Impact of coupling an ocean model to WRF nor'easter simulations

Rutgers University has made this article freely available. Please share how this access benefits you.
Your story matters. [https://rucore.libraries.rutgers.edu/rutgers-lib/47935/story/]

This work is an ACCEPTED MANUSCRIPT (AM)
This is the author's manuscript for a work that has been accepted for publication. Changes resulting from the publishing process, such as copyediting, final layout, and pagination, may not be reflected in this document. The publisher takes permanent responsibility for the work. Content and layout follow publisher's submission requirements.

Citation for this version and the definitive version are shown below.

Citation to Publisher Version: Nicholls, Stephen D. & Decker, Steven G.. Impact of coupling an ocean model to WRF nor'easter simulations. Monthly Weather Review Accepted for publication. http://journals.ametsoc.org/loi/mwre

Citation to this Version: Nicholls, Stephen D. & Decker, Steven G.. Impact of coupling an ocean model to WRF nor'easter simulations. Monthly Weather Review Accepted for publication. Retrieved from http://dx.doi.org/doi:10.7282/T3222WPW

Terms of Use: Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

Article begins on next page
Impact of coupling an ocean model to WRF nor’easter simulations

Stephen D. Nicholls* and Steven G. Decker

Department of Environmental Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ

(Revision submitted 15 July 2015)
(Second revision submitted 15 September 2015)

* Current affiliation: NASA Postdoctoral Program Fellow, NASA Goddard Space Flight Center, Code 612, Greenbelt, MD, 20771

Corresponding author: Steven G. Decker
Department of Environmental Sciences, Rutgers University,
14 College Farm Rd.,
New Brunswick, NJ, 08901
E-mail: decker@envsci.rutgers.edu
Abstract

The impact of ocean-atmosphere coupling and its possible seasonal dependence upon Weather Research and Forecasting (WRF) model simulations of seven, winter-time cyclone events was investigated. Model simulations were identical aside from the degree of ocean model coupling (static SSTs, 1D mixed-layer model, full-physics 3D ocean model). Both 1D and 3D ocean model coupling simulations show that SSTs following the passage of a nor’eastern did tend to cool more strongly during the early season (Oct-Dec) and were more likely to warm late in the season (Feb-Apr). Model simulations produce SST differences of up to 1.14 K, but this change did not lead to significant change in storm track (< 100 km), maximum 10 m winds (< 2 m s⁻¹), or minimum sea-level pressure(<= 5 hPa). Simulated precipitation showed little sensitivity to model coupling, but all simulations did tend to over-predict precipitation extent (bias > 1) and have low-to-moderate threat scores (0.31 – 0.59). Analysis of the storm environment and the overall simulation failed to reveal any statistically significant differences in model error attributable to ocean-atmosphere coupling. Despite this result, ocean model coupling can reduce dynamical field error at a single level by up to 20%, and this was slightly greater (1-2%) with 3D ocean model coupling as compared to 1D ocean model coupling. Thus, while 3D ocean model coupling tended to generally produce more realistic simulations, its impact would likely be more profound for longer-term simulations.
1. Introduction

Computational power increases have driven the development of increasingly complex numerical weather prediction models, such as the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008). Despite the complexity and land-surface coupling abilities of the WRF model, it lacked any direct ocean-atmosphere coupling functionality until April 2009 (WRF version 3.2). This coupling is vital for tropical cyclone (Sutyrin and Khain 1984; Bender et al. 1993; Emanuel 1999; Warner et al. 2010; Olabarrieta et al. 2012) and coastal, mid-latitude cyclone (Anthes et al. 1983; Kuo et al. 1991; Ren et al. 2004; Eckhardt and Stohl 2004; Knippertz and Wernli 2009) simulations because of those cyclones’ strong dependence upon ocean-atmosphere heat flux and momentum exchanges. Such dependencies motivate ocean-atmosphere coupled model development.

One early study (Bender et al. 1993) coupled the Geophysical Fluid Dynamics Laboratory tropical cyclone prediction model to a multilayer primitive equation ocean model to investigate wind-stress-induced sea surface temperature (SST) cooling during Hurricane Gloria. Their simulations only varied cyclone propagation speed (slow, medium, fast) and produced a maximum simulated SST cooling of 5.3°C, 3.5°C, and 1.8°C, respectively and reduced ocean-to-atmosphere heat fluxes accordingly. These changes decreased simulated cyclone intensity and maximum winds by up to 7.3 hPa and 2.7 m s⁻¹, respectively; storm track was not significantly altered.

More recently, Ren et al. (2004) coupled the Canadian Mesoscale Compressibility Community atmospheric model to the full-physics, 3D Princeton Oceanography Model to investigate ocean-atmosphere dynamics during the September 1998 extratropical transition of Hurricane Earl and an intense January 2000 cyclone dubbed “Superbomb”. These systems were
selected due to their comparable size and intensity, but the shallower mixed later depth during
Earl produced greater maximum SST reductions than for “Superbomb” (5 versus 1°C). Despite
this result, the slower propagation speed of “Superbomb” led to greater intensity and maximum
wind speed changes (4 hPa and 4 m s$^{-1}$, respectively [1 hPa and 1 m s$^{-1}$ more than Earl]).
Consistent with Bender et al. (1993), model coupling did not significantly alter storm track.

Using the Ren et al. (2004) coupled model, Yao et al. (2008) investigated 42 North
Atlantic October mid-latitude cyclones (locally dubbed “nor’easters”) off the U.S. East Coast. As
compared to Ren et al. (2004), their simulations produced larger SST cooling (~6°C) and greater
cyclone weakening (4-5 hPa), yet a similar decrease in 10 m winds (2-4 m s$^{-1}$). Larger changes in
Yao et al. (2008) were attributed to the thinner October ocean mixed-layer depth and the wider
case variety.

As of WRF version 3.6.1, three built-in SST options exist: Static SST, SST update, or 1D
ocean mixed-layer coupling. For static SST, initial SST data is ingested from an external data
source (e.g., 0.5° real-time global SSTs [RTGSST; Gemmill et al. 2007]) and is held constant.
For SST update, SST values are updated at user-prescribed intervals using lower boundary input
files and are otherwise constant. Finally, the 1D ocean mixed layer model option couples WRF
to the Pollard et al. (1972) 1D ocean mixed-layer model. At initialization, standard ocean
temperature profiles are affixed to SST values and a uniform, user-selected, mixed-layer depth is
prescribed over the entire model domain. During forward integration, atmospheric wind stress is
applied to each simulated ocean column, but neither horizontal advection nor transport processes
(e.g., Ekman transport) are simulated, which prevents accurate simulations of fine-scale SST
gradients (e.g., the Gulf Stream) and the corresponding heat flux and momentum exchanges
(Chelton et al. 2001; O’Neill et al. 2003; Tikinaga et al. 2005; LaCasse et al. 2008). Such
limitations motivate the development of fully coupled modeling systems, such as the Coupled-Ocean-Atmosphere-Wave-Sediment Transport (COAWST) modeling system (Warner et al. 2010; hereafter W10).

The COAWST modelling system allows WRF to be directly coupled to the Regional Ocean Modeling System (ROMS; Shchepetkin and McWilliams 2005; Haidvogel et al. 2008), the Simulating Waves Nearshore (SWAN; Booij et al. 1999) wave, and the Community Sediment Transport Modeling System (CSTMS; Warner et al. 2008) models via the Model Coupling Toolkit (MCT; Larson et al. 2004). As described by W10, COAWST is specifically designed “to better identify the significant processes affecting our coastlines and how those processes create coastal change.” Since its release, COAWST-related studies have focused primarily on two main themes: Tropical cyclones (W10; Olabarrieta et al. 2012), and coastal process (rip currents, surf zone) simulations (Kumar et al. 2011; Olabarrieta et al. 2011). Applications of COAWST to baroclinically driven mid-latitude cyclones (such as nor’easters) have yet to be addressed.

We define a “nor’easter” as a large (~2000 km), mid-latitude cyclone occurring between October and April, bringing punishing winds, copious precipitation, and potential coastal flooding to the Northeastern U.S. (Kocin and Uccellini 2004; Jacobs et al. 2005; Ashton et al. 2008). Total nor’easter-related damages can currently reach several billion U.S. dollars per event (NCDC 2008). With warmer temperatures and increased atmospheric water vapor content expected during the 21st Century (Trenberth et al. 2007), the risk of damaging nor’easters is also expected to increase, primarily from sea-level rise (Yin et al. 2009) and more intense precipitation (Lombardo et al. 2015). Since storm surge and precipitation crucially depend on
air-sea interactions in the nor’easter environment, it is important to simulate these interactions accurately.

However, do we need a fully coupled model (such as COAWST) to capture these interactions, or is a less computationally demanding 1D ocean model sufficient? Furthermore, the ocean mixed layer is shallow in the fall, but deepens throughout the winter (Ren et al. 2004). Does this seasonality in the ocean’s thermal structure lead to a seasonal dependence in the impact of model coupling on nor’easter simulations? These are the questions this study aims to address.

The remainder of this paper is divided into four sections. Section 2 explains the methodology and analysis methods. Section 3 shows the results. Section 4 provides the conclusions and their implications. Finally, section 5 will discuss areas of potential future research.

2. Methods

a. COAWST Overview

This study utilized COAWST revision 727 (May 2013 release), which offers four WRF-based configurations (one uncoupled, three coupled; see Fig. 1). Given our ocean-atmosphere focus, we used the WRF-only and WRF-ROMS configurations (Figs. 1a,b) and did not couple to either SWAN or CSTMS (Figs. 1c,d). The Spherical Coordinate Remapping Interpolation Package (SCRIP; Jones 1998) computed interpolation weights between the WRF (Fig. 2a) and ROMS (Fig. 2b) model grids needed to facilitate their coupling. COAWST did not allow for multi-domain coupling when conducting this study, so ROMS domain 1 was only directly coupled to WRF domain 2 given their similar horizontal grid spacings (10 and 15 km,
respectively) and spatial extent. Further WRF and ROMS model details are contained in the sections below.

**b) WRF configuration**

The Advanced Research WRF version 3.4 (hereafter W34) solves a set of fully-compressible, non-hydrostatic, Eulerian equations in terrain-following coordinates (Skamarock et al. 2008). The three model domain grids have one-way feedback, 45-, 15-, and 5-km horizontal grid spacing, respectively, and 61 vertical levels which afforded simulation of key precursor synoptic and meso-α scale phenomena (e.g., jet streaks, short- and long-wave troughs) on outer domains and smaller-scale phenomena (e.g., orographic forcing, condensational latent heating) on inner domains. Atmospheric lateral boundary conditions were derived from 1°×1° Global Forecast System (GFS) model analysis (GMA), and the model top was set to 50 hPa. As in W10, RTGSST data was ingested for all three WRF domains, but for domain 2 these SST values were overwritten with ROMS-simulated SSTs within the ROMS coverage area (Fig. 2b). The three WRF domains had time steps of 90, 30, and 10 seconds, respectively.

Parameterization options were consistent with the recent WRF studies of Shi et al. (2010) and Tao et al. (2011) and include the following:

- Longwave radiation: New Goddard Scheme (Chou and Suarez 1999; Shi et al. 2010)
- Shortwave radiation: New Goddard Scheme (Chou and Suarez 1999)
- Surface layer: Eta similarity (Monin and Obukhov 1954; Janjic 2002)
- Land surface: NOAH (Chen and Dudhia 2001)
- Cumulus parameterization: Kain-Fritsch (Kain 2004) (Not applied to domain 3)
Microphysics: Goddard (Lang et al. 2007)

c) ROMS configuration

ROMS release 455 (Shchepetkin and McWilliams 2005; Haidvogel et al. 2008) numerically integrates 3D Reynolds-averaged Navier–Stokes equations using hydrostatic and Boussinesq approximations along vertically stretched, terrain-following (sigma) coordinates. The ROMS domain (Fig. 2b; shaded) has an average 10-km grid spacing (needed for fine-scale SST gradients), 16 sigma levels, and a 50 m sigma-coordinate surface/bottom layer width. Time step size was fixed at 30 seconds to match WRF domain 2 and for consistency with W10.

Ocean initialization and open-boundary conditions required atmospheric, ocean, and tidal information. Atmospheric state variables are interpolated from WRF. Ocean state variables (currents, temperature, salinity, sea surface height) originate from Hybrid Coordinate Ocean Model (HyCOM; Bleck 2002). Tides were calculated from the Advanced Circulation (ADCIRC; Mukai et al. 2002) tidal dataset and included the five tidal constituents (M2, N2, S2, O1, and K1). All these data sources were consistent with W10.

d) Cases for analysis and model run description

Research work focused on a small, but diverse sample of seven subjectively chosen nor’easter cases which vary in both severity and time of year (Table 1). A modest sample size was needed to integrate computationally expensive, high-resolution COAWST simulations on available resources. To emphasize any potential seasonal dependencies at least one case from every month (October-March) was selected, and the cases are sorted by month and day rather than by chronological order. Case severity was determined from the Northeast Snowfall Impact
Scale (NESIS; Kocin and Uccellini 2004) which classifies nor’easters on a scale of 1 (notable) to 5 (extreme) based upon the population impacted, area affected, and snowfall severity. Early and late season storms (Cases 1, 2, and 7) brought large impacts through heavy rain, strong winds, and coastal flooding, rather than snow, so they did not receive NESIS ratings.

Simulations are integrated for 180 hours, starting roughly 72 hours prior to the first precipitation impacts in the highly populated Mid-Atlantic region. This starting point allowed sufficient model spin-up time and for the establishment of both strong baroclinicity and latent heat flux off the U.S. east coast. The high baroclinicity and heat fluxes are a byproduct of the clash between the cold continental air and the expansive (>1000 km) northern edge of the Gulf Stream, which is an interaction crucial for accurate nor’easter simulations (Kuo et al. 1991; Mote et al. 1997; Yao et al. 2008). The first nor’easter-related precipitation is defined as the first 0.5 mm (~0.02 inch) precipitation reading from the New Jersey Weather and Climate Network (Robinson 2005). A smaller threshold is not used so that we avoid capturing isolated showers well ahead of the primary precipitation shield. A New Jersey-centric approach was chosen due to its high population density (461.6/km$^2$), significant contribution ($473 billion) to the U.S. gross domestic product, and its central location in the Mid-Atlantic (United States Census Bureau 2012).

To investigate the importance of ocean-atmosphere coupling, three model runs were completed for each case within the COAWST modelling system: a WRF-only run (CW34), a WRF run utilizing the Pollard 1D mixed-layer model (CWPol), and a fully coupled WRF-ROMS simulation (CWR). All simulations were initialized identically, only varying by the ocean coupling option. Specifically, CW34 fixes SSTs, CWPol updates SST and mixed layer depth
(200 m initially) every time step, and CWR exchanges ocean and atmosphere data between the component models every ten minutes as in W10.

e) **Verification and analysis techniques**

Atmospheric model validation compared WRF model output to both GMA and 4-km resolution Stage IV precipitation data (Lin and Mitchell 2005). GMA data validated all model output (except precipitation) because of its extensive coverage, especially in data-sparse regions. Stage IV validated simulated precipitation because it is based from rain gauge and radar data, is gridded, and offers high resolution. Prior to any validations, all data were interpolated to the coarsest grid spacing. Given our atmospheric focus, ROMS model data was not validated, but we may do so in future research.

Model output analysis is comprised of several parts. First, skin temperatures (hereafter, TSK) and their differences are analyzed; TSK is essentially SST at oceanic grid points. SST values are not compared because CWPol simulations do not update it and CWR simulations overwrite it with RTGSST data at lower-boundary update times. Storm track and intensity were derived from local minima in sea-level pressure (SLP) via an objective, self-coded algorithm similar to that used at the Climate Prediction Center (Serreze 1995; Serreze et al. 1997). Using these tracks, minimum SLP, storm maximum wind speed, and average TSK change were calculated and compared to GMA. Next, precipitation patterns and their distribution were evaluated against Stage IV data and validated using bias and threat score calculations. Overall accuracy of the local-storm environment (i.e., within a 600-km wide, storm-centered box that captures the storm and minimizes background contamination) and the large-scale environment (i.e., entire model domain and all times) was evaluated using the non-hydrostatic, moist, total
energy norm (Kim and Jung 2009). Energy norm integrations were capped at ~100 hPa to limit 
large temperature errors near the model top.

\[
E_m = \iiint_{x,y} \frac{1}{2} \left[ u'^2 + v'^2 + w'^2 + \left( \frac{g}{N_r \theta_r} \right)^2 \theta'^2 + \left( \frac{1}{\rho_r c_s} \right)^2 p'^2 + \omega_q \frac{L^2}{c_p T_r} q'^2 \right] dx dy d\sigma
\]  

(1)

In (1), \(E_m\) is the moist total energy norm (J m\(^2\) kg\(^{-1}\)); \(u', v', \text{ and } w'\) are the zonal, 
meridional, and vertical wind perturbations (m s\(^{-1}\)), respectively; \(p'\) is the pressure perturbation 
(Pa); \(\theta'\) is the potential temperature perturbation (K); \(q'\) is the mixing ratio perturbation (kg kg\(^{-1}\)). 
\(N_r, \theta_r, \rho_r, T_r, \text{ and } c_s\) are the reference Brunt Väisälä frequency (0.0124 s\(^{-1}\)), reference potential 
temperature (270 K), reference air density (1.27 kg m\(^{-3}\)), reference air temperature (270 K), and 
speed of sound (329.31 m s\(^{-1}\)), respectively. Finally, \(c_p\) is the specific heat at constant pressure 
(1005 J kg\(^{-1}\) K\(^{-1}\)), \(g\) is the gravitational constant (9.8 m s\(^{-2}\)), and \(\omega_q\) is a scaling factor (0.1).

Local-storm environment energy norm calculations were split into two frameworks: 
model-centered and GMA-centered. The former centers the energy norm integration box at the 
model-indicated cyclone center location, whereas the latter integrates only at GMA-indicated 
nor’easter locations. For the large-scale environment analysis, the energy norm was calculated 
for each model domain and was complemented with root mean square error (RMSE) calculations 
of SLP, 850-hPa temperature, 500-hPa geopotential height, 300-hPa winds, and potential 
temperature on the dynamic tropopause (defined as the 2-PVU surface; 1 PVU = \(10^{-6}\) K m\(^2\) kg\(^{-1}\) 
s\(^{-1}\)). The first four variables were referenced in Kocin and Uccellini (2004) for nor’easter 
analysis. Our decision to focus on the energy norm was influenced by Buizza et al. (2005), who 
made a compelling case for its usage at ECMWF for model validation given its total model 
volume integration, lack of single-layer sensitivity, and inclusion of temperature, wind, pressure, 
and moisture errors. For both RMSE and the energy norm, smaller values denote less error.
3. Results

a) Simulated skin temperature comparison

Figure 3 displays TSK values from CW34 simulations at forecast hour 48 and the TSK differences for CWPol and CWR between forecast hour 48 and the end of the simulation (forecast hour 180) for cases 2, 4, and 7. These cases illustrate TSK variability during and following an early, mid, and late season nor’easter event and dramatically illustrates the greater TSK variability associated with 3D ocean-atmosphere coupling versus 1D coupling. Higher TSK differences in 3D ocean-atmosphere coupling is a product of shifts in the Gulf Stream and the loop current, wind-driven upwelling, and submerged warm water pockets as noted by Ren et al. (2004). To quantify how TSK varies in domain 2, Table 3 shows ocean TSK differences for all seven cases from GMA, CWPol, and CWR; differences for CW34 are exactly 0 K. Largely because the Gulf Stream’s location remains fixed (only temperature changes), average CWPol TSK differences in Table 2 vary by up to 1.18 K less than in CWR. Overall mean TSK differences vary up to 1.41 K (CWR, Case 1). Table 2 and Figure 3 also indicate a degree of seasonality in TSK differences. Early-season cases (Oct-Jan) favor stronger overall cooling, whereas late-season cases (Feb-Apr) favor less cooling and even warming. These trends are not noted in previous nor’easter modeling studies due to their limited case or time focus and are associated with seasonal mixed-layer depth (shallower mixed-layer depth early in the season; Ren et al. 2004) and solar heating variations. Notably, these trends are better handled by CWR than in CWPol because the former includes bathymetry rather than a set ocean depth (200 m), allows for advection between neighboring grid cells, and includes precipitation-based cooling.

b) Nor’easter track and property analysis
Figure 4 displays WRF- (colored lines) and GMA- (black lines) based storm tracks for each case. Figure 5 displays GMA-relative track errors (positive y-axis = GMA cyclone propagation direction) every six hours (small symbols, colors) and their mean (larger symbols, black) only when all simulations contained the nor’easter concurrently within domain 2. Average track error for all seven cases (km) is shown in Table 4. Figure 5 shows a wide spread (up to 490 km) in average track error between cases, yet the track variability between experiments for each particular case remains fairly consistent (differences less than 100 km), which is consistent with previous studies (Ren et al. 2004; Yao et al. 2008; W10). Following from Fig. 5, Table 3 shows that no experiment exhibits a definitive (>60%) GMA-relative track bias over all 70 storm track points (all cases). Even on a case-by-case basis, the track biases of the coupled runs only vary from CW34 at five (CWPol) and four (CWR) track points, and each is associated with SST variations along and near the Gulf Stream. Despite lacking definitive GMA-relative track biases, Fig. 5 does demonstrate that the accuracy and precision of storm tracks did vary between cases. For Cases 1, 4, and 6, which had relatively low track error, each simulated nor’easter formed in a region of stronger differential cyclonic vorticity advection (CVA), whereas the other cases had higher track errors, and these nor’easters (especially Case 7) underwent cyclogenesis in regions of less concentrated differential CVA.

In addition to average storm track error, Table 4 also displays various nor’easter properties including the lowest minimum SLP, highest sustained 10-m wind within 100 km of the cyclone center, and average TSK difference within 100 km of the cyclone center along the storm track for GMA and all 3 experiments. We compare each model to GMA because no in-situ dataset exists for the data sparse regions of the Western Atlantic. With the exception of Cases 2 and 5, simulated nor’easter intensity fell within 5 hPa of GMA. Notably for these two cases the
nor’easter developed ahead of a 500-hPa cutoff where differential CVA was weak as shown in Figure 6 with 500-hPa geopotential height and relative vorticity fields from Case 2. In GMA (not shown), weaker differential CVA values result in higher minimum SLP values. Relative to CW34, CWPol and CWR-based SLP values are all within 5 hPa of each other with the exception of Case 3. Further analysis of Case 3 revealed a region of enhanced surface baroclinicity along the stronger SST gradient and a region of enhanced latent heat flux just west of the cyclone center and co-located with the SST gradient (Fig. 6). For CW34 or CWPol both the baroclinicity and latent heat fluxes near the cyclone are weaker which likely explains why CWR produced a stronger cyclone. Despite the enhanced SLP falls for CWR in Case 3, for no case did the maximum 10 m wind between WRF simulations vary by more than 2 m s\(^{-1}\). Finally, similar to the SST differences averaged over domain 2, SST differences along each storm track also indicate a seasonal cycle, albeit generally more muted (lower magnitude) given the strong warm and cold SST pockets associated with the Gulf Stream.

\[ d) \text{Error analysis of the local-storm environment} \]

Evaluation of the local-storm environment will focus exclusively on the energy norm (when storm center is at least 300 km from the domain 2 boundary) and will contain two parts: Averaged energy norm total and its component parts at each model level (Fig. 7) and the vertical integrated energy norm at each relevant time (Fig. 8). Figure 7 shows the total energy norm (black) and its component parts (colors) for CW34 and then the percent change in the CW34 represented by the difference between it and CWPol and CWR for Cases 2–5. In Figure 7, negative percentage values denote a decrease in the energy norm. Note that model error is dominated by the horizontal wind (U, V), temperature (T), and moisture (Q) components of the
energy norm. These four cases were selected because they best illustrate situations where model coupling reduced model error the most. These cases also cover four separate months. Among the missing cases, model coupling reduces error slightly in Cases 1 and 7, but Case 6 is slightly worse on average. In Figure 7, model coupling is most beneficial to simulated water vapor (yellow line) and vertical velocity (brown line) values with up to a 59% and a 20% decrease in the CW34 energy norm for a given level, respectively. Consistent with Ren et al. (2004) and Yao et al. 2008), ocean-atmosphere coupling leads to simulated differences throughout the entire troposphere.

Figure 8 shows the percentage change in the CW34 energy norm of the non-CW34 runs in both the GMA- (thin lines) and model-centered (thick lines) frameworks for all cases. In Fig. 8, a -5% value indicates a non-CW34 experiment has an energy norm 5% less than CW34, and data gaps represent times when the data integration box extended beyond the domain 2 boundary for one or more experiments. CWPol and CWR energy norm differences range from 40% lower (CWPol, Case 7) to 40% higher (CWPol, Case 2) than the W34 energy norm and normally fall within 10% of the CW34 energy norm value. Taken as a whole, the average CWR energy norm is smaller than CW34 in all 7 cases and CWPol in 6 of 7 cases. Despite this success, on a time-by-time basis and combining all cases (64 total times), the bottom of Table 5 shows CWR only had model-relative energy norm values less than CW34 in 36 (56.3%) time periods, and less than CWPol in 40 (62.5%) time periods. In comparison, CWPol has energy norm values lower than CW34 in 32 (50.0%) time periods. Thus CWR simulations do tend to give slightly better forecasts on average and in most situations, but two-tailed T-tests did not reveal these energy norm differences as statistically significant either overall or for any case (minimum p-value 0.84;
CWR/CW34 comparison). Perhaps error in the GMA itself, which does not involve a fully-coupled ocean, penalizes the CWR enough in this metric to prevent statistical significance.

e) Stage IV precipitation analysis

WRF-simulated precipitation fields and their accuracy are strongly linked to storm track, forecast hour, and to a lesser degree ocean-atmosphere coupling. Figure 9 shows 72-hour precipitation accumulations (forecast hours 48 – 120) from Stage IV and CW34, CWPol, and CWR 72-hour precipitation accumulations differenced from CW34, and the precipitation probability density function (PDF) and cumulative distribution function (CDF) of 72-hour accumulated precipitation for Stage IV and all models from Case 4 and 6. These two cases were selected for their low storm track error and relatively minimal track error spread. To supplement Fig. 9 and reference the other five cases, Table 6 shows the bias and threat scores for all seven cases assuming a 12.5 mm (0.5”) reference value. We exclude Cases 2 and 3 from consideration because either the GMA or WRF model storm tracks lie outside the spatial coverage of Stage IV data, which results in extreme biases and low forecast skill.

Focusing on the remaining 5 cases, all WRF simulations tend to similarly overextend their precipitation extent from 1.17 (Case 7, CW34) to 2.13 (Case 4, CWPol) times that shown in Stage IV data. Threat score values show low (0.31, Case 7 CWR) to moderate (0.59, Case 5 all experiments) success in producing precipitation in the same locations as Stage IV. Of the five cases, only in Case 4 does both the bias and skill score improve slightly (relative to CW34) with model coupling. As seen in Fig. 9, this improvement is quite subtle and is visible as a small bump in the 10-15 mm bin on the CWR CDF curve and small reductions (10 mm) in precipitation in Eastern Massachusetts, near Washington D.C. and east of the North Carolina
coast. Figure 10 shows that these reductions are likely associated with changes within the lower troposphere given the similarity of the 500 hPa geopotential height and lower concentrations of differential CVA and other upper-level fields (i.e., 300 hPa winds and geopotential height). Instead cooler SSTs along the Mid-Atlantic coastline and a slight cooling (1-2K) of the warm water tongue off North Carolina decreased latent heat flux (50-100 W m\(^{-2}\)) compared to nearby areas (See “*” in Fig. 10), which increased local stability and in turn precipitation. The relatively minor improvements in QPF seen here are consistent with prior work that has shown QPF to be one of the most difficult forecast aspects for models to handle in a variety of situations (e.g., Fritsch and Carbone 2004; Wang and Clark 2010).

f) Error analysis over all WRF domains and times

Similar to Fig. 8, Fig. 11 displays the percentage change in the energy norm for the coupled experiments relative to CW34 (negative values denote reduced model error) for Cases 4 and 6 on all three model domains. These two cases represent situations with high (Case 4) and low (Case 6) energy norm variability. Until cyclogenesis occurs, energy norm values tend to vary negligibly (< 1%) due to a lack of strong surface wind stress and precipitation to drive SST changes. Once generated, energy norm differences vary more rapidly (< 5%) and sharply spike (up to 12%) when the nor’easter leaves the model domain due to varying fractions of the storm existing within the domain of interest until all simulations no longer contain the nor’easter within the model domain. As seen in Fig. 11, the percent change in the CW34 energy norm also tended to increase with high resolutions because the energy norm was summed over a smaller spatial extent so that small position errors in better resolved updrafts and banded structures associated with the nor’easter contributed more significantly to the energy norm. Amongst the seven cases,
3D ocean-atmosphere coupling reduced model error in all model domains for at least four (domain 3) and up to six (all but Case 6) out of seven cases. As compared to CWPol, CWR energy norm values were smaller in at least five (all but Cases 5 and 6) of the seven cases. CWPol energy norm values also compared favorably (4 or more cases with a lower energy norm) in all but domain 3, where CWPol energy norm values were often within 5% of CW34 as illustrated by Fig. 11, but more often than not slightly higher than CW34.

To help quantify whether ocean-atmosphere coupling generally benefitted the entire model simulation, Table 5 displays the total number and percentage of all 217 six-hour periods (31 six-hour periods per case, 7 cases) where the shown inequality is valid for both the energy norm and five RMSE calculations. If only considering the energy norm, 3D ocean model coupling certainly produces more accurate nor’easter simulations than both CW34 and CWPol for domains 1 and 2 and arguably for domain 3 too. For the former, more than 62.7% of all 217 analysis times periods exhibit a lower energy norm from CWR simulations than either CW34 or CWPol. For the latter, this value drops to 51.6% of all 217 analysis times. This may be a function of model resolution and also a consequence of SST values not being passed from domain 2 to domain 3 where far coarser RTGSST data would be used. In comparison, CWR RMSEs show considerably weaker tendencies where regions above 500 hPa generally show a marginal to slight benefit (50.2%-60.8%) and a marginal to slightly negative (36.9%-48.4%) response for 850-hPa temperature and SLP. In comparison to CWR, 1D ocean model coupling produces energy norm and RMSE results that are comparatively less variable than CWR, but only two RMSE tests (Domain 2, 300 hPa winds and Domain 1, 850-hPa temperatures) show any distinct benefit (> 55%) compared to the uncoupled simulation. As is, the RMSE results indicate that 1D ocean model coupling has a greater benefits than 3D ocean model coupling in the lower
atmosphere, yet in the mid and upper-atmosphere this situation reverses. However, overall model error (as indicated by the energy norm) indicates that 3D ocean model coupling produces better overall simulations than either 1D ocean model coupling or an uncoupled model. Despite the above results, two-tailed T-tests did not reveal any significant variability amongst the energy norm or any of the RMSE tests (minimum p-value 0.73; CWR/CWPol domain 3 SLP).

4. Conclusions

An investigation into the benefit of ocean-atmosphere coupling and its potential seasonality was evaluated using seven sets of nor’easter simulations. Each set was comprised of an uncoupled WRF (CW34), a coupled 1D ocean mixed layer model (CWPol), and a coupled 3D ocean model (CWR) experiment. All experiments were configured and initialized identically (except for ocean coupling option) and ran for 180 hours starting roughly 72 hours prior to the first nor’easter-related precipitation measurements in the Mid-Atlantic region. Model output was validated against Stage IV and GMA data. Area-averaged (100 km radius), CWR-simulated skin temperature differences varied up to 1.41 K and were in all but one case (Case 7) an order of magnitude larger than those from CWPol. CWR simulations (to a lesser extent CWPol) indicated a degree of seasonality in TSK differences where early-season cases (Oct-Jan) favor stronger overall cooling, whereas late-season (Feb-Apr) favor less cooling and even warming. This seasonality is consistent with GMA, and also matches expectations based on the shallower mixed layer observed early in the cold season (Ren et al. 2004).

Although TSK differences could exceed 1K, ocean model coupling did not produce significant inter-model changes to simulated storm track (< 100 km), maximum 10 m wind speed (<0.79 m s⁻¹), storm intensity (± 5 hPa, exception CWR Case 3, 5.25 hPa) or average, storm-
track relative TSK (< 1.27 K) or produce any distinctive track bias tendencies. These results are consistent with Ren et al. (2004), Yao et al. (2008), and W10. Simulated precipitation fields showed little sensitivity to both 1D and 3D ocean model coupling, but bias and threat scores did show that WRF tended to over predict precipitation extent and have low (0.31) to moderate (0.59) model forecasting skill. Similar to previous modelling studies (i.e., Ridout et al. 2005; Dravitzki and McGregor 2011), most WRF simulations generated excess light precipitation and a dearth of heavy precipitation with the notable exception of Cases 2 and 5 where differential CVA aloft was weak.

Analysis of the local-storm environment and the overall simulation revealed the main sources of simulation error are associated with temperature and horizontal wind fields both within the planetary boundary layer and in the upper troposphere. With the exception of Case 3, CWR simulations generated slightly lower energy norm values (20% reduction at a level) throughout the troposphere compared to corresponding CWPol simulations. Such reductions shed light upon how the error in CWR simulations decreased from CW34 in at least 34 out of 64 (56.3%) times evaluated in both the model and GMA-relative frameworks. RMSE and energy norm analysis of the entire simulation demonstrated that ocean-atmosphere coupling does slightly-to-moderately (50.1% - 71.9% of all analysis times) benefit from some degree of ocean model coupling. Benefits stemming from ocean-atmosphere coupling are not clear cut because neither CWR nor CWPol show clear improvements in the RMSE calculations; however, energy norm results do clearly show CWR to have the best overall model simulation. Despite this claim, the overall impact to simulation outcome is modest given that simulation differences are shown to be non-significant (minimum p-value 0.73; CWR/CWPol domain 3 SLP).
This study has investigated the merits of ocean-atmosphere coupling, compared 1D to 3D ocean modelling coupling, and determined whether any seasonal dependencies in SST differences exist. Although model coupling lacked any notable impact upon storm track, storm intensity, storm maximum winds, and statistically significant RMSE and energy norm differences were not found, atmosphere-ocean coupling still does produce small, but visible differences in temperature, moisture, and precipitation, among other dynamical and physical fields. Despite these relatively meager differences, both 3D and to a limited extent 1D ocean model coupling did show a seasonal dependence where SSTs more strongly cooled and warmed depending upon the time of year. Because SST differences from CWPol simulations tended to be an order of magnitude smaller than CWR simulations, this does underscore the importance of advection and realistic bathymetry for realistic SST simulations. Even with non-significant error analysis results (both near the storm and throughout the domains) this study still demonstrates that ocean-atmosphere coupling has tangible benefits (albeit small) to nor’easter simulations by reducing model error. What remains less clear is the importance of 3D ocean modelling. This study has shown that CWR simulations did produce lower overall error, yet with a 3 day lead time, its simulations were not much improved upon output from the 1D ocean model. Thus it is likely that with longer model integration times the impacts from 3D ocean model coupling would become more apparent and crucial for model simulations.

5. Future Work

This study investigated how ocean-atmosphere coupling impacts nor’easter simulations. Several expansions are possible. First, including the COAWST wave and sedimentation models (i.e., SWAN and CSTMS) would allow for further alterations in surface heat fluxes and wave-
based variability in surface roughness values. These changes may lead to more realistic precipitation and cloud cover simulations. Second, these same simulations could be re-examined with the recently released (Oct 2014) COAWST rev 828 or newer, which now includes multiple-nest coupling. Finally, this analysis could be expanded through comparing ROMS simulated currents to Coastal Ocean Dynamics Applications Radar (CODAR; Paduan et al. 2004) and outgoing longwave radiation data to ex-situ data from Clouds and Earth's Radiant Energy System (CERES; Wielicki et al. 1996).

6. Acknowledgements

This research was funded in part by the New Jersey Agricultural Experiment Station. We acknowledge both Dr. John Warner of the Woods Hole Oceanographic Institute and Dr. John Wilkin of the Institute of Marine and Coastal Sciences at Rutgers University for their invaluable and timely assistance. We also wish to thank the NASA Graduate Student Summer Program for the learning experience afforded by the first author’s internship at NASA Goddard Space Flight Center. We also acknowledge Dr. Wei-Kuo Tao of NASA Goddard Space Flight Center who generously supplied computational time for our research.
7. References


Mathematics and Computer Science Division, Argonne National Laboratory, 25 pp.


**Fig. 1:** All available COAWST configurations involving the WRF model and exchanged data fields. (A) WRF only, (B) WRF-ROMS coupling, (C) WRF-ROMS, ROMS-SWAN coupling, and (D) WRF-ROMS-SWAN-CSTMS coupling. Adapted from Fig. 5 of Warner et al. (2010).

**Fig. 2:** (a) WRF model configuration where domains 1, 2, and 3 have 45, 15, and 5 km grid spacing, respectively. (b) ROMS model configuration superimposed on WRF model domain 2 with 8 km average grid spacing.

**Fig. 3:** (top) Sea surface temperatures (SST; K) from CW34 at forecast hour 48 and (bottom) SST differences (180 hour - 48 hour SST) for CWPol and CWR from Cases 2, 4, and 7.

**Fig. 4:** Storm tracks from GMA and the model runs. Line legend is shown on the upper-left of each plot. Shown symbols indicate simulated storm position every six hours. White numbers indicate case number.

**Fig. 5:** WRF forecasted storm position bias as compared to GMA. The positive Y-axis in each panel denotes the GMA-relative storm motion direction. Shown symbols represent WRF position bias every six hours (smaller symbols, color) and their mean (large symbols, black). Case number is indicated with a white number.

**Fig. 6:** (top) 500 hPa positive relative vorticity (fills, $10^{-5}$ s$^{-1}$) and 500 hPa geopotential height (contours, dam) from Case 2 on 06 UTC 8 Nov. (middle) Sea surface temperatures (K, fills) from Case 3 on 18 UTC 20 Dec. 2009. (bottom) Latent heat release at the surface (fills, W m$^{-2}$) and sea-level pressure (contours, hPa) from Case 3 on 18 UTC 20 Dec. 2009.
Fig. 7: (left column) Model-relative energy norm values at each model sigma level for CW34 for cases 2, 3, 4, and 5 averaged with 600 km of the cyclone center and only when all models and GMA contained the cyclone. (center and right columns) Difference in model-relative energy norm (model – CW34) for CWPol and CWR, respectively. Shown values indicate the total energy norm (black) and each of the six components comprising the energy norm (colors) noted in equation 1.

Fig. 8: Percentage change in domain 2 energy norm values as compared to CW34 within 300 km of the GMA storm center (thin lines) and each model simulated storm center (thick lines). Negative values denote reduced model error as compared to CW34. Times on x-axis denote day and hour. Missing data denotes times where at least one of the model runs or GMA is within 300 km of the domain 2 boundary.

Fig. 9: (top) 72-hour total precipitation accumulation (mm; forecast hours 48 - 120) from Stage IV and CW34. (middle) 72-hour accumulated precipitation from CWPol and CWR differenced from CW34 (mm). (bottom) Probability density function (PDF) and cumulative distribution function (CDF) of 72-hour accumulated precipitation for Stage IV and all models. Left-hand panels are for Case 4 and right-hand panels are for Case 6.

Fig. 10: (top) Sea surface temperatures (K, fills) and sea-level pressure (hPa, contours), (middle) Latent heat flux at the surface (fills, W m-2) and sea-level pressure (contours, hPa), and (bottom) 500 hPa positive relative vorticity (fills, 10^{-5} s^{-1}) and 500 hPa geopotential height (contours, dam) from Case 4 on 18 UTC 27 Jan 2015.

Fig. 11: Percentage change in energy norm values relative to CW34 within the entirety of domains 1, 2, and 3 for Cases 4 and 6. Negative values denote error reduction as compared to CW34 energy norm values. Times on x-axis denote day and hour.
Table 1: Nor’easter case list. Column 3 lists the dates when each event impacted the Mid-Atlantic region; the last two columns denote the first and last times of each model simulation. All cases are sorted by seasonal order (i.e., by month and day, but not year).

<table>
<thead>
<tr>
<th>Case Number</th>
<th>NESIS</th>
<th>Event Dates</th>
<th>Model Run Start Date</th>
<th>Model Run End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N/A</td>
<td>15-16 Oct 2009</td>
<td>10/12 12UTC</td>
<td>10/20 00UTC</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
<td>07-09 Nov 2012</td>
<td>11/04 06UTC</td>
<td>11/11 18UTC</td>
</tr>
<tr>
<td>3</td>
<td>4.03</td>
<td>19-20 Dec 2009</td>
<td>12/16 06UTC</td>
<td>12/23 18UTC</td>
</tr>
<tr>
<td>4</td>
<td>2.62</td>
<td>26-28 Jan 2015</td>
<td>01/23 00UTC</td>
<td>01/30 12 UTC</td>
</tr>
<tr>
<td>5</td>
<td>4.38</td>
<td>04-07 Feb 2010</td>
<td>02/02 18UTC</td>
<td>02/10 06UTC</td>
</tr>
<tr>
<td>6</td>
<td>1.65</td>
<td>01-02 Mar 2009</td>
<td>02/26 12UTC</td>
<td>03/06 00UTC</td>
</tr>
<tr>
<td>7</td>
<td>N/A</td>
<td>12-14 Mar 2010</td>
<td>03/09 06UTC</td>
<td>03/16 18UTC</td>
</tr>
</tbody>
</table>
Table 2: Surface skin temperature (TSK) differences (K, forecast hour 180 – forecast hour 48) for all seven cases for GMA, CWPol and CWR. In CW34, TSK is fixed over the ocean, so all TSK differences in that experiment are 0K. Differences are averaged over all ocean points in Domain 2.

<table>
<thead>
<tr>
<th>ΔTSK (K)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMA</td>
<td>-0.85</td>
<td>-0.78</td>
<td>-0.55</td>
<td>-0.6</td>
<td>-0.34</td>
<td>0.22</td>
<td>0.48</td>
</tr>
<tr>
<td>CWPol</td>
<td>0</td>
<td>0.04</td>
<td>-0.09</td>
<td>0.01</td>
<td>0</td>
<td>-0.01</td>
<td>0.24</td>
</tr>
<tr>
<td>CWR</td>
<td>-1.41</td>
<td>-0.83</td>
<td>-1.17</td>
<td>-0.55</td>
<td>-0.5</td>
<td>-0.26</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Table 3: GMA-relative storm track bias analysis. Values indicate the percentage of total six-hour time periods (out of 70 total, includes all cases) where each simulation exhibited the shown GMA-relative track bias. Biases were only calculated for each case when all three model simulations and GMA contained a nor’easter concurrently.

<table>
<thead>
<tr>
<th>Key</th>
<th>Lead Left</th>
<th>Lead Right</th>
<th>Lag Left</th>
<th>Lag Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>CW34</td>
<td>25.7%</td>
<td>21.4%</td>
<td>34.3%</td>
<td>21.4%</td>
</tr>
<tr>
<td>CWPol</td>
<td>25.7%</td>
<td>22.9%</td>
<td>32.9%</td>
<td>22.9%</td>
</tr>
<tr>
<td>CWR</td>
<td>17.1%</td>
<td>22.9%</td>
<td>22.9%</td>
<td>21.4%</td>
</tr>
</tbody>
</table>
Table 4: Various nor’eastern characteristics and their differences. Shown values represent the minimum overall sea-level pressure (SLP, hPa), maximum overall sustained 10 m wind (m s\(^{-1}\)) within 100 km of the cyclone center, average sea surface temperature (SST) difference (K) within 100 km of the cyclone center, and average GMA-relative storm track error (km). Bolded values denote differences in SLP, maximum wind, SST, and track error greater than or equal to 5 hPa, 2 m s\(^{-1}\), 1K, and 300 km, respectively. All difference values are as indicated below.

<table>
<thead>
<tr>
<th>GMA</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min SLP (hPa)</td>
<td>991.38</td>
<td>989.12</td>
<td>972.89</td>
<td>980.18</td>
<td>979.54</td>
<td>1000.4</td>
<td>993.94</td>
</tr>
<tr>
<td>Max Wind (m s(^{-1}))</td>
<td>11.59</td>
<td>11.89</td>
<td>13.69</td>
<td>23.66</td>
<td>11.89</td>
<td>7.07</td>
<td>10.73</td>
</tr>
<tr>
<td>Avg SST diff (K; 180 hr - 48 hr)</td>
<td>-0.73</td>
<td>-0.32</td>
<td>-0.35</td>
<td>-0.55</td>
<td>-0.72</td>
<td><strong>-1.06</strong></td>
<td>0.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CW34</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min SLP (hPa)</td>
<td>993.99</td>
<td>976.56</td>
<td>975.51</td>
<td>983.93</td>
<td>986.51</td>
<td>998.13</td>
<td>998.29</td>
</tr>
<tr>
<td>Avg Diff SLP (CW34 - GMA)</td>
<td>2.61</td>
<td>-12.56</td>
<td>2.62</td>
<td>3.75</td>
<td><strong>6.97</strong></td>
<td>-2.27</td>
<td>4.35</td>
</tr>
<tr>
<td>Max Wind (m s(^{-1}))</td>
<td>8.33</td>
<td>10.92</td>
<td>12.36</td>
<td>23.66</td>
<td>7.01</td>
<td>6.66</td>
<td>11.2</td>
</tr>
<tr>
<td>Diff Max Wind (CW34 - GMA)</td>
<td><strong>-3.26</strong></td>
<td>-0.97</td>
<td>-1.33</td>
<td>0</td>
<td><strong>-4.88</strong></td>
<td>-0.41</td>
<td>0.47</td>
</tr>
<tr>
<td>Avg SST diff (K; 180 hr - 48 hr)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Avg Track Error (km)</td>
<td><strong>328</strong></td>
<td><strong>653</strong></td>
<td>284</td>
<td>179</td>
<td><strong>417</strong></td>
<td>290</td>
<td><strong>525</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CWPol</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min SLP (hPa)</td>
<td>993.1</td>
<td>976.91</td>
<td>975.51</td>
<td>984.53</td>
<td>986.22</td>
<td>998.424</td>
<td>998.75</td>
</tr>
<tr>
<td>Avg Diff SLP (CWPol - GMA)</td>
<td>1.72</td>
<td>-12.21</td>
<td>2.62</td>
<td>4.35</td>
<td><strong>6.68</strong></td>
<td>-1.976</td>
<td>4.81</td>
</tr>
<tr>
<td>Avg Diff SLP (CWPol - CW34)</td>
<td>-0.89</td>
<td>0.35</td>
<td>0</td>
<td>0.6</td>
<td>-0.29</td>
<td>0.294</td>
<td>0.46</td>
</tr>
<tr>
<td>Max Wind (m s(^{-1}))</td>
<td>8.47</td>
<td>10.98</td>
<td>12.31</td>
<td>22.88</td>
<td>7.8</td>
<td>6.66</td>
<td>11.35</td>
</tr>
<tr>
<td>Diff Max Wind (CWPol - GMA)</td>
<td><strong>-3.12</strong></td>
<td>-0.91</td>
<td>-1.38</td>
<td>-0.78</td>
<td><strong>-4.09</strong></td>
<td>-0.41</td>
<td>0.62</td>
</tr>
<tr>
<td>Diff Max Wind (CWPol - CW34)</td>
<td>0.14</td>
<td>0.06</td>
<td>-0.05</td>
<td>-0.78</td>
<td>0.79</td>
<td>0</td>
<td>0.15</td>
</tr>
<tr>
<td>Avg SST diff (K; 180 hr - 48 hr)</td>
<td>-0.02</td>
<td>-0.05</td>
<td>-0.09</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>Avg Track Error (km)</td>
<td><strong>339</strong></td>
<td><strong>660</strong></td>
<td>289</td>
<td>183</td>
<td><strong>444</strong></td>
<td>255</td>
<td><strong>551</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CWR</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min SLP (hPa)</td>
<td>993.18</td>
<td>978.69</td>
<td>970.26</td>
<td>984.43</td>
<td>986.5</td>
<td>997.96</td>
<td>996.34</td>
</tr>
<tr>
<td>Avg Diff SLP (Model - GMA)</td>
<td>1.8</td>
<td>-10.43</td>
<td>-2.63</td>
<td>4.25</td>
<td><strong>6.96</strong></td>
<td>-2.44</td>
<td>2.4</td>
</tr>
<tr>
<td>Avg Diff SLP (Model - CW34)</td>
<td>-0.81</td>
<td>2.13</td>
<td><strong>-5.25</strong></td>
<td>0.5</td>
<td>-0.01</td>
<td>-0.17</td>
<td>-1.95</td>
</tr>
<tr>
<td>Max Wind (m s(^{-1}))</td>
<td>8.43</td>
<td>10.72</td>
<td>12.37</td>
<td>23.67</td>
<td>7.15</td>
<td>6.66</td>
<td>11.81</td>
</tr>
<tr>
<td>Diff Max Wind (CWR - GMA)</td>
<td><strong>-3.16</strong></td>
<td>-1.17</td>
<td>-1.32</td>
<td>0.01</td>
<td><strong>-4.74</strong></td>
<td>-0.41</td>
<td>1.08</td>
</tr>
<tr>
<td>Diff Max Wind (CWR - CW34)</td>
<td>0.1</td>
<td>-0.2</td>
<td>0.01</td>
<td>0.14</td>
<td>0</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>Avg SST diff (K; 180 hr - 48 hr)</td>
<td>-0.63</td>
<td><strong>-1.27</strong></td>
<td>-0.6</td>
<td>-0.25</td>
<td><strong>-1.07</strong></td>
<td>-0.6</td>
<td>0.46</td>
</tr>
<tr>
<td>Avg Track Error (km)</td>
<td><strong>349</strong></td>
<td><strong>663</strong></td>
<td><strong>304</strong></td>
<td>173</td>
<td><strong>357</strong></td>
<td>280</td>
<td><strong>590</strong></td>
</tr>
</tbody>
</table>
Table 5: Error analysis summary for all three WRF model domains (D1, D2, D3) and the storm-performance analysis energy norm results (GMA and model-centered) summed over all cases. Displayed values indicate the number model output times (6-hour intervals) and the percentage of total times where the shown inequality is valid. Boldface denote instances of 50% or higher.

<table>
<thead>
<tr>
<th>D1 Performance Analysis (217 Times)</th>
<th>CWP pol &lt;= CW34</th>
<th>CWR &lt;= CW34</th>
<th>CWR &lt;= CWP pol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Norm</td>
<td>108 (49.8%)</td>
<td>142 (65.4%)</td>
<td>136 (62.7%)</td>
</tr>
<tr>
<td>RMSE 2 PVU Pot T</td>
<td>116 (53.5%)</td>
<td>119 (54.8%)</td>
<td>132 (60.8%)</td>
</tr>
<tr>
<td>RMSE 300-hPa Winds</td>
<td>112 (51.6%)</td>
<td>127 (58.5%)</td>
<td>120 (55.3%)</td>
</tr>
<tr>
<td>RMSE 500-hPa Geo Hght</td>
<td>108 (49.8%)</td>
<td>118 (54.4%)</td>
<td>117 (53.9%)</td>
</tr>
<tr>
<td>RMSE 850-hPa Temps</td>
<td>121 (55.8%)</td>
<td>103 (47.5%)</td>
<td>100 (46.1%)</td>
</tr>
<tr>
<td>RMSE SLP</td>
<td>99 (45.6%)</td>
<td>93 (42.9%)</td>
<td>99 (45.6%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D2 Performance Analysis (217 Times)</th>
<th>CWP pol &lt;= CW34</th>
<th>CWR &lt;= CW34</th>
<th>CWR &lt;= CWP pol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Norm</td>
<td>107 (49.3%)</td>
<td>152 (70.0%)</td>
<td>156 (71.9%)</td>
</tr>
<tr>
<td>RMSE 2 PVU Pot T</td>
<td>103 (47.5%)</td>
<td>113 (52.1%)</td>
<td>118 (54.4%)</td>
</tr>
<tr>
<td>RMSE 300-hPa Winds</td>
<td>116 (53.5%)</td>
<td>119 (54.8%)</td>
<td>109 (50.2%)</td>
</tr>
<tr>
<td>RMSE 500-hPa Geo Hght</td>
<td>100 (46.1%)</td>
<td>113 (52.1%)</td>
<td>120 (55.3%)</td>
</tr>
<tr>
<td>RMSE 850-hPa Temps</td>
<td>124 (57.1%)</td>
<td>100 (46.1%)</td>
<td>90 (41.5%)</td>
</tr>
<tr>
<td>RMSE SLP</td>
<td>114 (52.5%)</td>
<td>96 (44.2%)</td>
<td>80 (36.9%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D3 Performance Analysis (217 Times)</th>
<th>CWP pol &lt;= CW34</th>
<th>CWR &lt;= CW34</th>
<th>CWR &lt;= CWP pol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Norm</td>
<td>87 (40.1%)</td>
<td>112 (51.6%)</td>
<td>121 (55.8%)</td>
</tr>
<tr>
<td>RMSE 2 PVU Pot T</td>
<td>115 (53.0%)</td>
<td>113 (52.1%)</td>
<td>111 (51.2%)</td>
</tr>
<tr>
<td>RMSE 300-hPa Winds</td>
<td>114 (52.5%)</td>
<td>105 (48.4%)</td>
<td>103 (47.5%)</td>
</tr>
<tr>
<td>RMSE 500-hPa Geo Hght</td>
<td>94 (43.3%)</td>
<td>105 (48.4%)</td>
<td>109 (50.2%)</td>
</tr>
<tr>
<td>RMSE 850-hPa Temps</td>
<td>104 (47.9%)</td>
<td>97 (44.7%)</td>
<td>103 (47.5%)</td>
</tr>
<tr>
<td>RMSE SLP</td>
<td>107 (49.3%)</td>
<td>93 (42.9%)</td>
<td>84 (38.7%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D2 Storm Environment Analysis (64 Times)</th>
<th>CWP pol &lt;= CW34</th>
<th>CWR &lt;= CW34</th>
<th>CWR &lt;= CWP pol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Norm GMA Rel</td>
<td>29 (45.3%)</td>
<td>43 (67.2%)</td>
<td>43 (67.2%)</td>
</tr>
<tr>
<td>Energy Norm Model Rel</td>
<td>32 (50.0%)</td>
<td>36 (56.3%)</td>
<td>40 (62.5%)</td>
</tr>
</tbody>
</table>
Table 6: Stage IV-relative precipitation biases and threat scores. The below values assume a threshold value of 12.5 mm (0.5") and assume Stage IV data to be the validation data. Bias scores > 1 denotes the spatial extent of precipitation to be too large. Threat score values denote the spatial and temporal accuracy of forecasted precipitation and ranges between 0 (no skill) and 1 (perfect coverage).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CW34</td>
<td>1.47</td>
<td>4.93</td>
<td>0.82</td>
<td>2.14</td>
<td>1.5</td>
<td>1.36</td>
<td>1.17</td>
</tr>
<tr>
<td>CWPol</td>
<td>1.47</td>
<td>4.93</td>
<td>0.82</td>
<td>2.13</td>
<td>1.5</td>
<td>1.36</td>
<td>1.17</td>
</tr>
<tr>
<td>CWR</td>
<td>1.53</td>
<td>4.98</td>
<td>0.87</td>
<td>2.09</td>
<td>1.5</td>
<td>1.37</td>
<td>1.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Threat Score</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CW34</td>
<td>0.42</td>
<td>0.16</td>
<td>0.29</td>
<td>0.35</td>
<td>0.59</td>
<td>0.55</td>
<td>0.32</td>
</tr>
<tr>
<td>CWPol</td>
<td>0.41</td>
<td>0.16</td>
<td>0.29</td>
<td>0.36</td>
<td>0.59</td>
<td>0.55</td>
<td>0.32</td>
</tr>
<tr>
<td>CWR</td>
<td>0.4</td>
<td>0.16</td>
<td>0.28</td>
<td>0.36</td>
<td>0.59</td>
<td>0.55</td>
<td>0.31</td>
</tr>
</tbody>
</table>
Fig. 1: All available COAWST configurations involving the WRF model and exchanged data fields. (A) WRF only, (B) WRF-ROMS coupling, (C) WRF-ROMS, ROMS-SWAN coupling, and (D) WRF-ROMS-SWAN-CSTMS coupling. Adapted from Fig. 5 of Warner et al. (2010).
Fig. 2: (a) WRF model configuration where domains 1, 2, and 3 have 45, 15, and 5 km grid spacing, respectively. (b) ROMS model configuration superimposed on WRF model domain 2 with 8 km average grid spacing.
Fig. 3: (top) Sea surface temperatures (SST; K) from CW34 at forecast hour 48 and (bottom) SST differences (180 hour - 48 hour SST) for CWPol and CWR from Cases 2, 4, and 7.
Fig. 4: Storm tracks from GMA and the model runs. Line legend is shown on the upper-left of each plot. Shown symbols indicate simulated storm position every six hours. White numbers indicate case number.
**Fig. 5:** WRF forecasted storm position bias as compared to GMA. The positive Y-axis in each panel denotes the GMA-relative storm motion direction. Shown symbols represent WRF position bias every six hours (smaller symbols, color) and their mean (large symbols, black). Case number is indicated with a white number.
Fig. 6: (top) 500 hPa positive relative vorticity (fills, $10^{-5} \text{ s}^{-1}$) and 500 hPa geopotential height (contours, dam) from Case 2 on 06 UTC 8 Nov. (middle) Sea surface temperatures (K, fills) from Case 3 on 18 UTC 20 Dec. 2009. (bottom) Latent heat release at the surface (fills, W m$^{-2}$) and sea-level pressure (contours, hPa) from Case 3 on 18 UTC 20 Dec. 2009.
Fig. 7: (left column) Model-relative energy norm values at each model sigma level for CW34 for cases 2, 3, 4, and 5 averaged with 600 km of the cyclone center and only when all models and GMA contained the cyclone. (center and right columns) Difference in model-relative energy norm (model – CW34) for CWPol and CWR, respectively. Shown values indicate the total energy norm (black) and each of the six components comprising the energy norm (colors) noted in equation 1.
Fig. 8: Percentage change in domain 2 energy norm values as compared to CW34 within 300 km of the GMA storm center (thin lines) and each model simulated storm center (thick lines). Negative values denote reduced model error as compared to CW34. Times on x-axis denote day and hour. Missing data denotes times where at least one of the model runs or GMA is within 300 km of the domain 2 boundary.
Fig. 9: (top) 72-hour total precipitation accumulation (mm; forecast hours 48 - 120) from Stage IV and CW34. (middle) 72-hour accumulated precipitation from CWPol and CWR differenced from CW34 (mm). (bottom) Probability density function (PDF) and cumulative distribution function (CDF) of 72-hour accumulated precipitation for Stage IV and all models. Left-hand panels are for Case 4 and right-hand panels are for Case 6.
Fig. 10: (top) Sea surface temperatures (K, fills) and sea-level pressure (hPa, contours), (middle) Latent heat flux at the surface (fills, W m$^{-2}$) and sea-level pressure (contours, hPa), and (bottom) 500 hPa positive relative vorticity (fills, 10$^{-5}$ s$^{-1}$) and 500 hPa geopotential height (contours, dam) from Case 4 on 18 UTC 27 Jan 2015.
Fig. 11: Percentage change in energy norm values relative to CW34 within the entirety of domains 1, 2, and 3 for Cases 4 and 6. Negative values denote error reduction as compared to CW34 energy norm values. Times on x-axis denote day and hour.