

Impact of SST and Surface Waves on Hurricane Florence (2018): A Coupled

Modeling Investigation

Joseph B. Zambon* and Ruoying He

*Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University,
Raleigh, North Carolina.*

John C. Warner, Christie A. Hegermiller

United States Geological Survey, Woods Hole, Massachusetts

*Corresponding author: Joseph B. Zambon, jbzambon@ncsu.edu

Early Online Release: This preliminary version has been accepted for publication in *Weather and Forecasting*, may be fully cited, and has been assigned DOI 10.1175/WAF-D-20-0171.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT

Hurricane Florence (2018) devastated the coastal communities of the Carolinas through heavy rainfall that resulted in massive flooding. Florence was characterized by an abrupt reduction in intensity (Saffir-Simpson Category 4 to Category 1) just prior to landfall and synoptic-scale interactions that stalled the storm over the Carolinas for several days. We conducted a series of numerical modeling experiments in coupled and uncoupled configurations to examine the impact of sea surface temperature (SST) and ocean waves on storm characteristics. In addition to experiments using a fully coupled atmosphere-ocean-wave model, we introduced the capability of the atmospheric model to modulate wind stress and surface fluxes by ocean waves through data from an uncoupled wave model. We examined these experiments by comparing track, intensity, strength, SST, storm structure, wave height, surface roughness, heat fluxes, and precipitation in order to determine the impacts of resolving ocean conditions with varying degrees of coupling. We found differences in the storm's intensity and strength, with the best correlation coefficient of intensity ($r=0.89$) and strength ($r=0.95$) coming from the fully-coupled simulations. Further analysis into surface roughness parameterizations added to the atmospheric model revealed differences in the spatial distribution and magnitude of the largest roughness lengths. Adding ocean and wave features to the model further modified the fluxes due to more realistic cooling beneath the storm which in turn modified the precipitation field. Our experiments highlight significant differences in how air-sea processes impact hurricane modeling. The storm characteristics of track, intensity, strength, and precipitation at landfall are crucial to predictability and forecasting of future landfalling hurricanes.

Significance statement. Hurricane Florence (2018) was a major hurricane that weakened and eventually remained stationary after landfall over the coastal Carolinas, leading to devastating flooding. Atmospheric-only numerical models neglect the impact of Sea Surface Temperature (SST) and surface waves changing beneath these storms and tend to over-predict the intensity in some cases. We conducted experiments that include atmosphere-ocean interaction by tying in realistic coupled SST and surface waves beneath the storm. We employed a novel approach of including surface wave information to the atmospheric model. While examining the underlying features, we found different approaches resulted in drastic differences in the result, including a 27.2% difference in precipitation among our experiments. We found improvement in the numerical models with more advanced coupling to the ocean environment but further improvement could be achieved through data assimilation.

1. Introduction

Tropical cyclones (TCs) are large-scale, discrete events that can have drastic impacts on coastal communities. TCs have complex interactions with the ocean environment before, during, and after landfall. Within the United States (U.S.), where approximately 40% of the population lives near a coast, TCs are one of the costliest natural disasters and account for significant damage to infrastructure, injury, and loss of life (Emanuel 2005). As people continue to migrate to the coast, damage caused by these intense storms will continue to increase, due to more intense and more frequent events (Emanuel et al. 2004; Oey et al. 2007). In this work, we examine Hurricane Florence (2018), which led to \$24 Billion in damage (Stewart and Berg 2019), through several experiments using numerical models in increasing complexity of uncoupled and coupled states resolving atmosphere, ocean, and wave conditions.

Hurricane Florence was a long-track, Cape Verde TC that persisted in the Atlantic Ocean for 2.5 weeks before making landfall on the southeastern coast of North Carolina on 14 September 2018 at 1115 UTC. The storm reached peak intensity as a Category 4 hurricane just two days prior, but underwent significant weakening before landfall. This weakening was attributed to two factors: an eyewall replacement cycle, and a reduction in the depth of the ocean mixed layer as the storm reached the eastern U.S. Atlantic Shelf (Stewart and Berg 2019). While the sea surface temperature (SST) across the shelf remained around 29.5°C, the depth of the mixed layer decreased. Ocean heat content (OHC; Leipper and Volgenau 1972) decreased by more than half from the open ocean to the shelf, from 56 kJ cm⁻² to 25 kJ cm⁻² (Stewart and Berg 2019). The weakening of the hurricane due to the reduction of ocean heat content and interaction with larger-scale atmospheric circulation resulted in the storm remaining over the Carolinas for several days (Stewart and Berg 2019).

As the storm motion reduced to near-stationary, heavy precipitation persisted, resulting in massive flooding. Much of the Carolinas received over 50 cm of rainfall over a few days, with locally as much as 91 cm of rainfall measured outside of Elizabethtown, NC (Stewart and Berg 2019). Precipitation led to runoff into streams, tributaries, and rivers within the Cape Fear, Neuse, White Oak, Lumber, and Tar-Pamlico watersheds. Widespread reports of millions of gallons of raw sewage and other contaminants were reported to have been leaked into these watersheds, including 39 million gallons of sewage into the Cape Fear River basin (Henderson 2018).

One result of this extreme precipitation and runoff was a phenomenon known as compound flooding (Wahl et al. 2015; Pietrafesa et al. 2019), during which storm surge from the hurricane making landfall to the southeast combined with the Neuse River runoff flooding from the northwest. This put the city of New Bern, NC under a torrent of water (Pietrafesa et al. 2019; Stewart and Berg 2019) that resulted in over \$100M of damage (Bennett 2018).

Several studies discuss the importance of atmosphere-ocean-wave coupling on hurricanes (Bao et al. 2000; Bender and Ginis 2000; Khain et al. 2000; Bender et al. 2007; Liu et al. 2008; Yablonsky and Ginis 2009; Liu et al. 2010; Warner et al. 2010; Xie et al. 2010; Olabarrieta et al. 2011; Lee and Chen 2012; Liu et al. 2012a,b; Chen et al. 2013; Wada et al. 2013; Lee and Chen 2014; Zambon et al. 2014a,b; Prakash et al. 2019; Kumar et al. 2020). Atmosphere-ocean boundary dynamics are modulated by ocean roughness, which, during extreme forcing such as hurricanes, can be a function of ocean waves, rainfall, and spume and spray. Though WRF includes the expected upper limit of the drag coefficient under hurricane conditions (e.g., Davis et al., 2008), it uses the Charnock (1955) relation to describe variation of roughness with wind speed, and thus sea state. More complex parameterizations for ocean surface roughness have been introduced since to explicitly include factors such as wave steepness, age, or direction relative to the wind, but have not been implemented in WRF.

This study investigates, through increasing complexity of model inputs and coupling, the impact of resolving ocean surface conditions, such as SST and waves, on hurricane forecasting. We employ a novel approach in the atmospheric model by resolving the surface wave conditions in hindcast through use of ingested wave fields to calculate the surface roughness beneath the storm over the open ocean. Hurricane Florence exhibited several instances where ocean coupling was found to be of crucial importance in the predictability of the storm and hence is used as our test case. Using Hurricane Florence as our test case, we aim to demonstrate the importance of resolving the ocean state and its impact on storm characteristics, from track and intensity, through landfall and associated precipitation produced throughout.

2. Methodology

a. Model Configuration

We conducted a control and series of numerical modeling experiments of Hurricane Florence atmosphere, ocean, and wave conditions to examine the impacts of ocean surface conditions, such as SST and surface waves, on storm characteristics (Fig. 1, Table 1). The Weather Research and Forecasting (WRF) model version 4.1.3 (Skamarock et al. 2019) was used to simulate the atmosphere. In the uncoupled experiments, SST and surface wave conditions were supplied beneath the storm by analysed SST and hindcast wave fields. In the coupled experiments, we used the Coupled Ocean-Atmosphere-Wave-Sediment Transport (COAWST; Warner et al. 2010) model to simulate the ocean and wave states as well. COAWST exchanges relevant variables between WRF and simultaneously running ocean (Regional Ocean Modeling System, ROMS; Shchepetkin and McWilliams 2005) and wave (Simulating Waves Nearshore, SWAN; Booij et al. 1999) models, allowing real-time environmental coupling at user-configurable intervals. COAWST has been successfully used to model a number of different significant weather events, including strong hurricanes (Warner et al. 2010; Zambon et al. 2014a; Kumar et al. 2020), hurricanes undergoing extratropical transition (Olabarrieta et al. 2012; Zambon et al. 2014b), and extra-tropical winter cyclones (Nelson and He 2012; Nelson et al. 2014).

1) UNCOUPLED MODEL EXPERIMENTS

The control run used WRF, and the uncoupled lower-boundary SST condition was supplied by the Real-Time Global SST (RTG SST) analysis product (Gemmill et al. 2007). RTG SST is a globally gridded product with a 0.083° by 0.083° spatial grid and daily temporal resolution. We

input this data into the WRF model through a set of routines included in the WRF Preprocessing System publicly available source code distribution.

We conducted a temporal ensemble of experiments initialized every 6 hours from 0000 UTC 11 September 2018 to 0000 UTC 13 September 2018. The 9-member temporal ensemble, with a control and 5 additional experiments at each initialization point resulted in 56 total simulations. For the bulk of the study, we used the initialization with the furthest lead time as the incremental improvement from closer to landfall was minimal. Those experiments were simulated over 10 days of Hurricane Florence's impact on the southeastern U.S. from 0000 UTC 11 September 2018 to 0000 UTC 21 September 2018 (hereafter referred to as the 11Sept00Z runs).

The WRF domain (Fig. 2) was selected to enclose the ocean and wave domains at 9 km grid spacing (490 x 459 grid points) with a 48 s timestep. A smaller, vortex-following grid was used to resolve the complex small-scale features of the hurricane at 3 km grid spacing (151 x 151 grid points) with a 16 s timestep. Both grids have 48 vertical layers following hybrid-sigma coordinates (Skamarock et al. 2013). All WRF experiments use the following model physics: WSM 6-class microphysics from Hong and Lim (2006), Grell and Dévényi (2002) Cumulus parameterization on the 9 km domain only, short- and long-wave physics from the Rapid Radiative Transfer Model for Global (RRTMG) climate models (Iacono et al. 2008), planetary boundary layer scheme from the Mellor–Yamada Nakanishi Niino (MYNN) level 3 turbulent kinetic energy scheme (Nakanishi and Niino 2006, 2009; Olson et al. 2019), MYNN surface layer physics (Nakanishi and Niino 2006, 2009; Olson et al. 2019), and land surface physics from the Unified Noah Land Surface model (Tewari et al. 2004). The atmospheric model draws its initial and boundary conditions from the National Center for Environmental Prediction Global Data Assimilation System (NCEP GDAS; National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce 2018), with 0.25° grid spacing and updated every 3 hours.

To conduct uncoupled experiments with wave data, we implemented a number of changes to the WRF surface physics parameterizations. To modify sea surface roughness with modeled wave data, we introduced three wave roughness parameterizations from Taylor and Yelland (2001) derived from the Coupled Ocean Atmosphere Response Experiment (COARE), Drennan et al. (2005), and Oost et al. (2002) to the WRF surface physics schemes.

The default WRF configuration uses the Charnock (1955) relation to compute sea surface roughness length (z_0) as,

$$z_0 = z_{ch} \frac{u_*^2}{g},$$

where u_* is the friction velocity, g the acceleration due to gravity, and z_{ch} the Charnock parameter, an empirical nondimensional parameter that is either a constant or, here, dependent on the wind speed as,

$$z_{ch} = 0.011 + 0.007 \frac{U_{10} - 10}{8},$$

where U_{10} is the wind speed at 10 m above the ocean surface.

The parameterization from Taylor and Yelland (2001) computes sea surface roughness as a function of the wave steepness as,

$$z_0 = 1200 \frac{H_S^2}{L_W^{4.5}},$$

where H_S is the significant wave height and L_W is the wave length.

The parameterization from Drennan et al. (2005) computes sea surface roughness as a function of the wave age as,

$$z_0 = 3.35 H_S \left(\frac{u_*}{C_W} \right)^{3.4},$$

where C_W is wave celerity, expressed as,

$$C_W = \frac{L_W}{T_P},$$

where T_P is peak wave period.

Lastly, the parameterization from Oost et al. (2002) also computes sea surface roughness as a function of the wave age as,

$$z_o = 25 \frac{T_P}{\pi} \left(\frac{u_*}{C_W} \right)^{4.5}.$$

For the uncoupled lower-boundary wave condition, we used WAVEWATCHIII (Tolman 1991; Tolman et al. 2002; Tolman 2014) model hindcast solutions, which have global coverage at 0.5° grid spacing, available every 3 hours. Higher resolution data is available for some regional domains (e.g. 0.067° for the Gulf of Mexico and Northwest Atlantic); however, because the higher-resolution spatial data did not cover our entire domain of interest, we defaulted to the lower-resolution global model output.

WAVEWATCHIII hindcast data include significant wave height (H_S), peak wave period (T_P), wave direction, and bathymetry. Though wave direction is not currently used, it is input into the WRF registry and surface physics scheme for future implementation of direction-dependent sea surface roughness parameterizations (e.g., Porchetta et al. 2019). For the parameterizations dependent on wave age, we iteratively solved for wave length (L_W) using the dispersion relation.

Davis et al. (2008) found that using an alternate drag formulation based on high-wind wind-tunnel studies from Donelan et al. (2004) gave upper and lower limits of z_0 to be 0.125×10^{-6} and 2.85×10^{-3} , respectively. These limits are employed in all experiments. Further enhancement to the ocean fluxes can be found through inclusion of a sea spray parameterization as found in Prakash et al. (2019) and Liu et al. (2012a), but this is left to future investigation.

An investigation of Hurricane Ida (2009) by Olabarrieta et al. (2012) determined that the impact on TC intensity and strength varied considerably between these sea surface roughness parameterizations as the storm underwent extra-tropical transition to become a Nor'easter. Similarly, we analyzed the intensity and strength prediction of the uncoupled model simulations for Hurricane Florence to determine which parameterization to apply to the coupled experiments.

2) COUPLED MODEL EXPERIMENTS

The COAWST modeling system utilizes the Model Coupling Toolkit (MCT; Jacob et al. 2005; Larson et al. 2005; Warner et al. 2008, 2010) to exchange variables between the coupled models at user-defined intervals. We conducted experiments using the COAWST framework first with SST feedback from the ROMS ocean model, and then fully coupled to both ocean and wave models. The MCT utilizes the Spherical Coordinate Remapping Interpolation Package (SCRIP; Jones 1998) to develop the interpolation weights that are used to pass variables from model to model.

In our coupled experiments, SCRIP was used to convert between the WRF and ROMS or the WRF and co-located ROMS and SWAN domains. The nested WRF moving domain was interpolated to the ROMS and SWAN grids (and vice versa) by calculating the interpolation weights for every possible position of the moving domain and the weights are changed as the inner moving nest tracks along with the storm. MCT then determines the position and interpolation weights between the nested, moving WRF domain and the ROMS/SWAN grids.

The Regional Ocean Modeling System (ROMS; Shchepetkin and McWilliams 2005; Haidvogel et al. 2008) rev. 948 was used as the ocean model in the coupled experiments. ROMS is a free-surface, terrain-following numerical model that solves the three-dimensional Reynolds-averaged Navier-Stokes equations using hydrostatic and Boussinesq approximations. ROMS is fully parallelized and can be run using multiple advection schemes, turbulence models, lateral boundary conditions, and surface and bottom boundary layer schemes.

The ocean domain is slightly inside of the WRF domain (Fig. 2). This allows the ocean and wave models to have their surface conditions resolved everywhere by the regional WRF model at a higher resolution than global models. Neither the Pacific Ocean southwest of Mexico nor the Great Lakes are resolved by ROMS. There, SST is provided by RTG SST. This is demonstrated in

Fig. 2 as the Atlantic Ocean, Gulf of Mexico, and Caribbean Sea features fine-scale SST features which are absent in the southwest corner of the atmospheric domain. The SST in the extreme eastern edge of the WRF domain is also not resolved by ROMS (represented in Fig. 2 by the area to the east of the black line running near 60°W). This is to ensure a buffer between the edge of the ROMS/SWAN grids and the WRF grids to ensure the former is completely covered.

The ROMS grid is spaced approximately 7 km (402 x 482 grid points), with 36 vertical levels in a stretched, terrain-following configuration that allows for a finer vertical grid near the ocean surface. The stretched, terrain-following grid allows vertical points in the upper-ocean to be packed closer together in the coastal environment than in the deep ocean. A 120 s baroclinic time step was used. The eastern boundary is the only open boundary, along which we utilized the HYbrid Coordinate Ocean Model with Navy Coupled Ocean Data Assimilation (HYCOM/NCODA; Cummings 2005; Chassignet et al. 2007; Cummings and Smedstad 2013) version 3.1 for boundary conditions, updated daily.

We followed the scheme of Marchesiello et al. (2001), whereby Orlanski (1976) radiation conditions were used in conjunction with relaxation (with timescale of 0.5 day on inflow and 10 days on outflow) to pass HYCOM/NCODA tracer (salinity and temperature) and 3-d velocity fields to ROMS. For the free surface and depth-averaged velocity boundary conditions, we adopted the method of Flather (1976) with external values defined by HYCOM/NCODA, updated daily. We used the Generic Length Scale (GLS) turbulence closure model (Umlauf and Burchard 2003; Warner et al. 2005) to compute vertical turbulent mixing, and the quadratic drag formulation for bottom friction.

The Simulating WAVes Nearshore (SWAN; Booij et al. 1999) surface wave model version 40.80 was used to model ocean waves in the coupled experiments. SWAN is a spectral wave model that

solves for the evolution of wave action, is fully parallelized, and includes wind wave generation, propagation, and dissipation by whitecapping, wave breaking, and bottom friction in coastal waters.

As described above, the SWAN grid is co-located with the ROMS domain. Spectral space was discretized with 36 directional bins and 24 frequencies logarithmically spaced between 1 and 25 s. Nonlinear quadruplet wave-wave interactions were activated in the model to solve for energy redistribution with the spectrum. Wave dissipation due to bottom friction was parameterized using the Madsen et al. (1988) formulation, with an equivalent roughness length scale of 0.05 m. The depth-induced breaking constant was set to 0.73. Wind-wave growth was generated using the Komen and Hasselmann (1984) formulation. A backward-in-space, backward-in-time advection scheme was used for iteration. The surface wave condition was initialized from a flat sea surface (i.e. a cold start). Boundary conditions were derived from the global WAVEWATCHIII surface wave model, updated every 3 hours.

Fields among the coupled models were exchanged every 20 minutes. Several of our previous experiments into Hurricanes Isabel (2003; Warner et al. 2010), Ivan (2004; Zambon et al. 2014a), and Sandy (2012; Zambon et al. 2014b) show that this coupling interval is sufficient to resolve the temporal changes in exchanged variables throughout the coupling system.

A summary of the experiments, SST, wave model, surface roughness parameterization, and required wave components are found in Table 1. Note that we shorten WAVEWATCHIII to WW3 and the COARE-derived (Taylor and Yelland 2001) to CTY in our experiment descriptors and figures. Fig. 1 shows the fields that are exchanged in the coupled model configurations. In the WRF-ROMS-WW3 case, SST was exchanged from ROMS to WRF to calculate surface heat fluxes, which were then exchanged back to ROMS to modify ocean temperature. In addition, WRF provided surface stress (τ) to ROMS, modulated by surface wave roughness calculated using WAVEWATCHIII wave parameters as described above. In the fully coupled, WRF-ROMS-SWAN

experiment, SWAN provided significant wave height, wave length, and peak wave period to WRF to compute sea surface roughness and surface stress. 10 m wind from WRF were used to drive wind-wave generation in SWAN. Although different surface roughness parameterizations will use different variables as detailed above. Wave parameters were passed from SWAN and ROMS to exchange momentum between waves and ocean currents, which were returned to SWAN for wave-current interaction. The vortex force formalism (Kumar et al. 2012) was used to include the wave effect on currents, with depth-limited wave breaking dissipation from SWAN. Sea surface height was passed from ROMS to SWAN to modify depth-limited breaking and bottom friction.

3. Results and Discussion

1) HURRICANE CHARACTERISTICS

To determine TC track, we derived modeled surface pressure to sea level pressure (SLP) within the vortex-following nest. We then record the location of the lowest SLP (track), the value of the lowest SLP (intensity), and the value of the highest 10 m wind speed (strength). This simple derivation of track works well for an intense storm and/or a storm tracked by the WRF vortex tracker, both of which are the case with our simulations. However, this simple approximation begins to break down as the storm weakens and the SLP gradients between the storm and surrounding environment are reduced. For simulations that do not allow for this approximation, we suggest using the GFDL Vortex Tracker (Biswas et al. 2018) for analysis. Verification is provided by the National Hurricane Center (NHC) best track from Stewart and Berg (2019).

Variations in Hurricane Florence's track between the coupling schemes (Fig. 2, also Fig. 7) are relatively minor (typically within 30km), indicating that the hurricane track is largely dependent on large-scale atmospheric circulation processes and less influenced by ocean-atmosphere interaction on the time and spatial scales of the models (Moon et al. 2007; Wada et al. 2013, 2014; Zambon

et al. 2014a,b). Compared to validation, all of the modeled storm tracks are shifted slightly (approximately 60km) to the south just prior to landfall (Fig. 2, also Fig. 7). Their landfall locations are all within a few tens of kilometers of each other and validation (Fig. 2, also Fig. 7). All of the simulations predicted landfall to occur approximately 12 hours earlier than observed.

The first 96 hours are used in the analysis because after this point, it becomes difficult to discern the exact storm position as it weakens over the Carolinas. In addition, the storm spreads out rapidly after landfall, impacting larger areas of the Carolinas. Therefore, the exact storm center is not significant for our post-landfall analysis. We calculated the Root Mean Squared Error (RMSE) for the first 96 hours of the simulation from storm track at 6-hourly positions provided by NHC Best Track (Stewart and Berg 2019). The first column of Table 2 shows the RMSE track results from the 11Sept00Z runs. We found that the control performed the best, but all simulations performed very similarly (i.e. within 22km of each other; Table 2). The first column of Table 3 shows the RMSE track results from the temporal ensemble. To simplify further analysis, we will not dive into individual members of the 56-member temporal ensemble experiments but rather their average. The temporal ensemble had lower track RMSE values across the board, as expected with shorter lead times improving track error before landfall. As in 11Sept00Z runs, the Control had the best track RMSE, but all were within approximately 10 km of each other.

Fig. 3 demonstrates the simulated storm intensity (hPa) and strength (m s^{-1}) of the 11Sept00Z Control and experiments from model initialization through landfall and storm dissipation (approximately at F144, or 0000 UTC 17 September 2018). As shown from the NHC Best Track (Stewart and Berg 2019), after reaching a peak intensity (strength) of 937 hPa (67 m s^{-1}) at F018 (1800 UTC 11 September 2018), the storm began to weaken, reducing in intensity by 20 hPa and in strength by 25.7 m s^{-1} .

In the 11Sept00Z Control and all of the experiments, the initial analysis of the storm strength and intensity suffered. The discrepancy between the intensity and strength of the TC at initialization is due to the assimilation of surface winds without surface pressure used in NCEP GDAS (National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce 2018). Previous versions of GDAS utilized a bogus vortex to assimilate both winds and surface pressure, however this has been removed from the analysis as of July 2017. The NCDP GDAS development document can be found online (https://www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html). The best approach around this would be to utilize data assimilation (Pu and Braun 2001) for both surface winds and pressure. Surface winds can be assimilated from satellite observations (as in GDAS), however surface pressure data is too sparse at this location to be assimilated. As this is a model-coupling exercise and because we initialize our storm far enough in advance (3 days before landfall), the authors believe this to be sufficient to avoid having to employ data assimilation techniques for initialization.

There was a rapid intensification and strengthening after initialization (Fig. 3), owing to the weak initialization. As discussed with regard to the Track analysis above, landfall (and associated reduction in intensity and strength) was approximately 12 hours early. In the Control, and all of the experiments, the storm rapidly intensified from the bogus vortex initialization and maintained roughly the same intensity and strength until landfall with minor fluctuations, as was observed. One exception being, the second small peak of intensity (15hPa) and strength (8 m s^{-1}) from F012 to F030 was missed by all of the runs.

The coupled experiments were slightly weaker than the uncoupled (Control, WRF-WW3) experiments; however, their intensity trend better matched observation. Correlation coefficient (r) for intensity was best among the coupled models in the 11Sept00Z runs (Table 2) as well as in

the temporal ensemble (Table 3). Although we stated for simplicity that we would not delve into individual temporal ensemble member results, it is worth noting that of the 9 temporal ensemble series of experiments, the WRF-ROMS-SWAN case had the best intensity correlation coefficient in 8 of them. The lone outlier belongs to the other ocean-coupled case (WRF-ROMS-WW3). Similarly, correlation coefficients compared to observed strength also demonstrated the most skill in the WRF-ROMS-SWAN coupled case in the temporal ensemble (Table 3) and was within a statistically negligible amount (0.0002) of the best in the 11Sept00Z runs (Table 2).

Owing in part to the weak initialization that the models had to recover from, the best RMSE of intensity and strength were found in uncoupled cases (11Sept00Z runs in Table 2, and temporal ensemble in Table 3). These uncoupled cases had the highest SSTs and allowed the storm to strengthen more relative to the coupled cases. We wanted to use the model performance metrics of Stow et al. (2009) to further examine our results. These metrics include Reliability Index (RI; Leggett and Williams 1981; perfect score is 1), Average Error (AE; perfect score is 0), Absolute Average Error (AAE; perfect score is 0), and Modelling Efficiency (MEF; Nash and Sutcliffe 1970; Loague and Green 1991; perfect score is 1). Not surprisingly, due to the weak initialization and warmer SSTs described above, in both the 11Sept00Z runs (Table 4) and the temporal ensemble (Table 5), the uncoupled models performed best. Our novel approach of introducing wave fields to the WRF model improved the Stow et al. (2009) model metrics in the temporal ensemble in every category except Strength AAE (where the Control performed best) and Strength MEF (where WRF-ROMS-WW3 performed best).

In contrast to Olabarrieta et al. (2012), we found minimal difference among wave roughness parameterizations. This is likely due to the storm track of Hurricane Florence becoming stationary over land whereas Hurricane Ida stalled out over the water of the Mid-Atlantic Bight. To simplify

the remainder of the study, we will consider only the 11Sept00Z runs and we will use only the Oost et al. (2002) scheme in the WRF-ROMS-WW3 and WRF-ROMS-SWAN coupled experiments.

Next, we looked at TC characteristics to compare the speed and location of the RMW. Fig. 4 shows a radial-height cross section of azimuthally-averaged tangential winds (shaded and contoured) extending from the storm center to a radius of 300km around the time of maximum strength (approximately F072). We found among all of the cases that the RMW extended to approximately 50-60km from the TC center. The 10 m wind strength (Fig. 3) showed that the WRF-ROMS-SWAN case was at the low-end, 8 m s^{-1} below the peak experiment (WRF-WW3-CTY). Fig. 4 demonstrates a 60 m s^{-1} contour in the WRF-WW3-CTY case that is roughly 100 m closer to the surface than found in the Control, WRF-WW3-CTY, and WRF-ROMS-SWAN experiments, likely resulting in that higher 10 m wind. This 60 m s^{-1} contour was missing entirely in the WRF-WW3-Drennan and WRF-WW3-ROMS cases.

To examine the TC structure further, Fig. 5 shows a radial-height cross section of azimuthally-averaged absolute vorticity (shaded) and vertical wind (contoured every 0.2 m s^{-1} from 0 to 1 m s^{-1}) extending from the storm center to a radius of 100km around the time of maximum strength (approximately F072). The WRF-WW3-Oost and WRF-ROMS-SWAN experiments demonstrated the most coherent updraft maximum extending from approximately 1000m to over 8000m in height. In this same updraft region, there is an ascending pillar of strong absolute vorticity colocated with the maximum winds in Fig. 4. This resulted in a much narrower contour of maximum winds in the WRF-WW3-Oost and WRF-WW3-SWAN versus the other cases. The Control, WRF-WW3-CTY, and WRF-WW3-Drennan had broader areas of maximum winds and absolute vorticity with a weaker updraft. At this point in the simulation, the WRF-ROMS-WW3 case had the maximum winds and absolute vorticity in the shallowest layer, and the most incoherent updraft. In all, this demonstrates a boundary-layer sensitivity to both complexity of coupling as well as the choice

of surface roughness parameterization, which should be examined more in depth but beyond the scope of this paper.

2) OCEAN CONDITION

To examine why the intensity differed with the coupled model configurations, we analyzed the solution of the fully-coupled (WRF-ROMS-SWAN) experiment just prior to landfall (approximately F072) in Fig. 6. Sea level pressure (Fig. 6a) shows the hurricane with an intensity of 962 hPa about to make landfall over Carolina coast. The strength of the storm (Fig. 6c) is approximately 49 m s^{-1} with a radius of maximum wind of approximately 50 km (or roughly 16 grid cells in the vortex-following 3 km grid), which is found in Figs. 4 and 5. In the SST field (Fig. 6b) the right-side SST cooling bias that has been demonstrated in numerous papers (Price et al. 1978; Price 1981; Price et al. 1994; Chen et al. 2007; Wada et al. 2014; Zambon et al. 2014a,b) is evident. The Gulf Stream on the coastal shelf appears to continue to advect warm water and bisect across this cold wake, despite the presence of the hurricane. The strong rainfall in this storm is apparent in the 3 hr precipitation (Fig. 6d) with rainfall exceeding 2.5 in h^{-1} and simulated radar echoes above 50 dBZ (Fig. 6e). Finally, through the fully-coupled wave model, the impact of the storm on the wave environment (Fig. 6f) results in waves in excess of 13 m just prior to landfall.

To examine why the intensity differed between the control and experiments, we analyzed SST at four National Data Buoy Center (NDBC) buoys located near the hurricane track: deep-ocean buoy 41048, continental slope buoy 41002, and buoys 41004 and 41013 adjacent to the coastline near the point of landfall (Fig. 7). These buoys, along with observed track, modeled tracks, and the WRF-ROMS-SWAN experiment, reported SST immediately prior to landfall (approximately F072, 0000 UTC 14 September). Coastal upwelling in the post-storm environment (characterized by northeastward winds) impacted the coastal buoys (41004 and 41013) more than the offshore

buoys. SST at the offshore buoys had minimal change over the first 144 hours of the model solution (Fig. 8): the farthest offshore buoy (41048) had a SST reduction of less than 2°C and the shelf buoy (41002) was reduced by less than 0.5 °C. The RTG SST analysis (used in the control and WRF-WW3 experiments) and the coupled experiments resolved this within approximately 1°C through this period.

By the end of the simulation, both of the coastal buoys reported SST cooling of 2.5 °C (41004 and 41013) due to a combination of TC passage and coastal upwelling. The analyzed (RTG SST) and coupled SSTs arrived at values within 1 °C of the observed SST at these sites. However, throughout the integration of the model, particularly at the critical point where the storm was observed to have weakened, the RTG SST was higher at these coastal stations by over 1 °C. The RTG SST was also consistently higher than the coupled SSTs at these stations before landfall. Partly owing to this artificial increase of surface heat to drive storm fluxes, the storms simulated using the RTG SST product were stronger (resulting in a better RMSE, RI, AE, AAE, MEF; Tables 2, 3, 4, 5) but did not demonstrate the weakening observed (resulting in a worse correlation coefficient shown in Tables 2 and 3).

In addition to these in situ NDBC buoy measurements, we compiled a series of Advanced Very High Resolution Radiometer (AVHRR) daily data to determine the Pre- and Post-Storm environment as well as the difference between them. As AVHRR data is not available beneath cloud cover, we had to average daily data from 1 day before to 1 day after. Therefore, the Pre-Storm SST was averaged AVHRR data from 10-12 September 2018. Likewise, the Post-Storm SST was averaged AVHRR data from 20-22 September 2018.

The spatial features in the AVHRR Observations are largely absent from the RTG SST and therefore the uncoupled experiments. While the Δ SST shows a large swath of cooler water in the RTG SST, there is unrealistically warm water (confirmed in Fig. 8; buoy 41004) near the coastline,

where the storm slowed down before landfall and coastal upwelling dominated the Post-Storm environment. Meanwhile, the coupled models show similar features to those observed including coastal upwelling and the warm intrusion of the Gulf Stream across the SST cooling wake (with a strong right-side cooling bias). Enhanced mixing in the SST from wave-coupling was found in only one buoy, 41048 which was furthest offshore; this is also demonstrated in Fig. 9 with minimal difference in the WRF-ROMS-WW3 and WRF-ROMS-SWAN SST. Improvement in the SST, and the resultant storm intensity and strength in the coupled cases could be improved with data assimilation of in situ and remote observations (e.g. satellite radiometers and altimeters, buoys, gliders, moorings, etc).

We used the same NDBC buoys to observe significant wave height and investigate skill in the uncoupled WAVEWATCHIII and coupled SWAN models (Fig. 10). Offshore (buoy 41048), neither the WAVEWATCHIII nor the SWAN models accurately represented the wave height. We attribute the wave height underprediction during the peak of the storm in the WAVEWATCHIII solution to coarse (0.5°) model grid spacing. We attribute the deficiency in the higher resolution SWAN model to model spinup as the storm was closest to this buoy at initialization, and the SWAN model was initialized from a cold start. The WRF model had to first bring the storm up in strength/intensity (Fig. 3) and then the SWAN model had to build up the wave field from those surface winds.

Further along the track, on the continental shelf, the comparison at buoy 41002 was much better with SWAN and WAVEWATCHIII having accurately represented the magnitude and timing of the highest waves. The two buoys inshore (41004 and 41013) also accurately represent the wave fields, with minimal difference between the WAVEWATCHIII hindcast and coupled SWAN runs. To further the point, the agreement between the wave fields of WAVEWATCHIII and SWAN are all-the-more impressive considering the atmospheric surface conditions and resolution driving them

are wholly different. Caution should be placed in the choice of uncoupled wave fields to input into an atmospheric model as the surface forcing may not agree with the lower boundary information provided by the wave model (e.g. a displacement in TC track between the atmospheric and wave models).

We then looked at the different parameterizations of surface wave roughness derived from the uncoupled WAVEWATCHIII and coupled SWAN wave fields (Fig. 11). We chose F060 (1200 UTC 13 September 2018) as this was immediately prior to landfall. The top row (Figs. 11a, b, c), uncoupled WAVEWATCHIII wave components, show the relatively smooth fields of significant wave height (Fig. 11a), wave length (Fig. 11b), and peak wave period (Fig. 11c) from the global WAVEWATCHIII hindcast. The resultant surface wave roughness lengths are shown in the middle row (Figs. 11d, e, f). The maximum surface wave roughness defined by Davis et al. (2008) of 2.85×10^{-3} is reached within 100-200 km of the storm center. The broadest area of enhanced sea surface roughness was found in the Oost et al. (2002) parameterization (Fig. 11f), followed by Taylor and Yelland (2001, Fig. 11d), and then Drennan et al., (2005; Fig. 11e). While the Drennan et al., (2005; Fig. 11e) and Oost et al., (2002; Fig. 11f) surface roughness fields show radial coherence, this is not found in the Taylor and Yelland, (2001; Fig. 11d) case. Investigating the Taylor and Yelland (2001) surface roughness, wave steepness varies around the hurricane due to swell and sea generation and propagation.

The bottom row (Figs. 11g, h, i, j) shows the sea surface roughness as parameterized by Oost et al. (2002) in the fully-coupled model. The wave celerity, and thus sea surface roughness (Fig. 11i), varies spatially artificially, consistent with the Garden Sprinkler Effect (Tolman 2002). The numerical artifacts caused by the Garden Sprinkler Effect can be reduced by increasing the spectral grid spacing of the wave model, at significant computational expense. A sensitivity experiment

doubling the number of directional and frequency bins (not shown) improved the numerical artifacts but otherwise resulted in minimal change to the simulation.

The latent and sensible heat fluxes are of utmost importance to hurricane development and intensity (Emanuel et al. 2004; Rotunno and Emanuel 1987). They are also intrinsic to the surface roughness parameterization of surface roughness (Olabarrieta et al. 2012; Prakash et al. 2019) which we vary during our experiments (Fig. 11). Therefore, we investigate them next in our study, using the same temporal data point as Fig. 11, F060 (1200 UTC 13 September 2018), immediately prior to landfall. The uncoupled fields show intense waves of heat flux wrapping cyclonically around the around the eye (Figs. 12a, b, c, d, g, h, i, j). This is to be expected as the uncoupled models resolved a more intense and stronger storm. This is accentuated when involving a lower-boundary wave condition as shown in the WRF-WW3 panels with different roughness parameterizations. The maximum latent heat flux in the moving nest are within 7.8% of each other among the uncoupled cases (Figs. 12a, b, c, d). The maximum sensible heat flux in the moving nest are within 8.6% of each other among the uncoupled cases (Figs. 12g, h, i, j). Considering only the WRF model modified by a lower-boundary wave condition, the maximum latent heat flux in the moving nest are within 7.3% (Figs. 12b, c, d); the maximum sensible heat flux in the moving nest are within 8.6% (Figs. 12h, i, j). The ocean-coupled with WAVEWATCHIII case reduces latent heat from the maximum uncoupled configurations by 11.0% (Figs. 12a, e); sensible heat from the maximum uncoupled configurations by 13.5% (Figs. 12j, k). The fully-coupled experiment further reduced both the latent heat from the maximum uncoupled configurations by 20.0% (Figs. 12a, f); sensible heat from the maximum uncoupled configurations by 21.6% (Figs 12j, l).

Note that as the WRF nest is tracking the center of circulation, the difference of track makes modest changes to this result as the approach of land to the northwest varies among the experiments (Fig. 12). Given that the difference between latent and sensible heat fluxes over this small area

of land at this temporal point in the simulation are orders of magnitude weaker, they are of little consequence.

The simulation that represents the heat flux best is debatable, as heat flux is intrinsic to both surface roughness and sea surface temperature. The comparison of the least coupled case (Control) and the case with the greatest coupling (WRF-ROMS-SWAN) feature lowest RMSE versus best correlation coefficient (Tables 2, 3, 4, 5). Similarly, the Control run had the intensification benefit of higher SST values nearshore. Future model studies into grid spacing, surface physics parameterizations with the inclusion of sea spray (Prakash et al. 2019; Liu et al. 2012a), data assimilation, and boundary layer physics are undoubtedly needed.

3) PRECIPITATION IMPACTS

We hypothesized that the reduction in surface sensible and latent heat fluxes to the hurricane, which weakened the hurricane in intensity and strength, would also have reduced the amount of precipitation from the storm. We used NOAA's Advanced Hydrologic Prediction Service (AHPS) precipitation data product as validation. The precipitation data are quality-controlled, multi-sensor (radar and rain gauge) precipitation estimates obtained from National Weather Service (NWS) River Forecast Centers (RFCs). The final mosaic product is produced by NCEP. For the best comparison, we interpolated this AHPS precipitation data product to our model grid using a 2-D cubic interpolation. We then took a subset of our model output from the southeast U.S. on the coarse (9 km) grid configured to align with the availability of interpolated AHPS data (Fig. 13e). The coarse domain extends eastward into the Atlantic beyond radar and rain gauge precipitation estimates, necessitating this subset. As a result, we were also limited to comparing precipitation during the landfall and dissipation phases of the storm, after its weakening.

The validation field (Fig. 13e) shows a relatively cohesive structure of precipitation over land, where rain gauge data can be used to supplement radar data. Extending offshore are areas of radial gradients of precipitation. Some of this could be explained by the circular structure of the TC's heavy rain bands moving shoreward and slowing down, which is also demonstrated in the precipitation fields of the experiments (Figs. 13a, b, c, d). These features might also be partly due to lack of data at distances hundreds of kilometers from the source stations. Radial features are present in the control and experiments, though they do not extend as far offshore as in the validation.

Comparison of validation to the control and experiments in Fig. 13 shows a large amount of precipitation falling along the NC side of the Carolinas southeastern border. As expected, in the uncoupled experiments using the RTG SST analysis, the embiggened fluxes from higher SSTs inshore resulted in a more intense and stronger storm that in turn produced more precipitation over a larger area (Figs. 13a, b). The weakening trend observed in Stewart and Berg (2019) and modeled in the coupled experiments due to cooler inshore SSTs resulted in a less intense and weaker storm with less precipitation (Figs. 13c, d). A very slight movement just prior to landfall in the control (Fig. 13a) and WRF-WW3 (Fig. 13b) experiments would explain the slight displacement of precipitation towards the north in those experiments. Overall the region of heaviest precipitation was found to be within the error caused by the less-than-30 km track deviation between the experiments.

To quantify the amount of precipitation that fell over the region in the AHPS validation product and compare that to our experiments, we calculated the volume of water as the area of the domain multiplied by the grid-resolved (Hong and Lim 2006) and parameterized (Grell and Dévényi 2002) total integrated precipitation that would fall over it (Fig. 14). To avoid areas of no precipitation from the AHPS Validation product, we integrated over the sub-domain visualized in Fig. 13.

The period of peak domain-integrated precipitation (13-September) shows a reduction of 22.0% in total precipitation between the validation product and the fully-coupled WRF-ROMS-SWAN experiment. At this time, the storm would have been located far offshore, so verification values may be inaccurate due to fewer measurements. The greatest spread among rainfall predictions by the experiments occurs just after landfall on 15-September. On that date, the amount of precipitation that fell over the region varied by 27.2% among the simulations. The inclusion of surface wave conditions on the TC solution between the uncoupled (control vs WRF-WW3) resulted in an decrease in precipitation of approximately 2.3%, likely owing to the reduction in Latent and Sensible heat fluxes prior to landfall (Fig. 12). The difference between coupled simulations was also small (at most, 2.8%), which could also be explained by the heat flux differences in Fig. 12. The addition of a surface wave model with the Oost et al. (2002) surface roughness parameterization slightly reduced the precipitation predicted for the region.

4. Summary and Conclusion

We examined Hurricane Florence, a storm that brought compound flooding (Wahl et al. 2015; Pietrafesa et al. 2019), a devastating combination of precipitation and storm surge, to the Carolinas in 2018. A useful contribution we made as a part of this study is the development of three wave parameterizations that can directly utilize operational wave model output for the community WRF model. The surface ocean condition under the current WRF code distribution is limited to a time-variant SST field from the RTG SST product Gemmill et al. (2007) implemented by Davis et al. (2008) and the Charnock (1955) parameterization of surface roughness as a function of wind speed. We introduced time-variant surface wave fields from WAVEWATCHIII (Tolman 1991; Tolman et al. 2002; Tolman 2014) and SWAN (Warner et al. 2010; Booij et al. 1999), including wave height, period, and direction to modulate wind stress using three surface wave roughness

parameterizations added to the WRF surface physics scheme, including Taylor and Yelland (2001), Drennan et al. (2005), and Oost et al. (2002). To these routines, we included the surface roughness limiter from Davis et al. (2008). All parameterizations and associated functions are able to be changed by the user at runtime through the WRF namelist file.

A series of experiments were conducted using Hurricane Florence to demonstrate and analyze the impact of wave coupling on track, intensity, strength, SST, storm structure, wave height, surface roughness, heat fluxes, and precipitation. These experiments included a control, with a time-variant SST condition; three uncoupled WRF-only experiments, with a time-variant wave field and three surface parameterizations; a coupled COAWST WRF-ROMS-WW3, with a dynamic ocean driving SST change and a time-variant but uncoupled wave field; and a fully coupled COWAST WRF-ROMS-SWAN, with feedback between the atmosphere, ocean, and wave environments. We added to the robustness of our results by running all 6 experiments across a 9-member temporal ensemble (54 total simulations), to ensure the results are not constrained by any one choice of initialization time.

TC track changed little between the control and the experiments. However, choice in sea surface roughness parameterizations demonstrated some significant changes in TC characteristics. To simplify the comparison of the coupled simulations, we used the best parameterization for this particular case, Oost et al. (2002), in the coupled experiments.

The experiments elucidated the importance of including a fully coupled ocean in modeling TC intensity and strength and therefore precipitation. Hurricane Florence had a drop in intensity and strength before landfall, owing partly to a shallower mixed layer that reduced OHC by more than half (Stewart and Berg 2019). This drop in intensity was resolved best by the coupled experiments, with the highest correlation coefficients of the experiments. This result demonstrated robustness with analysis of a temporal ensemble of 9 initialization times. In all of those 9 temporal ensemble

series of experiments, the best intensity correlation was found in the coupled models: 8 in the WRF-ROMS-SWAN case, and 1 in the WRF-ROMS-WW3 case. The best strength correlation coefficient of the temporal ensemble also belonged to the WRF-ROMS-SWAN experiments and in the 11Sept00Z case, the difference to the experiment with the best correlation coefficient was negligible (0.0002).

Owing to the simulation's weak initialization of the TC combined with the increased SST in the RTG SST field, the uncoupled simulations were able to get closer to the observed intensity and strength before landfall. We utilized a number of model-evaluation metrics in addition to correlation and foreRMSE including RI, AE, AAE, and MEF. Partly owing to an artificial increase of surface heat to drive storm fluxes, the storms simulated using the RTG SST product were stronger (resulting in a better RMSE, RI, AE, AAE, MEF). Our novel approach of introducing wave fields to the WRF model improved these model metrics in the temporal ensemble in nearly every metric except correlation coefficient where the coupled models scored highest. Overall, the simulations varied by as much as 10hPa in intensity and 15 m s^{-1} in strength.

We confirmed that the SST in the uncoupled simulations was higher than in situ (buoy) and satellite remote (AVHRR) measurements just prior to storm landfall. The coupled cases demonstrated a few features missing in the coarser, uncoupled RTG SST analysis including the bisection of the cold wake by the Gulf Stream as well as coastal upwelling due to the northeastward coastal winds as the storm remained stalled over the Carolinas well after landfall. In situ measurements of wave height from the same buoys demonstrated that poor agreement in the offshore buoy from both the coupled and uncoupled wave models, due in the former to the model initialization and in the latter to coarse grid spacing. As the storm environment evolved and the TC approached the coast, the comparison of wave heights resolved by both the coupled and uncoupled wave models were found to be cromulent with the magnitude and timing of the largest waves.

We further examined the storm structure. We looked at the radial-height cross section of azimuthally-averaged tangential winds, absolute vorticity, and vertical winds extending from the storm center at a time when the TC was most intense and strongest. We found that while the WRF-ROMS-SWAN case had weaker 10 m winds than the uncoupled cases, the core of winds in that case was displaced just a few hundred meters above the surface. A similar signature was found in the uncoupled case using the same surface roughness parameterization. In addition, the vertical velocity and absolute vorticity in this radial-height cross section were very similar between the coupled and uncoupled cases using the same surface roughness parameterization.

Examining the surface roughness parameterizations resulted in changes in the spatial fields based on the different calculations from Charnock (1955), Taylor and Yelland (2001), Drennan et al. (2005), and Oost et al. (2002). These differences then translated to differences in the spatial distribution, magnitude, and extent of latent and sensible heat fluxes. For the uncoupled (WRF-only) cases, the difference in the maximum magnitudes of the latent and sensible heat fluxes were as much as 7.8% and 8.6%, respectively. When adding ocean coupling, the latent and sensible heat fluxes were reduced against the uncoupled cases by at most 11.0% and 13.5%, respectively. In the fully coupled atmosphere-ocean-wave case, the latent and sensible heat fluxes were further reduced to a maximum of 20.0% and 21.6%, respectively.

The precipitation impacts of the experiments over a southeast U.S. sub-region were more accurately predicted by the coupled models. Compared to a verification dataset using the NOAA AHPS precipitation product, all the experiments did well at resolving the breadth of precipitation. The higher SSTs in the uncoupled simulations resulted in increased fluxes and therefore increased predicted precipitation by almost 22.9% at landfall, and the difference between the highest and lowest precipitation among the experiments was over 27.2%.

The air-sea interaction beneath hurricanes is a complicated set of processes that are difficult to model. By evaluating modeling experiments of increasing ocean and wave coupling, we found significant differences in resultant intensity, strength, SST, storm structure, wave height, surface roughness, heat fluxes, and precipitation. As a result, at the very least, an uncoupled atmospheric model should include time-variant SST and surface waves in simulating a TC. A fully-coupled atmosphere-ocean-wave model is best able to resolve these complex interactive processes. However, the addition of data assimilation from various oceanographic platforms (e.g. satellites, buoys, gliders, moorings, etc) to the fully-coupled model would be of enormous benefit to provide detailed data on the vertical structure of heat in the ocean. This would help ensure there is not excessive SST cooling which will reduce the vital fluxes of heat and moisture to the TC. In addition to this, future work should be devoted to adding sea spray parameterizations to both coupled and uncoupled models (e.g. Prakash et al. 2019), as well as evaluating coupling to a hydrology model (e.g. WRF-Hydro; Gochis et al. 2015) to capture processes like compound flooding during landfalling hurricanes.

Acknowledgments. This work has been supported by the U.S. Geological Survey Coastal/Marine Hazards and Resources Program, and by Congressional appropriations through the Additional Supplemental Appropriations for Disaster Relief Act of 2019 (H.R. 2157). The authors also wish to acknowledge research support through NSF Grant OCE-1559178 and NOAA grant NA16NOS0120028. We also wish to thank Chris Sherwood from the U.S. Geological Survey for his help in deriving wave length from WAVEWATCHIII data. Most of our experiments and analysis were completed using the Renaissance Computing Institute (RENCI), a collaboration involving the State of North Carolina, University of North Carolina at Chapel Hill (UNC-CH), Duke University, and North Carolina State University.

References

- Bao, J. W., J. M. Wilczak, J. K. Choi, and L. H. Kantha, 2000: Numerical simulations of air–sea interaction under high wind conditions using a coupled model: A study of hurricane development. *Monthly Weather Review*, **128** (7), 2190–2210.
- Bender, M. A., and I. Ginis, 2000: Real-case simulations of hurricane–ocean interaction using a high-resolution coupled model: Effects on hurricane intensity. *Monthly Weather Review*, **128**, 917–946.
- Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The Operational GFDL Coupled Hurricane–Ocean Prediction System and a Summary of Its Performance. *Monthly Weather Review*, **135** (12), 3965–3989.
- Bennett, A., 2018: New bern is counting up the damage from hurricane florence. it's at \$100 million so far. *The News and Observer*, URL <https://www.newsobserver.com/news/local/article218889660.html>.
- Biswas, M. K., D. Stark, and L. Carson, 2018: GFDL Vortex Tracker Users Guide Version 3.9a. National Center for Atmospheric Research and Developmental Testbed Center, URL https://dtcenter.org/sites/default/files/community-code/gfdl/standalone_tracker_UG_v3.9a.pdf.
- Booij, N., R. C. Ris, and L. H. Holthuijsen, 1999: A third-generation wave model for coastal regions: 1. Model description and validation. *Journal of Geophysical Research*, **104** (C4), 7649.
- Charnock, H., 1955: Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc.*, **81** (350), 639–640.
- Chassignet, E. P., H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, P. J. Hogan, A. J. Wallcraft, R. Baraille, and R. Bleck, 2007: The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *Marine Environmental Monitoring and Prediction Selected papers from the 36th*

International Liège Colloquium on Ocean Dynamics 36th International Liège Colloquium on Ocean Dynamics, **65** (1–4), 60–83.

Chen, S. S., W. Zhao, M. A. Donelan, and H. L. Tolman, 2013: Directional Wind–Wave Coupling in Fully Coupled Atmosphere–Wave–Ocean Models: Results from CBLAST-Hurricane. *Journal of the Atmospheric Sciences*, **70** (10), 3198–3215.

Chen, S. S., W. Zhao, Dong, J. Price, E. Walsh, T. Sanford, and H. Tolman, 2007: The CBLAST-Hurricane Program and the Next-Generation Fully Coupled Atmosphere–Wave–Ocean Models for Hurricane Research and Prediction. *Bull. Amer. Meteor. Soc.*, **88** (3), 311–317.

Cummings, J. A., 2005: Operational multivariate ocean data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131** (613), 3583–3604.

Cummings, J. A., and O. M. Smedstad, 2013: Variational Data Assimilation for the Global Ocean. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*, Springer, Berlin, Heidelberg, Berlin, Heidelberg, 303–343.

Davis, C., and Coauthors, 2008: Prediction of Landfalling Hurricanes with the Advanced Hurricane WRF Model. *Mon. Wea. Rev.*, **136** (6), 1990–2005.

Donelan, M. A., B. K. Haus, N. Reul, W. J. Plant, M. Stiassnie, H. C. Graber, O. B. Brown, and E. S. Saltzman, 2004: On the limiting aerodynamic roughness of the ocean in very strong winds. *Geophys. Res. Lett.*, **31** (18), L18 306.

Drennan, W. M., P. K. Taylor, and M. J. Yelland, 2005: Parameterizing the Sea Surface Roughness. *J. Phys. Oceanogr.*, **35** (5), 835–848.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436** (7051), 686–688.

- Emanuel, K., C. DesAutels, C. Holloway, and R. Korty, 2004: Environmental Control of Tropical Cyclone Intensity. *J. Atmos. Sci.*, **61** (7), 843–858.
- Flather, R. A., 1976: A tidal model of the northwest European continental shelf. *Mem. Soc. Roy. Sci. Liege*, **10** (6), 141–164.
- Gemmill, W., B. Katz, and X. Li, 2007: Daily Real Time Global Sea Surface Temperature - High Resolution Analysis at NOAA/NCEP. Tech. Rep. 260, NOAA/NCEP.
- Gochis, D. J., W. Yu, and D. N. Yates, 2015: The WRF-Hydro Model Technical Description and User's Guide, Version 3.0. NCAR Technical Document.
- Grell, G. A., and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29** (14), 38–1–38–4.
- Haidvogel, D. B., and Coauthors, 2008: Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System. *J. Comput. Phys.*, **227** (7), 3595–3624.
- Henderson, J., 2018: 39 million gallons of sewage: What florence left in the cape fear. *The Fayetteville Observer*, URL <https://www.fayobserver.com/news/20181013/39-million-gallons-of-sewage-what-florence-left-in-cape-fear>.
- Hong, S. Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113** (D13).

- Jacob, R., J. Larson, and E. Ong, 2005: M X N Communication and Parallel Interpolation in Community Climate System Model Version 3 Using the Model Coupling Toolkit: *The International Journal of High Performance Computing Applications*, **19** (3), 293–307.
- Jones, P. W., 1998: A user's guide for SCRIP: A Spherical Coordinate Remapping and Interpolation Package. Los Alamos National Laboratory, URL <http://climate.lanl.gov/Software/SCRIP/>.
- Khain, A., M. Frumin, I. Ginis, and A. Falkovich, 2000: Interaction of binary tropical cyclones in a coupled tropical cyclone-ocean model. *Journal of Geophysical Research: Oceans* (1978–2012), **105**, 22.
- Komen, G. J., and K. Hasselmann, 1984: On the Existence of a Fully Developed Wind-Sea Spectrum. *J. Phys. Oceanogr.*, **14** (8), 1271–1285.
- Kumar, N., G. Voulgaris, J. C. Warner, and M. Olabarrieta, 2012: Implementation of the vortex force formalism in the coupled ocean-atmosphere-wave-sediment transport (COAWST) modeling system for inner shelf and surf zone applications. *Ocean Modelling*, **47**, 65–95.
- Kumar, R. R., B. Sandeepan, and D. M. Holland, 2020: Impact of different sea surface roughness on surface gravity waves using a coupled atmosphere–wave model: a case of Hurricane Isaac (2012). *Ocean Dynamics*, **70** (3), 421–433.
- Larson, J., R. Jacob, and E. Ong, 2005: The Model Coupling Toolkit: A New Fortran90 Toolkit for Building Multiphysics Parallel Coupled Models: *The International Journal of High Performance Computing Applications*, **19** (3), 277–292.
- Lee, C.-Y., and S. S. Chen, 2012: Symmetric and Asymmetric Structures of Hurricane Boundary Layer in Coupled Atmosphere–Wave–Ocean Models and Observations. *Journal of the Atmospheric Sciences*, **69** (12), 3576–3594.

- Lee, C.-Y., and S. S. Chen, 2014: Stable Boundary Layer and Its Impact on Tropical Cyclone Structure in a Coupled Atmosphere–Ocean Model. *Monthly Weather Review*, **142** (5), 1927–1944.
- Leggett, R. W., and L. R. Williams, 1981: A reliability index for models. *Ecological Modelling*, **13** (4), 303–312, doi:10.1016/0304-3800(81)90034-x.
- Leipper, D. F., and D. Volgenau, 1972: Hurricane Heat Potential of the Gulf of Mexico. *J. Phys. Oceanogr.*, **2** (3), 218–224.
- Liu, B., C. Guan, and L. Xie, 2008: Investigating the impacts of wave state and sea spray on typhoon via a coupled atmosphere-wave system: the idealized case. *28th Conference on Hurricanes and Tropical Meteorology*.
- Liu, B., C. Guan, L. Xie, and D. Zhao, 2012a: An investigation of the effects of wave state and sea spray on an idealized typhoon using an air-sea coupled modeling system. *Advances in Atmospheric Sciences*, **29** (2), 391–406.
- Liu, B., H. Liu, L. Xie, C. GUAN, and D. Zhao, 2010: A Coupled Atmosphere–Wave–Ocean Modeling System: Simulation of the Intensity of an Idealized Tropical Cyclone. *Mon. Wea. Rev.*, **139** (1), 132–152.
- Liu, L., J. Fei, X. Cheng, and X. Huang, 2012b: Effect of wind-current interaction on ocean response during Typhoon Kaemi (2006). *Science China Earth Sciences*, **56**, 418–433.
- Loague, K., and R. E. Green, 1991: Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology*, **7** (1-2), 51–73, doi: 10.1016/0169-7722(91)90038-3.

- Madsen, O. S., Y.-K. Poon, and H. C. Graber, 1988: Spectral wave attenuation by bottom friction: Theory. *Coastal Engineering Proceedings*, B. L. Edge, Ed., American Society of Civil Engineers, Torremolinos, Spain, 492–504.
- Marchesiello, P., J. C. McWilliams, and A. Shchepetkin, 2001: Open boundary conditions for long-term integration of regional oceanic models. *Ocean Modelling*, **3**, 1–20.
- Moon, I.-J., I. Ginis, T. Hara, and B. Thomas, 2007: A Physics-Based Parameterization of Air–Sea Momentum Flux at High Wind Speeds and Its Impact on Hurricane Intensity Predictions. *Mon. Wea. Rev.*, **135** (8), 2869–2878.
- Nakanishi, M., and H. Niino, 2006: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog. *Boundary-Layer Meteorology*, **119** (2), 397–407.
- Nakanishi, M., and H. Niino, 2009: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *J. Meteor. Soc. Japan*, **87** (5), 895–912.
- Nash, J., and J. Sutcliffe, 1970: River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, **10** (3), 282–290, doi:10.1016/0022-1694(70)90255-6.
- National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2018: Global Data Assimilation System (GDAS). NOAA/NCEP.
- Nelson, J., and R. He, 2012: Effect of the Gulf Stream on winter extratropical cyclone outbreaks. *Atmospheric Science Letters*, **13** (4), 311–316.
- Nelson, J., R. He, J. C. Warner, and J. Bane, 2014: Air–sea interactions during strong winter extratropical storms. *Ocean Dynamics*, **64**, 1233–1246.

- Oey, L. Y., T. Ezer, D. P. Wang, X. Q. Yin, and S. J. Fan, 2007: Hurricane-induced motions and interaction with ocean currents. *Cont. Shelf Res.*, **27** (9), 1249–1263.
- Olabarrieta, M., J. C. Warner, B. Armstrong, J. B. Zambon, and R. He, 2012: Ocean–atmosphere dynamics during Hurricane Ida and Nor’Ida: An application of the coupled ocean–atmosphere–wave–sediment transport (COAWST) modeling system. *Ocean Modelling*, **43–44**, 112–137.
- Olabarrieta, M., J. C. Warner, and N. Kumar, 2011: Wave-current interaction in Willapa Bay. *J. Geophys. Res.*, **116** (C12), C12 014.
- Olson, J. B., J. S. Kenyon, W. A. Angevine, J. M. Brown, M. Pagowski, and K. Sušelj, 2019: A Description of the MYNN-EDMF Scheme and the Coupling to Other Components in WRF–ARW. *NOAA Technical Memorandum*, 37.
- Oost, W. A., G. J. Komen, C. M. J. Jacobs, and C. van Oort, 2002: New evidence for a relation between wind stress and wave age from measurements during ASGAMAGE. *Boundary-Layer Meteorology*, **103** (3), 409–438.
- Orlanski, I., 1976: A simple boundary condition for unbounded hyperbolic flows. *J. Comput. Phys.*, **21** (3), 251–269.
- Pietrafesa, L. J., H. Zhang, S. Bao, P. T. Gayes, and J. O. Hallstrom, 2019: Coastal Flooding and Inundation and Inland Flooding due to Downstream Blocking. *Journal of Marine Science and Engineering*, **7** (10), 336.
- Porchetta, S., O. Temel, J. C. Warner, D. Muñoz-Esparza, J. van Beeck, and N. van Lipzig, 2019: Evaluation of a roughness length parameterization accounting for wind-wave alignment in a coupled atmosphere-wave model, note.

- Prakash, K. R., V. Pant, and T. Nigam, 2019: Effects of the Sea Surface Roughness and Sea Spray-Induced Flux Parameterization on the Simulations of a Tropical Cyclone. *Journal of Geophysical Research: Atmospheres*, **124** (24), 14 037–14 058, doi:10.1029/2018jd029760.
- Price, J. F., 1981: Upper Ocean Response to a Hurricane. *J. Phys. Oceanogr.*, **11** (2), 153–175.
- Price, J. F., C. N. K. Mooers, and J. C. Van Leer, 1978: Observation and simulation of storm-induced mixed-layer deepening. *J. Phys. Oceanogr.*, **8** (4), 582–599.
- Price, J. F., T. B. Sanford, and G. Z. Forristall, 1994: Forced Stage Response to a Moving Hurricane. *J. Phys. Oceanogr.*, **24** (2), 233–260.
- Pu, Z.-X., and S. A. Braun, 2001: Evaluation of Bogus Vortex Techniques with Four-Dimensional Variational Data Assimilation. *Monthly Weather Review*, **129** (8), 2023–2039, doi:10.1175/1520-0493(2001)129<2023:eobvtw>2.0.co;2.
- Rotunno, R., and K. A. Emanuel, 1987: An air–sea interaction theory for tropical cyclones. Part II: Evolutionary study using a nonhydrostatic axisymmetric numerical model. *Journal of the Atmospheric Sciences*, **44** (3), 542 – 561.
- Shchepetkin, A. F., and J. C. McWilliams, 2005: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, **9** (4), 347–404.
- Skamarock, C., and Coauthors, 2019: A Description of the Advanced Research WRF Model Version 4. *NCAR Technical Note*, 1–162.
- Skamarock, W. C., J. B. Klemp, S.-H. Park, L. D. Fowler, and M. G. Duda, 2013: Evaluation of Global Atmospheric Solvers Using Extensions of the Jablonowski and Williamson Baroclinic Wave Test Case. *Mon. Wea. Rev.*, **141** (9), 3116–3129.

- Stewart, S. R., and R. J. Berg, 2019: National Hurricane Center Tropical Cyclone Report: Hurricane Florence (2018). *National Hurricane Center*, 98.
- Stow, C. A., J. Jolliff, D. J. McGillicuddy, S. C. Doney, J. I. Allen, M. A. Friedrichs, K. A. Rose, and P. Wallhead, 2009: Skill assessment for coupled biological/physical models of marine systems. *Journal of Marine Systems*, **76** (1-2), 4–15, doi:10.1016/j.jmarsys.2008.03.011.
- Taylor, P. K., and M. J. Yelland, 2001: The dependence of sea surface roughness on the height and steepness of the waves. *J. Phys. Oceanogr.*, **31** (2), 572–590.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land surface model in the WRF model. *20th conference on weather analysis and forecasting / 16th conference on numerical weather prediction*.
- Tolman, H. L., 1991: A Third-Generation Model for Wind Waves on Slowly Varying, Unsteady, and Inhomogeneous Depths and Currents. *J. Phys. Oceanogr.*, **21** (6), 782–797.
- Tolman, H. L., 2002: Alleviating the garden sprinkler effect in wind wave models. *Ocean Modelling*, **4** (3), 269 – 289, doi:https://doi.org/10.1016/S1463-5003(02)00004-5, URL <http://www.sciencedirect.com/science/article/pii/S1463500302000045>.
- Tolman, H. L., 2014: User manual and system documentation of WAVEWATCH III version 4.18. *NOAA NWS NCEP MMAB Technical Note*, **316**, 194.
- Tolman, H. L., B. Balasubramanian, 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** (2), 311–333.
- Umlauf, L., and H. Burchard, 2003: A generic length-scale equation for geophysical turbulence models. *J. Mar. Res.*, **61** (2), 235–265.

- Wada, A., M. F. Cronin, A. J. Sutton, Y. Kawai, and M. Ishii, 2013: Numerical simulations of oceanic pCO₂ variations and interactions between Typhoon Choi-wan (0914) and the ocean. *J. Geophys. Res.: Oceans*, **118** (5), 2667–2684.
- Wada, A., T. Uehara, and S. Ishizaki, 2014: Typhoon-induced sea surface cooling during the 2011 and 2012 typhoon seasons: observational evidence and numerical investigations of the sea surface cooling effect using typhoon simulations. *Progress in Earth and Planetary Science*, **64** (10), 3562–3578.
- Wahl, T., S. Jain, J. Bender, S. D. Meyers, and M. E. Luther, 2015: Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, **5** (12), 1093–1097.
- Warner, J. C., B. Armstrong, R. He, and J. B. Zambon, 2010: Development of a Coupled Ocean–Atmosphere–Wave–Sediment Transport (COAWST) Modeling System. *Ocean Modelling*, **35** (3), 230–244.
- Warner, J. C., C. R. Sherwood, H. G. Arango, and R. P. Signell, 2005: Performance of four turbulence closure models implemented using a generic length scale method. *Ocean Modelling*, **8** (1-2), 81–113.
- Warner, J. C., C. R. Sherwood, R. P. Signell, C. K. Harris, and H. G. Arango, 2008: Development of a three-dimensional, regional, coupled wave, current, and sediment-transport model. *Comput. Geosci.*, **34** (10), 1284–1306.
- Xie, L., B. Liu, H. Liu, and C. GUAN, 2010: Numerical simulation of tropical cyclone intensity using an air-sea-wave coupled prediction system. *Advances in Geosciences*, **18**, 19–43.

- Yablonsky, R. M., and I. Ginis, 2009: Limitation of One-Dimensional Ocean Models for Coupled Hurricane–Ocean Model Forecasts. *Monthly Weather Review*, **137** (12), 4410–4419.
- Zambon, J. B., R. He, and J. C. Warner, 2014a: Investigation of hurricane Ivan using the coupled ocean–atmosphere–wave–sediment transport (COAWST) model. *Ocean Dynamics*, **64** (11), 1535–1554.
- Zambon, J. B., R. He, and J. C. Warner, 2014b: Tropical to extratropical: Marine environmental changes associated with Superstorm Sandy prior to its landfall. *Geophys. Res. Lett.*, 2014GL061357.

LIST OF TABLES

Table 1.	Table showing the configuration of our experiments. The SST condition used is either RTG-SST in hindcast, or ROMS in a fully-coupled forecast. The wave model used is either none, WAVEWATCHIII in hindcast, or SWAN in a fully-coupled forecast. The surface roughness parameterizations, described in Section 2. a. 1) used include the Charnock (1955) formulation with no wave input, or wave-dependent formulations which require specified wave fields listed in the final 3 columns. These formulations include COARE-Taylor-Yelland (CTY; requiring Wave Height, and Wave Length), Drennan (requiring Wave Height, Wave Length, and Wave Period), and Oost (requiring Wave Length, and Wave Period).	40
Table 2.	Table showing the track, intensity, and strength of model results initialized from 11-September 2018 00Z compared to NHC Best Track. Track is based on RMSE; intensity and strength are based on correlation coefficient (r) and RMSE. Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.	41
Table 3.	Table showing the track, intensity, and strength of averaged temporal ensemble compared to NHC Best Track. Track is based on RMSE; intensity and strength are based on correlation coefficient (r) and RMSE. Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.	42
Table 4.	Table showing the intensity, and strength of model results initialized from 11-September 2018 00Z compared to NHC Best Track. These additional intensity and strength statistical metrics are based on Reliability Index (RI; Stow et al. 2009; Leggett and Williams 1981), Average Error (AE), Absolute Average Error (AAE), and Modelling Efficiency (MEF; Stow et al. 2009; Nash and Sutcliffe 1970; Loague and Green 1991). Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.	43
Table 5.	Table showing the intensity, and strength of averaged temporal ensemble compared to NHC Best Track. These additional intensity and strength statistical metrics are based on Reliability Index (RI; Stow et al. 2009; Leggett and Williams 1981), Average Error (AE), Absolute Average Error (AAE), and Modelling Efficiency (MEF; Stow et al. 2009; Nash and Sutcliffe 1970; Loague and Green 1991). Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.	44

TABLE 1. Table showing the configuration of our experiments. The SST condition used is either RTG-SST in hindcast, or ROMS in a fully-coupled forecast. The wave model used is either none, WAVEWATCHIII in hindcast, or SWAN in a fully-coupled forecast. The surface roughness parameterizations, described in Section 2. a. 1) used include the Charnock (1955) formulation with no wave input, or wave-dependent formulations which require specified wave fields listed in the final 3 columns. These formulations include COARE-Taylor-Yelland (CTY; requiring Wave Height, and Wave Length), Drennan (requiring Wave Height, Wave Length, and Wave Period), and Oost (requiring Wave Length, and Wave Period).

Experiment	SST Condition	Wave Model	Surface Roughness Scheme	Wave Components		
				Height	Length	Period
Control	RTG-SST	None	Charnock	N	N	N
WRF-WW3-CTY	RTG-SST	WAVEWATCHIII	COARE-Taylor-Yelland	Y	Y	N
WRF-WW3-Drennan	RTG-SST	WAVEWATCHIII	Drennan	Y	Y	Y
WRF-WW3-Oost	RTG-SST	WAVEWATCHIII	Oost	N	Y	Y
WRF-ROMS-WW3	ROMS	WAVEWATCHIII	Charnock	N	Y	Y
WRF-ROMS-SWAN	ROMS	SWAN	Charnock	N	Y	Y

TABLE 2. Table showing the track, intensity, and strength of model results initialized from 11-September 2018 00Z compared to NHC Best Track. Track is based on RMSE; intensity and strength are based on correlation coefficient (r) and RMSE. Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.

Experiment	Track	Intensity		Strength	
	RMSE	r	RMSE	r	RMSE
Control	47.18	0.8720	14.51	0.9398	8.85
WRF-WW3-CTY	60.56	0.8700	15.07*	0.8917	9.58
WRF-WW3-Drennan	63.98	0.8871*	16.47	0.9526*	9.49
WRF-WW3-Oost	52.87*	0.8715	15.18	0.9102	9.26*
WRF-ROMS-WW3	61.32	0.8896	17.24	0.9328	10.06
WRF-ROMS-SWAN	68.66	0.8924	16.04	0.9524	9.21

TABLE 3. Table showing the track, intensity, and strength of averaged temporal ensemble compared to NHC Best Track. Track is based on RMSE; intensity and strength are based on correlation coefficient (r) and RMSE. Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.

Experiment	Track	Intensity		Strength	
	RMSE	r	RMSE	r	RMSE
Control	36.17	0.8351	13.24	0.9670	6.27
WRF-WW3-CTY	38.48*	0.8259	13.44	0.9240	6.57
WRF-WW3-Drennan	41.26	0.8406*	13.24*	0.9379*	6.40
WRF-WW3-Oost	38.56	0.8308	13.35	0.9328	6.33*
WRF-ROMS-WW3	46.43	0.8593	14.24	0.9692	7.76
WRF-ROMS-SWAN	46.15	0.8740	13.92	0.9704	7.39

TABLE 4. Table showing the intensity, and strength of model results initialized from 11-September 2018 00Z compared to NHC Best Track. These additional intensity and strength statistical metrics are based on Reliability Index (RI; Stow et al. 2009; Leggett and Williams 1981), Average Error (AE), Absolute Average Error (AAE), and Modelling Efficiency (MEF; Stow et al. 2009; Nash and Sutcliffe 1970; Loague and Green 1991). Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.

Experiment	Intensity				Strength			
	RI	AE	AAE	MEF	RI	AE	AAE	MEF
Control	1.0152	8.88	9.12	0.3392	1.3134	-6.74	7.44	0.1588
WRF-WW3-CTY	1.0158*	9.64*	9.88*	0.3604	1.3041*	-5.86*	7.90	0.1729
WRF-WW3-Drennan	1.0172	12.24	12.24	0.3913*	1.3517	-7.86	8.00	0.2575*
WRF-WW3-Oost	1.0159	9.80	10.04	0.3862	1.3258	-6.14	7.52*	0.1819
WRF-ROMS-WW3	1.0180	13.36	13.36	0.3812	1.3811	-8.02	8.34	0.2044
WRF-ROMS-SWAN	1.0167	12.12	12.12	0.2454	1.3781	-7.66	7.97	0.0943

TABLE 5. Table showing the intensity, and strength of averaged temporal ensemble compared to NHC Best Track. These additional intensity and strength statistical metrics are based on Reliability Index (RI; Stow et al. 2009; Leggett and Williams 1981), Average Error (AE), Absolute Average Error (AAE), and Modelling Efficiency (MEF; Stow et al. 2009; Nash and Sutcliffe 1970; Loague and Green 1991). Bold denotes best score among all runs (Control and Experiments); asterisk denotes the best score among WRF-WW3 uncoupled experiments.

Experiment	Intensity				Strength			
	RI	AE	AAE	MEF	RI	AE	AAE	MEF
Control	1.0137	3.12	8.57	0.4147	1.1949	-4.06	5.12	0.2763
WRF-WW3-CTY	1.0139	2.55*	8.90	0.4343*	1.1982	-1.26*	5.46	0.2689
WRF-WW3-Drennan	1.0137*	3.38	8.48*	0.3909	1.2056	-2.26	5.35*	0.2669
WRF-WW3-Oost	1.0138	2.84	8.57	0.4178	1.1919*	5.27	7.52	0.2734*
WRF-ROMS-WW3	1.0148	7.50	9.17	0.4281	1.2620	-6.23	6.51	0.2908
WRF-ROMS-SWAN	1.0145	8.08	9.24	0.3411	1.2511	-5.79	6.17	0.2758

LIST OF FIGURES

- Fig. 1.** Control and experiment design for uncoupled (shaded) and coupled (non-shaded) experiments. In the control setup, WRF ingests SST from the Real-Time Global SST (RTG SST) product. In the three WRF-WW3 experiments, WRF uses surface wave fields from WAVEWATCHIII to compute sea surface roughness. The WRF-ROMS-WW3 experiment uses a coupled model where the atmosphere and ocean provide feedback to each other. WRF-ROMS-SWAN is the fully-coupled atmosphere-ocean-wave configuration. 47
- Fig. 2.** Simulated tracks and SST: Simulated tracks of control and experiments through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). SST from the WRF-ROMS-SWAN experiment was taken from F072 (0000 UTC 14 September 2018), immediately prior to landfall. ROMS/SWAN domains defined by the Atlantic Ocean west of the thick black line at approximately 60°W, SST of water points outside of ROMS domain (e.g. Great Lakes, eastern Pacific Ocean along Central America) represented by RTG SST. 48
- Fig. 3.** Simulated intensity (top; hPa) and strength (bottom; m s^{-1}) of control and experiments from initialization through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). 49
- Fig. 4.** Radial-height cross section of azimuthally-averaged tangential winds extending from the storm center to a radius of 300km. Control and experiments shown immediately prior to landfall where intensity and strength are greatest (F072, 0000UTC 14 September 2018). Shading shows wind speed (m s^{-1}) with contours every 10 m s^{-1} 50
- Fig. 5.** Radial-height cross section of azimuthally-averaged absolute vorticity (shaded) and upward velocity (contours) extending from the storm center to a radius of 100km. Control and experiments shown immediately prior to landfall where intensity and strength is greatest (F072, 0000UTC 14 September 2018). Shading shows absolute vorticity in 10^{-5} s^{-1} . Contours show upward velocity every 0.2 m s^{-1} 51
- Fig. 6.** WRF-ROMS-SWAN Simulation at F072 0000 UTC 14 September 2018 over the Southeast US coast and Sargasso Sea: Solution of the fully coupled, WRF-ROMS-SWAN experiment taken from F072 (0000 UTC 14 September 2018), immediately prior to landfall where intensity and strength are greatest. Top row: sea level pressure (hPa; a), SST ($^{\circ}\text{C}$; b), wind speed (m s^{-1} ; c) and direction (vector; c). Bottom row: 3-hour precipitation (inches; d), simulated radar reflectivity (dBZ; e), and significant wave height (m; f). 52
- Fig. 7.** Simulated Tracks and SST: Simulated tracks of control and experiments through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). SST is from the WRF-ROMS-SWAN experiment taken from F072 (0000 UTC 14 September 2018), immediately prior to landfall. Locations of NDBC Buoys used for SST analysis are included. NHC Best Track storm locations and times to right of black track. The approximate experimental track forecast times (F042-F090) located along the experimental tracks. 53
- Fig. 8.** Buoy SST Comparisons: Time-series comparisons of SST from NDBC buoys (black) and model solutions through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). Control and WRF-WW3 SST condition from RTG SST analysis are shown in red. Coupled WRF-ROMS-WW3 experiments shown in Pink. Fully coupled WRF-ROMS-SWAN experiment shown in green. 54
- Fig. 9.** Spatial SST comparisons of AVHRR remotely sensed SST observations and experiments' SST fields in $^{\circ}\text{C}$. Left column (a, d, g, j): Pre-Storm SST, the average of daily AVHRR data for

10-12 September 2018 or experiments' SST at initialization (11-September 2018). Middle column (b, e, h, k): Post-Storm SST, the average of daily AVHRR data for 20-22 September 2018 or experiments' SST at termination (21-September 2018). Right column (c, f, i, l): the difference between the Pre- and Post- Storm SST. Top row (a, b, c): AVHRR remote observations. Second row (d, e, f): Control and WRF-WW3 experiments (RTG SST). Third row (g, h, i): WRF-ROMS-WW3 experiment. Bottom row (j, k, l): WRF-ROMS-SWAN experiment. 55

Fig. 10. Time-series comparisons of Wave Height from NDBC buoys (black) and model solutions through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). Control, WRF-WW3, and WRF-ROMS-WW3 Wave Height condition from WAVEWATCHIII analysis are shown in Pink. Fully coupled WRF-ROMS-SWAN experiment shown in green. 56

Fig. 11. Components of wave fields and parameterizations of sea surface roughness immediately prior to landfall (F060, 1200 UTC 13 September 2018). Top row shows uncoupled wave components from interpolated WAVEWATCHIII fields as ingested by the WRF model for surface roughness parameterization: Wave Height (m; a), Wave Length (m; b), Wave Period (s; c). Second row shows resultant sea surface roughness parameterizations through the Taylor and Yelland (2001; d), Drennan et al. (2005; e), and Oost et al. (2002; f) schemes. Third and bottom rows show the coupled SWAN wave model components of Wave Height (m; g)*, Wave Length (m; h), and Wave Period (s; i) ingested into the Oost et al. (2002) sea surface roughness parameterization with the resultant sea surface roughness field (m; j). *Note that the Wave Height in (g) is not used in the Oost et al. (2002) parameterization but rather included for completeness of demonstrating the SWAN model fields. 57

Fig. 12. Heat Flux (W m^{-2}) at F060 1200UTC 13 September 2018. Top row represents Latent Heat Flux: Control (W m^{-2} ; a), WRF-WW3 with Taylor and Yelland (2001; W m^{-2} ; b), WRF-WW3 with Drennan et al. (2005; W m^{-2} ; c), WRF-WW3 with Oost et al. (2002; W m^{-2} ; d), WRF-ROMS-WW3 (W m^{-2} ; e), and WRF-ROMS-SWAN (W m^{-2} ; f) experiments. Bottom row represents Sensible Heat Flux: Control (W m^{-2} ; g), WRF-WW3 with Taylor and Yelland (2001; W m^{-2} ; h), WRF-WW3 with Drennan et al. (2005; W m^{-2} ; i), WRF-WW3 with Oost et al. (2002; W m^{-2} ; j), WRF-ROMS-WW3 (W m^{-2} ; k), and WRF-ROMS-SWAN (W m^{-2} ; l) experiments. Text inset of each image shows the maximum value in the region shown. 58

Fig. 13. 24-Hr Rainfall (mm) 16 September 2018: Twenty four hour rainfall in mm through 0000 UTC 16 September 2018 after Hurricane Florence has made landfall. Top row: Control (mm; a) and WRF-WW3 (Oost et al., 2002 parameterization) experiment (mm; a). Middle row: WRF-ROMS-WW3 (mm; c) and WRF-ROMS-SWAN (mm; d) experiments. Bottom row: validation taken from NOAA's Advanced Hydrologic Prediction Service precipitation product, interpolated to the model domain grid (mm; e). 59

Fig. 14. 24-Hr Domain-Integrated Precipitation (m^3): Time series comparisons of precipitation integrated over the model sub-grids shown in Fig. 7 through 10-day model iterations (0000 UTC 11 September 2018 through 0000 UTC 21 September 2018). Control run in red. WRF-WW3 (Oost et al., 2002 parameterization) experiment in yellow. Coupled WRF-ROMS-WW3 experiment in pink. Fully coupled WRF-ROMS-SWAN experiment in green. 60

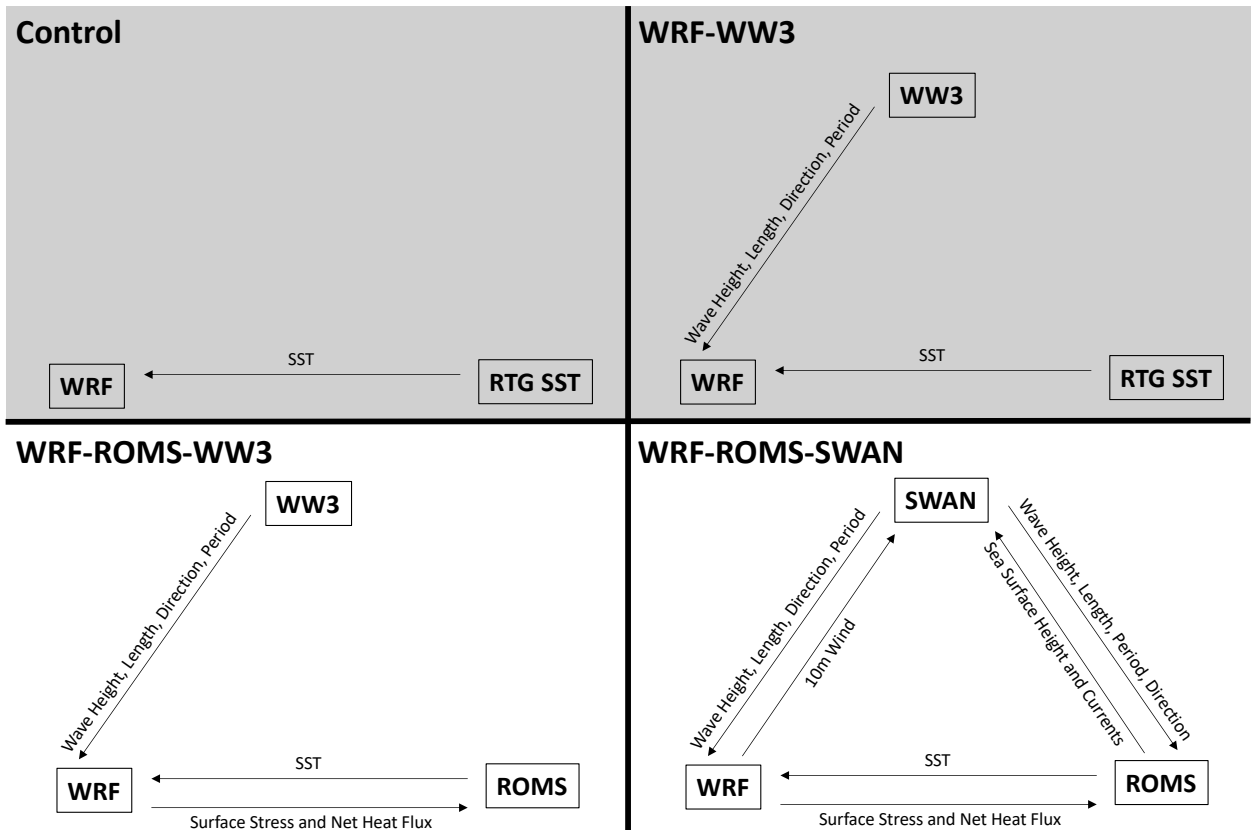


Fig. 1. Control and experiment design for uncoupled (shaded) and coupled (non-shaded) experiments. In the control setup, WRF ingests SST from the Real-Time Global SST (RTG SST) product. In the three WRF-WW3 experiments, WRF uses surface wave fields from WAVEWATCHIII to compute sea surface roughness. The WRF-ROMS-WW3 experiment uses a coupled model where the atmosphere and ocean provide feedback to each other. WRF-ROMS-SWAN is the fully-coupled atmosphere-ocean-wave configuration.

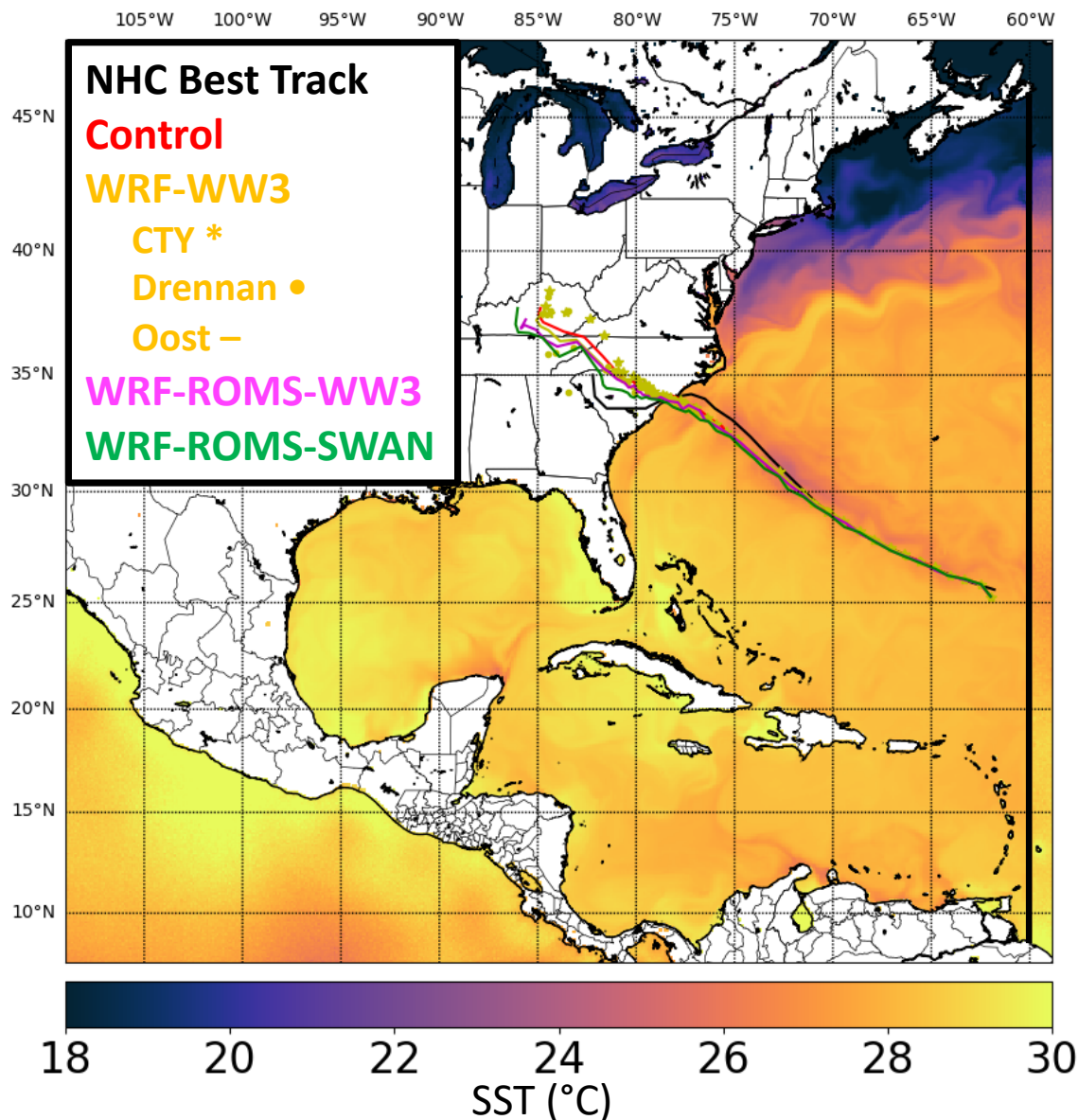


Fig. 2. Simulated tracks and SST: Simulated tracks of control and experiments through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). SST from the WRF-ROMS-SWAN experiment was taken from F072 (0000 UTC 14 September 2018), immediately prior to landfall. ROMS/SWAN domains defined by the Atlantic Ocean west of the thick black line at approximately 60°W, SST of water points outside of ROMS domain (e.g. Great Lakes, eastern Pacific Ocean along Central America) represented by RTG SST.

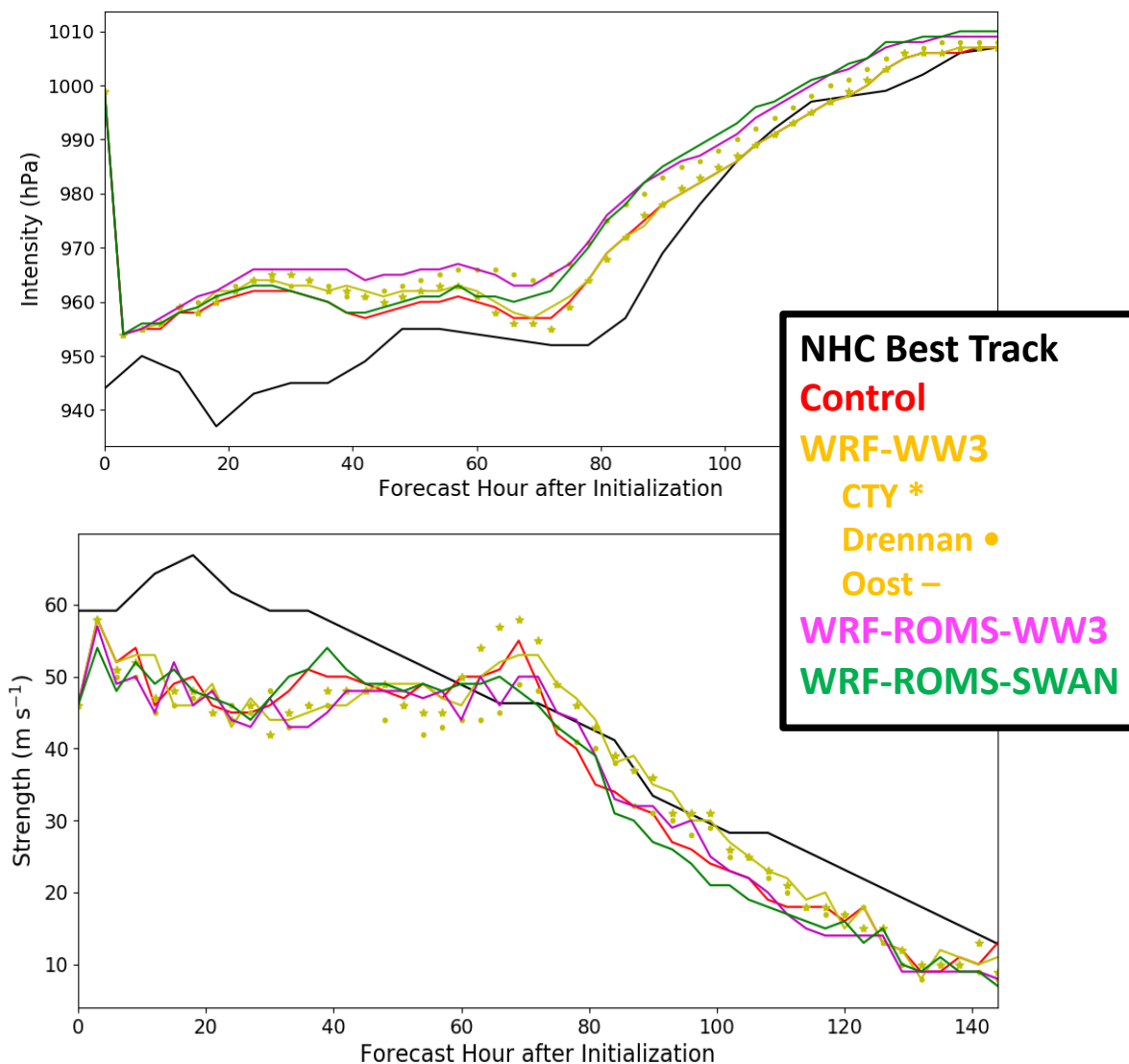


Fig. 3. Simulated intensity (top; hPa) and strength (bottom; m s⁻¹) of control and experiments from initialization through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018).

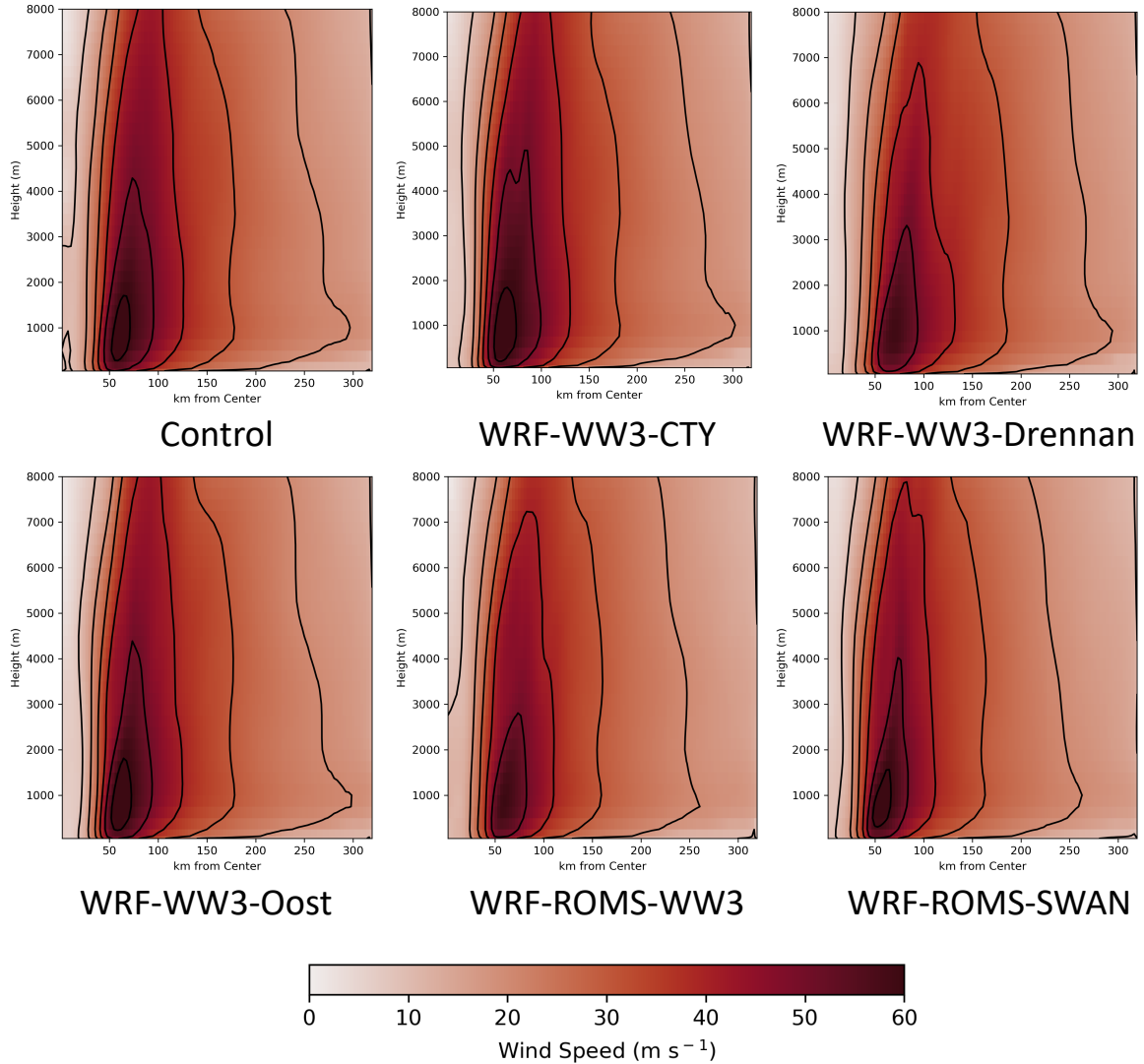


Fig. 4. Radial-height cross section of azimuthally-averaged tangential winds extending from the storm center to a radius of 300km. Control and experiments shown immediately prior to landfall where intensity and strength are greatest (F072, 0000UTC 14 September 2018). Shading shows wind speed (m s^{-1}) with contours every 10 m s^{-1} .

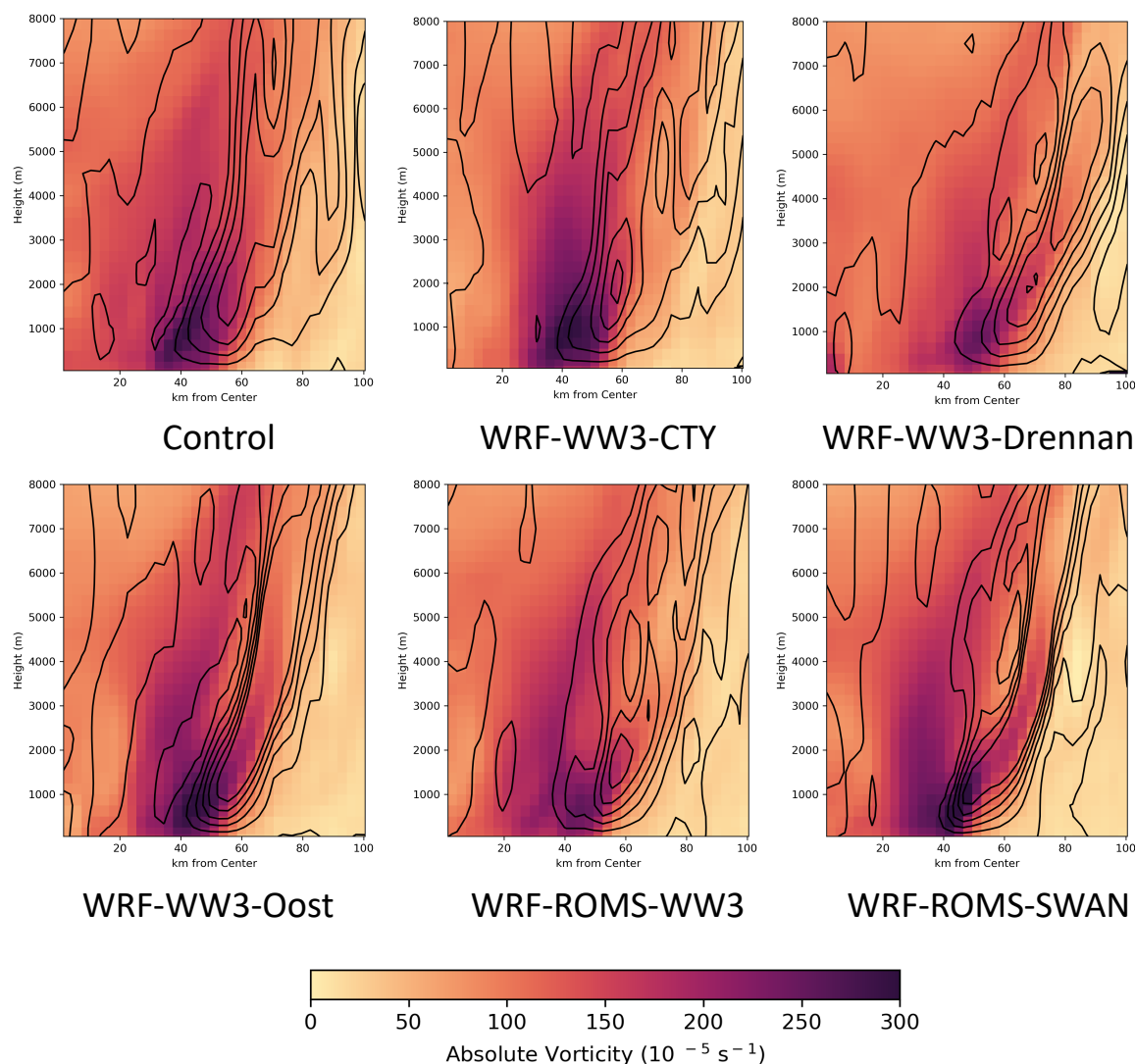


Fig. 5. Radial-height cross section of azimuthally-averaged absolute vorticity (shaded) and upward velocity (contours) extending from the storm center to a radius of 100km. Control and experiments shown immediately prior to landfall where intensity and strength is greatest (F072, 0000UTC 14 September 2018). Shading shows absolute vorticity in 10^{-5} s^{-1} . Contours show upward velocity every 0.2 m s^{-1} .

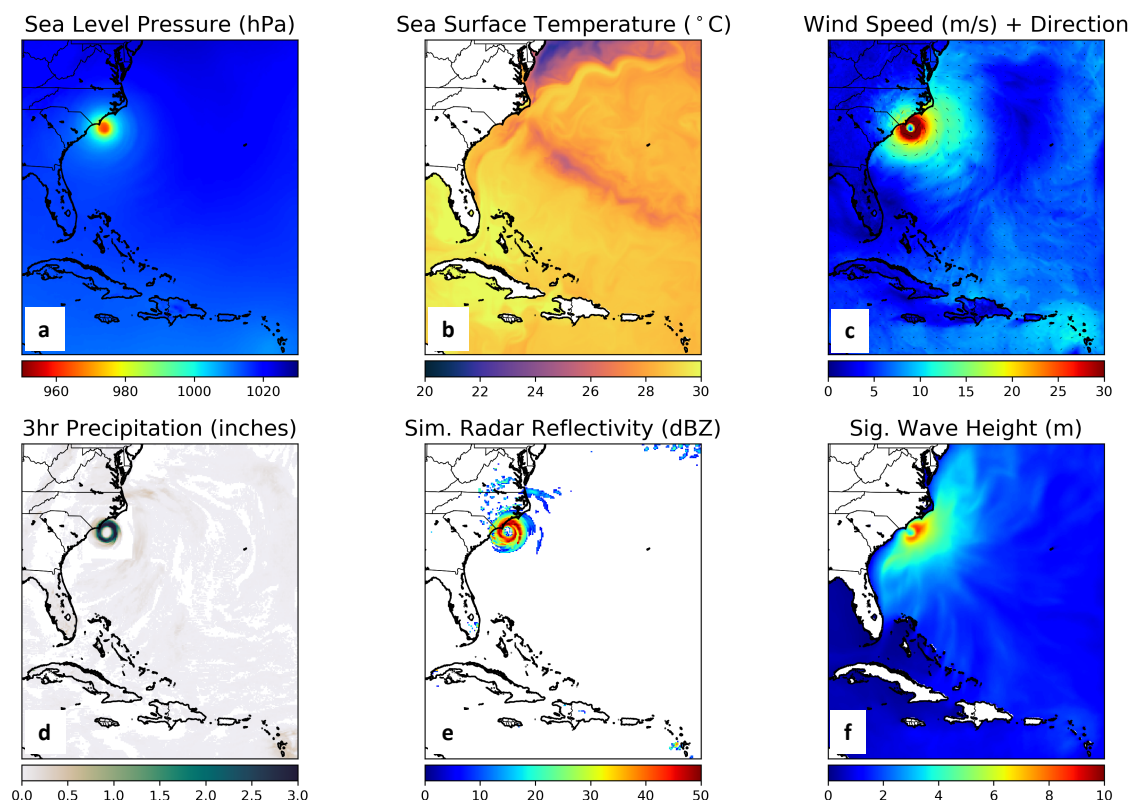


Fig. 6. WRF-ROMS-SWAN Simulation at F072 0000 UTC 14 September 2018 over the Southeast US coast and Sargasso Sea: Solution of the fully coupled, WRF-ROMS-SWAN experiment taken from F072 (0000 UTC 14 September 2018), immediately prior to landfall where intensity and strength are greatest. Top row: sea level pressure (hPa; a), SST ($^{\circ}\text{C}$; b), wind speed (m s^{-1} ; c) and direction (vector; c). Bottom row: 3-hour precipitation (inches; d), simulated radar reflectivity (dBZ; e), and significant wave height (m; f).

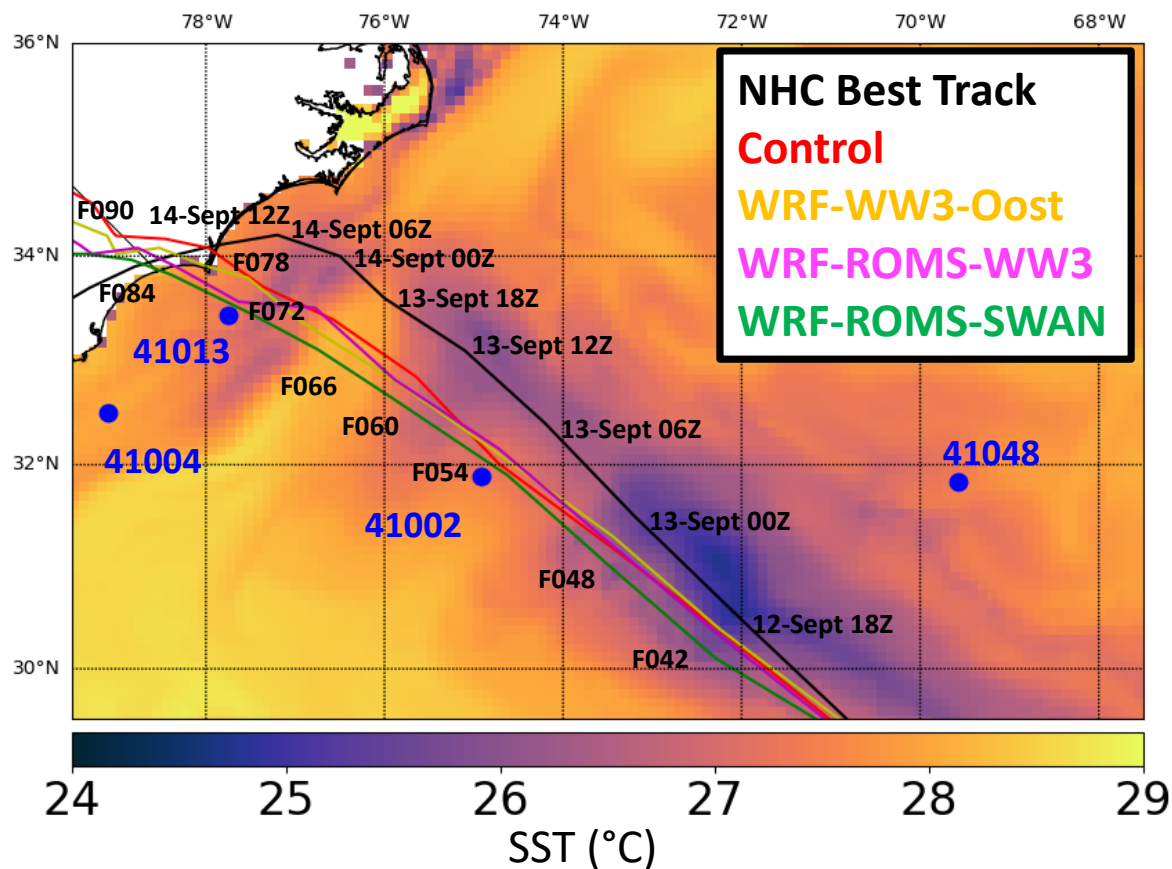
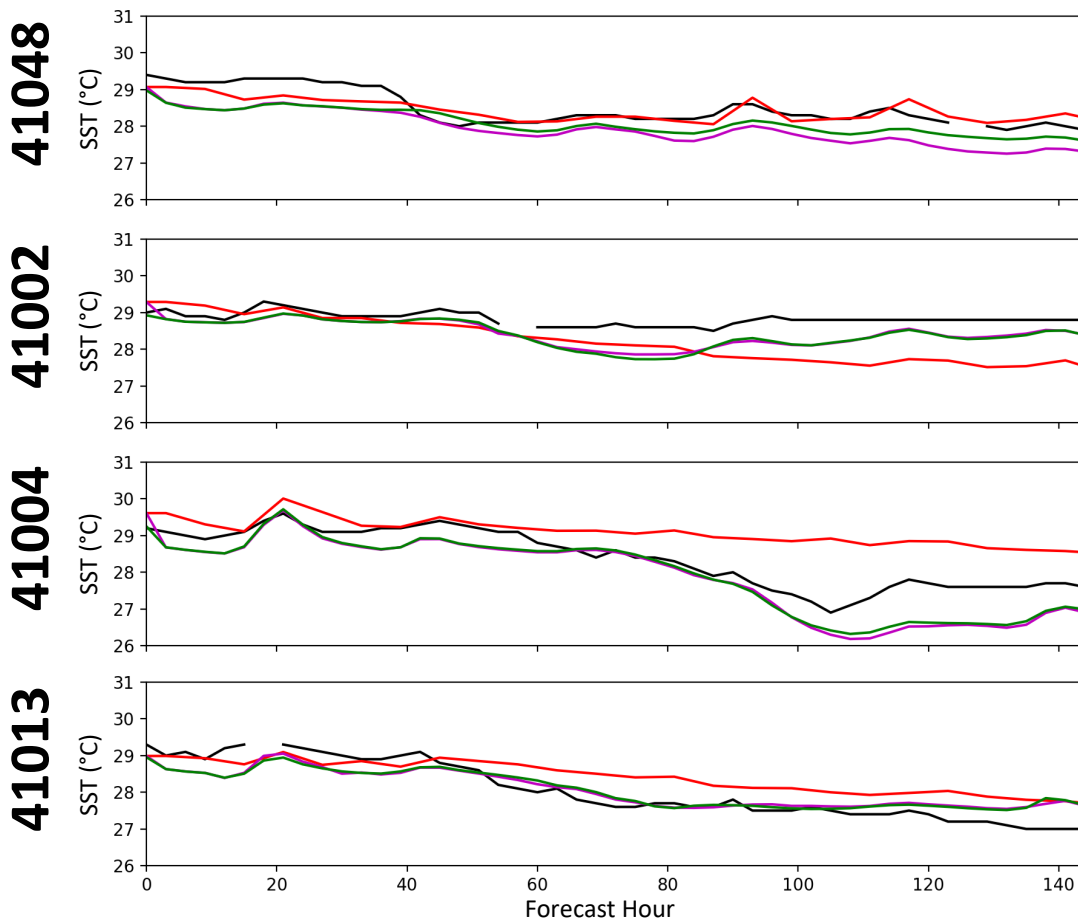


Fig. 7. Simulated Tracks and SST: Simulated tracks of control and experiments through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). SST is from the WRF-ROMS-SWAN experiment taken from F072 (0000 UTC 14 September 2018), immediately prior to landfall. Locations of NDBC Buoys used for SST analysis are included. NHC Best Track storm locations and times to right of black track. The approximate experimental track forecast times (F042-F090) located along the experimental tracks.



Validation **Control / WRF-WW3** **WRF-ROMS-WW3** **WRF-ROMS-SWAN**

Fig. 8. Buoy SST Comparisons: Time-series comparisons of SST from NDBC buoys (black) and model solutions through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). Control and WRF-WW3 SST condition from RTG SST analysis are shown in red. Coupled WRF-ROMS-WW3 experiments shown in Pink. Fully coupled WRF-ROMS-SWAN experiment shown in green.

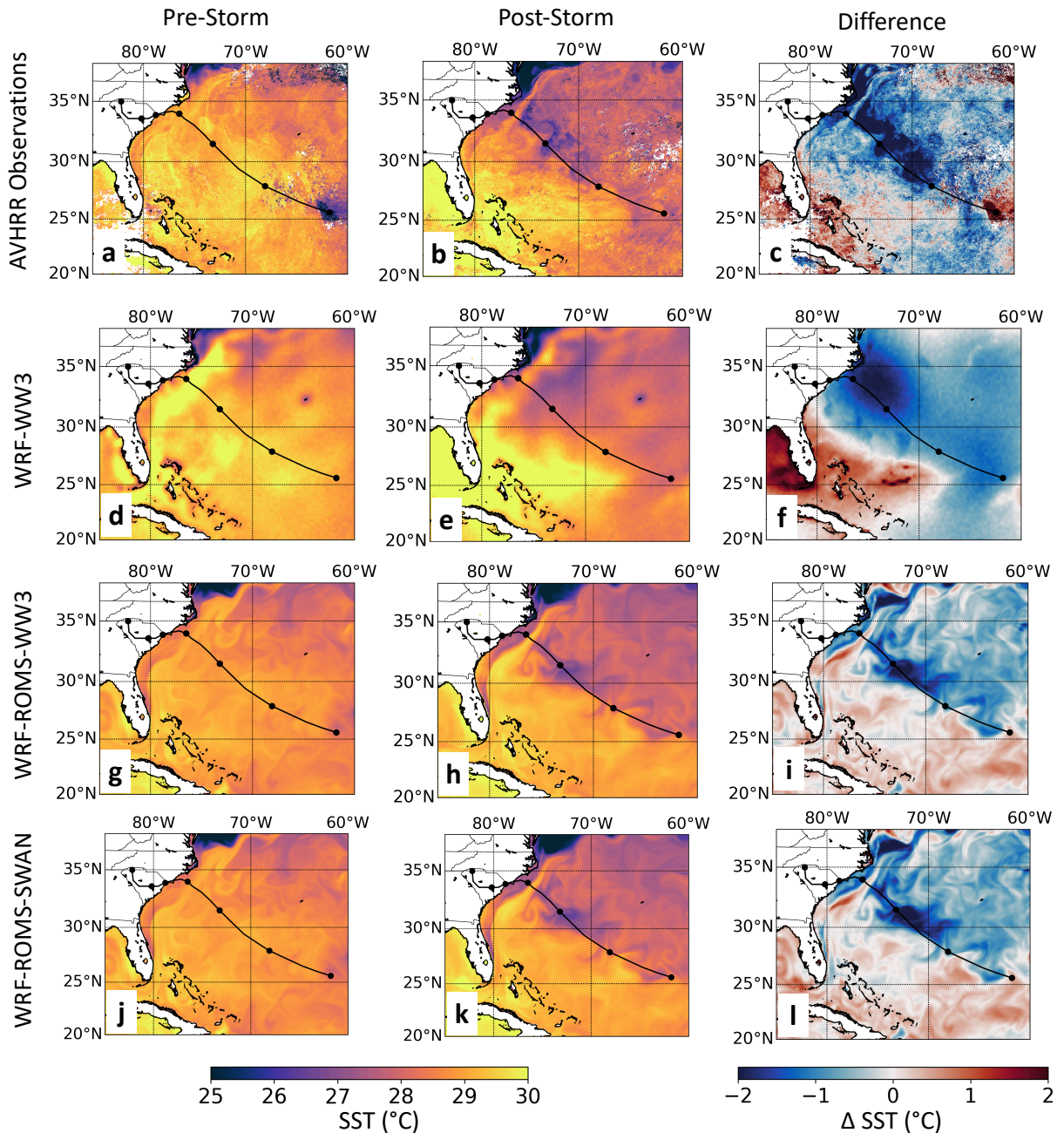
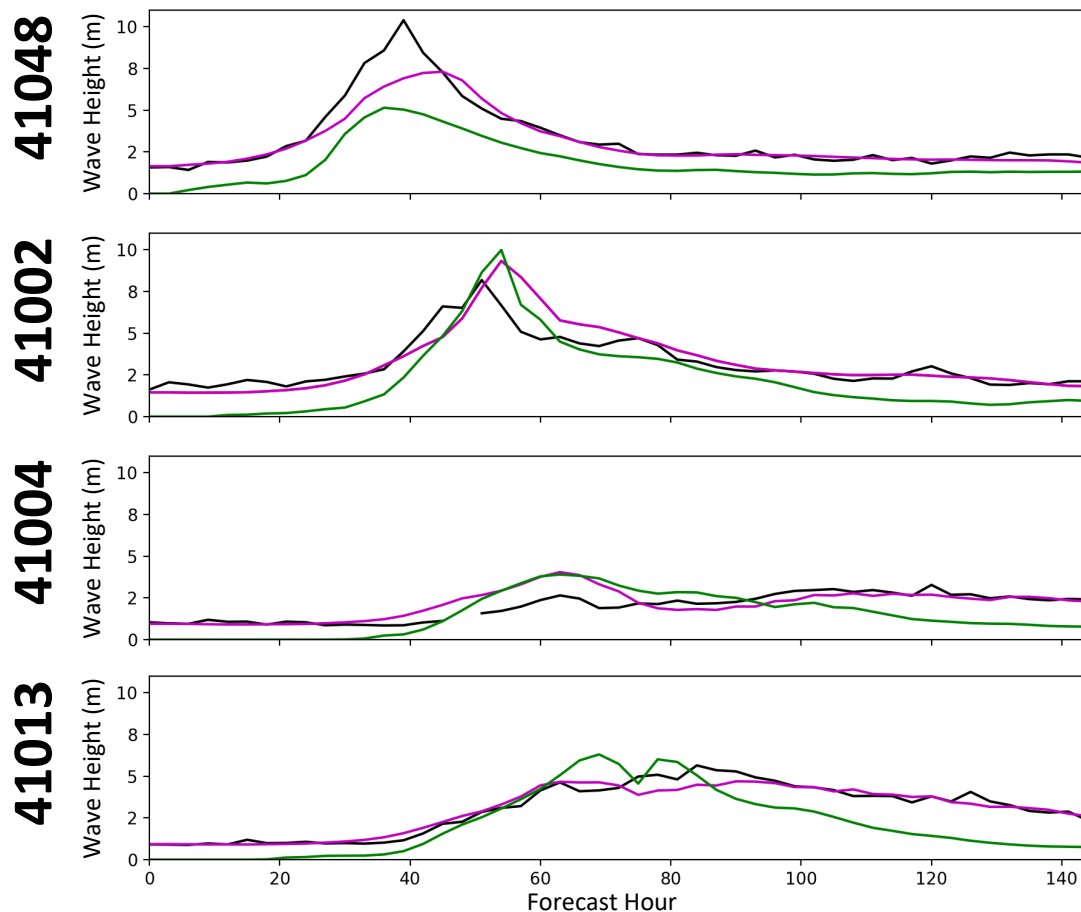


Fig. 9. Spatial SST comparisons of AVHRR remotely sensed SST observations and experiments' SST fields in °C. Left column (a, d, g, j): Pre-Storm SST, the average of daily AVHRR data for 10-12 September 2018 or experiments' SST at initialization (11-September 2018). Middle column (b, e, h, k): Post-Storm SST, the average of daily AVHRR data for 20-22 September 2018 or experiments' SST at termination (21-September 2018). Right column (c, f, i, l): the difference between the Pre- and Post- Storm SST. Top row (a, b, c): AVHRR remote observations. Second row (d, e, f): Control and WRF-WW3 experiments (RTG SST). Third row (g, h, i): WRF-ROMS-WW3 experiment. Bottom row (j, k, l): WRF-ROMS-SWAN experiment.



Validation **WRF-ROMS-WW3** **WRF-ROMS-SWAN**

Fig. 10. Time-series comparisons of Wave Height from NDBC buoys (black) and model solutions through landfall and storm dissipation (approximately F144, 0000 UTC 17 September 2018). Control, WRF-WW3, and WRF-ROMS-WW3 Wave Height condition from WAVEWATCHIII analysis are shown in Pink. Fully coupled WRF-ROMS-SWAN experiment shown in green.

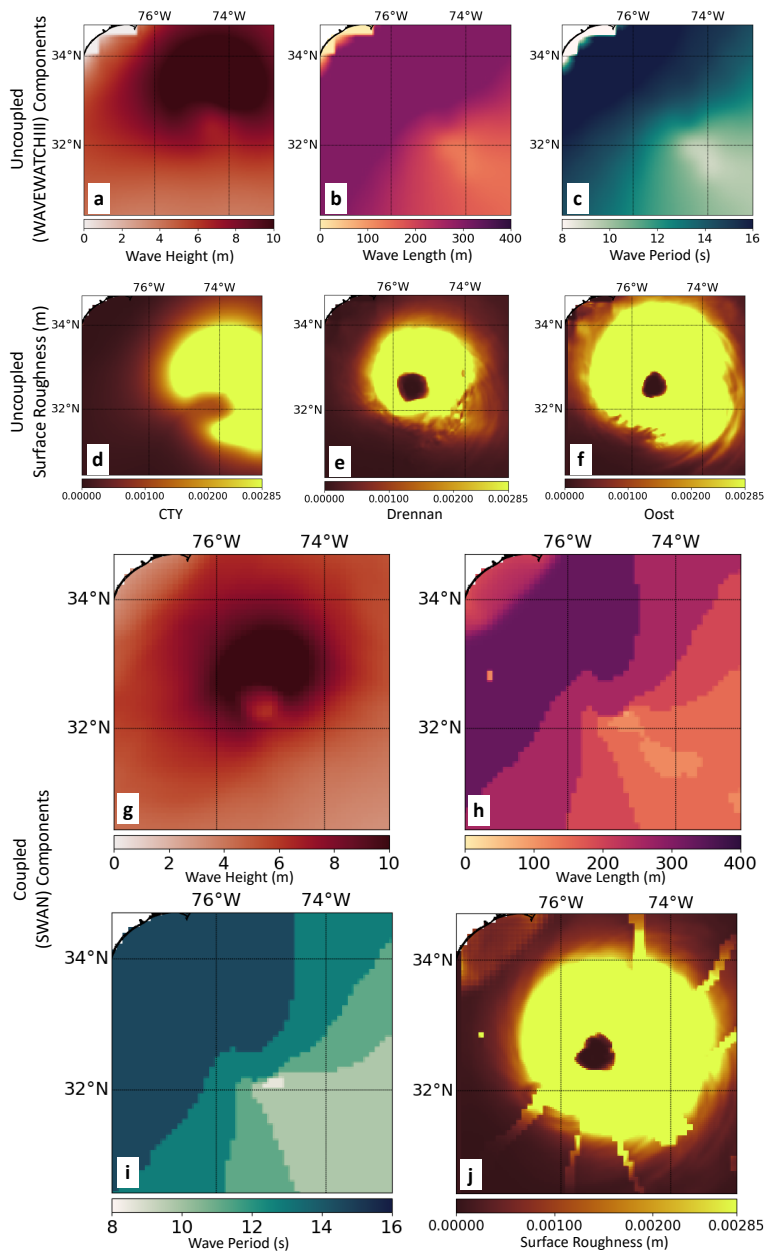


Fig. 11. Components of wave fields and parameterizations of sea surface roughness immediately prior to landfall (F060, 1200 UTC 13 September 2018). Top row shows uncoupled wave components from interpolated WAVEWATCHIII fields as ingested by the WRF model for surface roughness parameterization: Wave Height (m; a), Wave Length (m; b), Wave Period (s; c). Second row shows resultant sea surface roughness parameterizations through the Taylor and Yelland (2001; d), Drennan et al. (2005; e), and Oost et al. (2002; f) schemes. Third and bottom rows show the coupled SWAN wave model components of Wave Height (m; g)*, Wave Length (m; h), and Wave Period (s; i) ingested into the Oost et al. (2002) sea surface roughness parameterization with the resultant sea surface roughness field (m; j). *Note that the Wave Height in (g) is not used in the Oost et al. (2002) parameterization but rather included for completeness of demonstrating the SWAN model fields.

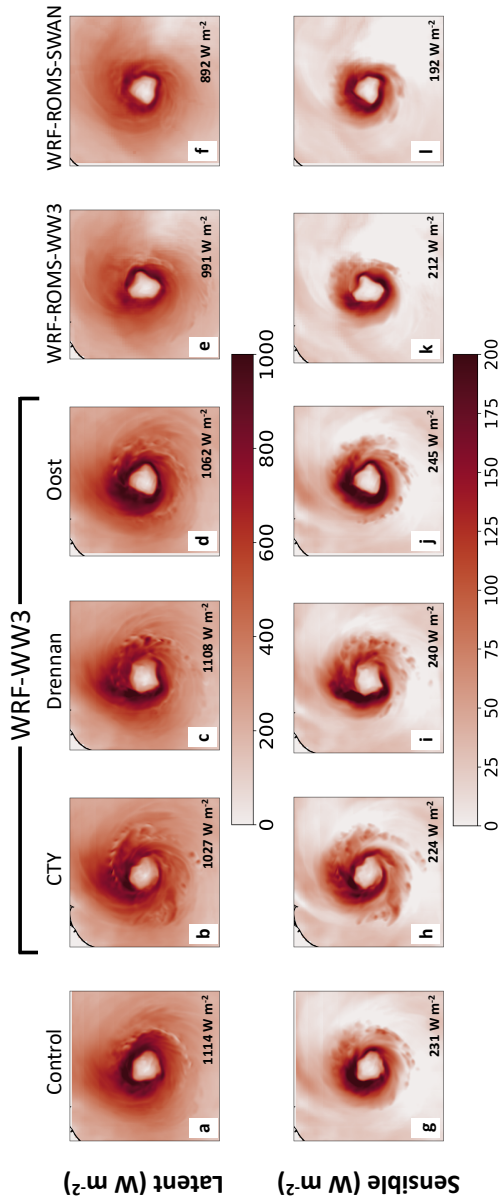


Fig. 12. Heat Flux (W m^{-2}) at F060 1200UTC 13 September 2018. Top row represents Latent Heat Flux: Control (W m^{-2} ; a), WRF-WW3 with Taylor and Yelland (2001; W m^{-2} ; b), WRF-WW3 with Drennan et al. (2005; W m^{-2} ; c), WRF-WW3 with Oost et al. (2002; W m^{-2} ; d), WRF-ROMS-WW3 (W m^{-2} ; e), and WRF-ROMS-SWAN (W m^{-2} ; f) experiments. Bottom row represents Sensible Heat Flux: Control (W m^{-2} ; g), WRF-WW3 with Taylor and Yelland (2001; W m^{-2} ; h), WRF-WW3 with Drennan et al. (2005; W m^{-2} ; i), WRF-WW3 with Oost et al. (2002; W m^{-2} ; j), WRF-ROMS-WW3 (W m^{-2} ; k), and WRF-ROMS-SWAN (W m^{-2} ; l) experiments. Text inset of each image shows the maximum value in the region shown.

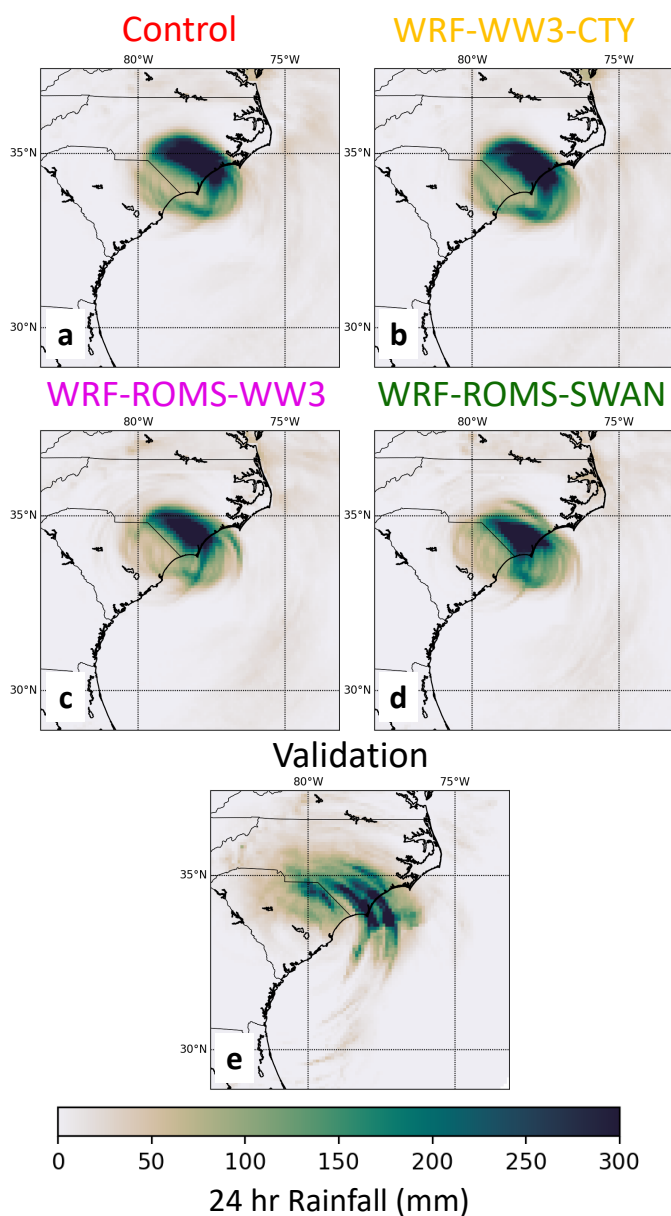


Fig. 13. 24-Hr Rainfall (mm) 16 September 2018: Twenty four hour rainfall in mm through 0000 UTC 16 September 2018 after Hurricane Florence has made landfall. Top row: Control (mm; a) and WRF-WW3 (Oost et al., 2002 parameterization) experiment (mm; a). Middle row: WRF-ROMS-WW3 (mm; c) and WRF-ROMS-SWAN (mm; d) experiments. Bottom row: validation taken from NOAA's Advanced Hydrologic Prediction Service precipitation product, interpolated to the model domain grid (mm; e).

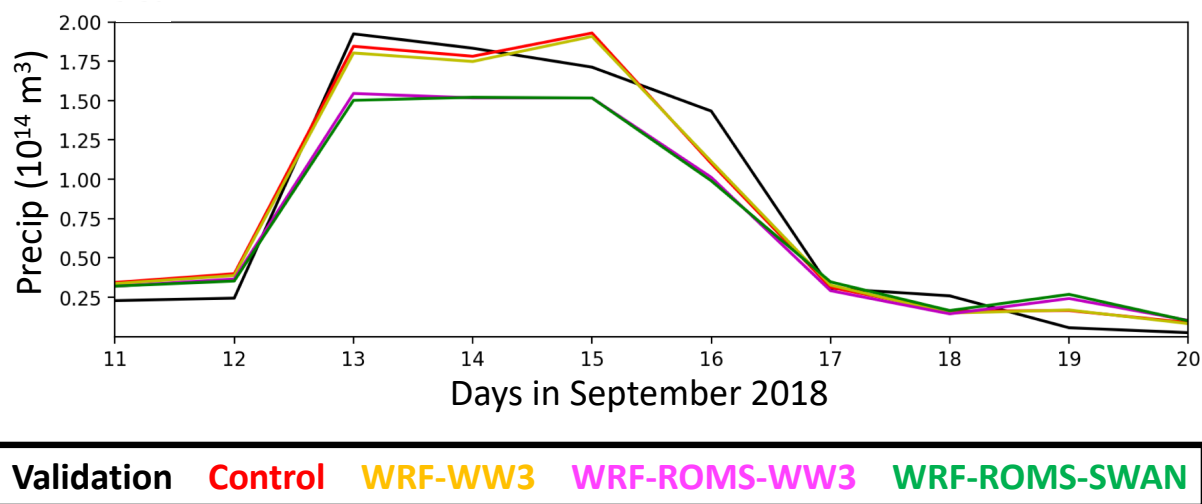


Fig. 14. 24-Hr Domain-Integrated Precipitation (m^3): Time series comparisons of precipitation integrated over the model sub-grids shown in Fig. 7 through 10-day model iterations (0000 UTC 11 September 2018 through 0000 UTC 21 September 2018). Control run in red. WRF-WW3 (Oost et al., 2002 parameterization) experiment in yellow. Coupled WRF-ROMS-WW3 experiment in pink. Fully coupled WRF-ROMS-SWAN experiment in green.