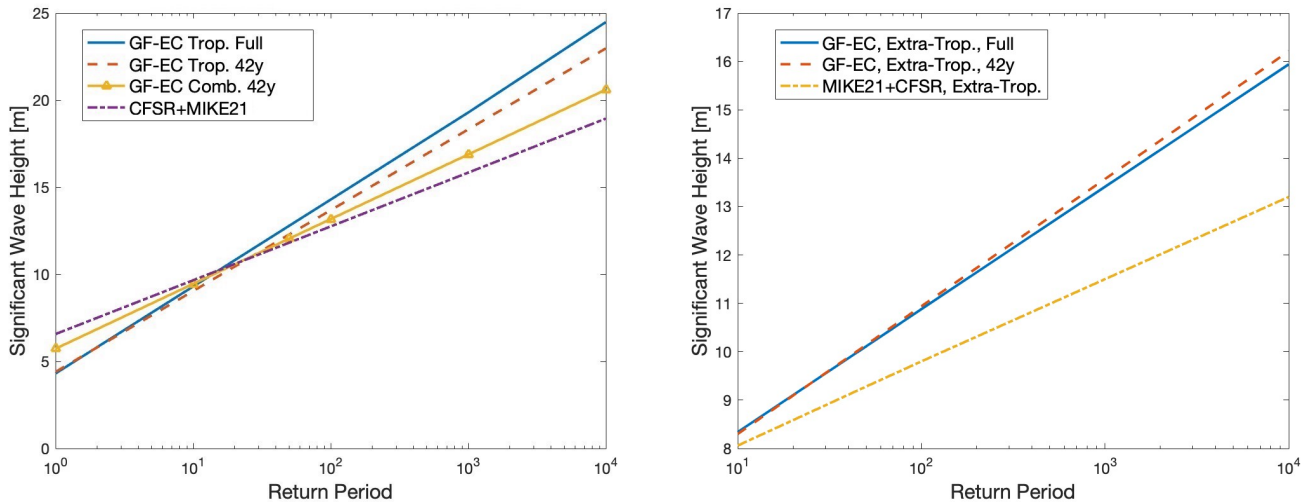
(b) GF-EC Tropical and NA high-resolution normalized wind, squared, outside of the storm radius (fetch) of "Hurricane Dorian"

Figure 12. Effective normalized wind stress on ocean surface during two tropical cyclones at the North Atlantic site.

peak storm winds delivered 48 % larger ocean stress; the assimilated waves in the GF-EC Tropical dataset were 10 % lower than in the high-resolution model. Drawing from the discrepancies between peak winds in the two models, relying on calibration to compensate for under-representation by the model of wave heights may not be sufficient without methodical selection of observations from within and from outside of the tropical cyclone fetch to calibrate against. Additionally, given the differences between wave heights derived from coupled wind forcing and wave heights derived from reanalysis, the treatment of momentum transfer between coupled wind and wave model components during tropical cyclone events over multi-decade simulations should be investigated further.

As stress on the ocean surface is a function of the square of wind velocity (equation 4), the height of non-fully-developed wind waves generated by tropical cyclones are sensitive to the square of the error in peak modeled wind speed. Under-representations of these values, as noted in Campos et al. (2022) and Caires and Sterl (2005), in conjunction with characterization of points solely outside of the storm fetch, can lead to the under-representation of modeled wave growth and significant wave heights.

3.2 Influence of Statistical Choices

In the North Atlantic, where extra-tropical cyclones are an annual event, the 10-year return value for the high-resolution model shows a characteristic of mixed species analysis: the influence of these events lifts the overall return value over the tropical-cyclone-only values (see the intersection of the "CFSR+MIKE21" and "GF-EC Trop. Full" trend lines after the 10-year mark, Figure 13a). However, the opposite is observed at longer return periods: the influence of the less-intense storm type appears to

(a) Tropical-only return values from the GF-EC Tropical, high-resolution, and post-processed GF-EC model data sets.

(b) Extra-Tropical-only return values from the GF-EC Extra-Tropical, high-resolution, and post-processed GF-EC model data sets.

Figure 13. Return values for single-storm conditions at the North Atlantic site from the high-resolution and post-processed GF-EC data sets.

reduce the overall return value, leading to increasing divergence between the "GF-EC Trop. Full" and "CFSR+MIKE21" trend lines.

When assessing return value sensitivity to dataset length, or the number of extreme events represented in the data, the GF-EC Tropical and Extra-Tropical datasets are shortened to 42 years (for the North Atlantic) and 30 years (for the Mid-Atlantic
360 Bight). The 10- to 10,000-year return values reduced slightly in the North Atlantic case (Figure 13a) and increased slightly in the Mid-Atlantic case (Figure 14a). In neither scenario does this shortened dataset explain the trend and magnitude differences between the two sets of results. On the contrary, the increase in return values observed in the Mid-Atlantic case is influenced by higher-than-average tropical cyclone activity in the 1990's and 2000's, and misses the lower-than-average period of the 1970's and 1980's.

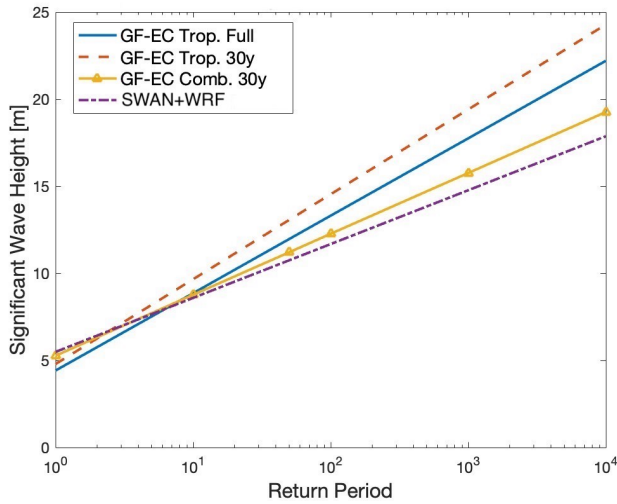
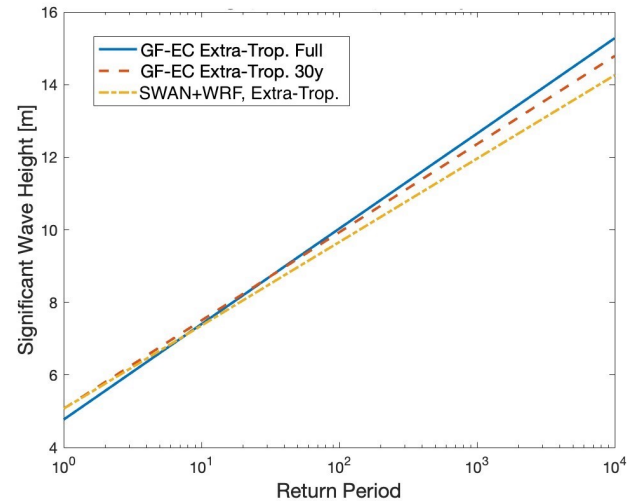
365 Similarly, analysis of the post-processed high-resolution data into a single storm type data did not mitigate differences between the GF-EC Tropical- and high-resolution-derived return values in Figures 13a and 14a; in fact, these values are ~~further reduced~~ observed in Figure 5 to be further reduced (red dashed line) from the original high-resolution data set (blue solid line). In both locations, these lower values suggest that the periods of 1989–2019 (Mid-Atlantic) and 1979–2020 (North Atlantic) do not present a sufficient basis for fully characterizing extremes due to tropical cyclones. As previously mentioned, the under-
370 representation in the high-resolution model of one of the largest peaks in the dataset, "Hurricane Bob", also likely contributes to the lower projected return values in the North Atlantic case. To mitigate this effect due to limited dataset duration of the high-resolution models, grid point "pooling" proposed by Heideman and Mitchell (2009) may be ~~worthwhile~~appropriate.

(a) Tropical-only return values from the GF-EC Tropical, high-resolution, and post-processed GF-EC model data sets.

(b) Extra-Tropical-only return values from the GF-EC Extra-Tropical, high-resolution, and post-processed GF-EC model data sets.

Figure 14. Return values for single-storm conditions at the Mid-Atlantic Bight site from the high-resolution and post-processed GF-EC data sets.

This trend is not observed in either location for extreme values due to extra-tropical cyclones (Figure 13b, North Atlantic, and Figure 14b, Mid-Atlantic). Differences of 1.2 %–1.5 % were observed in the 1,000- and 10,000-year cases (respectively) for the North Atlantic, and of 2.3 %–3.3 % for the Mid-Atlantic Bight. The numerical and statistical representation of extra-tropical extremes by the high-resolution model duration, configuration, and boundary conditions appears sufficient in both the North Atlantic and Mid-Atlantic cases.

When the GF-EC Tropical and GF-EC Extra-Tropical model results are combined into a single dataset and analyzed together in one distribution, return values reduce to values close to those derived from the high-resolution models (see Figure 13a, GF-EC "Combined" 42-year ~~dataset~~line with triangle, and Figure 14a, GF-EC "Combined" 30-year ~~dataset~~line with triangle). The cumulative effect of a reduced storm sample (shortening the GF-EC Tropical data from 100 to 42 and 30 years) and single statistical analysis the mixed storm types explains much of the differences between return values derived from the high-resolution model and from the GROW-Fine tropical models in Figures 5a and 5b: extreme values calculated from this "Combined" dataset are up to 15.9 % lower than the Tropical-only case in the North Atlantic and up to 13.2 % lower than the Tropical-only case in the Mid-Atlantic. These differences occur similarly in both locations, despite differences in tropical storm frequency and peak storm intensities observed in the historical record.

4 Conclusions

In this study, the factors influencing extreme significant wave height estimation due to tropical cyclones were assessed for the relative influence of model and statistical choices at two locations on the US Atlantic coast. The performance of two high-resolution models with differing calibration methods were assessed alongside reanalyzed ocean wave values, forced by reconstructed storm winds, during a period spanning from 30 to 100 years. Choices leading to the statistical distribution of extreme events with the block maxima method were also investigated.

Overall, return values due to extra-tropical events are shown in this study to be well-resolved by established methods of metocean modeling with CFSR- or WRF-generated winds and ~~ERA5 boundary~~ ERA5 wave boundary conditions, as shown by the model performance and directional wave spreading compared to observations and by similar return values generated from all models investigated. However, the differences in extreme values calculated in this study when only considering tropical cyclones suggests that under-representation of peak wave parameters by the high-resolution models can not be mitigated by calibration alone, due in part to the model temporal resolution of wind-generated waves, ~~and in part to the limited number of observations of storms from a range of locations~~ model treatment of the momentum flux transfer between coupled wind and waves, and in the limited treatment by calibration factors of spatially-dependent features within and outside of the ~~storm fetch~~ fetch of a tropical cyclone. While both high-resolution models, with different wind forcing, model design, boundary conditions, and calibration techniques captured tropical cyclone peak significant wave heights within 7 % on average for both the NA and MAB high-resolution time series, ~~there were larger differences~~ larger differences were observed for the largest storm waves.

In addition to numerical representation, statistical choices ~~made during extreme value analysis were a major factor that contributed to the difference between calculated~~ are a major contributing factor to significant wave height return values. ~~The sensitivity of values derived from the GF-EC Tropical datasets when reduced to the high-resolution time periods, in addition to the reduction of values when the high-resolution datasets were post-processed to single-storm time series suggests that further assessment of the number of minimum sufficient number of storms is required for~~ Regardless of the ~~proper characterization of a site, and that this dataset length requirement differs between tropical and extra-tropical assessment, due to observed variations in storm size and intensity. The subsequent calculation of extreme values based on a single distribution of these two storm types over a smaller number of storm events resulted in the under-estimation of 100-, 1,000- and~~ storm type driving the largest 2 - 3 storms in a dataset, analysis of decades of storms as a combined set of types resulted in an overall reduction of derived return values for periods 10 ~~,000-year design values~~ years and greater in both the North Atlantic and Mid-Atlantic locations. Adequate sample size is also necessary (but insufficient alone) to characterize extreme wave heights on the US Atlantic coast due to these storms. The combination of sufficient sample size and analysis by single storm type leads to the notable differences in estimated return values between the datasets presented in this study~~for both~~. Were an offshore substation deck height to be calculated from the North Atlantic ~~(18 %,~~ "high-resolution" data, which is based on the maximum wave crest height, it would be 20% shorter than that from the GF-EC dataset if following API guidance (based on a 1,000-year ~~Hs) and Mid-Atlantic Bight (17 % , 1,~~maximum crest height) and 26% shorter if following standard DNV-ST-0145 (10,000-year ~~Hs)locations between the GROW-Fine East Coast and high-resolution models~~maximum crest height).

For more ~~accurate determination~~ representative estimation of return values for offshore infrastructure design in areas with tropical cyclone activity, the following is recommended:

- Analysis of a point or range of ~~points within the storm wind radius : proximity to the storm eye of 200km or less, depending on storm size~~ocean locations within and outside of characteristic storm radius at the site.

- Storm type has a has a non-trivial effect on estimated significant wave height return values. Extreme value analysis of significant wave height and should be carried out on single-storm-type datasets, consistent with IEC 61400-1 Annex J, which is concerned with separated analysis of synthetic storm wind fields.

- A 30- to 40-year hindcast period is sufficient to characterize extra-tropical extremes in the North Atlantic and Mid-Atlantic locations investigated.

- A 30- to 40-year hindcast period is not sufficient in the investigated areas to characterize tropical cyclone extremes, even with model calibration to a number of tropical cyclone events. If a longer data period is not available, grid point pooling is ~~necessary~~recommended.

Appendix A: Comparing ~~Extreme Value Method~~ Return Values by Method

The sensitivity of ~~extreme~~ return values to the ~~Gumbel distribution and block maxima~~ chosen distribution and extreme value method employed in this study was conducted for a ~~subset of the high-resolution model results with the~~ number of distributions and methods. For block maxima, sensitivity to the Generalized Extreme Value Distribution and Gumbel Distributions was tested; for peaks-over-threshold~~method. The POT analysis sequence is based on a~~, sensitivity to the method using Generalized Pareto distribution ~~cumulative distribution function, where u describes the selected threshold, ξ and σ describe the shape and scale of the distribution, respectively, and and λ is the number of annual exceedences, such that $x > u$:~~was inspected.

While the Generalized Extreme Value Distribution is the most comprehensive distribution for the Block Maxima method, the Gumbel distribution was chosen for its treatment of the tail of the data, the small or neglient values of shape factor parameter over all investigated data sets, and for overall goodness of fit as assessed by the Kolmogorov-Smirnov one-sample test (Figure A1).

$$1 - p = exp\left(-\lambda\left(1 + \xi\frac{x_p - u}{\sigma}\right)^{-1/\xi}\right), \quad for \quad \xi \neq 0$$

~~(from Jonathan and Ewans (2013)). The statistical basis is composed of values that exceed a physically realistic threshold, which can be subjective to user choice~~

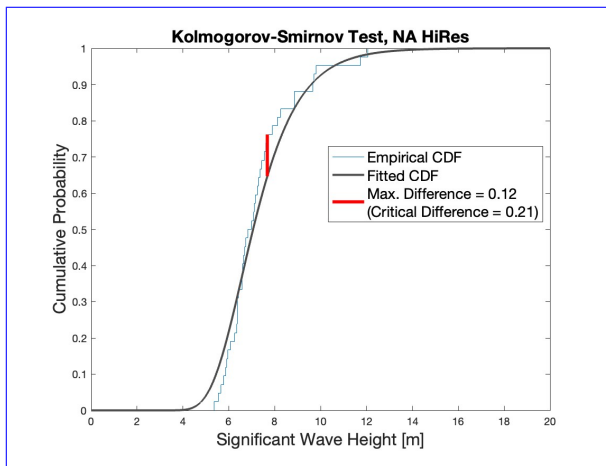
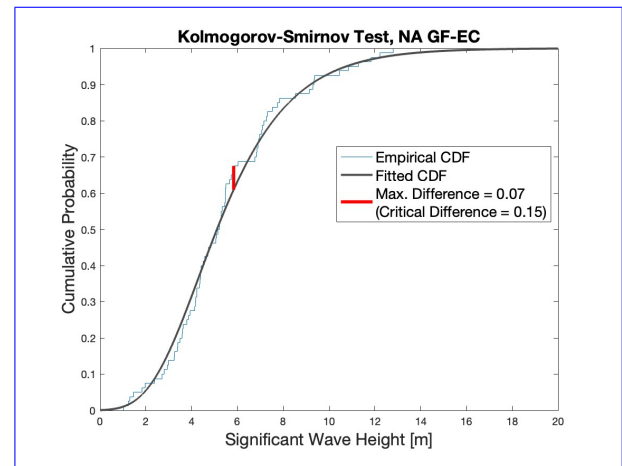
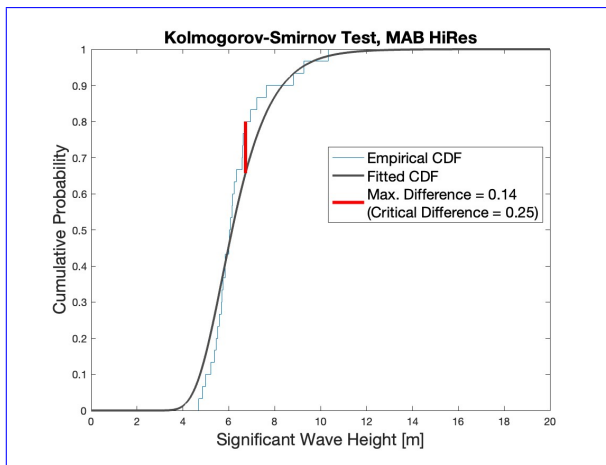
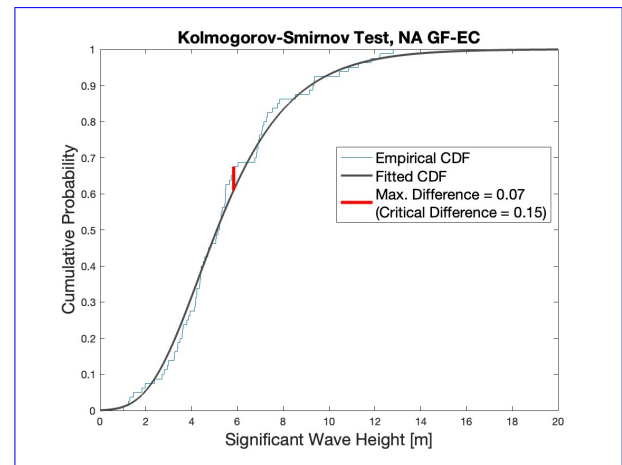
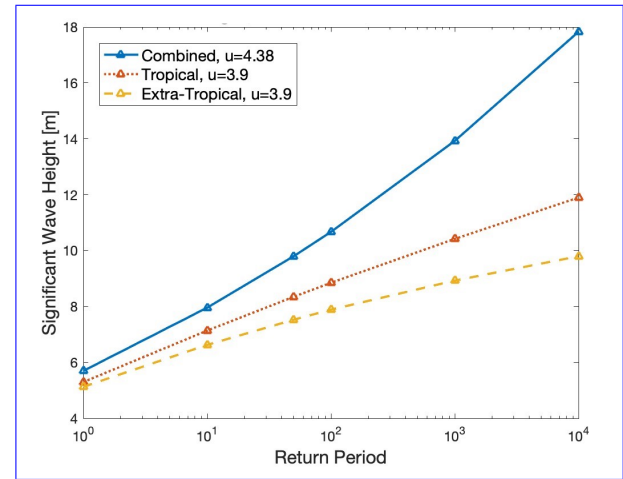
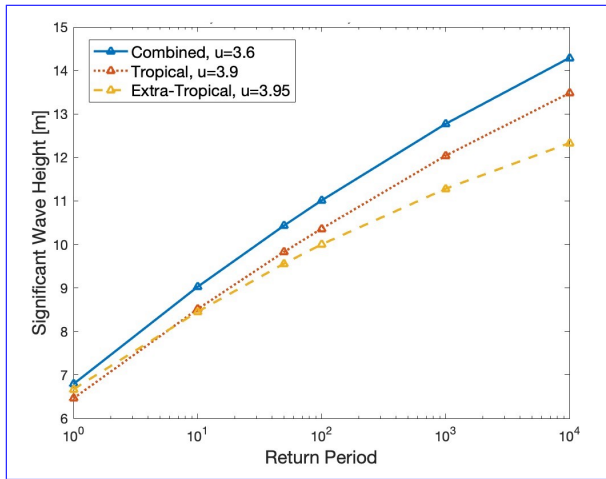
The influence of the chosen method was also investigated. Return values are displayed in Figure A2 for both the NA and MAB high-resolution data with Peaks-Over-Threshold, fit with a Generalized Pareto distribution.

(a) NA Kolmogorov-Smirnov

(b) NA GF-EC Kolmogorov-Smirnov

(c) MAB Kolmogorov-Smirnov

(d) MAB GF-EC Kolmogorov-Smirnov

Figure A1. Gumbel distribution goodness of fit tests

450 ~~Minor differences due to data distribution fit are observed at the North Atlantic site and remain constant for longer return periods. Minor differences due to data distribution, which increase with return period, are observed at the Mid-Atlantic Bight site, resulting in a maximum difference of 8 % in the 10~~While there are differences between absolute return values between extreme value analysis by POT and by BM (BM-Gumbel returned values 3-4m higher than for POT-GP of the same data at both sites), similar layered trends by dataset are observed, ~~000-year case. Overall, the chosen fit is therefore not considered to influence the statistical trends investigated in~~ which is the focus of this study.

455 ~~When comparing values derived from the two extreme value methods, the largest discrepancy~~ At the 1,000-year period (Figure A2), an interesting and opposite trend is observed between the ~~post-processed "tropical" results. Peaks-over-threshold returns larger 10,000-year values than by block maxima in the Mid-Atlantic Bight (10 % larger, POT-Weibull; 3 % larger,~~

(a) Return values at the North Atlantic site from the high-resolution model ~~results~~ ("Combined") and two post-processed subsets of the high-resolution model ~~results~~ based on storm type.

(b) Return values at the Mid-Atlantic Bight site from the high-resolution model ~~results~~ ("Combined") and two post-processed subsets of the high-resolution model ~~results~~ based on storm type.

Figure A2. Return values by Peaks-Over-Threshold for the "high-resolution~~model results when fit~~" return estimates with ~~Gumbel~~ a Generalized Pareto distribution and ~~Weibull distributions~~ selected thresholds, u.

~~POT-Gumbel). In the~~ two sites: in the North Atlantic, ~~analysis by peaks-over-threshold returned smaller 10,000-year values than by block maxima (9.3 % smaller, POT-Weibull; 8.2 % smaller, POT-Gumbel)~~ values trend lower in a non-linear fashion, whereas in the Mid-Atlantic, they trend higher non-linearly. Given that there is no significant bias between the two locations ~~of~~ toward one method or the other, the block maxima method ~~, which is less sensitive to user choices, is~~ is considered suitable for ~~the study~~ this investigation.

Appendix B: Tabulated Extreme Values

Appendix C: Calibration and Tropical Cyclone Events During the Hindcast Period

~~The~~ All tropical cyclone events that occurred during the high-resolution ~~model skill was assessed at the nearest buoy to turbine analysis locations in the North Atlantic and Mid-Atlantic Bight project areas, presented as quantile-quantile plots in Figures ?? - ??. Calibration of~~ forecast and buoy-based observations of these events are provided in Table C3. The largest 16 storms in the 2009–2020 period were used for calibration, four of which were tropical cyclones; there were six tropical cyclones in the North Atlantic ~~high-resolution model was conducted based on observations (Table ??)~~ throughout the model domain. All ~~tropical cyclone events that occurred during the high-resolution forecast and buoy-based observations~~ during this time period.

The Mid-Atlantic model was calibrated against LiDAR buoy observations of ~~these events~~ "Hurricane Isais" in July 2020. The friction coefficient, whitecapping parameterization (Westhuysen or Komen), time step, and wave boundary conditions

Model	50	100	1000	10,000	μ	σ	k
NA HiRes 95% Confidence Intervals	9.35–11.85	9.89–12.72	11.67–15.59	13.44–18.46			
NA HiRes (BM Gumb)	10.60	11.30	13.63	15.95	6.66	1.01	0
NA HiRes (BM GEV)					6.56	0.92	0.192
NA GF-EC Trop. 95% Confidence Intervals	10.58–14.30	11.79–16.01	15.79–21.64	19.78–27.26			
NA ~~high-res~~GF-EC Trop. (BM Gumb)	~~11.82~~12.44	~~12.75~~13.90	~~15.82~~18.71	~~18.94~~23.52	~~1.34~~4.30	~~6.57~~2.09	0
NA GF-EC Trop. (BM GEV)					~~2.17~~4.39	~~4.31~~2.12	-0.0498
NA GF-EC Extra-Trop. 95% Confidence Intervals	9.32–12.72	10.03–13.89	12.37–17.72	14.71–21.55			
NA GF-EC Extra-Trop. (BM Gumb)	~~10.11~~10.11	~~10.88~~10.88	~~13.41~~13.41	~~15.94~~15.94	~~1.10~~5.79	~~5.86~~1.34	0
NA GF-EC Extra-Trop. (BM GEV)					1.72	1.036	0.22
MAB HiRes 95% Confidence Intervals	7.87–10.35	8.30–11.11	9.71–13.60	11.11–16.09			
MAB high-res. (BM Gumb)	~~10.00~~10.75	~~10.76~~11.68	~~14.77~~11.65	~~15.74~~17.87	~~1.34~~5.82	~~5.49~~0.85	0
~~MAB~~MAB HiRes (BM GEV)					5.76	0.799	0.135
MAB GF-EC Trop. 95% Confidence Intervals	10.53–14.17	11.71–15.84	15.62–21.35	19.52–26.85			
MAB GF-EC Trop. (BM Gumb)	~~11.97~~12.35	~~13.31~~13.77	~~17.75~~18.48	~~22.22~~23.18	~~1.93~~4.38	~~4.42~~2.04	0
MAB GF-EC Trop. (BM GEV)					4.62	2.53	-0.22
MAB GF-EC Extra-Trop. 95% Confidence Intervals	7.91–10.93	8.54–11.95	10.62–15.35	12.69–18.74			
MAB GF-EC Extra-Trop. (BM Gumb)	~~9.24~~9.42	~~10.03~~10.25	~~12.65~~12.99	~~15.28~~15.72	~~1.51~~4.79	~~4.65~~1.19	0
~~height~~MAB GF-EC Extra-Trop. (BM GEV)					1.83	0.95	0.04

height

Table B1. Selected return period extreme significant wave heights [m] and associated fit parameters for Block Maxima (BM) analysis at the North Atlantic (NA) and at the Mid-Atlantic Bight (MAB) locations. Fit parameters of the data with the Generalized Extreme Value distribution are also provided, showing a shape factor, k, near 0.

were adjusted during the calibration process. The boundary condition calibration (ERA5 wave forcing on three boundaries, or ERA5 waves on the eastern boundary with open condition on the remaining boundaries) resulted in notable differences

475 in the mean wave direction compared to observations during hurricane passage. This is attributed to improved treatment of wind-driven waves during this event, and ERA5 wave forcing only on the eastern boundary was chosen for the final configuration. Significant wave heights were similar to observations for both boundary condition configurations. The final calibration values are provided in Table ~~C3.~~ C1. Validation statistics are provided in Table C2.

Whitecapping / ~~ID~~	~~Buoy Coordinates~~	Time step	~~Overlapping Period~~	Chezy coefficient
~~NDBC 44008~~height Calibration Range	~~40.496N, 69.250W~~	Westhuysen or Komen	~~March 26, 2003 – December 31, 2020~~	2min–0.2 min
~~NDBC 44017~~Selected	~~40.693N, 72.049W~~	Westhuysen	~~April 30, 2008 – December 31, 2020~~	Variable, reduced from 2 min during certain eve…
~~NDBC 44020~~height				

Table C1. Calibration values for the MAB "high-resolution" model.

Location	~~41.497N, 70.283W~~ RMSE	~~March 10, 2009 – December 31, 2020~~ Mean Absolute Error	Index of Agreement
LiDAR Buoy height 06/2020 - 07/2020	0.26m	0.20m	0.66
NDBC44099 height ~~NDBC 44097~~ 2009–2019	~~40.967N, 71.124W~~0.25m	~~September 17, 2009 – December 31, 2020~~0.18m	~

Table C2. ~~Buoy observations from the National Data Buoy Center (NDBC) used~~ Validation statistics for ~~North Atlantic~~ the MAB "high-resolution" model~~calibration~~.

480 ~~NA high-resolution model skill for significant wave height (Hs) at 40.86N,70.81W, compared to concurrent NDBC buoy 44097 observations. NA high-resolution model skill for peak period (Tp) at 40.86N,70.81W, compared to concurrent NDBC buoy 44097 observations. MAB high-resolution model skill for significant wave height (Hs) at 36.38N, 75.00W, compared to concurrent NDBC buoy 44014 observations. MAB high-resolution model skill for peak period (Tp) at 36.38N, 75.00W, compared to concurrent NDBC buoy 44014 observations. Quantile-Quantile comparison of calibrated high-resolution model wave results with measurements. Blue dots represent data in the overlapping time period, and red circles represent the quantiles~~

485 ~~in the associated regime. The black diagonal line represents a 1:1 correspondence between model and measurement data.~~

Storm (Approx. day of peak)	North Atlantic Observations	Mid-Atlantic Bight Observations	In GF-EC NA model?	In GF-EC MAB model?
August 19, 1991		NDBC 44014	X	X
October 30, 1991		NDBC 44014	X	X
September 1, 1993		NDBC 44014		X
November 18, 1994		NDBC 44014 NDBC 44019		X
August 17, 1995		NDBC 44014		X
July 13, 1996		NDBC 44014	X	X
September 6, 1996		NDBC 44014		X
October 8, 1996		NDBC 44014		X
August 28, 1998		NDBC 44014		X
September 1, 1999		NDBC 44014		X
September 16, 1999		NDBC 44014	X	
October 18, 1999		NDBC 44014		X
September 10, 2002		NDBC 44014		X
September 28, 2003		NDBC 44014	X	X
August 3, 2004		NDBC 44014	X	X
October 25, 2005		NDBC 44014		X
November 3, 2007		NDBC 44014 NDBC 44056	X	X
September 6, 2008		NDBC 44014 NDBC 44056 NDBC 44100		X
September 3, 2010	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100	X	X
August 27, 2011	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100	X	X
October 29, 2012	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100	X	X
July 4, 2014	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100		X
October 9, 2016	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100	X	X
September 19, 2017	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100	X	X
September 13, 2018	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100 NDBC44086	X	X
October 12, 2018	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100 NDBC44086		X
September 6, 2019	NDBC 44097	NDBC 44014 NDBC 44056 NDBC 44100 NDBC44086	X	X

Table C3. Available observations of hurricane events in the modeled time period

Data availability. Original high-resolution wave model time series at 40.8°N, 70.7°W and 36.2°N, 75.0°W are available at https://doi.org/10.5281/zenodo.13884957. Due to intellectual property protection, the GROW-Fine EC data is not publicly available.

Author contributions. Sarah McElman contributed data, designed the study methodology, and conducted numerical/statistical analysis. Amrit Verma and Andrew Goupee contributed paper review, editing, and guidance.

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

Acknowledgements. This work was possible with the valuable insight and perspectives of Louis Bowers, Cyril Frelin, Chan Kwon Jeong, and Gregory Gerbi. The ~~author~~ authors would also like to thank the Oceanweather team (Andrew Cox, Erin Harris) for their feedback and for supporting the use of the GROW-Fine model for academic work. This research was partially funded by the US National Science Foundation (CBET-2343210 to Dr. Verma and Dr. Goupee)

References

Barthelmie, R. J., Dantuono, K. E., Renner, E. J., Letson, F. L., and Pryor, S. C.: Extreme wind and waves in U.S. east coast offshore wind energy lease areas, Energies, 14, https://doi.org/10.3390/en14041053, 2021.

Bhaskaran, S., Verma, A. S., Goupee, A. J., Bhattacharya, S., Nejad, A. R., and Shi, W.: Comparison of Extreme Wind and Waves Using Different Statistical Methods in 40 Offshore Wind Energy Lease Areas Worldwide, Energies, 16, https://doi.org/10.3390/en16196935, 2023.

Caires, S. and Sterl, A.: American Meteorological Society 100-Year Return Value Estimates for Ocean Wind Speed and Significant Wave Height from the ERA-40 Data, Source: Journal of Climate, 18, 1032–1048, https://doi.org/10.2307/26253445, 2005.

Campos, R. M., Gramcianinov, C. B., de Camargo, R., and da Silva Dias, P. L.: Assessment and Calibration of ERA5 Severe Winds in the Atlantic Ocean Using Satellite Data, Remote Sensing, 14, https://doi.org/10.3390/rs14194918, 2022.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag, London, 2001.

Dolan, R. and Davis, R. E.: An Intensity Scale for Atlantic Coast Northeast Storms, Journal of Coastal Research, 8, 840–853, http://www.jstor.org/stable/4298040, 1992.

Fonseca, R. B., Gonçalves, M., and Guedes Soares, C.: Comparing the Performance of Spectral Wave Models for Coastal Areas, Journal of Coastal Research, 33, 331–346, https://doi.org/10.2112/JCOASTRES-D-15-00200.1, 2016.

Forristall, G. Z. and Ewans, K. C.: Worldwide Measurements of Directional Wave Spreading, Journal of Atmospheric and Oceanic Technology, 15, 440 – 469, https://doi.org/10.1175/1520-0426(1998)015<0440:WMODWS>2.0.CO;2, 1998.

Gandoin, R. and Garza, J.: Underestimation of strong wind speeds offshore in ERA5: evidence, discussion and correction, Wind Energy Science, 9, 1727–1745, https://doi.org/10.5194/wes-9-1727-2024, 2024.

Gomes, L. and Vickery, B. J.: Extreme wind speeds in mixed wind climates, Journal of Wind Engineering and Industrial Aerodynamics, 2, 331–344, https://doi.org/10.1016/0167-6105(78)90018-1, 1978.

Gomez, M. S., Lundquist, J. K., Deskos, G., Arwade, S. R., Myers, A. T., and Hajjar, J. F.: Wind Fields in Category 1–3 Tropical Cyclones Are Not Fully Represented in Wind Turbine Design Standards, Journal of Geophysical Research: Atmospheres, 128, https://doi.org/10.1029/2023JD039233, 2023.

Groll, N. and Weisse, R.: A multi-decadal wind-wave hindcast for the North Sea 1949-2014: CoastDat2, https://doi.org/10.5194/essd-9-955-2017, 2017.

Haselsteiner, A. F. and Thoben, K. D.: Predicting wave heights for marine design by prioritizing extreme events in a global model, Renewable Energy, 156, 1146–1157, https://doi.org/10.1016/j.renene.2020.04.112, 2020.

Heideman, J. C. and Mitchell, D. A.: Grid Point Pooling in Extreme Value Analysis of Hurricane Hindcast Data, Journal of Waterway, Port, Coastal, and Ocean Engineering, 135, 31–37, https://doi.org/10.1061/ASCE0733-950X2009135:231, 2009.

Hwang, P. A. and Walsh, E. J.: Estimating maximum significant wave height and dominant wave period inside tropical cyclones, Weather and Forecasting, 33, 955–966, https://doi.org/10.1175/WAF-D-17-0186.1, 2018.

Jonathan, P. and Ewans, K.: Statistical modelling of extreme ocean environments for marine design: A review, https://doi.org/10.1016/j.oceaneng.2013.01.004, 2013.

Jones, I.: Wind Stress Over the Ocean, Cambridge University Press, Cambridge, 2011.

Kim, E. and Manuel, L.: Simulation of Wind, Waves, and Currents During Hurricane Sandy for Planned Assessment of Offshore Wind Turbines, Journal of Offshore Mechanics and Arctic Engineering, 141, https://doi.org/10.1115/1.4043777, 2019.

Kresning, B., Hashemi, M. R., Shirvani, A., and Hashemi, J.: Uncertainty of extreme wind and wave loads for marine renewable energy farms in hurricane-prone regions, Renewable Energy, 220, https://doi.org/10.1016/j.renene.2023.119570, 2024.

MacAfee, A. W. and Wong, S. W.: Extreme value analysis of tropical cyclone trapped-fetch waves, Journal of Applied Meteorology and Climatology, 46, 1501–1522, https://doi.org/10.1175/JAM2555.1, 2007.

Mackay, E. and Johanning, L.: A generalised equivalent storm model for long-term statistics of ocean waves, Coastal Engineering, 140, 411–428, https://doi.org/10.1016/j.coastaleng.2018.06.001, 2018.

Neary, V. S., Ahn, S., Seng, B. E., Allahdadi, M. N., Wang, T., Yang, Z., and He, R.: Characterization of Extreme Wave Conditions for Wave Energy Converter Design and Project Risk Assessment, Journal of Marine Science and Engineering, 8, https://doi.org/10.3390/JMSE8040289, 2020.

Northrop, P. J., Attalides, N., and Jonathan, P.: Cross-validatory extreme value threshold selection and uncertainty with application to ocean storm severity, https://academic.oup.com/jrsssc/article/66/1/93/7068122, 2017.

Oceanweather: GROW-FINE U.S. East Coast: Global Reanalysis of Ocean Waves-U.S. East Coast, 2022.

O'Grady, J. G., Stephenson, A. G., and McInnes, K. L.: Gauging mixed climate extreme value distributions in tropical cyclone regions, Scientific Reports, 12, https://doi.org/10.1038/s41598-022-08382-y, 2022.

Padilla-Hernández, R., Perrie, W., Toulany, B., and Smith, P. C.: Modeling of two northwest Atlantic storms with third-generation wave models, Weather and Forecasting, 22, 1229–1242, https://doi.org/10.1175/2007WAF2005104.1, 2007.

Portilla-Yandun, J., Ocampo Torres, F., and Monbaliu, J.: Spectral Partitioning and Identification of Wind Sea and Swell, Journal of Atmospheric and Oceanic Technology - J ATMOS OCEAN TECHNOL, 26, https://doi.org/10.1175/2008JTECHO609.1, 2009.

Simpson, R.: The hurricane disaster–potential scale, Weatherwise, 27, 1974.

Stephens, S. A. and Gorman, R. M.: Extreme wave predictions around New Zealand from hindcast data, New Zealand Journal of Marine and Freshwater Research, 40, 399–411, https://doi.org/10.1080/00288330.2006.9517431, 2006.