Development of a Wave Model Component in the First Coupled Global Ensemble Forecast System at NOAA

JOSE-HENRIQUE ALVES,[a] ROBERTO PADILLA-HERNANDEZ,[b] DEANNA SPINDLER,[c] WALTER KOLCZYNSKI,[d] BHAVANI RAJAN,[g] TODD SPINDLER,[c] ALI ABDOLALI,[e] RICARDO CAMPOS,[f] SAEIDEH BANIHASHEMI,[c] AND JESSICA MEIXNER[g]

[a] *Weather Program Office, NOAA/Research, Silver Spring, Maryland*
[b] *Fleet Numerical Meteorology and Oceanography Center, U.S. Navy, Stennis Space Center, Mississippi*
[c] *NOAA/Lynker at National Centers for Environmental Prediction, College Park, Maryland*
[d] *SAIC at Environmental Modeling Center, NOAA/National Centers for Environmental Prediction, College Park, Maryland*
[e] *U.S. Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory, Vicksburg, Mississippi*
[f] *NOAA/Cooperative Institute for Marine and Atmospheric Studies (University of Miami) at Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida*
[g] *Environmental Modeling Center, NOAA/National Centers for Environmental Prediction, College Park, Maryland*

(Manuscript received 14 March 2024, in final form 25 August 2024, accepted 29 August 2024)

ABSTRACT: We describe the development of the wave component in the first global-scale coupled operational forecast system using the Unified Forecasting System (UFS) at the National Oceanic and Atmospheric Administration (NOAA), part of the U.S. National Weather Service (NWS) operational forecasting suite. The operational implementation of the atmosphere–wave coupled Global Ensemble Forecast System, version 12 (GEFSv12), in September 2020 was a critical step in NOAA's transition to the broader community-based UFS framework. GEFSv12 represents a significant advancement, extending forecast ranges and empowering the NWS to deliver advanced weather predictions with extended lead times for high-impact events. The integration of a coupled wave component with higher spatial and temporal resolution and optimized physics parameterizations notably enhanced forecast skill and predictability, particularly benefiting winter storm predictions of wave heights and peak wave periods. This successful endeavor encountered challenges that were addressed by the simultaneous development of new features that enhanced wave model forecast skill and product quality and facilitated by a multidisciplinary team collaborating with NOAA's operational forecasting centers. The GEFSv12 upgrade marks a pivotal shift in NOAA's global forecasting capabilities, setting a new standard in wave prediction. We also describe the coupled GEFSv12-Wave component impacts on NOAA operational forecasts and ongoing experimental enhancements, which altogether represent a substantial contribution to NOAA's transition to the fully coupled UFS framework.

KEYWORDS: Wind waves; Ensembles; Forecast verification/skill; Coupled models; Numerical weather prediction/forecasting; Other artificial intelligence/machine learning

1. Introduction

The implementation of atmosphere–wave coupling in the National Oceanic and Atmospheric Administration (NOAA) Global Ensemble Forecast System, version 12 (GEFSv12), was a first step toward expanding the coupled forecast system paradigm in the planned transition of NOAA's operational modeling systems to the community-based Unified Forecasting System (UFS). This paper recounts successes achieved and lessons learned while establishing the wind-wave component in what became the first global-scale coupled operational forecast system at NOAA.

GEFSv12 became operational on 23 September 2020. The upgrade incorporated the innovative concept of atmosphere–wave and atmosphere–aerosols coupling, fueled by contributions from the UFS community modeling framework. The new system expanded the prior atmospheric and wave forecast durations, transitioning from a 10-day range to ranges up to 16 days (high resolution) and then out to 35 days (lower resolution). This enhancement enabled the National Weather Service (NWS) to provide numerical weather prediction (NWP) model guidance with an extended lead time of up to 4 weeks (Zhou et al. 2022). The enhancement offered invaluable lead time for preparing and responding to high-impact weather events and expanded the system's capabilities toward a probabilistic subseasonal product, currently in development.

The GEFSv12 upgrade integrated one-way coupling with the global wave ensemble component (henceforth GEFSv12-Wave), featuring one-way coupling from the GFDL finite-volume cubed-sphere (FV3) atmospheric model component in the UFS. Relative to a previous global operational ensemble system at NOAA, the higher-resolution coupled GEFSv12-Wave component products exhibit a marked increase in forecast skill, also extending the forecast range from 10 to 16 days and augmenting the ensemble size from 21 to 31 members [30 perturbations and 1 control (Ctrl) model], among other benefits discussed below. Notably, GEFSv12-Wave provides wave forecasts with reduced biases that were prominent during the challenging-to-forecast winter season, holding significant implications for

Alves, Padilla-Hernandez, D. Spindler, and T. Spindler's former affiliation: Environmental Modeling Center, NOAA/NCEP, College Park, Maryland.

Corresponding author: Jose-Henrique Alves, henrique.alves@noaa.gov

DOI: 10.1175/WAF-D-24-0048.1

regions frequently impacted by severe and small-scale, fast-moving, and rapidly changing winter storms, such as the Alaska nearshore region. Enhanced NWP skill, particularly on day 7 and beyond, was also found to improve wave guidance associated with tropical cyclones represented in the ensemble members.

Among the notable advances that warranted the system implementation were the higher agreement of the wave-height ensemble mean with observations, as explored in more detail below and in a companion paper (Campos et al. 2024a), and greater agreement between predicted peak wave periods with measured data. For both wave heights and peak wave periods, root-mean-square error (RMSE) and bias of the ensemble mean exhibited enhancements, more notably during Northern Hemisphere winters, further highlighting the system's operational value. For all validation parameters, no systematic degradation was detected.

Several challenges were addressed that resulted in a robust wave modeling system with greater reliability. The latter included adaptations of the model package to the new coupling paradigm, modifications of source terms to adapt to changes in the atmospheric model dynamic core, new strategies to achieve deadlines at a faster pace, and a significant redesign of infrastructure to allow meeting simultaneously computer hardware resource demands and operational product-delivery requirements. To ensure implementation timelines were satisfied, the challenge of developing a new coupling infrastructure was tackled in tandem with adding new features intrinsic to the wave model component, targeting the improvement of forecast skill and product quality. Some examples were the implementation of higher-resolution grids, updating product definitions to align with current WMO standards, and rolling out experimental enhanced product generation using neural networks/artificial intelligence (AI)-based techniques.

Challenges were successfully overcome, thanks to the establishment of a team with a multidisciplinary background and its close collaboration with key stakeholders from NOAA operational forecasting centers. The outcome was a system that has set a gold standard in wave prediction and is still relevant since its implementation in 2020. The developmental effort and results associated with GEFSv12-Wave have seeded new products at NOAA, such as the 20-yr global wave ensemble reforecast and ongoing efforts at NOAA to develop AI/machine learning (ML)-based probabilistic wave-height products (Campos et al. 2024b).

The GEFSv12 upgrade inaugurated a new era of NOAA's global forecast capabilities leveraging the UFS community modeling framework. It was the first step in the planned NWS transition of its suite of stand-alone forecasting systems to the coupled UFS framework. This paper summarizes the development effort behind the operational implementation of the wave component in the coupled GEFSv12 and is structured as follows. Section 2 reviews global wave models at NOAA and their transition to the coupled UFS paradigm. Details of the GEFSv12-Wave component development and optimization efforts are provided in section 3, while section 4 describes the verification and validation process, as well as performance metrics. In section 5, we discuss how the coupled wave component implementation impacted forecasting efforts at NOAA, as well as recent experimental product developments in wave modeling efforts associated with these initial efforts. Concluding remarks are made in section 6.

2. Overview of NOAA's wave ensemble systems

NOAA has used operational wave models to provide forecast guidance to the National Weather Service since 1956 (Hubert 1957). Early models produced a single wave height and period at each grid point, using a direct relation between the local wind speed and the wave height and period. This simple prediction approach, generally identified as the "representative wave" approach, remained operational at the NWS until 1985. Since then, NOAA operational wave systems have used spectral models to produce forecasts (Tolman et al. 2002). Spectral wind-wave models have used a spectral representation of the ocean surface, breaking it down into a sum of waves with different frequencies and directions. By solving the wave action balance equation for each wave spectral component, the models can predict the behavior of the wave field over time and space.

The first spectral model used operationally at NCEP [then known as the National Meteorological Center (NMC)] was based on the second-generation "SAIL" model (Cardone and Ross 1979). The NMC version of the SAIL model, named the NOAA operational wave model (NOW), became operational in 1985 (Feit 1986). In 1994, an implementation of the Wave Modeling Group (WAM) model (WAMDI Group 1988; Komen et al. 1994) replaced the global NOW model as the operational model at NCEP and resulted in a major improvement in the quality of the numerical forecasts of significant wave heights (SWHs) (Chen 1995).

Limitations in the WAM model, discussed in detail by Tolman et al. (2002), led NOAA to start developing a new model in 1993. The new model, called NOAA WAVEWATCH III[1], or NWW3, formally became operational for global applications on 9 March 2000 (Chen et al. 1999). The first version of WAVEWATCH III (WW3) was made available to the general public several months before its first operational implementation, a step that facilitated evolving WW3 from an institutional model to an open-source package to, currently, a community modeling package that inspired NOAA-wide changes in its model development approach (Alves et al. 2023). The latest WW3 package includes state-of-the-art scientific advancements in wind-wave modeling and dynamics. WW3 is now available on GitHub and has become a community model with several thousand users in more than 100 countries. A comprehensive description of WW3 features is provided in its user manual (WW3DG 2019).

NOAA's first operational global wave ensemble forecasting system was implemented in 2004 (Chen 2006). The new system had 11 WW3 model members and was run twice daily using a single domain with a 2.5° × 2.5° latitude–longitude resolution. The implemented ensemble forecast system was found to have better forecast skill than the deterministic operational forecast

[1] WAVEWATCH is an acronym originally representing wave height, water depth, and current hindcasting. The expanded version has been deprecated by use.

system, as discussed in Cao et al. (2007). The initial version used input and dissipation source terms ST2 (Tolman and Chalikov 1996), nonlinear wave–wave interactions based on discrete interaction approximations (DIAs) by Hasselmann and Hasselmann (1985), and propagation employing a third-order-accurate scheme (Leonard 1991). New physics parameterizations added to the underlying WW3 model included the linear growth parameterization of Cavaleri and Rizzoli (1981) and a depth-induced breaking term from Battjes and Janssen (1978).

In 2008, NOAA's global wave ensemble system was upgraded from a 10- to a 20-member ensemble running 4 cycles per day with a forecast horizon out to 10 days. Upgrades included a higher-resolution global grid with a $1° \times 1°$ latitude–longitude resolution and changes to initial conditions that increased ensemble variability. The upgrades resulted in greater forecast skill and spread in wave-height data, providing better statistical estimates and probabilistic forecasts (Alves et al. 2015a). The upgraded system provided high-quality deep-water wave forecasts and expanded the applicability of NOAA's wave ensemble system for nearshore areas and under hurricane-forcing conditions, as demonstrated in the evaluation of its performance during the 2012 Tropical/Posttropical Storm Sandy (Alves et al. 2015b).

The U.S. Navy joined forces with NOAA in 2011 to implement the first global multicenter system to forecast wave heights (Alves et al. 2013). The system combined two independent wave ensemble systems, resulting in a more accurate probabilistic forecast of wave heights. The new system outperformed deterministic and ensemble wave models from the two forecast centers, which led to expanding collaborative efforts in wave forecasting to other operational centers in North America, including the Canadian Meteorological Centre (Bernier et al. 2016).

Upgrades to the global wave ensemble system (GWES) made in July 2014 introduced the Ardhuin et al. (2010) physics and a global grid with an increased $1/2° \times 1/2°$ latitude–longitude resolution. The changes improved the model's overall forecast performance (Alves et al. 2015a) and the probabilistic forecast skill in extreme wave conditions (Campos et al. 2018). Case studies demonstrating the high performance of NOAA's wave ensemble forecasts in hurricanes are presented by Alves et al. (2015b) for Hurricane Sandy (2012) and Campos et al. (2022) for Hurricane Lorenzo (2019). After 2014, the GWES remained largely unchanged until transitioning from a stand-alone model to the global coupled atmosphere–wave system GEFSv12 in 2020. This paper describes several efforts in support of improvements in wave models toward the implementation of NOAA's global coupled atmosphere–wave forecast systems, particularly GEFSv12.

3. Coupled GEFSv12-wave system development

The GEFSv12 upgrade on 23 September 2020 brought a substantial enhancement in scientific and technical capabilities that improved NOAA's probabilistic wave forecast skill, significantly reducing biases and increasing ensemble spread. GEFSv12 scientific advancements included the replacement of its spectral atmospheric dynamical core, the Global Spectral Model (GSM: Sela 1980), with the FV3 model, the incorporation of new perturbation techniques, and the enhancement of physical parameterization schemes, including the introduction of new microphysics schemes (Zhou et al. 2022).

The GEFSv12 upgrade expanded the ensemble size from 21 to 31 members, integrated a one-way coupling with the global wave ensemble component, and introduced an additional control member with coupling with global aerosols. The atmospheric component's horizontal resolution was increased to approximately 25 km, whereas 16-day forecasts were enhanced with extended weeks 3–4 forecast step with a resolution of ~25 km.

GEFSv12-Wave featured an expanded ensemble size from 21 to 31 members and extended the forecast duration from 10 to 16 days, matching data availability from the atmospheric component. In addition, the atmospheric wind field intake stride was reduced from 3 to 1 h forcing. The spatial domain was extended by replacing the single GWES 0.5° grid with a multigrid mosaic with three grid domains, consisting of a core low/midlatitude belt with a $^1/_4°$ resolution and two high-latitude $^1/_3°$ resolution grids covering the Arctic and Antarctic regions. The improvements culminated in a significant boost in forecast skill, achieved through physics optimization tailored to the updated atmospheric forcing using an objective framework (see below).

Development of wave model systems at NCEP in the last few years focused on transitioning operational systems from the stand-alone wave forecasting paradigm to the coupled systems approach within the UFS framework. The effort led both global-scale wave models operational at NOAA, namely, the multigrid deterministic system (Multi-1) and the ensemble system (GWES), to become coupled to the NCEP's Global Forecast System, version 16 (GFSv16), and GEFSv12, respectively, within the UFS infrastructure.

The transition of stand-alone wave models to coupled UFS components faced numerous challenges, including adaptations of the model package to the new coupling paradigm, modifications of source terms to adapt to changes in the atmospheric model dynamic core, and a full redesign of infrastructure to allow meeting simultaneously computer hardware resource limitations and operational product-delivery requirements. Additional challenges were faced since coupling-related efforts had to be made in tandem with developing new features intrinsic to the wave model component, targeting the improvement of wave model forecast skill and product quality. The latter included the implementation of higher-resolution grids, updating product definitions to align with current WMO standards, and experimentally improved product generation using neural networks/AI-based techniques.

a. *Coupling infrastructure*

The coupling infrastructure for GEFSv12 employs the NOAA Environmental Modeling System (NEMS), which relies on the National Unified Operational Prediction Capability (NUOPC) layer Earth System Modeling Framework (ESMF) detailed in Theurich et al. (2016). Each component has a "cap" which communicates with the coupling infrastructure. An advantage of the NUOPC layer is the interoperability with other NUOPC-based systems, thus allowing a shared cap for multiple institutions and coupled systems. The WW3 cap was

TABLE 1. List of wave model source term parameters optimized using Cyclops v1.0.

Parameter	Code variable	Initial value	Final value
βmax	BETAMAX	1.33	1.3068
s_u	TAUWSHELTER	1.00	0.9626
s_0	SWELLFPAR	0.80	0.7921
s_2	SWELLF	−0.018	−0.018 32
s_3	SWELLF3	0.015	0.0149
s_4	SWELLF4	1×10^5	1×10^5
s_5	SWELLF5	1.20	1.20
s_7	SWELLF7	2.3×10^5	2.303×10^5
C	NLPROP	2.5×10^7	2.501×10^7
Sds	FXFM3	2.5	2.497
C_ds^sat	SDSC2	-2.2×10^{-5}	-2.197×10^{-5}
C_cu	SDSCUM	−0.4034	−0.4033
cturb	SDSC5	0.0	0.0
δ_d	SDSC6	0.30	0.2972
Br	SDSBR	9×10^{-4}	9.042×10^{-4}
C_ds^HCK	SDSHCK	1.5	1.5
sB	SDSCOS	2.00	2.0
$\epsilon_{(c,0)}$	ICE0	0.25	0.2503
$\epsilon_{(c,n)}$	ICEN	0.25	0.7505

originally developed at NRL and is used in their coupled system as well. Another advantage of the coupling infrastructure is that it allows for more frequent exchange of information between components. This feature enabled GEFSv12 to increase the frequency of wind input to the wave component, which combined with other improvements led to improved wave forecast skill.

b. Source-term optimization

Source terms are parameterizations in wave models representing growth, decay, and energy rebalancing, governing the transformation of wind-wave spectra as they evolve over space and time. The development of the coupled GEFSv12-Wave system included an optimization process consisting of refining wave model source term parameters by systematically minimizing error metrics in model predictions. The automated model optimization used the Cylc workflow engine (Cyclops v1.0) described in Gorman and Oliver (2018). The approach allowed the identification of parameter configurations that reduced error metrics both globally and locally within the model simulations.

The Cyclops optimization engine was adjusted to use the WW3 model using the GEFSv12-Wave model configuration. Optimization consisted of minimizing the RMSE of significant wave height (Hs; defined as the mean wave height of the highest third of the waves) against global altimeter data from multiple satellites over a period of 2 years, using a bound-constrained optimization algorithm described in Powell (2009). A total of 19 adjustable source-term parameters were included in the optimization, including wave model parameterizations of the following physical processes: wind-driven growth, whitecap dissipation, swell decay, nonlinear wave–wave interactions, and ice blocking.

Table 1 provides the initial and final values of each chosen parameter. Formulas for the source terms containing parameters listed in Table 1 are provided in WW3DG (2019). In the data availability statement, we provide links to the public GitHub

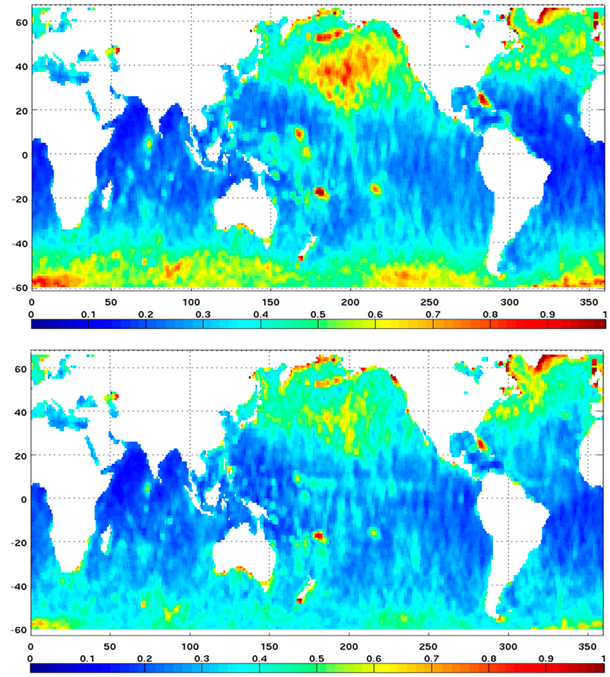

FIG. 1. Comparison of SWH (Hs) RMSE in meters from a 2-yr simulation period, showing improved metrics for (bottom) the optimized model output relative to the (top) initial manually tuned data. Global mean RMSE values reduced from 0.14 to 0.12 m.

repository with a complete set of WW3 input files, including parameter settings, model and grid configurations, used in the operational GEFSv12 system.

Optimal configurations of source-term parameters for GEFSv12-Wave converged after 44 Cyclops iterations. The tuning of WW3 physics parameterizations yielded significant improvements in Hs error metrics, as shown in Figs. 1 and 2. The automated optimization led to a substantial reduction of approximately 15% in the global RMSE, from 0.14 to 0.12 m. However, a closer look at Fig. 1 reveals remarkably higher local improvements in the RMSE, particularly in high-latitude hazard-prone regions, such as the North Atlantic and Pacific Ocean basins and the Southern Ocean. The global bias underwent a significant change, dropping from 0.12 to −0.06 m, representing a 50% reduction in the absolute bias level, as shown in Fig. 2.

Improved results were primarily driven by relatively small alterations in source-term parameters, which collectively had a great impact on the quality of model simulations, especially in high-latitude areas. Such accuracy in parameter tuning would likely not be possible to obtain using manual, trial-and-error optimization. Our results strongly encourage the establishment of the automatic calibration of wave-model physics in response to future atmospheric model upgrades to maintain and possibly elevate the quality of wave predictions.

4. Performance, verification, and validation

a. Scientific evaluation for operational implementation

Scientific evaluation of the GEFSv12-Wave system consisted of a retrospective analysis covering the period from

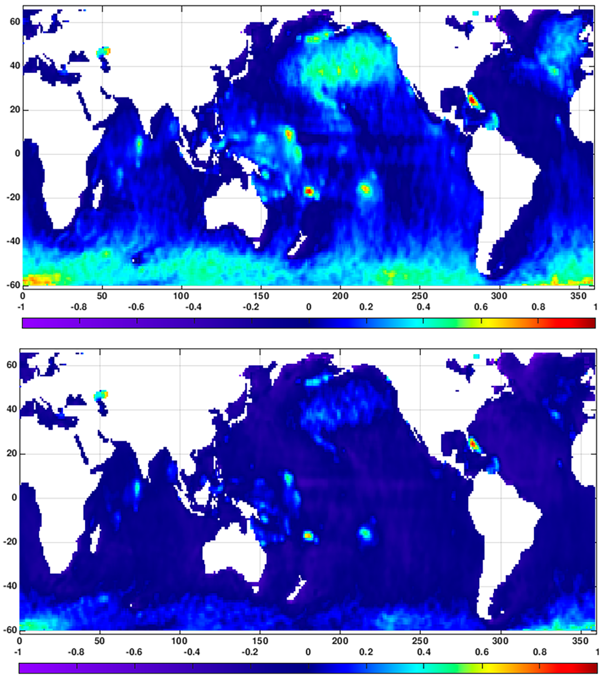

FIG. 2. Comparison of SWH (Hs) bias in meters from a 2-yr simulation period, showing improved metrics for (bottom) optimized model output relative to the (top) initial manually tuned data. Global mean biases improved from 0.12 to −0.06 m.

1 November 2018 to 31 October 2019, focusing on outputs from the 0000 UTC cycle, using data from all 31 members over a 16-day period. The primary objective of the scientific evaluation was to validate the accuracy of predictions for Hs and peak period (Tp). This validation employed standard metrics including bias, RMSE, and continuous ranked probability score (CRPS) and was the basis for decision-making at NOAA toward the operational implementation of GEFSv12-Wave.

The validation dataset consisted of observations collected from approximately 150 buoys from the National Data Buoy Center (NDBC) network, along with data acquired from the *Jason-3* and *SARAL* altimeters. A comparative analysis was also made consisting of a cross-reference with both the operational GWES and the deterministic stand-alone wave model Multi-1. The latter was considered to be the flagship operational global wave guidance system at NOAA at the time of evaluation.

Figure 3 illustrates the monthly Hs statistics for days 1 and 5, derived from altimeter data, indicating notable trends. Smaller Hs RMSE and bias were consistently observed throughout the year for both short- and long-range forecasts. The same behavior is observed for bias values. This reduction in RMSE and bias underscores the improved forecast skill of the upgraded GEFSv12-Wave system for predictions of Hs.

Values of the 95th percentile illustrate a remarkable improvement in the upgraded system's predictive capability for storm-wave events. Agreement with observations indicates a consistent enhancement of storm-wave forecast skill over the course of the year, encompassing both short- and long-forecasting ranges. This enhanced predictability for storm waves, combined with other bulk validation statistics shown in Fig. 3, suggests the robustness of GEFSv12-Wave in representing accurately wave conditions over a broad dynamic range of interest for forecasting applications.

The buoy data validation results shown in Fig. 4 confirm the trends in altimeter-derived predictions for monthly Hs statistics on days 1 and 5. This alignment between buoy and altimeter data underscores the consistency/robustness of outputs from the new system and its capacity to capture effectively and accurately observed Hs characteristics. Another remarkable feature of both validation sets is the enhanced performance of the wave ensemble forecast systems relative to the higher-resolution deterministic system Multi-1.

For the purposes of illustrating improvements in predictability achieved with the GEFSv12-Wave upgrade, Fig. 4 includes model values of Hs ensemble spread alongside plots of RMSE (top). Zhu (2005) argues that a skillful ensemble forecast is expected to produce spread and RMSE with comparable magnitude, reflecting a better representation of forecast uncertainty. A closer agreement between GEFSv12-Wave values of spread and RMSE indicates improvements in predictability.

The improvements in predictability are confirmed in Fig. 5, comparing the CRPS values for GEFSv12-Wave relative to GWES data. CRPS values are a useful verification tool for probabilistic forecasts (Hersbach 2000), providing a concise representation of the reliability and resolution of an ensemble system. At all forecast times, the GEFSv12-Wave predictions show increased predictability relative to the corresponding GWES forecasts. Repeating other validation statistics, the GWESv12-Wave ensemble data consistently provided a more reliable estimate of Hs on the basis of monthly statistics and at all forecast ranges.

Additional validation was performed using fuzzy verification based on Ebert (2008), expanding the validation through the use of additional metrics and plots, such as receiver operating characteristic (ROC) curves. Figure 6 displays ROC curves computed using operational GEFSv12-Wave data, focusing on the Pacific Ocean, which is exposed to the broadest range of different sea states and intensities. The two ROC curves are associated with detecting events above 4 and 6 m within the 2-week range. Perfect, ideal results have a true positive rate (y axis) of 1 and a false positive rate (x axis) of 0, aligning with the top-left corner of the graph. This point represents the optimal outcome.

The ROC curves demonstrate a skillful forecast capability of GEFSv12-Wave in predicting events in the week 2 range, showing a high hit rate and minimal false positives. The area under the curve (AUC) decreases as the threshold shifts from 4 to 6 m, highlighting the challenges in detecting certain extreme events. Examining specific cases individually, it has been observed that GEFSv12-Wave properly forecasts extratropical cyclones and their associated extreme waves in the week 2 range. In terms of recent events, successful predictions of GEFSv12 forecast for week 2 have been confirmed for the Extratropical Storm Ciarán (United Kingdom's Met Office), the Hurricanes Franklin (category 4) and Lee (category 5), whereas unsuccessful events not captured in the probability

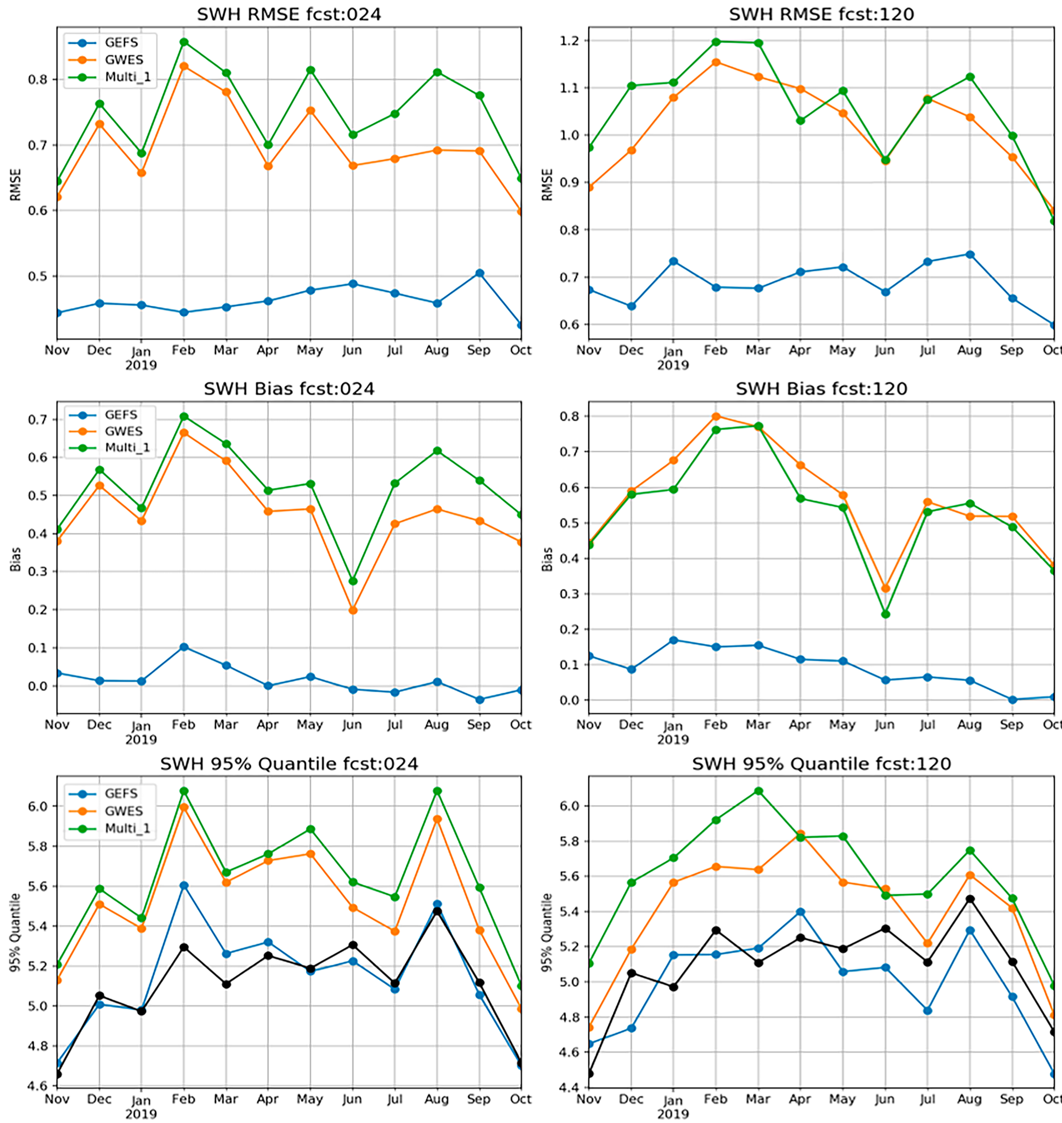

FIG. 3. Monthly SWH (Hs) validation statistics over the 1-yr evaluation period: (top) RMSE, (middle) bias, and (bottom) 95% quantile for the GEFSv12-Wave system (blue), the GWES system (orange), and the deterministic Multi-1 system (green). All values are in meters. Observed 95% percentile Hs are represented in black. Validation relative to altimeter data for November 2018–October 2019.

maps were Hurricanes Don (category 1), Idalia (category 3), and Otis (category 5).

The scientific evaluation of the GEFSv12-Wave system relative to the previous operational wave ensemble system GWES demonstrates a remarkable improvement in the quality of NOAA operational wave forecasts, which supported its operational implementation. The upgraded system showed improved forecast skill in terms of standard validation statistics and predictability in a broad range of wave conditions, including more severe storm-wave events. The alignment between altimeter and buoy data confirms the system's consistency and robustness in capturing observed Hs characteristics. For the purposes of reference, since it is expected that ensemble mean products will generally score better, a qualitative comparison of GEFSv12-Wave data with the operational Multi-1 (deterministic) wave forecasting system is provided, further highlighting its enhanced performance.

Additional validation employing fuzzy verification and ROC curves reaffirms the system's ability to skillfully predict extreme wave events, revealing promising forecast capabilities even out to the week 2 range. These findings provide compelling evidence that the operational implementation of the GEFSv12-Wave system represented a significant stride in advancing wave forecasting accuracy for NOAA's operational system capabilities.

b. *Performance assessment using a 20-yr wave ensemble reforecast*

A 20-yr global wave reforecast, from 1 January 2000 to 31 December 2019, has recently been generated, utilizing the

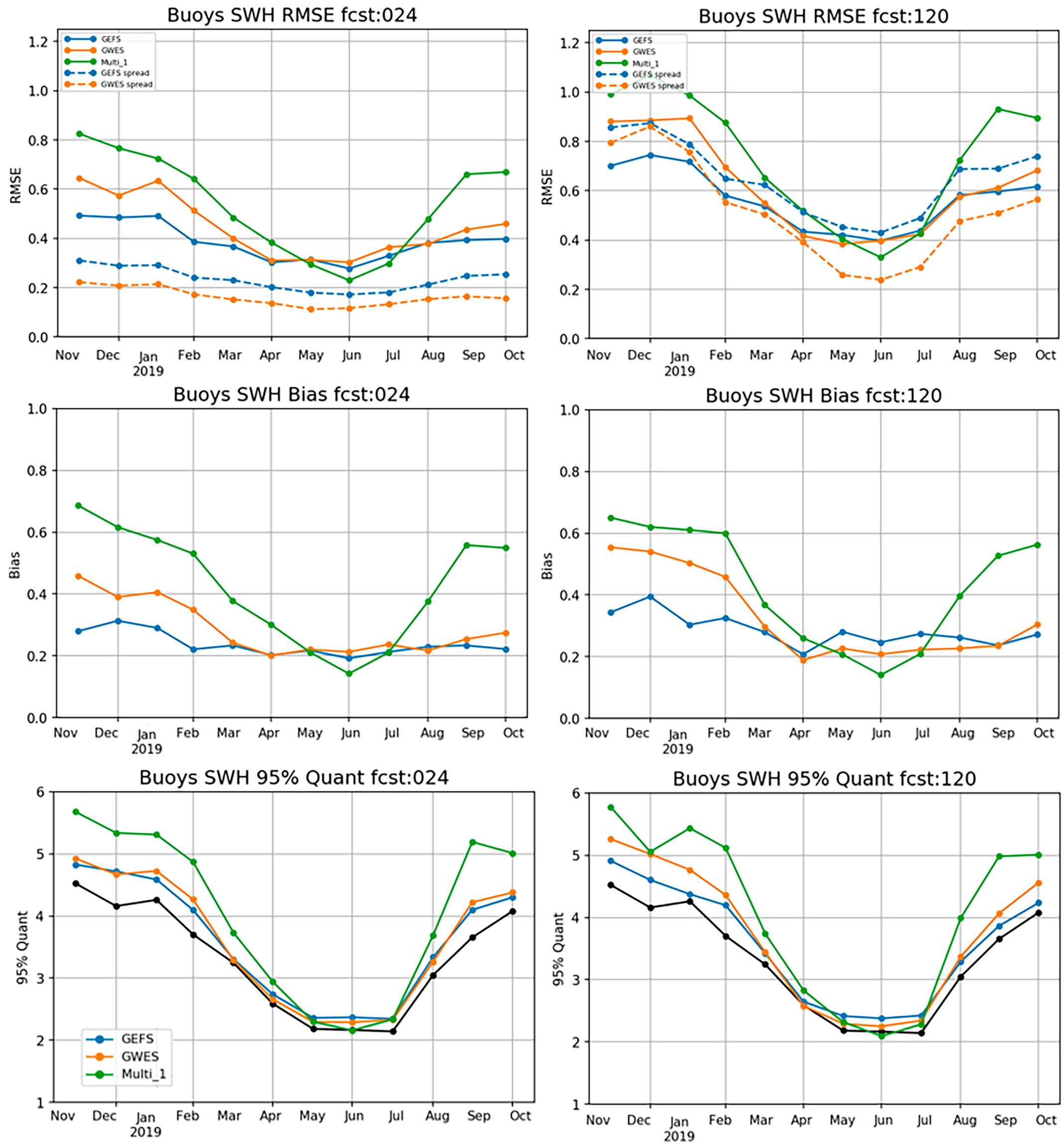

FIG. 4. Monthly SWH (Hs) validation statistics over the 1-yr evaluation period: (top) RMSE, (middle) bias, and (bottom) 95% quantile for the GEFSv12-Wave system (blue), the GWES system (orange), and the deterministic Multi-1 system (green). All values are in meters. Observed 95% percentile Hs are represented in black. Validation relative to buoy data for November 2018–October 2019.

GEFSv12-Wave operational configuration, marking the first publicly available global wave ensemble reforecast dataset with such a long duration. A reforecast mimics the operational runs for each forecast cycle, applied to past conditions, which significantly expands the dataset for validation and further analyses.

The main difference between this new reforecast product and the GEFSv12-Wave operational runs is that the reforecast consists of five members with just 1 cycle per day, as opposed to 31 members with 4 cycles per day. Once a week, the reforecast expands the forecast range from 16 up to 35 days with 11 members. Additionally, the reforecast was run with 3-hourly wind inputs, while the operational system has 1 h of temporal resolution. More information on this new ensemble reforecast product based on GEFSv12-Wave can be found in a companion paper, Campos et al. (2024a). A description of reforecasts using the atmospheric component of GEFSv12 is provided by Hamill et al. (2022) and Zhou et al. (2022).

Validation of the GEFSv12-Wave reforecast was conducted using satellite data from the Australian Ocean Data Network (AODN) altimeter database (Ribal and Young 2019), which provides observations of Hs and 10-m wind speed (U10). The matchups of model/altimeters were built following the methodology described in Campos (2023) and the error metrics computed based on Mentaschi et al. (2013) for deep waters only.

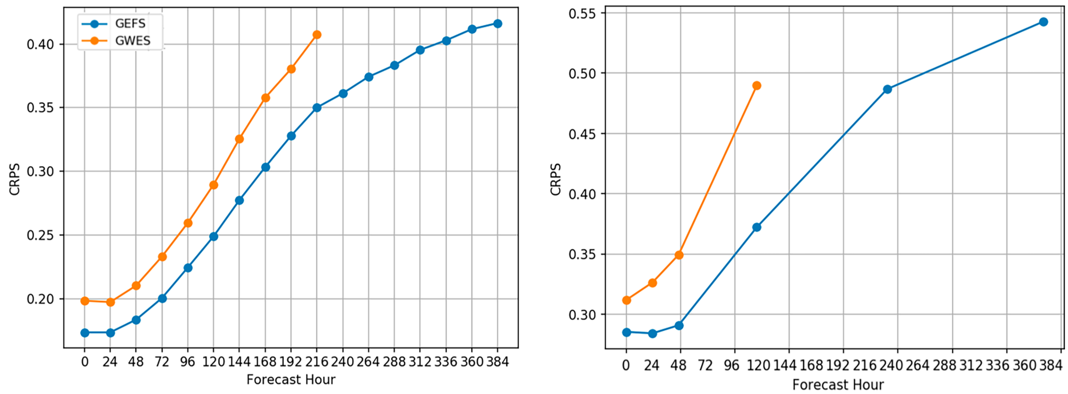

FIG. 5. CRPS values for GEFSv12-Wave and GWES across forecast times. CRPS represents the reliability and resolution of an ensemble system.

The evolution of the normalized RMSE (NRMSE) and the correlation coefficient (CC) with the forecast time is presented in Fig. 7 for Hs and expanding to U10. It is interesting to note the improved performance of the ensemble mean compared to the control member, which highlights the relevance of the ensemble forecast system. Figure 5 confirms the high performance of the wave prediction during the first forecast days, proving the success of the optimization of the wave model described previously.

The reforecast nowcast and forecast day-1 outputs have an RMSE below 10% and a correlation coefficient of 0.96 for Hs. The scatter error increases significantly from forecast day 3 to 10, accompanied by the drop of correlation coefficient. This pattern is observed in both U10 and Hs, with better results for Hs. This is expected since ocean waves act as a low-pass band filter relative to the forcing winds, exhibiting lower variance compared to U10. Beyond 10 days, the validation indicates a decrease in performance, primarily due to the requirement for exact match in time and space between the model and observed data considered during the validation.

The results confirm the remarkable performance of GEFSv12-Wave products at the weather forecasting scales. Performance is still good at longer scales but suggests that improvements can be achieved through the development of probabilistic products that employ temporal and spatial analysis to generate more informative and useful operational wave forecasts. The latter is discussed below.

5. Impacts to future NOAA forecast systems development

The establishment of an operational coupled UFS implementation has propelled the NWS to expand its coupled forecasting capabilities through ample collaboration with NOAA Laboratories, research centers, and academia. The next step in that direction is to enhance GEFS predictability in the longer forecast range up to day 35. The ongoing development aims to provide improved week 2 weather forecasts and weeks 3 and 4 skillful probabilistic guidance. In alignment with the UFS Medium-Range Weather (MRW) application, the upcoming system, GEFSv13, will be an ensemble system including the fully coupled atmosphere, land, wave, ocean, ice, and aerosol components. The wave model component in GEFSv13 will inherit an increase in the forecast length from 16 to 35 days.

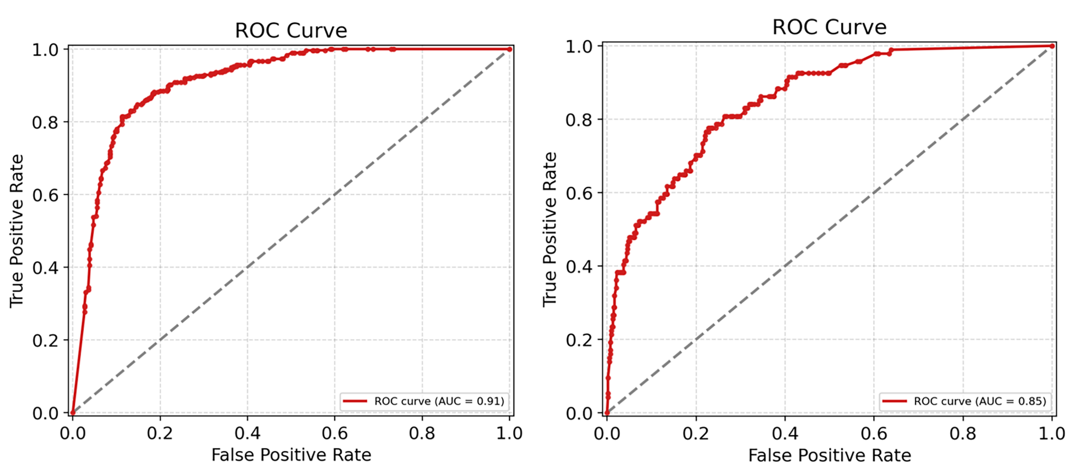

FIG. 6. ROC curves and associated AUC for the week 2 forecast of Hs greater than (left) 4 and (right) 6 m.

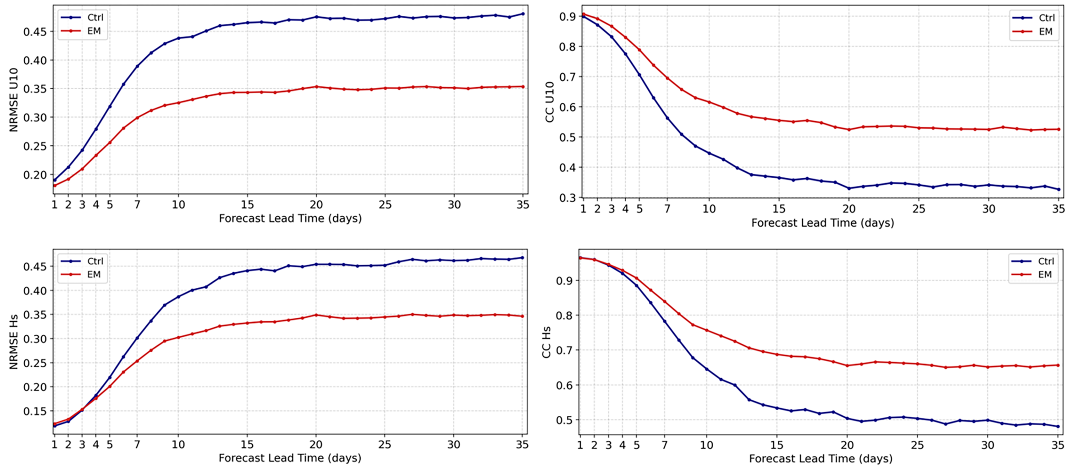

FIG. 7. Analysis of the forecast performance as a function of the forecast lead time, for (top) U10 and (bottom) Hs, calculated from the 20-yr reforecast dataset. The Ctrl members (blue) are compared with the EM (red). (left) NRMSE. (right) CC.

Coupling time steps between the wave and atmosphere will also change from 1 h to 30 min.

Beyond direct impacts on the upcoming upgrades to NOAA's global ensemble forecasting systems, the GEFSv12-Wave results reported above have inspired several other initiatives with positive impacts on NOAA's ability to issue more accurate wave guidance and develop useful forecast products. We explore three more notable initiatives next, including the development of probabilistic forecast products using GEFSv12-Wave data, the exploration of experimental AI/ML applications, and the possibilities of further improvements from using unstructured grids.

a. *Probabilistic forecasts and performance for week 2+ forecasts*

Errors beyond 10 days are notably high in GEFSv12-Wave data, which limits the practical application of forecasts for longer forecast periods. The deterioration in prediction is associated with three main effects: (i) a misalignment between forecast and observed occurrences, leading to storms peaking at different times; (ii) inaccuracies in intensity, causing the model to overestimate or underestimate events; and (iii) an increasing misplacement of meteorological systems, evident in errors such as cyclone track deviations.

The use of probabilistic information, considering a temporal and spatial interval, provides a recourse to attenuate the significant impact of these effects on wind and wave forecasts in week 2 and beyond. On that basis, new products are being developed using outputs from GEFSv12 operational forecasts, including probability maps of Hs and U10 which have been generated experimentally since June 2023. Notable improvements obtained since then have inspired the development of a prototype probabilistic wave product suite scheduled for a future implementation.

Following group meetings with NOAA forecasters and considering the Beaufort scale and Saffir–Simpson hurricane scale, probabilistic wave products will become available at NOAA's Ocean Prediction Center (OPC) website (https://ocean.weather.gov) this year at the following levels:

- U10: 34, 48, and 64 kt (1 kt $\approx$ 0.51 m s^{-1}),
- Hs: 4, 6, 9, and 14 m.

Figure 8 shows an example of probability maps for week 2 associated with the recent Hurricane Lee, where it is possible to confirm that the probabilities calculated using the operational GEFSv12 could successfully indicate the position and intensity of the extreme event on the East Coast of the United States. Results obtained with GEFSv12-Wave data and its expansion through reforecast have made possible the development of probabilistic wave forecasts for week 2. Future developments in this direction will include adding more observations to the validation and expanding the statistical analyses, designing new strategies for probabilistic hurricane wave forecasts, and developing algorithms using machine learning techniques. Some developments in the latter are discussed next.

b. *AI/ML applications and other experimental products*

While GEFSv12 provides reliable wave forecasts, particularly within the first week, there are still biases and significant scatter errors observed in forecast ranges beyond 10 days. These issues can be attenuated through postprocessing based on data-driven algorithms. This topic started to be further explored in 2017 using neural networks applied to the then-operational GWES wave ensemble forecast system (Campos et al. 2017). The authors presented an initial strategy to improve the ensemble mean, whereby one location in the Atlantic Ocean and one in the Pacific Ocean were used to train and test multilayer perceptron neural networks (MLP-NNs) at single points. Campos et al. (2019) advanced the hybrid modeling based on NNs in the Gulf of Mexico, finding that shallow networks with a small number of neurons effectively reduce bias, while 35–50 neurons are required to provide a reduction in both the scatter and system-

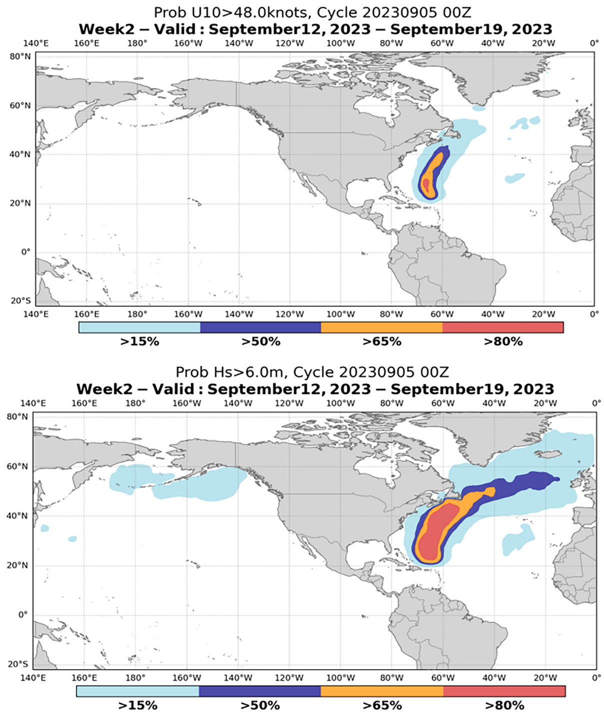

FIG. 8. Probability maps for (top) U10 and (bottom) Hs generated with GEFSv12 forecast cycle 5 Sep 2023. The maps show extreme conditions expected for the week 2 forecast, between 12 and 19 Sep, associated with Hurricane Lee and its transition into an extratropical cyclone. The colors represent the probabilities of U10 greater than 48 kt and Hs greater than 6 m.

atic errors. The authors concluded that the principal advantage of employing MLP-NNs lies not in short-range forecasts but in longer forecast ranges extending beyond 4 days.

A significant step in increasing the domain and training set has been achieved by Campos et al. (2020), also using GWES and satellite data. Expanding the methodology of Campos et al. (2019), based on MLP-NNs, to the whole globe and training with altimeter observations, NN models were constructed using six variables sourced from 21 ensemble members, plus latitude, sin/cos of longitude, sin/cos of time, forecast lead time, and GWES cycle. The NN outputs are the residues of U10 and Hs, the two variables contained in the altimeter dataset, which are then added to the GWES ensemble mean (EM) to compose the final forecast, from nowcast to 10 days. Figure 9 illustrates that the GWES bias could be reduced to values close to zero. Moreover, the RMSE of day-10 forecasts from the NN simulations indicated a gain of 2 days in predictability when compared to the EM.

The hybrid modeling using MLP-NNs presented has not yet been integrated into operational use. Current efforts are focused on expanding bias correction algorithms, considering various methods (Campos et al. 2024b), ranging from simple univariate linear bias correction to more complex multivariate models, including the quantile mapping method, and more sophisticated nonlinear models that consider the entire ensemble forecast, such as long short-term memory (LSTM) networks, gradient boosting (XGBoost), and support vector regression (SVR). These model developments are currently under evaluation. Candidate solutions are expected to be incorporated in future versions of GEFS.

c. *Advanced spatial grids and numerical approaches*

The performance analysis of GEFSv12-Wave and other NOAA global wave modeling systems in operational use, along with the necessity for improvements in spatial and spectral resolution in future updates, has identified limitations of the existing wave model configuration, leading to independent investigations of alternative approaches to address spatial

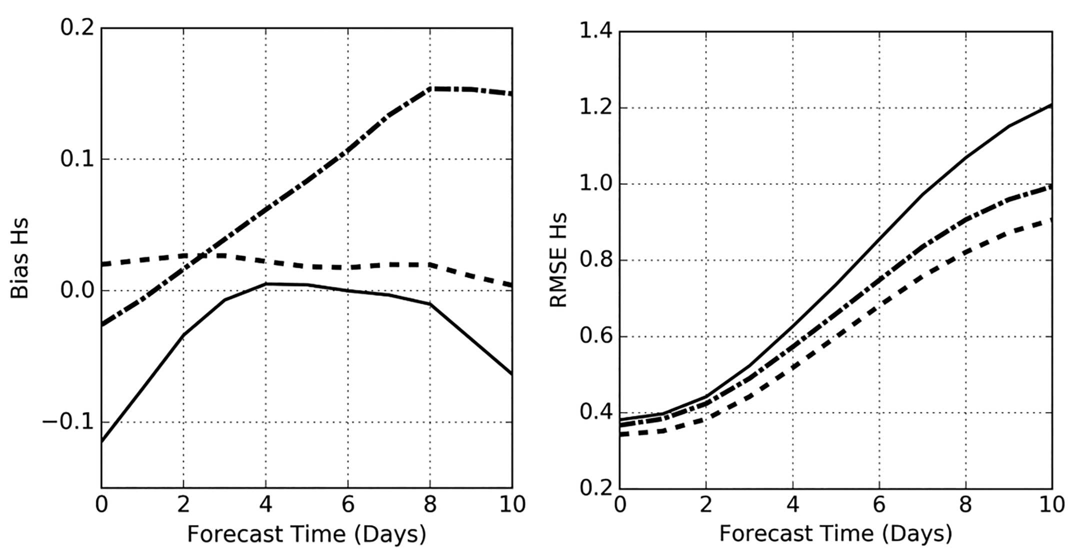

FIG. 9. (left) Bias and (right) RMSE in meters of Hs vs forecast lead time, considering the independent validation with global coverage. Solid curves show the deterministic run, dashed–dotted shows the EM, and dashed curves show the nonlinear NN ensemble averaging.

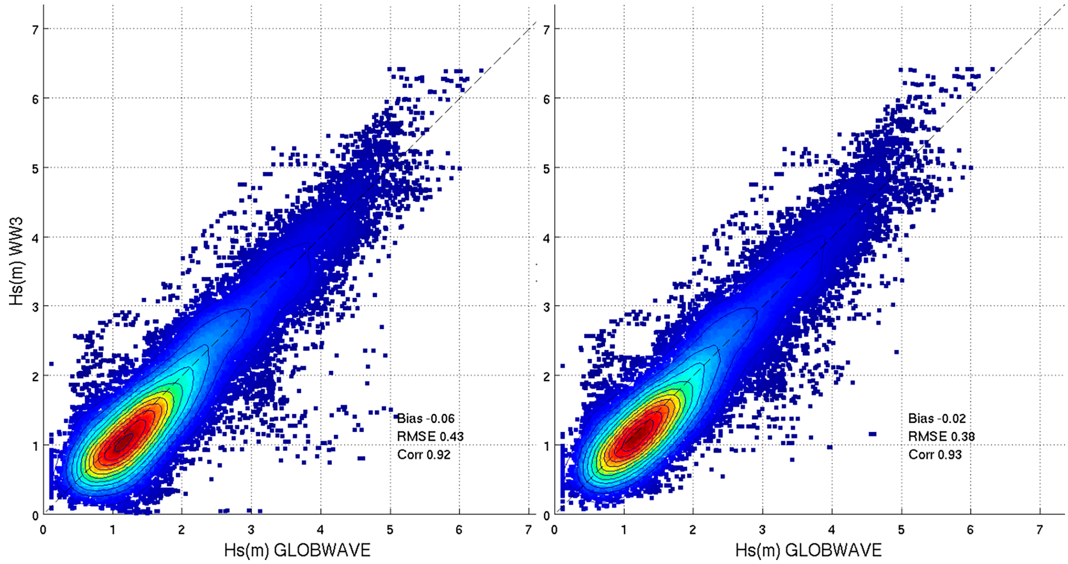

FIG. 10. Scatterplots showing wave heights in the Arctic predicted by Multi-1 (vertical) vs measurements from the GlobWave database made by three altimeters (*Jason-2*, *CryoSat-2*, and *SARAL/Altika*) in September 2014. The diagrams show results for two versions of Multi-1: (left) rectangular grid limited at 82°N and (right) curvilinear grid upgrade with full Arctic coverage for oceanic regions within a 60°–82°N band. Color contours illustrate lower (blue) to highest (red) observation densities.

domain discretization needs, numerical scheme limitations, and opportunities for improving performance. We will depict here results from two lines of development that are being tested for future improvements in the existing systems or useful for scientific applications.

1) CURVILINEAR GRIDS FOR HIGHER LATITUDES

Rectangular latitude–longitude grids in models using explicit numerical schemes face strong limitations in high latitudes. Meridians converge toward the North Pole, leading to increasingly small spacing in physical/geographical space. Therefore, the increasingly smaller spacing requires vanishingly small time steps, resulting in a critical increase in computer resource demands for solving governing equations numerically due to constraints imposed by the Courant–Friedrichs–Lewy (CFL) condition. Therefore, for computational efficiency needed in operational applications, global wave model grids are typically limited to maximum latitudes near 80°N and/or 80°S.

The first successful attempt to deal with that limitation in WW3 was to develop curvilinear grid capabilities (Rogers and Campbell 2009). The technology was used successfully at NOAA in two modeling systems, the Great Lakes Wave (GLW) prediction system and the stand-alone global multigrid deterministic wave model (Multi-1). A description of the curvilinear grid implementation for the GLW system is provided by Alves and Chawla (2015). The implementation of curvilinear grids in Multi-1 is described by J. H. Alves (2017, unpublished manuscript). In view of its unpublished status, the latter will be briefly described here for future reference and to ensure its public availability. Succinct considerations are also made of its impact on more recent upgrades to wave model systems, as they became components in NOAA's coupled global forecast systems.

The Multi-1 system was upgraded in 2017 to include a curvilinear Arctic grid using the polar stereographic projection, with a resolution of approximately 18 km. The enhancement overcame numerical time stepping constraints, allowing for accurate representation of wave dynamics beyond 82°N, the previous northernmost limit of global wave models at NOAA. Comparisons with buoys and altimeter wave-height data for a 3-yr period between 2014 and 2016 showed significant skill improvement. This is illustrated in Fig. 10, where scatterplots of WW3 predictions compared to altimeter measurement from the GlobWave database (Farquhar et al. 2013) show the higher skill obtained with the curvilinear grid upgrade (full Arctic coverage) relative to the previous rectangular grid implementation (limited at 82°N). With the Arctic grid, biases were reduced from 6 to 2 cm, RMSEs were reduced from 43 to 38 cm, and correlation coefficients improved slightly from 0.92 to 0.93.

The successful implementation of a curvilinear grid in Multi-1 ensured that NOAA's global operational systems would be ready to provide wave guidance in the Arctic regions traditionally covered in ice, but that progressively become open ocean as a consequence of climate change. The Arctic domain is defined as the region north of 60°N. As the global wave model transitioned from a stand-alone system to a component in a coupled system, however, the need for sharing computer resources with the atmospheric component as well as limitations in the coupler that preclude using more than one grid resulted in software engineering limitations that are still being overcome.

As a consequence of the limitations above, the coupled wave component in GEFSv12 reverted back to rectilinear grids, with the extent limited to 82°N and 82°S. Investigations are ongoing on strategies to overcome such limitations. One option is to improve the performance of the UFS coupling framework. The other is experimenting with more efficient spatial grid configurations and advanced numerical methods. Although the former is yet to produce conclusive results, the inclusion in WW3 of unstructured grids and implicit numerical schemes has provided a potential solution for overcoming the limitations of rectangular grids in representing the global ocean, as discussed briefly next.

2) IMPACT OF IMPLICIT SCHEMES AND UNSTRUCTURED GRIDS

Traditionally, the prevailing parallelization method within WW3 numerical propagation schemes, referred to as the "card deck," was optimized to be scalable up to the number of spectral bands multiplied by directional bands (NSPEC), effectively setting the model's maximum processor-count capacity to the latter size. Moreover, the use of explicit schemes meant the model's internal time steps were dictated by the CFL condition, which constrained model performance in high latitudes, leading to inefficiencies, as discussed above.

To address scalability issues arising from curvilinear grids, multigrid mosaic option in WW3 (Chawla et al. 2013) served as a workaround currently used in all NOAA operational wave models, enabling semigeographical domain decomposition and facilitating higher resolutions in regions of interest. Simultaneously, for coastal applications, a new parallelization algorithm called "domain decomposition" was introduced in WW3 in association with new unstructured mesh capabilities. Such innovation was aimed at overcoming the card-deck limitations of parallel processing communicator size to NSPEC, by introducing an optional implicit numerical propagation scheme to bypass the CFL constraints, as described by Abdolali et al. (2020a).

The success of this approach for operational wave modeling applications at NOAA was shown in the operational implementation of the GLW unstructured (GLWU) model reported in Alves et al. (2023) and Abdolali et al. (2024). In 2018, the coastal and regional capabilities of unstructured meshes extended the model's coverage to a global scale and were successfully tested (Abdolali et al. 2020b). This enhancement enabled the wave model to run on a single unstructured mesh with variable resolution, controlled by physical distance. The approach met the UFS mesh cap requirement of a single export/import grid, also solving computer resource limitations found when rectilinear or curvilinear grids are used in the UFS coupling framework used in models such as GEFSv12-Wave.

An example of the application of unstructured grids for addressing needs of coupling WW3 to the atmospheric and other components in the UFS is illustrated in Fig. 11. Using a regular grid, the domain has to be masked beyond 82° to prevent the vanishingly small time steps between 82° and 90° (shaded area and dotted lines in Fig. 11), a limitation arising from the use of regular grids—due to the convergence of meridians near the pole—and the associated explicit numerical solver used in WW3 for this grid type. One possible solution would be to use curvilinear grids as previously implemented in NOAA's stand-alone operational global wave models.

As explained in the previous section, using curvilinear grids near the poles would require using a multigrid approach, which is currently not allowed by the existing UFS coupler, which requires a single exchange grid. The multiple grid approach would also need computer resources that are not currently available in an operational setting. To attend the requirement from the UFS coupler for using only one import and one export grid for each model component, upcoming upgrades to NOAA's global coupled system, such as GFSv17 and GEFSv13, are constrained to a single exchange mesh for enabling a two-way coupling between components. Given all existing constraints, an optimal solution for using the current capabilities in WW3 is to use a single global unstructured mesh.

Unstructured grid capabilities in WW3 come equipped with an implicit numerical solver that is not available for regular grids. Therefore, a single unstructured mesh would solve the UFS coupler requirements and also attenuate computer resource limitations, as the implicit solution permits maintaining high resolutions within a single mesh with time steps that are not constrained at high latitudes, as illustrated in Fig. 12. Therefore, near the poles and in scenarios where coastal regions necessitate resolution enhancement, the implicit scheme can be employed without incurring computational penalties, as it is free from the CFL criteria constraint. The solution is currently under investigation for use in coupled global UFS systems being developed at NOAA for future upgrades.

6. Concluding remarks

The successful integration of the wave component into the GEFSv12 operational forecast system marks a significant milestone in NOAA's pursuit of enhanced global forecasting capabilities. Through meticulous development efforts and close collaboration with stakeholders, NOAA has achieved a state-of-the-art wave prediction system that sets a new standard in forecast reliability and extends lead times for critical weather events.

With GEFSv12, the transition to the coupled UFS framework not only enhances forecast skill but also paves the way for future advancements in probabilistic forecasting and product development. As NOAA continues to innovate and refine its forecasting systems, the implementation of a coupled wave component in GEFSv12 also stands as a testament to the agency's commitment to advancing weather prediction on a global scale using a community modeling paradigm emphasizing community collaboration and forecaster priorities.

The development of GEFSv12-Wave, with its increased resolution, extended forecast duration, improved atmospheric coupling, and demonstrated remarkable progress in wave modeling. The optimization of source-term parameters using automated techniques resulted in substantial enhancements in forecast skill, particularly in high-latitude regions prone to hazardous conditions. Leveraging such automated calibration

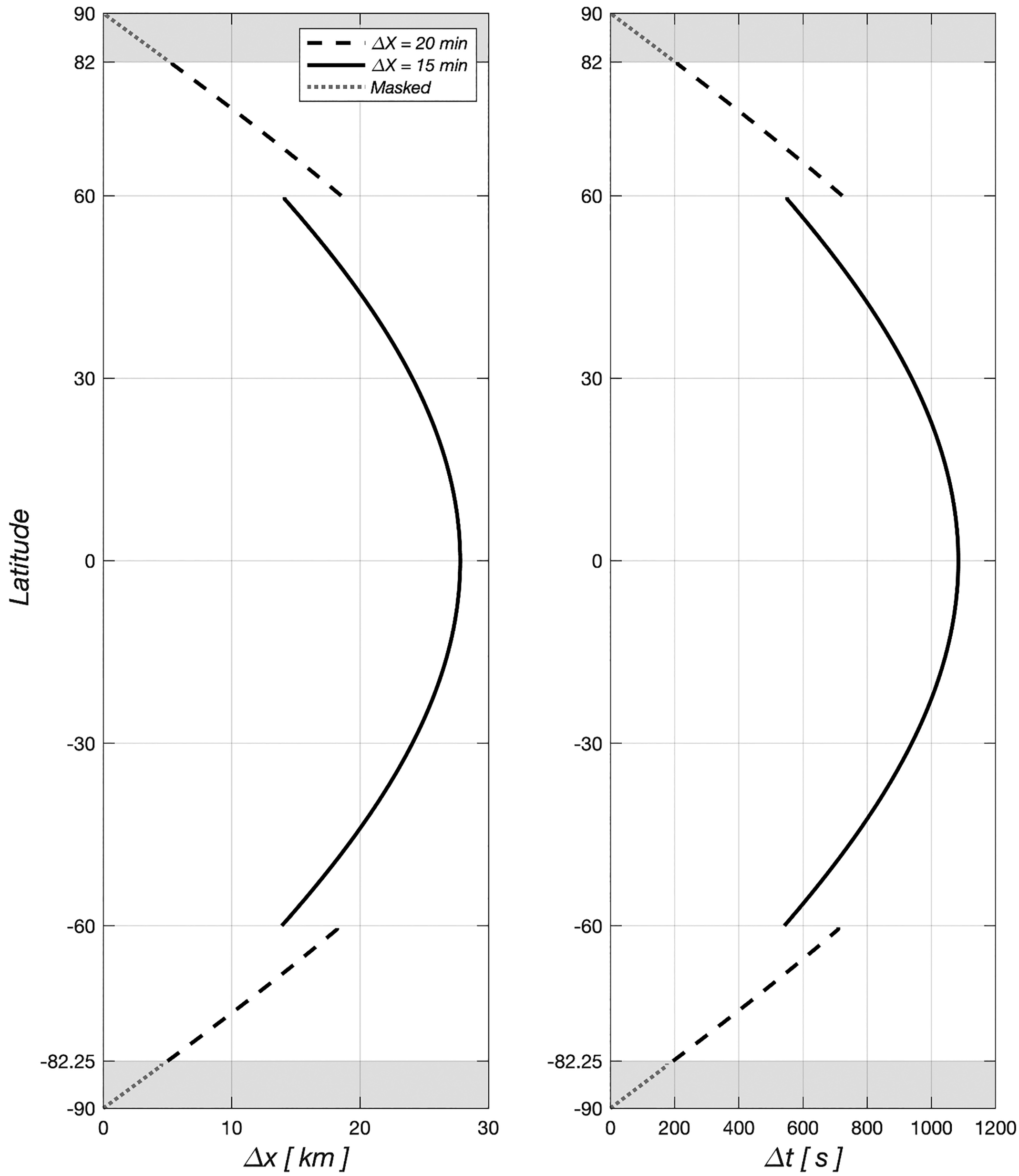

FIG. 11. (left) GEFSv12-Wave multigrid resolution in physical space or distance. (right) Minimum spatial time step required to meet CFL criteria.

techniques has the potential to further enhance the accuracy and reliability of wave predictions, with benefits to mitigating the impacts of extreme weather events on coastal regions and maritime activities.

The scientific evaluation of GEFSv12-Wave consisting of validation against buoy and altimeter data demonstrated improved accuracy and predictability in forecasts of Hs and Tp. The reduction in error and bias, particularly for storm-wave events, emphasizes the enhanced precision and reliability of the GEFSv12-Wave system. In addition, the recent generation of a 20-yr global wave ensemble reforecast dataset based on the GEFSv12-Wave implementation has provided valuable insights into the long-term performance of the system. The reforecast has confirmed the high performance of the wave predictions, especially in the short range, highlighting the success of the optimization efforts undertaken.

The evaluation of the GEFSv12-Wave system highlights successes in NOAA's wave forecasting capabilities, particularly in short-to-mid-range forecasts in the first and second weeks. Beyond 10 days, new probabilistic wave products are being developed on the basis of GEFSv12-Wave data, showing promise in enhancing forecast precision and reliability. In addition, AI/ML applications offer further opportunities for improvement toward mitigating errors for longer forecast

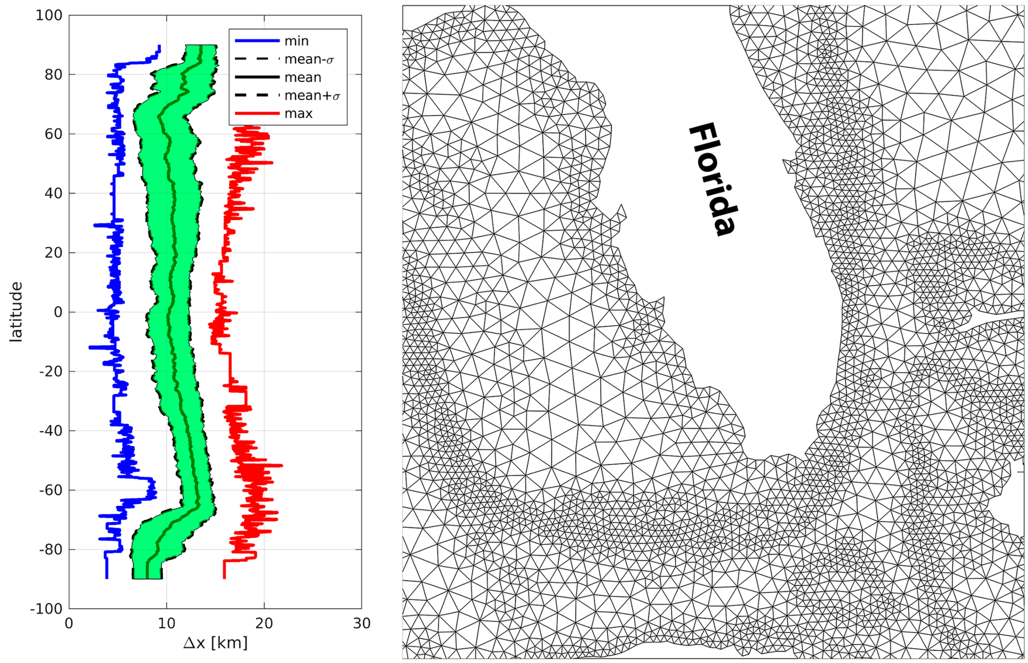

FIG. 12. (left) Distribution of the global mesh resolution as a function of latitude, highlighting a 4-km coastal resolution and a maximum resolution of nearly 20 km offshore. The mesh is a function of topography, bed slope, and distance from coast. (right) A zoom-in view of mesh in the Florida Peninsula.

ranges, and advancements in unstructured grid approaches present possibilities for enhancing spatial and spectral resolution in future updates of NOAA's wave model global coupled system components.

Finally, the development of GEFSv12-Wave as part of a larger effort within NOAA has supported the investigation of new numerical approaches to overcome several challenges, including computer resource limitations for operational applications of coupled systems. For instance, the use of curvilinear and unstructured grids helped overcome restrictions of rectangular grids in high latitudes. Unstructured grid technology furthered the benefits in addressing the needs of a single grid for coupling within the UFS and at the same time minimizing computer resource needs by using implicit numerical solvers.

Altogether, these developments underscore NOAA's dedication to continuously improving forecasting systems to better serve marine and coastal communities while advancing the science of wave prediction.

Acknowledgments. The work described in this paper is part of NOAA's Environmental Modeling Center (EMC) responsibilities for transitioning scientific and technical advances into NOAA operational weather forecasting systems. The authors want to thank EMC support personnel and leadership, in particular, the Global Ensemble Forecasting System (GEFS) team, for their support. We also acknowledge the invaluable support provided by the WAVEWATCH III community, including developers from the U.K. Met Office and the U.S. Navy, as well as the Earth System Modeling Framework (ESMF) and the National Unified Operational Prediction Capability (NUOPC) development teams at NCAR, NOAA, NASA, and the U.S. Navy. We thank Richard Gorman, from the New Zealand National Institute of Water and Atmospheric Research, for his help with the wave model source-term optimization. R. M. Campos is funded by NOAA's Cooperative Institute for Marine and Atmospheric Studies (CIMAS) at the University of Miami Agreement NA20OAR4320472. Finally, we want to thank NOAA's National Data Buoy Center (NDBC), the French Institute for Marine Research (Ifremer), and the Australian Ocean Data Network (AODN/IMOS) for providing publicly available observations utilized in this paper.

Data availability statement. WAVEWATCH III repository: https://github.com/NOAA-EMC/WW3. WW3 Configuration files used in the GEFSv12 implementation: https://github.com/NOAA-EMC/WW3/tree/feb333dd60ff4bd8a7b9d04ed6d6d7b4a00f58b7/regtests/ww3_ufs1.3/input. GEFSv12 20-year wave reforecast data: https://registry.opendata.aws/noaa-wave-ensemble-reforecast/ and https://noaa-nws-gefswaves-reforecast-pds.s3.amazonaws.com/index.html. GEFSv12 20-yr atmospheric reforecast data: https://registry.opendata.aws/noaa-gefs-reforecast/ and https://noaa-gefs-retrospective.s3.amazonaws.com/index.html. Australian Ocean Data Network (AODN) altimeter database: https://help.aodn.org.au/user-guide-introduction/aodn-portal/data-use-acknowledgement/ and http://thredds.aodn.org.au/thredds/catalog/IMOS/SRS/Surface-Waves/Wave-Wind-Altimetry-DM00/catalog.html. Repository containing the scripts used for the validation of the reforecast data: https://github.com/NOAA-EMC/WW3-tools.

REFERENCES

Abdolali, A., A. Roland, A. Van Der Westhuysen, J. Meixner, A. Chawla, T. J. Hesser, J. M. Smith, and M. Dutour Sikiric, 2020a: Large-scale hurricane modeling using domain decomposition parallelization and implicit scheme implemented in WAVEWATCH III Wave Model. *Coastal Eng.*, **157**, 103656, https://doi.org/10.1016/j.coastaleng.2020.103656.

——, W. Pringle, A. Roland, and A. Mehra, 2020b: Assessment of global wave models on unstructured domains. ESS Open Archive, 17 pp., https://doi.org/10.1002/essoar.10505107.1.

——, S. Banihashemi, J. H. Alves, A. Roland, T. J. Hesser, M. Anderson Bryant, and J. McKee Smith, 2024: Great Lakes wave forecast system on high-resolution unstructured meshes. *Geosci. Model Dev.*, **17**, 1023–1039, https://doi.org/10.5194/gmd-17-1023-2024.

Alves, J. H. G. M., and A. Chawla, 2015: Forecasting Wind-Waves at the North American Great Lakes. WGNE 2015 BLUE BOOK, https://www.wcrp-climate.org/WGNE/BlueBook/2015/individual-articles/08_Alves_Jose-Henrique_Chawla_Arun-Great-Lakes-Wave-model.pdf.

——, and Coauthors, 2013: The NCEP–FNMOC combined wave ensemble product: Expanding benefits of interagency probabilistic forecasts to the oceanic environment. *Bull. Amer. Meteor. Soc.*, **94**, 1893–1905, https://doi.org/10.1175/BAMS-D-12-00032.1.

Alves, J.-H. G. M., N. Bernier, A. Chawla, P. Etala, and P. Wittmann, 2015a: Probabilistic wave forecasting and ensemble-based data assimilation at the US National Weather Service. Research activities in atmospheric and oceanic modeling. CAS/JSC Working Group on Numerical Experimentation Rep. 12. WCRP Rep. 12/2015, 2 pp., https://www.wcrp-climate.org/WGNE/BlueBook/2015/individual-articles/08_Alves_Jose-Henrique_etal_wave-ensembles.pdf.

——, S. Stripling, A. Chawla, H. Tolman, and A. van der Westhuysen, 2015b: Operational wave guidance at the U.S. National Weather Service during tropical/post–Tropical Storm Sandy, October 2012. *Mon. Wea. Rev.*, **143**, 1687–1702, https://doi.org/10.1175/MWR-D-14-00143.1.

——, H. Tolman, A. Roland, A. Abdolali, F. Ardhuin, G. Mann, A. Chawla, and J. Smith, 2023: NOAA's Great Lakes wave prediction system: A successful framework for accelerating the transition of innovations to operations. *Bull. Amer. Meteor. Soc.*, **104**, E837–E850, https://doi.org/10.1175/BAMS-D-22-0094.1.

Ardhuin, F., and Coauthors, 2010: Semiempirical dissipation source functions for ocean waves. Part I: Definition, calibration, and validation. *J. Phys. Oceanogr.*, **40**, 1917–1941, https://doi.org/10.1175/2010JPO4324.1.

Battjes, J. A., and J. P. F. M. Janssen, 1978: Energy loss and set-up due to breaking of random waves. *Proc. 16th Int. Conf. on Coastal Engineering*, Hamburg, Germany, Amer. Soc. Civil Engineers, 569–578.

Bernier, N. B., and Coauthors, 2016: Operational wave prediction system at Environment Canada: Going global to improve regional forecast skill. *Wea. Forecasting*, **31**, 353–370, https://doi.org/10.1175/WAF-D-15-0087.1.

Campos, R. M., 2023: Analysis of spatial and temporal criteria for altimeter collocation of significant wave height and wind speed data in deep waters. *Remote Sens.*, **15**, 2203, https://doi.org/10.3390/rs15082203.

——, V. Krasnopolsky, J.-H. Alves, and S. Penny, 2017: Improving NCEP's probabilistic wave height forecasts using neural networks: A pilot study using buoy data. NCEP Office Note 490, 23 pp., https://doi.org/10.7289/V5/ON-NCEP-490.

——, J.-H. G. M. Alves, S. G. Penny, and V. Krasnopolsky, 2018: Assessments of surface winds and waves from NCEP Ensemble Forecast System. *Wea. Forecasting*, **33**, 1533–1546, https://doi.org/10.1175/WAF-D-18-0086.1.

——, V. Krasnopolsky, J.-H. G. M. Alves, and S. G. Penny, 2019: Nonlinear wave ensemble averaging in the Gulf of Mexico using neural networks. *J. Atmos. Oceanic Technol.*, **36**, 113–127, https://doi.org/10.1175/JTECH-D-18-0099.1.

——, J.-H. G. M. Alves, S. G. Penny, and V. Krasnopolsky, 2020: Global assessments of the NCEP Ensemble Forecast System using altimeter data. *Ocean Dyn.*, **70**, 405–419, https://doi.org/10.1007/s10236-019-01329-4.

——, M. Bernardino, M. Gonçalves, and C. Guedes Soares, 2022: Assessment of metocean forecasts for Hurricane Lorenzo in the Azores Archipelago. *Ocean Eng.*, **243**, 110292, https://doi.org/10.1016/j.oceaneng.2021.110292.

——, and Coauthors, 2024a: Development and validation of NOAA's 20-year global wave ensemble reforecast. Wea. Forecasting, https://doi.org/10.1175/WAF-D-24-0043.1, in press.

——, D. Figureskey, and A. Mehra, 2024b: Experiments on machine learning post-processing models applied to probabilistic wave forecasts. *Third Symp. on Community Modeling and Innovation*, Baltimore, MD, Amer. Meteor. Soc., 2.1, https://ams.confex.com/ams/104ANNUAL/meetingapp.cgi/Paper/439932.

Cao, D., H. S. Chen, and H. L. Tolman, 2007: Verification of ocean wave ensemble forecasts at NCEP. *Proc. 10th Int. Workshop on Wave Hindcasting and Forecasting and First Coastal Hazards Symp.*, Oahu, HI, Environment Canada, G1, https://polar.ncep.noaa.gov/mmab/papers/tn261/mmab261.pdf.

Cardone, V. J., and D. B. Ross, 1979: State-of-the-art wave prediction methods and data requirements. *Ocean Wave Climate*, Plenum Press, 61–91.

Cavaleri, L., and P. M. Rizzoli, 1981: Wind wave prediction in shallow water: Theory and applications. *J. Geophys. Res.*, **86**, 10 961–10 973, https://doi.org/10.1029/JC086iC11p10961.

Chawla, A., and Coauthors, 2013: A multigrid wave forecasting model: A new paradigm in operational wave forecasting. *Wea. Forecasting*, **28**, 1057–1078, https://doi.org/10.1175/WAF-D-12-00007.1.

Chen, H. S., 1995: Ocean surface waves. NOAA/NWS Tech. Procedures Bull. 426, 17 pp., https://polar.ncep.noaa.gov/mmab/tpbs/operational.tpbs/tpb494/tpb494.pdf.

——, 2006: Ensemble prediction of ocean waves at NCEP. *Proc. 28th Ocean Engineering Conf.*, Taipei, Taiwan, National Sun Yat-Sen University, 25–37, https://polar.ncep.noaa.gov/waves/ensemble/nsysu0611paper.pdf.

——, L. D. Burroughs, and H. L. Tolman, 1999: Ocean surface waves. NOAA/NWS Tech. Procedures Bull. 453, https://polar.ncep.noaa.gov/mmab/tpbs/obsolete.tpbs/nww3tpb/nww3tpb.html.

Ebert, E. E., 2008: Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. *Meteor. Appl.*, **15**, 51–64, https://doi.org/10.1002/met.25.

Farquhar, C. R., and Coauthors, 2013: GlobWave: A global wave data portal. 8 pp., https://www-cdn.eumetsat.int/files/2020-04/pdf_conf_p61_s6_02_deighton_v.pdf.

Feit, D. M., 1986: Compendium of marine meteorological and oceanographic products of the Ocean Products Center. NOAA Technical Memorandum NWS/NMC 68, 93 pp., https://repository.library.noaa.gov/view/noaa/13495/noaa_13495_DS1.pdf.

Gorman, R. M., and H. J. Oliver, 2018: Automated model optimisation using the Cylc workflow engine (Cyclops v1.0). *Geosci. Model Dev.*, **11**, 2153–2173, https://doi.org/10.5194/gmd-11-2153-2018.

Hamill, T. M., and Coauthors, 2022: The reanalysis for the Global Ensemble Forecast System, version 12. *Mon. Wea. Rev.*, **150**, 59–79, https://doi.org/10.1175/MWR-D-21-0023.1.

Hasselmann, S., and K. Hasselmann, 1985: Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part I: A new method for efficient computations of the exact nonlinear transfer integral. *J. Phys. Oceanogr.*, **15**, 1369–1377, https://doi.org/10.1175/1520-0485(1985)015<1369:CAPOTN>2.0.CO;2.

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea. Forecasting*, **15**, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Hubert, W. E., 1957: A preliminary report on numerical sea conditions forecast. *Mon. Wea. Rev.*, **85**, 200–204, https://doi.org/10.1175/1520-0493(1957)085<0200:APRONS>2.0.CO;2.

Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, 1994: *Dynamics and Modeling of Ocean Waves*. Cambridge University Press, 532 pp.

Leonard, B. P., 1991: The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Comput. Methods Appl. Mech. Eng.*, **88**, 17–74, https://doi.org/10.1016/0045-7825(91)90232-U.

Mentaschi, L., G. Besio, F. Cassola, and A. Mazzino, 2013: Problems in RMSE-based wave model validations. *Ocean Modell.*, **72**, 53–58, https://doi.org/10.1016/j.ocemod.2013.08.003.

Powell, M. J. D., 2009: The BOBYQA algorithm for bound constrained optimization without derivatives. University of Cambridge Rep. DAMTP 2009/NA06, 39 pp.

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

Rogers, W. E., and T. J. Campbell, 2009: Implementation of curvilinear coordinate system in the WAVEWATCH-III model. NRL Memo. Rep. NRL/MR/7320-09-9193, 48 pp., https://apps.dtic.mil/sti/pdfs/ADA507120.pdf.

Sela, J. G., 1980: Spectral modeling at the National Meteorological Center. *Mon. Wea. Rev.*, **108**, 1279–1292, https://doi.org/10.1175/1520-0493(1980)108<1279:SMATNM>2.0.CO;2.

Theurich, G., and Coauthors, 2016: The Earth System Prediction Suite: Toward a coordinated U.S. modeling capability. *Bull. Amer. Meteor. Soc.*, **97**, 1229–1247, https://doi.org/10.1175/BAMS-D-14-00164.1.

Tolman, H. L., and D. Chalikov, 1996: Source terms in a third-generation wind wave model. *J. Phys. Oceanogr.*, **26**, 2497–2518, https://doi.org/10.1175/1520-0485(1996)026<2497:STIATG>2.0.CO;2.

——, B. Balasubramaniyan, L. D. Burroughs, D. V. Chalikov, Y. Y. Chao, H. S. Chen, and V. M. Gerald, 2002: Development and implementation of wind-generated ocean surface wave models at NCEP. *Wea. Forecasting*, **17**, 311–333, https://doi.org/10.1175/1520-0434(2002)017<0311:DAIOWG>2.0.CO;2.

WAMDI Group, 1988: The WAM model—A third generation ocean wave prediction model. *J. Phys. Oceanogr.*, **18**, 1775–1810, https://doi.org/10.1175/1520-0485(1988)018<1775:TWMTGO>2.0.CO;2.

WW3DG, 2019: User manual and system documentation of WAVEWATCH III version 6.07. Tech. Note 333, NOAA/NWS/NCEP/MMAB, 326 pp., https://raw.githubusercontent.com/wiki/NOAA-EMC/WW3/files/manual.pdf.

Zhou, X., and Coauthors, 2022: The development of the NCEP Global Ensemble Forecast System version 12. *Wea. Forecasting*, **37**, 1069–1084, https://doi.org/10.1175/WAF-D-21-0112.1.

Zhu, Y., 2005: Ensemble forecast: A new approach to uncertainty and predictability. *Adv. Atmos. Sci.*, **22**, 781–788, https://doi.org/10.1007/BF02918678.