P2.543 IMPACT OF CYCLED 3DVAR ASSIMILATION OF COSMIC OBSERVATIONS ON NOR'EASTER SIMULATIONS

Stephen D. Nicholls* and Steven G. Decker
Department of Environmental Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

1. INTRODUCTION

The Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) mission is a joint operation of the National Space Organization (NSPO) of Taiwan and the University Corporation for Atmospheric Research (UCAR) that produces 1800+ Global Positioning System (GPS) radio occultation (RO) profiles daily. To achieve global coverage, the six polar-orbiting microsatellites (which make up COSMIC) measure and record the phase and amplitude of occulted GPS signals which are then used to generate vertical profiles of bending angle and refractivity. These profiles are then inverted via an Abel transform to generate vertical profiles of refractivity and bending angle as a function of height (Cucurull et al. 2008). Specifically, we can relate the refractivity to temperature, moisture, and pressure via

$$N = 77.6\left(\frac{P}{T}\right) + 3.73*10^5\left(\frac{P_v}{T^2}\right). \quad \textbf{(1)}$$

Eq. (1) was adapted from Cucurull et al. (2008). In this equation, N is refractivity, P is pressure (hPa), T is temperature (K), and P_v is the partial pressure of water vapor (hPa). Given this mathematical relationship, GPS RO data can be assimilated by a numerical weather prediction (NWP) models either directly or indirectly. For direct assimilation, raw refractivity values are directly ingested into the model and temperature, pressure, and moisture data is extracted via an adjoint model of (1). In contrast, indirect assimilation involves an intermediate model that pre-processes raw refractivity data and outputs pressure, temperature, and vapor pressure profiles, which are then later assimilated by the user's model.

To those in the atmospheric modeling community, the high potential of COSMIC in NWP stems from its global coverage and ability to acquire vertical temperature and moisture profiles in both data-dense and especially data-sparse regions. These qualities have motivated both ECMWF and NCEP to fund research studies into its application in their own in-house operational weather models (Healy and Thépaut 2006, Cucurull et al. 2008). Their results demonstrated a reduction in stratospheric temperature root mean square error (RMSE) and increased 200 hPa-temperature anomaly correlation scores for both the ECMWF and GFS models, respectively, once COMSIC was assimilated. Motivated by these results, our study will determine the impact of 3D cycled COSMIC RO data assimilation on Weather Research and Forecasting (WRF) Advanced Research WRF (WRF-ARW) model (hereafter shortened to WRF) simulations of eight intense, wintertime cyclone events (nor'easters) impacting the Northeastern United States (NEUS). Additionally, the relative importance of COSMIC RO data assimilation to the pre-cyclogenesis phase of our nor'easter simulations will involve a data denial experiment.

Corresponding author address: Stephen D. Nicholls, Rutgers University, 14 College Farm Rd., New Brunswick, NJ, 08901;
email: nicholls@envsci.rutgers.edu

2. METHODOLOGY

This WRF-ARW-based study will solely focus on nor'easters, which are themselves a byproduct of complex dynamical and thermodynamical interactions. Specifically for this study, we chose eight nor'easter cases that occurred between October to April of the years 2006 to 2010 (See Table 1). This period was selected for both its consistency with the nor'easter season defined by Jacobs et al. (2005) and because COSMIC has only been in operation since April 2006 (Anthes et al. 2008). By using the entire season, we aim to diversify the number of unique weather situations where a nor'easter later develops.

2.1 WRF MODEL SETUP AND CONFIGURATION

Each WRF model run was configured to allow it to simulate its assigned cyclone from its genesis through its passage east of Newfoundland. To provide focus on the NEUS and to make the best use of our computational resources, we chose a WRF model setup comprised of three nested domains (graphically shown in Figure 1) with two-way interaction and 27 vertical levels. Horizontal resolution on these three model domains were 45, 15, and 5 km for domains 1, 2, and 3, respectively. Our selected horizontal and vertical resolutions are consistent with our GFS 003 model forecast-derived input and boundary condition files, which have 28 vertical levels and 1-degree horizontal resolution. Finally, we selected a model top of 50 hPa (~20 km) to avoid known errors in WRF COSMIC RO data assimilation above 30 km (Cucurull et al. 2008).

To examine how COSMIC RO data assimilation affects WRF nor'easter simulations, we ran 40 WRF simulations. Our simulations included 8 control cases (WRF 3.2 with no COSMIC RO assimilation), and 32 experimental runs (WRF 3.2 run with 3D cycled COSMIC RO data assimilation [hereafter shorted to WRF-COSMIC]). All forty cases utilized the same WRF model parameterizations and GFS 003 model forecasts for their initializations. Our 32 WRF-COSMIC runs we sub-divided into 4 different WRF-COSMIC runs per case, and only varied with respect to the number of assimilation periods. For these runs, we assimilated all COSMIC RO data within a ± 1 hour window every three hours (for consistency with model input data) within a set number of assimilation periods (1, 9, 17, and 61 times used in this study). Combining the three-hour interval with the number of assimilation periods, this denotes a 0-hour, 24-hour, 48-hour, and 180-hour time duration, respectively. When referring to these cases in this work, we will reference each case by the number of COSMIC RO data assimilation hours.

Table 1: NESIS scores, date of impact in the Mid-Atlantic and Northeast United States, and model runtime specifics for the eight nor'easter cases studied. NESIS denotes the Northeast Snowfall Impact Scale (Kocin and Uccellini, 2004).

Case Number	NESIS	Dates of Event	Model Start	Model End
1	N/A	22-24 Nov 2006	11/19 12Z	11/27 00Z
2	2.55	15-17 Mar 2007	3/12 18Z	3/20 06Z
3	N/A	15-17 Apr 2007	4/12 06Z	4/19 18Z
4	1.65	1-2 Mar 2009	2/26 12Z	3/6 00Z
5	N/A	15-16 Oct 2009	10/12 12Z	10/20 00Z
6	4.03	19-20 Dec 2009	12/16 06Z	12/23 18Z
7	4.3	4-7 Feb 2010	2/2 18Z	2/10 6Z
8	N/A	12-14 Mar 2010	3/9 06Z	3/16 18Z

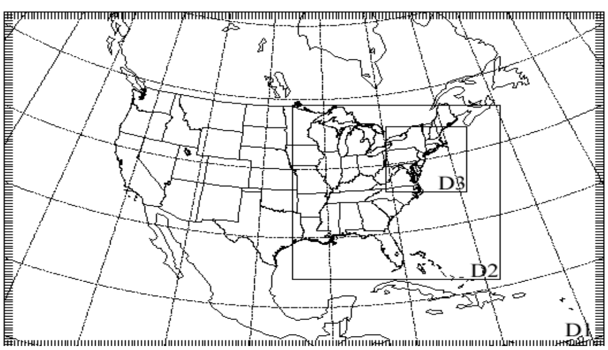

Figure 1: WRF-ARW model domains used in this study. The model resolutions for domain 1 (D1), domain 2 (D2), and domain 3 (D3) are 45, 15, and 5 km, respectively.

For this work, we assimilated 'wetprf' (wet profile) COSMIC RO data, which was obtained from the COSMIC Data Analysis and Archival Center (CDAAC) website (http://www.cosmic.ucar.edu/). All files in this dataset contain actual temperature (dry temperature accounting for moisture), and vapor pressure among others variables. To obtain these data profiles, raw COSMIC RO refractivity data was first pre-processed and assimilated by a 1D ECMWF model. From this model, the temperature and vapor pressure profiles were extracted and stored in 'wetprf' files on 100 m height intervals (CDAAC 2011). We choose to utilize the 'wetprf' dataset because of its ease of use and because WRF lacks the adjoint model necessary to assimilate refractivity data directly.

Before running WRF, we also had to choose when to initialize our model runs and over what time duration. Given the dependence of nor'easter development on both upper-level vorticity advection (Jacobs et al. 2005) and pre-existing surface fluxes (Kuo et al. 1991; Ren et al. 2004), we initialized our runs prior to rapid cyclogenesis. Given our interest in the NEUS region, and specifically New Jersey, we initialized our cases 72 hours prior to first 0.02" official NJ precipitation measurement attributed to a case. The 72-hour lead time was chosen for several reasons. First, Kuo et al. (1991) noted that sensible and latent heat fluxes were best able to influence the resulting cyclogenesis during its formative stages (about 24-48 hours prior to the onset of rapid cyclogenesis). Second, a more recent study by Mote et al. (1997) successfully applied a 72-hour lead time in their heavy snowstorm composite analysis study. Finally, we argue that a 72-hour lead time provides WRF with enough spin-up time, yet leaves sufficient forecast uncertainty to make simulated differences between our WRF runs more apparent. With this lead time, we needed a 180-hour WRF simulation duration to cover each case fully. The specific model run times are noted in Table 1.

2.2 EVALUATION

Evaluation of WRF model performance involved both an inter-WRF comparison and another to the GFS 003 model analysis, which we used as *ground truth*. To generate our ground truth dataset, GFS 003 model analysis data was pre-processed by WRF and then interpolated to our WRF model grid. Our decision to use the GFS 003 model analysis as our ground truth was based upon our usage of GFS 003 model forecasts to initialize our WRF model runs.

We evaluated WRF model performance using two methods: storm track and the energy norm. Storm tracks for both our ground truth and WRF model runs were derived via a custom algorithm that utilizes sea-level pressure (SLP) fields. This algorithm determines storm tracks in three steps. First, it determines whether a candidate grid point has the lowest sea-level pressure minimum within a five degree radius of its location at a given time. If it is the minimum, its location and value is noted, else it is ignored. Second, the uniqueness of each candidate point is determined by comparing its position and time information to that of the last stored point of existing storm track. Only if a candidate point was within five degrees distance and six hours to a previous storm position was it then appended to an existing storm track. Otherwise, it designated as a new storm track. Finally, if a stored storm track was comprised of two or more points, we compared each new candidate point instead to an estimated storm position at the current time rather than at the last storm track position. We derived this position by adding the difference in distance between the last two points in a storm track to the most recent storm position in that array. Given the nature of these data, we will only evaluate and validate the storm track qualitatively with respect to the location and timing.

For our main quantitative analysis, we used the energy norm described in Rabier et al. (1996), which is shown below.

$$\langle \Delta x, \Delta x \rangle = -\frac{1}{2}\int_{p_{sfc}}^{p_{top}}\int\int_{A}\left(\Delta u^2 + \Delta v^2 + R_d T_r * (\Delta \ln(p_s))^2 + \frac{c_p}{T_r}\Delta T^2\right)dx\,dy\,dp \quad \textbf{(2)}$$

In Eq. (2), $\langle \Delta x, \Delta x \rangle$ is the energy norm (J m s^{-2}), u is the zonal wind (m s^{-1}), v is the meridional wind (m s^{-1}), R_d is the dry air gas constant, T_r is the mean surface temperature (K), P_s is the surface pressure (Pa), c_p is heat capacity at constant pressure, and T is air temperature (K). We primarily focused on the energy norm rather than RMSE because RMSE represents only a slice of the atmosphere, whereas the energy norm represents the entire model volume. Additionally, RMSE is typically applied for one variable at a time (e.g., 500-hPa geopotential height, 850-hPa temperature, etc) and the energy norm represents a combined error from four meteorological variables at once. Despite their mathematical differences, both RMSE and the energy norm can be interpreted similarly. For both indices, higher values denote greater error and lower values denote lower error. In our results, we will primarily discuss model performance with respect to the energy norm, but we will also briefly discuss our RMSE analysis for comparison.

3. RESULTS

3.1 STORM TRACK ANALYSIS

Storm tracks for all eight cases from our five WRF simulations and the ground truth are shown in Figure 2. Looking solely at the ground truth (black lines), we were able to classify each event as a Type A or Type B cyclone using the Miller Classification System (Miller 1946). Under this system, a Type A cyclone typically develops over or near the Gulf of Mexico and then propagates up the eastern US seaboard. In contrast, a Type B cyclone propagates west of the Appalachian Mountains and then coastal redevelopment later occurs. Using these classifications, we determined that all of our cases sans the February 2010 case were Type A cyclones. It should be noted here that the April 2007 case was deemed to a type A cyclone only because no coastal redevelopment occurred.

After classifying each ground truth storm track, we classified each of our WRF simulations using Figure 2 and compared the results. In Figure 2, storm tracks for WRF 3.2 are yellow, WRF-COSMIC 0 hr are magenta, WRF-COSMIC 24 hr are brown, WRF-COSMIC 48 hr are green, and WRF-COSMIC 180 hr are grey. Our analysis of WRF storm track revealed only the March 2007 case classification was not consistent with the ground truth. In all WRF simulations, except for the WRF-COSMIC 24 and 48 hr simulations, our March 2007 cyclone originates over the Ohio Valley and then later redevelops off the Delmarva Peninsula and would therefore be classified as a Type B cyclone. For the WRF-COSMIC 24 and 48 hours, the main storm track however, never emerges off the eastern seaboard due to a sharp, negatively titled trough axis that develops over Wisconsin on 16 March (not shown). Due to this simulated trough, the main steering flow for these two cases kept the main surface cyclone inland. Thus, it could not be classified under the Miller scheme. For the remaining WRF runs, their type discrepancy relative to the ground truth occurs because the remaining WRF simulations under-predicted a 500-hPa shortwave height trough over Louisiana and Mississippi at 12 UTC 15 March (see Figure 3, top panels). To illustrate its significance, we complement our geopotential height analysis with a sea-level pressure analysis 12 hours later at 00 UTC 16 March 16 in the bottom two panels of Figure 3. These two panels showed that WRF-COSMIC 180 hr overextended the 1012 hPa surface into the Ohio Valley whereas WRF 3.2 has no 1012 hPa contour. In reality, the ground truth data shows a weak area of low pressure over Mississippi and Alabama just eastward of where our 500-hPa trough was previously located and thus via the QG omega equation this would suggest enhanced upward vertical motion that would not be present in any of the WRF runs.

In our analysis of minimum sea-level pressure, we found no clear and consistent over- or under-intensification bias as compared to the ground truth. Between the various WRF cases, the minimum sea-level pressure amongst the various cases did not vary

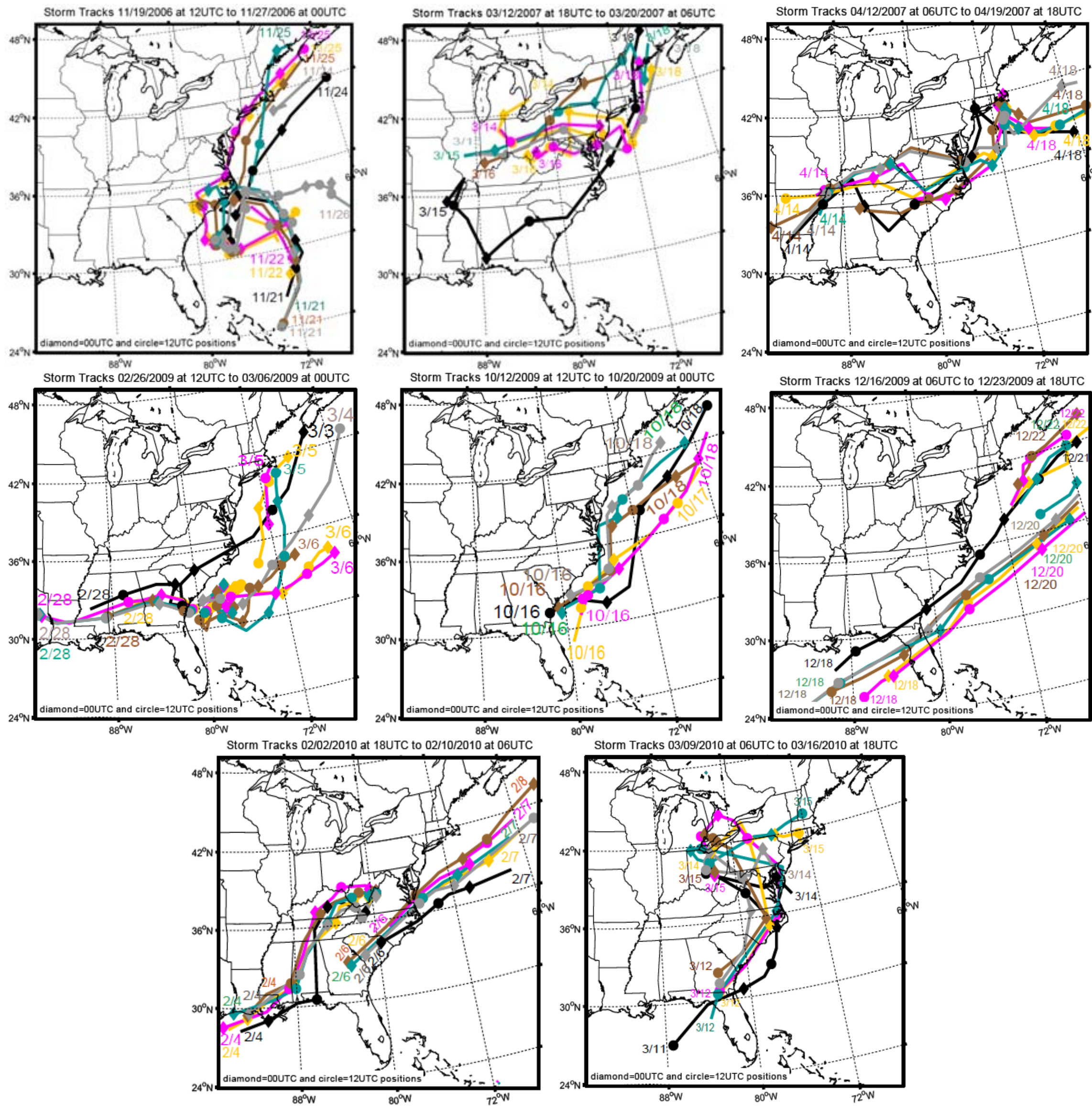

Figure 2: Sea-level pressure-based storm tracks for all eight nor'easter cases with the start date and end date of each track noted. Here we show the GFS 003 model analysis ground truth (black), WRF 3.2 (yellow), WRF 3.2 with 1 assimilation (magenta), WRF 3.2 with 9 assimilations (brown), WRF 3.2 with 17 assimilations (green), WRF 3.2 with 61 assimilations (grey).

by 5 hPa or greater. Therefore, all our WRF cases reached similar maximum intensities. Comparing the minimum sea-level pressure values simulated by WRF 3.2 to the ground truth revealed that in two cases (October 2009 and December 2009) WRF 3.2 under-intensified the cyclone by 5 hPa and 20 hPa, respectively. Comparatively, WRF-COSMIC 180 hr also under-intensified the main cyclone in two cases (February 2009 and December 2009) by 6 and 18 hPa, respectively. With respect to over-intensification, WRF 3.2 did this for 3 cases (November 2006, April 2007, and February 2010) by 8, 8, and 6 hPa, respectively. The WRF-COSMIC 180 hr model over-intensified only the April 2007 case by 9 hPa. By having fewer non-correct cases, this suggests that the additional moisture and temperature information assimilated from COSMIC

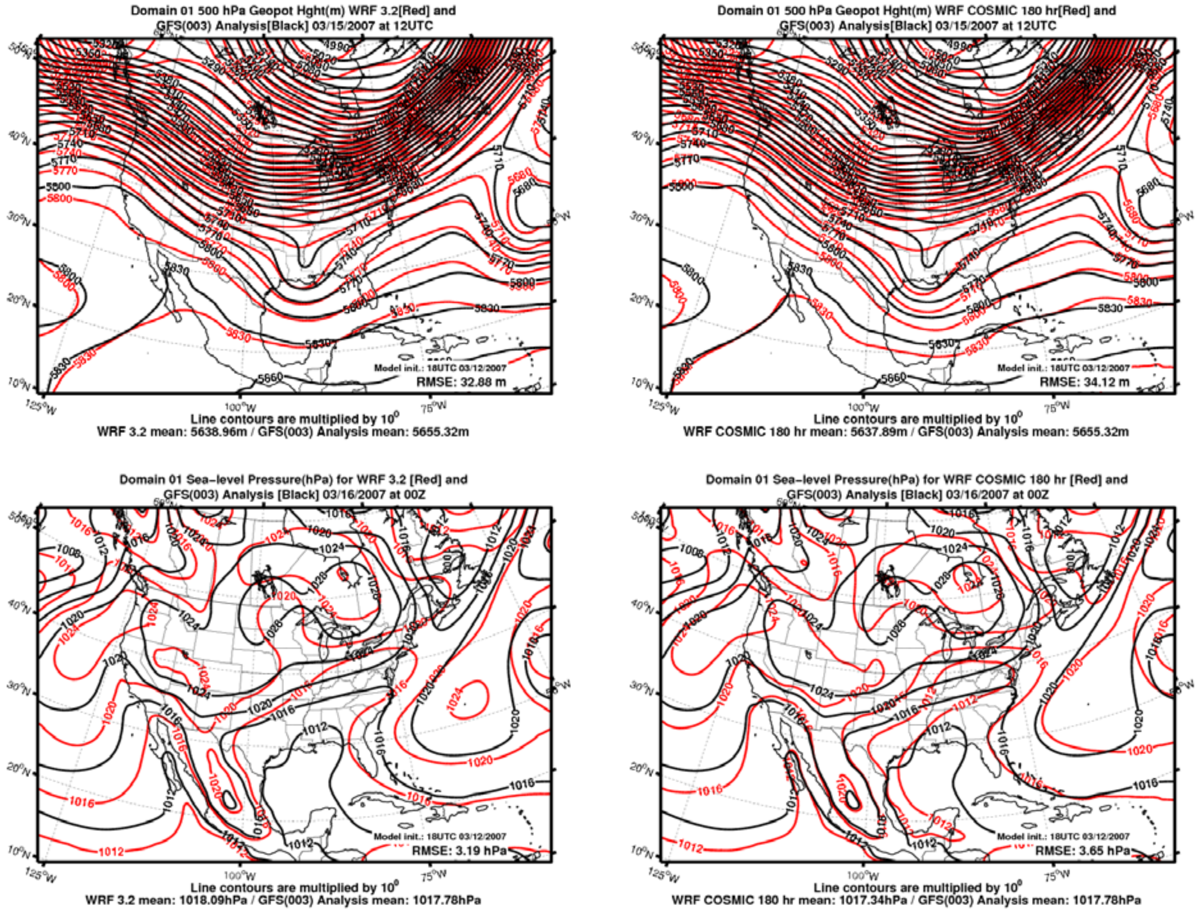

Figure 3: 500-hPa geopotential height (top panels) and sea-level pressure (lower panels) from the ground truth (black lines) and two WRF simulations (red lines) at 12 UTC 15 March 2007 and 00 UTC 16 March 2007, respectively. The shown WRF simulation include WRF 3.2 with no COSMIC RO (left panels), WRF-COSMIC 180 hours (right panels).

improves WRF sea-level pressure simulations. Our best evidence supporting this claim arises from another finding where we compared the minimum sea-level pressure error for our WRF-COSMIC runs to that from WRF 3.2. In our findings, we noted that WRF-COSMIC 48 hr produced more accurate minimum sea-level pressure forecasts than WRF 3.2 in 4 of 8 cases, and WRF-COSMIC 180 hrs did so in 5 of 8 cases.

The final piece of our storm track analysis found notable position and time biases between our WRF runs and the ground truth as well in our inter-WRF comparison. Unlike a parallel study which concerned solely different WRF version, we easily observed noticeable and large variations in storm tracks and sometimes their timing amongst the various WRF runs. Most notable of these was the February 2009 case. At first glance, it is readily apparent that all WRF model runs dipped too far south by upwards of hundreds of kilometers on March 2. What is more striking is the 48-hour time lag in WRF 3.2 when it exits domain 2 on 6 March. Comparing WRF 3.2 to the various model runs reveals that longer COSMIC RO data assimilation period acted to shorten this time lag drastically (36 hours) in the WRF-COSMIC 180 hr model run. 500-hPa geopotential height field analysis revealed the source of the poor performance of WRF 3.2 to its incorrect development of a cut-off 500-hPa height minimum over the Southeastern US as seen in Figure 4. Each panel of Figure 4 shows both ground truth geopotential height fields in black, and a WRF model run in red on March 3, 2009 at 12 UTC. Performance-wise, this cut-off height trough hurts our model forecast because it meanders in weak steering flow over Florida for 2-3 days instead of propagating up the US seaboard as it does in the ground truth. Despite this meandering geopotential height trough, Figure 4 shows clear evidence demonstrating how COSMIC RO assimilation leads to dramatic alterations to both geopotential height fields

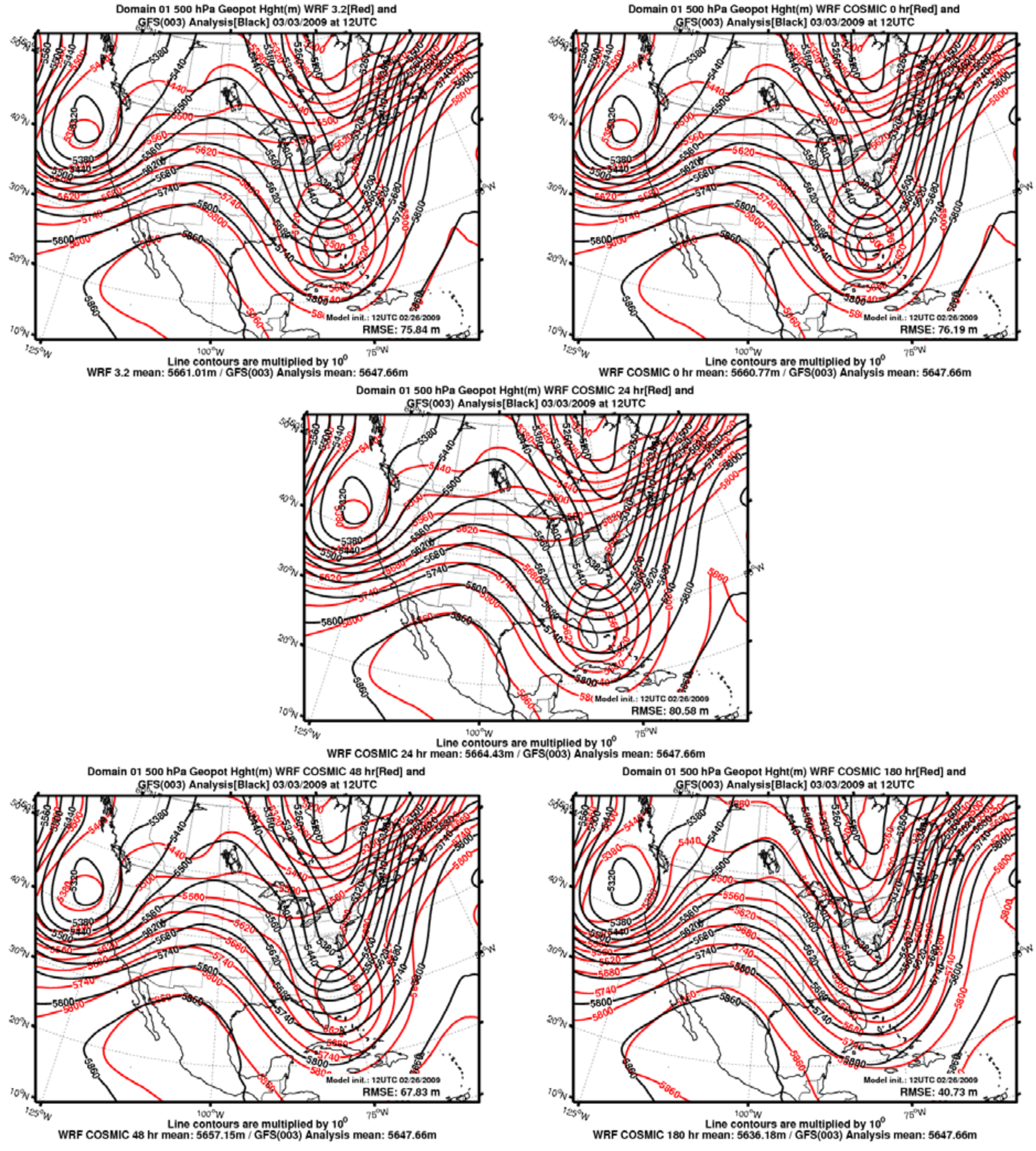

Figure 4: 500-hPa geopotential height from the ground truth (black lines) and WRF simulations (red lines) at 12 UTC 3 March 2009. The WRF simulation include WRF 3.2 with no COSMIC RO (top left), WRF-COSMIC 0 hours (top right), WRF-COSMIC 24 hours (middle), WRF-COSMIC 48 hours (lower left), and WRF-COSMIC 180 hours (lower right).

and storm tracks (from Figure 2) over the original WRF 3.2 model. Given how this experiment is configured, this improvement can only be attributed to the moisture and temperature data from COSMIC.

More generally speaking we found that WRF 3.2 lagged the ground truth in 4 cases (November 2007, February 2007, December 2009, and February 2010) by 24, 48, 24, and 12 hours, respectively. WRF 3.2 only lead the ground truth in one case (April 2007) by 12 hours. Six of eight WRF-COSMIC 180 hr cases (all except March 2007 and February 2010) lagged the by anywhere between 6 to 48 hours. Most times this difference was less than 12 hours. Directly comparing the WRF 3.2 to WRF-COSMIC 180 hr storm tracks reveals that WRF-COSMIC lead WRF 3.2 by 6 - 36 hours in 6 of 8 cases (all except April 2007, and October 2009). Finally, we researched the potential for storm track biases between each WRF model and the ground truth. More often than not (5 of 8 cases) WRF 3.2 showed a leftward storm track bias relative to the ground truth. WRF-COSMIC 180 hr storm tracks tended to be leftward (6 of 8 cases) of the ground truth, and rightward (5 of 8 cases) of WRF 3.2. Upon qualitative analysis of Figure 3, we noted that in 6 of 8 cases WRF-COSMIC 180 hr runs show storm tracks closer to the ground truth than WRF 3.2.

The above spatial and temporal differences we hypothesize exist due to three main causes. First, all our WRF model runs were cold-start runs, and therefore required 6-12 hours to spin up. Such a spin-up may cause a systematic time lag because existing weather features may develop more slowly in WRF than in the ground truth. Second, our WRF boundary conditions files were generated from GFS 003 model forecasts initialized at the times specified in Table 1. Therefore, our model simulations were influenced by GFS model errors propagating through our domains. Finally, our WRF simulations may mishandle the development of synoptic and mesoscale weather features, which are crucial to the development of each case.

3.2 ENERGY NORM ANALYSIS

Figure 5 shows that the energy norm between each WRF model and the ground truth appears similar, yet there are distinct and clear differences. In this figure, the energy norm and its mean are shown for WRF 3,2 (yellow), WRF-COSMIC 0 hr (magenta), WRF-COSMIC 24 hr (brown), WRF-COSMIC 48 hr (green), and WRF-COSMIC 180 hr (grey), respectively. From Figure 5, we found three key findings. First, all WRF energy norms showed a clear diurnal signal and generally increased with time. Second, the energy norm is not always inversely proportional to the number of COSMIC RO assimilations. Finally, our energy norm results conflict with our RMSE results from our WRF-COSMIC 180 hr model run. The former shows COSMIC RO to be more hurtful than helpful, whereas the latter clearly indicates the opposite.

The increasing energy norm and its diurnal signal were striking features on Figure 5. Energy norm values typically increased with time due to accumulating model errors as model run time increased. What came as more of a surprise to us was the strong diurnal signal in each WRF simulation. Originally, we hypothesized that this signal may have been a product of the RRTM longwave radiation scheme, which was later disproven in our parallel study. Ultimately, we broke down the energy norm into its components and found the diurnal cycle originated primarily from the temperature with a far smaller contribution (1 order of magnitude) from surface pressure.

Figure 5 displays the energy norm for each WRF run, and shows notable, but statistically insignificant differences. To prove non-significance, we show p-values from a two-tailed Student T-test assuming unequal variance and a 0.05 significance level between the WRF-COSMIC run and WRF 3.2 energy norms in Table 2. In this table, no WRF-COSMIC model run achieved a p-value of 0.05 or less which would denote a significant change from WRF 3.2. Out of all the runs and cases, the WRF-COSMIC 180 hr run comes closest to statistical significance with a 0.32 p-value, whereas the WRF-COSMIC 0 hr run for this same case has the least significance (p-value = 0.98). One shock in Table 2 was that the p-value did not always decrease with an increasing number of COSMIC RO assimilations. We originally hypothesized that when more COSMIC RO data was assimilated this would lead to increased nor'easter simulation divergence. This divergence trend however, only appeared in 3 of 8 cases (March 2007, October 2009, and March 2010). We theorize that this finding may be a result of compensating perturbations from WRF-COSMIC, which would bring WRF-COSMIC runs closer to WRF 3.2 and may originate from the large horizontal distances that a COSMIC RO profile covers with height.

Table 3 summarizes the energy norm-based performance for each model run as compared to WRF 3.2. The values shown on the upper table denote the number of time steps where the energy norm for each WRF-COSMIC run was lower than WRF 3.2. Percentage values in the bottom panel of Table 3 indicate the corresponding percentage of total time steps values in the upper table represent. Overall, the number of time steps in WRF-COSMIC runs having a lower energy norm than WRF 3.2 increased as the number of data assimilation periods increased. Despite this increase, no one WRF-COSMIC case overall had an outright (> 50% of time steps) improved energy norm performance versus WRF 3.2. On a case-by-case however, WRF-COSMIC 48 hr bests WRF 3.2 in the February 2009 case, and WRF-COSMIC 180 hr bests WRF 3.2 in three cases (November 2006, February 2009, and March 2010). In each of these cases, this improvement appears to stem from better realization of

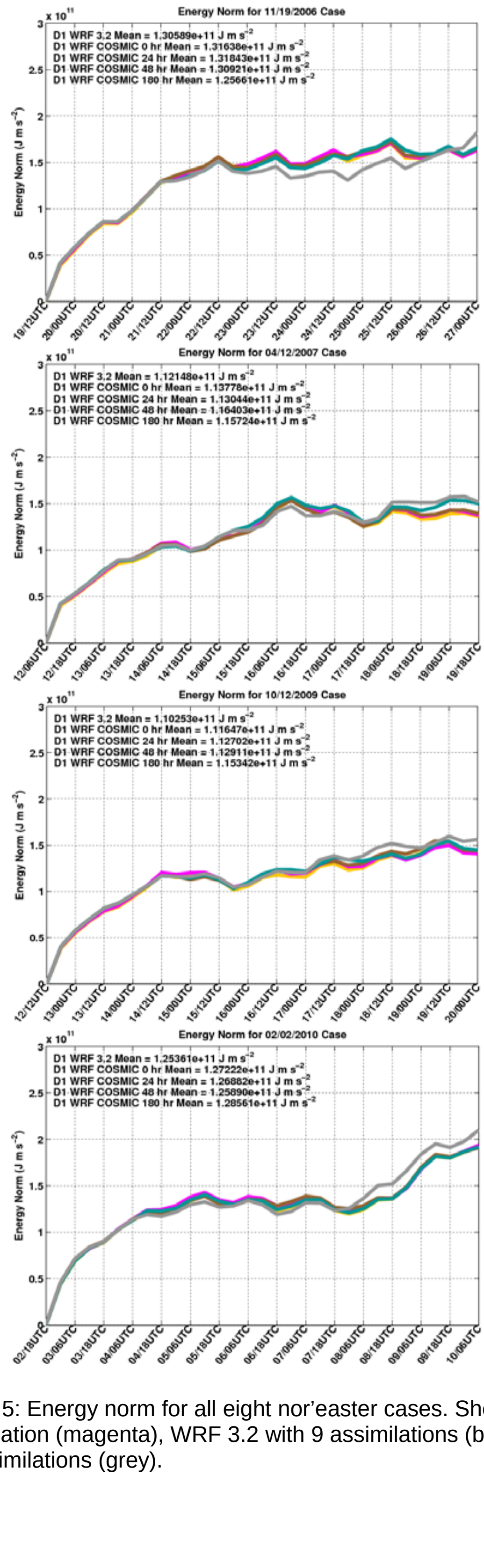

Figure 5: Energy norm for all eight nor'easter cases. Shown line colors denote WRF 3.2 (yellow), WRF 3.2 with 1 assimilation (magenta), WRF 3.2 with 9 assimilations (brown), WRF 3.2 with 17 assimilations (green), WRF 3.2 with 61 assimilations (grey).

Table 2: P-values for a two-tailed Student's T-test of unequal variance based upon a comparison between data in all four WRF-COSMIC runs to WRF 3.2.

Two-tailed T-test (P value)	11/19/2006	3/12/2007	4/12/2007	2/26/2009	10/12/2009	12/16/2009	2/2/2010	3/9/2010	All
WRF 3.2 / WRF COSMIC 0 hr	0.92	0.98	0.86	0.96	0.87	0.86	0.85	0.94	0.75
WRF 3.2 / WRF COSMIC 24 hr	0.91	0.51	0.92	0.87	0.78	0.79	0.88	0.69	0.49
WRF 3.2 / WRF COSMIC 48 hr	0.98	0.36	0.65	0.96	0.76	0.78	0.96	0.70	0.47
WRF 3.2 / WRF COSMIC 180 hr	0.64	0.32	0.71	0.62	0.57	0.91	0.76	0.85	0.79

Table 3: Energy norm-based performance of WRF 3.2 versus each WRF-COSMIC run. The above table denotes the total number of time steps per case and overall where the energy norm was lowest for each shown comparison. The bottom table lists the corresponding percentage of total time steps the upper table represents.

Physical Count	11/19/2006	3/12/2007	4/12/2007	2/26/2009	10/12/2009	12/16/2009	2/2/2010	3/9/2010	All
WRF 3.2 < WRF-COSMIC 180 hr	11	30	23	11	29	20	17	12	153
WRF COSMIC 0 hr < WRF 3.2	4	9	4	2	6	0	2	4	31
WRF COSMIC 24 hr < WRF 3.2	8	0	12	3	4	5	6	0	38
WRF COSMIC 48 hr < WRF 3.2	12	0	2	16	6	5	7	4	52
WRF COSMIC 180 hr < WRF 3.2	20	1	8	20	2	11	14	19	95
Total Obs	31	31	31	31	31	31	31	31	248

Percentage	11/19/2006	3/12/2007	4/12/2007	2/26/2009	10/12/2009	12/16/2009	2/2/2010	3/9/2010	All
WRF 3.2 < WRF-COSMIC 180 hr	35.48%	96.77%	74.19%	35.48%	93.55%	64.52%	54.84%	38.71%	61.69%
WRF COSMIC 0 hr < WRF 3.2	12.90%	29.03%	12.90%	6.45%	19.35%	0.00%	6.45%	12.90%	12.50%
WRF COSMIC 24 hr < WRF 3.2	25.81%	0.00%	38.71%	9.68%	12.90%	16.13%	19.35%	0.00%	15.32%
WRF COSMIC 48 hr < WRF 3.2	38.71%	0.00%	6.45%	51.61%	19.35%	16.13%	22.58%	12.90%	20.97%
WRF COSMIC 180 hr < WRF 3.2	64.52%	3.23%	25.81%	64.52%	6.45%	35.48%	45.16%	61.29%	38.31%

the actual storm track as compared to WRF 3.2 (see Figure 2). Overall, a comparison of WRF 3.2 to the best COSMIC run (WRF-COSMIC 180 hr) shows that full COSMIC RO data assimilation by WRF improves energy norm performance In 95 of 248 (38.3%) total time steps. Given the strong inter-case variability of the WRF-COSMIC model energy performance, we hypothesize that COSMIC-based improvement of WRF may depend upon both existing case uncertainty and properties of the COSMIC RO data, namely the density and location of these data.

To better visualize the inter-WRF energy norm variability with time, we generated panels depicting the difference in energy norm for each case in Figure 6. The shown differences are WRF 3.2-WRF—COSMIC 0 hr (magenta), WRF 3.2-WRF—COSMIC 24 hr (brown), WRF 3.2-WRF—COSMIC 48 hr (green), and WRF 3.2-WRF—COSMIC 180 hr (grey). Each case has only relatively small energy norm differences during the first 48 hours and then afterward these differences vary widely from case to case and WRF to WRF. The rise in energy norm differences after 48 hours exists given how each WRF run handles the development and maintenance of our nor'easter cases and other weather systems present in our model domains. Generally, the amount of deviation appears proportional to the number of data assimilation periods, which was not unexpected given our earlier p-values in Table 2. From these difference data, we note that WRF-COSMIC 0 hr, was all but identical to the WRF 3.2 given its near zero difference in the energy norm with time. The WRF-COSMIC 24 hr run deviated from zero, but its mean energy norm differences were are always negative. WRF-COSMIC 48 hr shows a positive mean energy

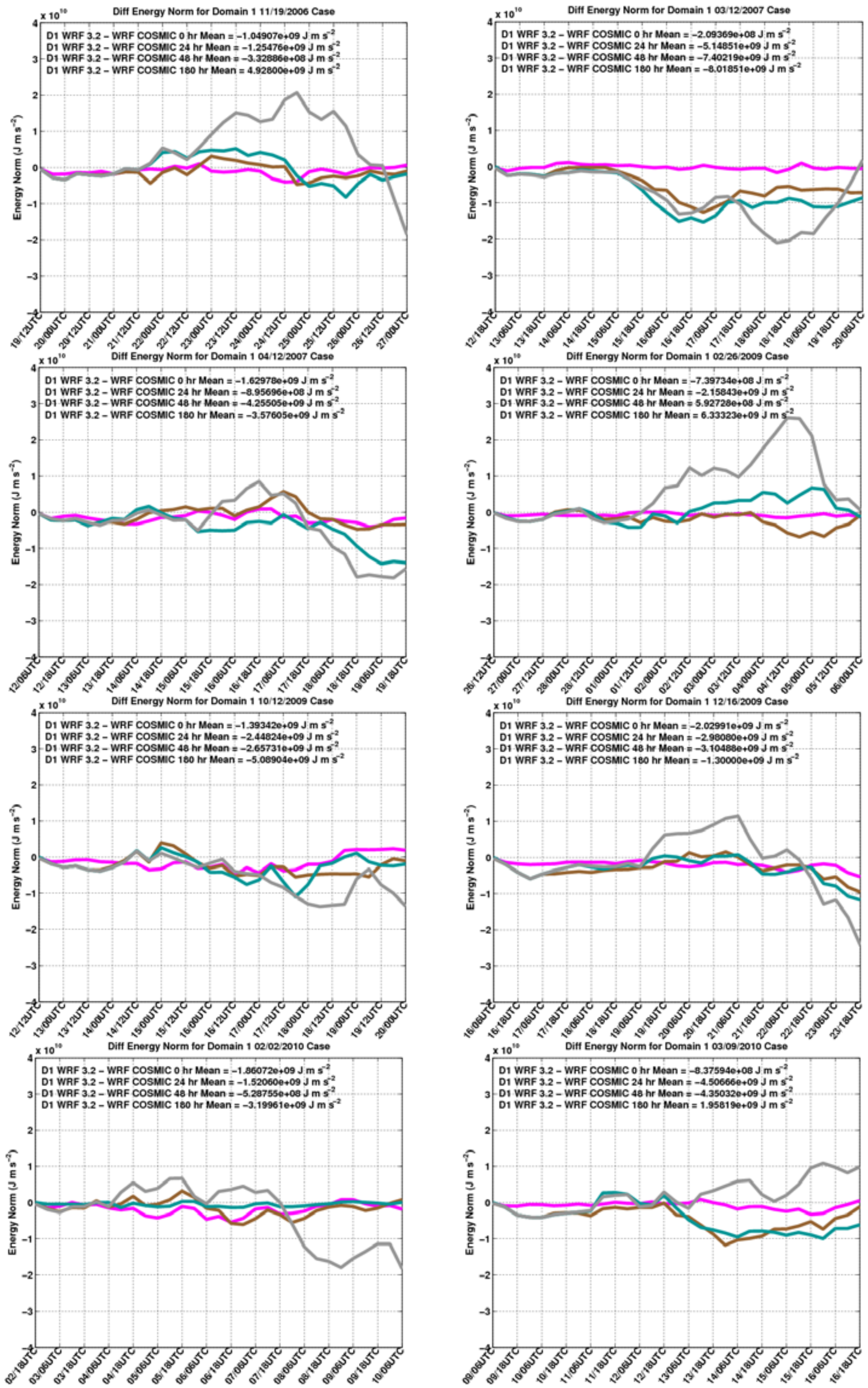

Figure 6: Difference in energy norm for all eight nor'easter cases. Shown line colors denote WRF 3.2 - WRF COSMIC (1 assimilations) [magenta], WRF 3.2 - WRF COSMIC (9 assimilations) [brown], WRF 3.2 - WRF COSMIC (17 assimilations) [green], WRF 3.2 - WRF COSMIC (61 assimilations) [grey].

norm difference in one case (February 2009), but it has an overall negative contribution of -2.436×10^9 J m s^{-2} to the energy norm. Finally, WRF-COSMIC 180 hr betters our simulations in 3 of 8 cases on average, but similar to all other COSMIC RO runs it too increased the energy norm relative to WRF 3.2 (7.959×10^9 J m s^{-2} increase on average). Despite these negative results, if we focus solely on the main nor'easter event times (forecast hours 72-120), WRF COSMIC 180 hr shows a visible reduction in the energy norm in 6 of 8 cases, while WRF-COSMIC 24 hrs shows improvement in 2 cases. Better performance during this time frame we believe can be related to improved surface fluxes realized during the critical pre-cyclogenesis phase suggested by Kuo et al. (1991) in these WRF-COSMIC runs.

For sake of completeness and comparison to our energy norm results, we determined how WRF 3.2 and WRF-COSMIC 180 hr compared to each other via an RMSE-based analysis, which is shown in Table 4. This table indicates the number of instances where the WRF-COSMIC 180 hr RMSE values of sea-level pressure, 850-hPa temperatures, and 500-hPa geopotential heights for each case, domain, and overall were lower than those from WRF 3.2. Most striking to us was that the overall results stand in complete contradiction to our energy norm results. As seen in this table, the WRF-COMSIC 180 hr model runs have an overall, lower RMSE values in 6 of 8 cases and reduced RMSE on domain 1 by a variable-averaged 178 of 248 (71.6%) time steps on domain 1. For the other two model domains, similar results are seen. Based upon the mathematical differences between the energy norm and RMSE, we can propose two hypotheses, which would require further work to verify. First, Cucurull et al. (2008) noted that COSMIC RO data assimilation errors become large above 30 km, but started to occur at 200 hPa. Hence it would be possible for data assimilation errors to cause significant negative contributions to the energy norm above 200 hPa that would not propagate downward to 500 hPa and strongly impact RMSE within the model run time. Finally, our RMSE error metrics do not involve the zonal and meridional winds, which are the largest contributors to the energy norm. Therefore, it is possible that perturbations to these winds would not be fully resolved by the RMSE given that it only represents a slice of the atmosphere.

4. SUMMARY AND CONCLUSIONS

We assessed and compared WRF 3.2 model performance both with and without 3D cycled COSMIC RO data assimilation during eight nor'easter cases that occurred between October and April of 2006 to 2010. To fully assess the performance implications of COSMIC RO data assimilation, we ran four WRF model simulations with varied assimilation periods (0 hr, 24 hrs, 48 hrs, and 180 hours). Other than the amount of COSMIC RO data assimilation events, all WRF runs utilized identical model parameters and physics options.

Performance evaluation amongst our various cases involved both inter-WRF and 'ground truth' comparisons of sea-level pressure-based storm track analysis, the energy norm, and RMSE. Our inter-WRF comparison of storm tracks revealed that it varied widely from case-to-case and that WRF-COSMIC runs (especially 180 hrs) tended to lead WRF 3.2 in 6 of 8 cases (all except April 2007, and October 2009) by 6 to 36 hours. As compared to the ground truth, both WRF 3.2 and WRF-COSMIC lagged behind, but only one WRF 3.2 case (April 2007) ever lead the ground truth. In terms of absolute time, WRF-COSMIC lagged the ground truth by typically less than 12 hours, while WRF 3.2 averaged just slightly longer. These observed biases and the inter-WRF differences were shown to be directly attributed to simulated differences in synoptic and mesoscale dynamical fields, most notably geopotential height.

Our energy norm analysis failed to reveal any statistically significant differences amongst the various WRF model runs and noted a mostly negative contribution of COSMIC RO data assimilation to the performance of WRF simulations. Despite the lack of statistically different energy norms between our WRF-COSMIC and WRF 3.2 runs, we did note several useful findings. First, only in 3 of 8 cases did p-values from our two-tailed Student's T-test decrease with an increasing number of COSMIC RO assimilations. Such results may be a product result from data position errors given the assumption that each COSMIC RO profile is vertical, yet each profile slants greatly with height. Overall, the WRF COSMIC runs had generally worse energy norm scores as compared to WRF 3.2. For the WRF-COSMIC 48 hr run, it was noted to decrease the energy norm in 58 of 248 (20.97%) total time steps, and on average increased the energy norm by 2.436×10^9 J m s^{-2}. The WRF-COSMIC 180 hrs run decreased the energy norm in 95 of 248 (38.31%) total time steps, and on average increased the energy norm by 7.959×10^9 J m s^{-2}. When considering these WRF-COSMIC cases solely during the forecast hours 72 and 120, when each nor'easter rapidly intensified, the energy norm values in 6 of 8 cases were smaller versus WRF 3.2 during this limited time period. Finally, our RMSE analysis of sea-level pressure, 850-hPa temperature and 500-hPa geopotential height for the WRF-COSMIC 180 hr case contradicts our energy norm results. In these results, RMSE was reduced in our WRF-COSMIC 180 hr run in a variable-averaged 178 of 248 (71.6%) total time steps, thus denoting a strong positive contribution to model performance by COSMIC RO assimilation into WRF. The opposite findings of our RMSE and energy norm results may potentially be attributed to COSMIC RO data assimilation errors by WRF above 200 hPa, which would affect our energy norm, but may not influence

Case	Sea-level Pressure (hPa)		850-hPa Temperature (K)		500 hPa Geopotential Height (m)	
11/19/2006			Total obs per box 31			
Domain 1	24	77.42%	26	83.87%	24	77.42%
Domain 2	28	90.32%	27	87.10%	28	90.32%
Domain 3	26	83.87%	25	80.65%	26	83.87%
3/12/2007			Total obs per box 31			
Domain 1	9	29.03%	26	83.87%	22	70.97%
Domain 2	13	41.94%	26	83.87%	12	38.71%
Domain 3	8	25.81%	18	58.06%	11	35.48%
4/12/2007			Total obs per box 31			
Domain 1	26	83.87%	27	87.10%	24	77.42%
Domain 2	27	87.10%	26	83.87%	26	83.87%
Domain 3	25	80.65%	29	93.55%	25	80.65%
2/26/2009			Total obs per box 31			
Domain 1	24	77.42%	26	83.87%	26	83.87%
Domain 2	23	74.19%	26	83.87%	26	83.87%
Domain 3	18	58.06%	26	83.87%	20	64.52%
10/12/2009			Total obs per box 31			
Domain 1	2	6.45%	23	74.19%	21	67.74%
Domain 2	20	64.52%	19	61.29%	26	83.87%
Domain 3	16	51.61%	14	45.16%	20	64.52%
12/16/2009			Total obs per box 31			
Domain 1	21	67.74%	21	67.74%	21	67.74%
Domain 2	28	90.32%	26	83.87%	27	87.10%
Domain 3	19	61.29%	23	74.19%	24	77.42%
2/2/2010			Total obs per box 31			
Domain 1	24	77.42%	24	77.42%	24	77.42%
Domain 2	20	64.52%	26	83.87%	24	77.42%
Domain 3	23	74.19%	26	83.87%	25	80.65%
3/9/2010			Total obs per box 31			
Domain 1	20	64.52%	26	83.87%	22	70.97%
Domain 2	19	61.29%	30	96.77%	23	74.19%
Domain 3	20	64.52%	25	80.65%	19	61.29%
By Domain			Total obs per box 248			
Domain 1	150	60.48%	199	80.24%	184	74.19%
Domain 2	178	71.77%	206	83.06%	192	77.42%
Domain 3	155	62.50%	186	75.00%	170	68.55%
Overall			Total obs per box 744			
All cases	483	64.92%	591	79.44%	546	73.39%

Table 4: RMSE analysis of WRF 3.2 versus WRF 3.2-COSMIC 180 hrs for sea-level pressure, 850-hPa temperature, and 500-hPa geopotential height for each case (31 time steps), domain (248 total time steps) and overall (744 total time steps). Above numbers and percentages reflect the instances where WRF 3.2-COSMIC had a lower RMSE than WRF 3.2.

regions below 500 hPa during the model integration time.

5. REFERENCES

Anthes, R. A., and Coauthors, 2008: The COSMIC/FORMOSAT-3 mission: Early results. *Bull. Amer. Meteor. Soc.*, **89**, 313–333. doi: 10.1175/BAMS-89-3-313

COSMIC Data Analysis and Archival Center (CDAAC), 2011: CDAAC descriptions. [Available online at http://cosmic-io.cosmic.ucar.edu/cdaac/cgi_bin/fileFormats.cgi?type=wetPrf].

Cucurull, L., J. C. Derber, R. Treadon, and R. J. Purser, 2008: Preliminary impact studies using Global Positioning System radio occultation profiles at NCEP. *Mon. Wea. Rev.*, **136**, 1865-1877.

Healy, S., and J.-N. Thépaut, 2006: Assimilation experiments with CHAMP GPS radio occultation measurements. *Quart. J. Roy. Meteor. Soc.,* **132**, 605-623.

Jacobs, N. A., G. M. Lackmann, and S. Raman, 2005: The combined effects of Gulf Stream-induced baroclinicity and upper-level vorticity on U.S. East Coast extratropical cyclogenesis. *Mon. Wea. Rev.*, **133**, 2494-2501.

Kocin, P. J., and L. W. Uccellini, 2004: A snowfall impact scale derived from Northeast storm snowfall distributions. *Bull. Amer. Meteor. Soc.*, **85**, 177–194.

Kuo, Y-H., R. J. Reed, and S. Low-Nam, 1991: Effects of surface energy fluxes during the early development and rapid intensification stages of seven explosive cyclones in the western Atlantic. *Mon. Wea. Rev.*, **119**, 457-476.

Miller, J. E., 1946: Cyclogenesis in the Atlantic coastal region of the United States. *J. Meteor.,* **3,** 31–44.

Mote, T. L., D. W. Gamble, S. J. Underwood, M. L. Bentley, 1997: Synoptic-scale features common to heavy snowstorms in the Southeast United States. *Wea. Forecasting*, **12**, 5-23.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.*, **122**, 121-150

Ren, X., W. Perrie, Z. Long, and J. Gyakum, 2004: Atmosphere-ocean coupled dynamics of cyclones in the midlatitudes. *Mon. Wea. Rev.,* **132**, 2432-2451.

6. ACKNOWLEDGEMENTS

We wish to thank the National Center for Atmospheric Research (NCAR) for its management and hosting the WRF-ARW model. Additionally, we express our gratitude to both the University Corporation for Atmospheric Research (UCAR) and the National Space Organization (NSPO) of Taiwan for their work in developing COSMIC and making its data available. We also thank the COSMIC Data Analysis and Archival Center (CDAAC) for maintaining the COSMIC RO dataset used in this study. I (SN) acknowledge and thank Rutgers University for their financial support of my teaching assistantship and their contribution towards attending the 2011 AMS Annual Meeting.