

2006

Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts

William A. Gallus Jr.

Iowa State University, wgallus@iastate.edu

James F. Bresch

National Center for Atmospheric Research

Follow this and additional works at: http://lib.dr.iastate.edu/ge_at_pubs



Part of the [Atmospheric Sciences Commons](#), and the [Geology Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/ge_at_pubs/48. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Geological and Atmospheric Sciences at Iowa State University Digital Repository. It has been accepted for inclusion in Geological and Atmospheric Sciences Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts

Abstract

A series of simulations for 15 events occurring during August 2002 were performed using the Weather Research and Forecasting (WRF) model over a domain encompassing most of the central United States to compare the sensitivity of warm season rainfall forecasts with changes in model physics, dynamics, and initial conditions. Most simulations were run with 8-km grid spacing. The Advanced Research WRF (ARW) and the nonhydrostatic mesoscale model (NMM) dynamic cores were used. One physics package (denoted NCEP) used the Betts–Miller–Janjic convective scheme with the Mellor–Yamada–Janjic planetary boundary layer (PBL) scheme and GFDL radiation package; the other package (denoted NCAR) used the Kain–Fritsch convective scheme with the Yonsei University PBL scheme and the Dudhia rapid radiative transfer model radiation. Other physical schemes were the same (e.g., the Noah land surface model, Ferrier et al. microphysics) in all runs. Simulations suggest that the sensitivity of the model to changes in physics is a function of which the dynamic core is used, and the sensitivity to the dynamic core is a function of the physics used. The greatest sensitivity in general is associated with a change in physics packages when the NMM core is used. Sensitivity to a change in physics when the ARW core is used is noticeably less. For light rainfall, the spread in the rainfall forecasts when physics are changed under the ARW core is actually less at most times than when the dynamic core is changed while NCAR physics are used. For light rainfall, the WRF model using NCAR physics is much more sensitive to a change in dynamic core than the WRF model using NCEP physics. For heavier rainfall, the opposite is true with a greater sensitivity occurring when NCEP physics is used. Sensitivity to initial conditions (Eta versus the Rapid Update Cycle with an accompanying small change in grid spacing) is generally less substantial than the sensitivity to changes in dynamic core or physics, except in the first 6–12 h of the forecast when it is comparable. As might be expected for warm season rainfall, the finescale structure of rainfall forecasts is more affected by the physics used than the dynamic core used. Surprisingly, however, the overall areal coverage and rain volume within the domain may be more influenced by the dynamic core choice than the physics used.

Keywords

climatology, dynamics, physics, radiation, rain, sensitivity analysis, Mellor–Yamada–Janjic planetary boundary layer, nonhydrostatic mesoscale model (NMM) dynamic cores, radiative transfer model radiation, warm season rainfall forecasts, weather research and forecasting (WRF) model, weather forecasting, mesoscale meteorology, rainfall, simulation

Disciplines

Atmospheric Sciences | Geology

Comments

This article is from *Monthly Weather Review* 134 (2006): 2632, doi: [10.1175/MWR3198.1](https://doi.org/10.1175/MWR3198.1). Posted with permission.

NOTES AND CORRESPONDENCE

Comparison of Impacts of WRF Dynamic Core, Physics Package, and Initial Conditions on Warm Season Rainfall Forecasts

WILLIAM A. GALLUS JR.

Department of Geological and Atmospheric Science, Iowa State University, Ames, Iowa

JAMES F. BRESCH

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 1 September 2005, in final form 14 December 2005)

ABSTRACT

A series of simulations for 15 events occurring during August 2002 were performed using the Weather Research and Forecasting (WRF) model over a domain encompassing most of the central United States to compare the sensitivity of warm season rainfall forecasts with changes in model physics, dynamics, and initial conditions. Most simulations were run with 8-km grid spacing. The Advanced Research WRF (ARW) and the nonhydrostatic mesoscale model (NMM) dynamic cores were used. One physics package (denoted NCEP) used the Betts–Miller–Janjic convective scheme with the Mellor–Yamada–Janjic planetary boundary layer (PBL) scheme and GFDL radiation package; the other package (denoted NCAR) used the Kain–Fritsch convective scheme with the Yonsei University PBL scheme and the Dudhia rapid radiative transfer model radiation. Other physical schemes were the same (e.g., the Noah land surface model, Ferrier et al. microphysics) in all runs. Simulations suggest that the sensitivity of the model to changes in physics is a function of which the dynamic core is used, and the sensitivity to the dynamic core is a function of the physics used. The greatest sensitivity in general is associated with a change in physics packages when the NMM core is used. Sensitivity to a change in physics when the ARW core is used is noticeably less. For light rainfall, the spread in the rainfall forecasts when physics are changed under the ARW core is actually less at most times than when the dynamic core is changed while NCAR physics are used. For light rainfall, the WRF model using NCAR physics is much more sensitive to a change in dynamic core than the WRF model using NCEP physics. For heavier rainfall, the opposite is true with a greater sensitivity occurring when NCEP physics is used. Sensitivity to initial conditions (Eta versus the Rapid Update Cycle with an accompanying small change in grid spacing) is generally less substantial than the sensitivity to changes in dynamic core or physics, except in the first 6–12 h of the forecast when it is comparable. As might be expected for warm season rainfall, the finescale structure of rainfall forecasts is more affected by the physics used than the dynamic core used. Surprisingly, however, the overall areal coverage and rain volume within the domain may be more influenced by the dynamic core choice than the physics used.

1. Introduction

Warm season rainfall forecasting is a difficult challenge (e.g., Olson et al. 1995) with little evidence that significant improvements in objective skill measures

can be obtained in deterministic forecasts by adjustments to initial conditions (Gallus and Segal 2001) or through the use of a particular model (McBride and Ebert 2000; Gallus et al. 2005) or group of physical parameterizations (Jankov et al. 2005). Therefore, ensemble techniques are increasingly being investigated for short-term rainfall forecasting (e.g., Wandishin et al. 2001; Alhamed et al. 2002).

Early research in ensemble forecasting generally made use of coarse grid models designed for medium-

Corresponding author address: Dr. William Gallus, Department of Geological and Atmospheric Science, Iowa State University, 3025 Agronomy Hall, Ames, IA 50011.
E-mail: wgallus@iastate.edu

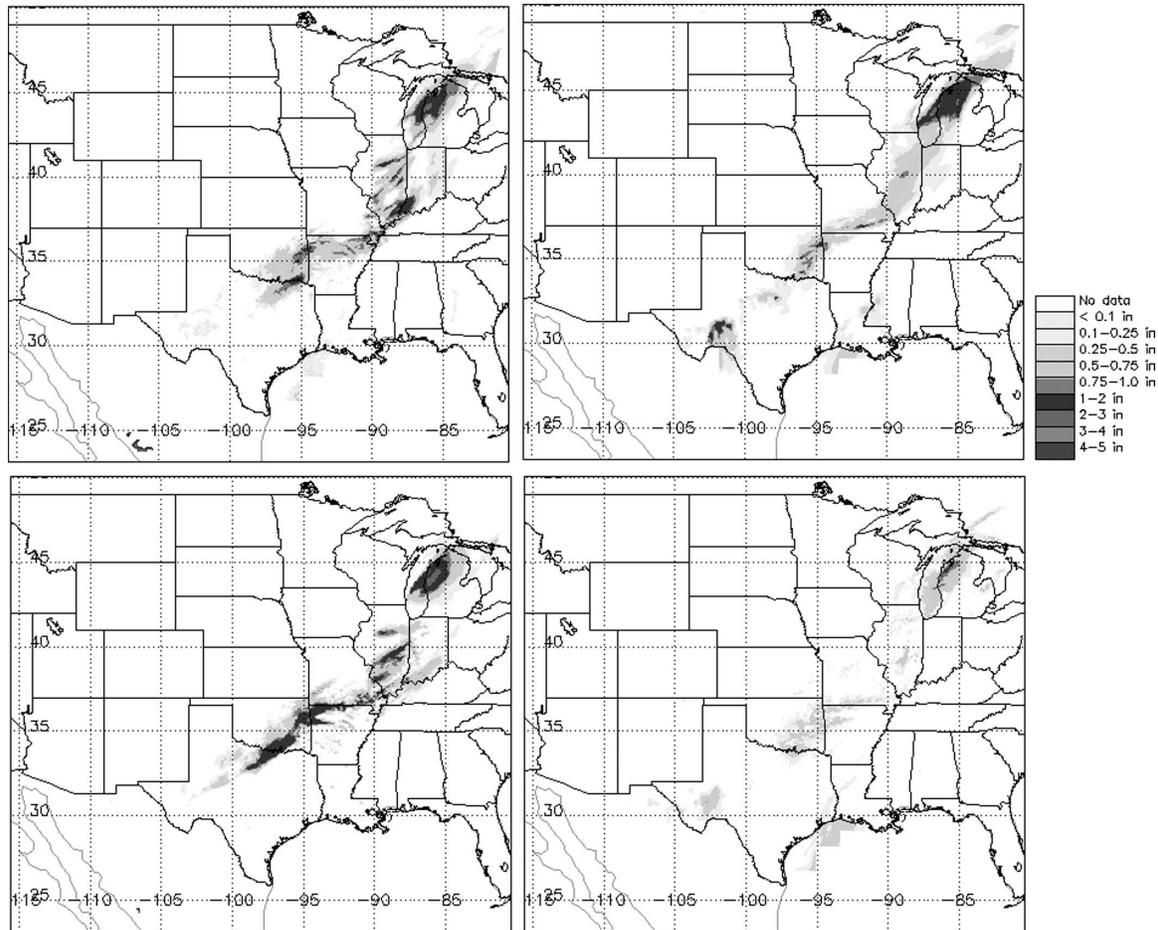


FIG. 1. Rainfall in the 12–18-h forecast period (0000–0600 UTC 14 Aug 2002) for a case initialized at 1200 UTC 13 Aug 2002. Runs using (top) ARW dynamic core and (bottom) NMM; (left) NCAR physics and (right) NCEP.

range forecasting and focused on ensemble members created through the use of perturbed initial conditions (e.g., Toth and Kalnay 1993; Buizza and Palmer 1995; Houtekamer and Derome 1995). This type of ensemble was found, however, to suffer from insufficient spread when applied to shorter-range forecasts, and it was found that spread could be increased by varying model physics (e.g., Buizza et al. 1999; Stensrud et al. 2000) or using multiple models (Alhamed et al. 2002). Alhamed et al. (2002) found in a study using 25 model members run with roughly 30-km grid spacing during the Storm and Mesoscale Ensemble Experiment (SAMEx) that members tend to cluster first by model, next by physics, and last by initial conditions in a mixed model, physics, and initial condition ensemble.

The present study extends the results of Alhamed et al. (2002) to much finer grid spacing (8–10 km) and examines if changes in dynamic core within the same model [e.g., the Weather Research and Forecasting (WRF) model] result in similar impacts to those present

from changes in models, physical parameterizations, and initial conditions. The 8-km grid spacing is more refined than that used in previously discussed ensembles, so that the findings of the present study may help influence design of future ensembles that can use finer grid spacings as computer power increases.

2. Data and methodology

The WRF model was run with 8-km grid spacing using 1200 UTC Eta 40-km Gridded Binary (GRIB) output for initialization and lateral boundary condition information for 15 cases during August 2002 in which substantial convective system rainfall occurred in the model domain (see Fig. 1 for depiction of the domain). All simulations were integrated for 48 h with lateral boundary conditions updated every 3 h. These simulations used 60 vertical layers to match the resolution used in WRF runs with the Nonhydrostatic Mesoscale Model (NMM) dynamic core performed previously by

the WRF Developmental Test bed Center (DTC; e.g., Seaman et al. 2004; Bernardet et al. 2004). For each case, four different WRF runs were examined with 8-km grid spacing and common initial and lateral boundary condition input. These runs used both dynamic cores available in the WRF model as of summer 2004: the Advanced Research WRF (ARW) and the NMM. In addition, two different physics packages were used. The National Center for Atmospheric Research (NCAR) used the Kain–Fritsch (KF; Kain and Fritsch 1992) convective scheme, with the Yonsei University (YSU; Noh et al. 2003) planetary boundary layer (PBL) scheme, and Dudhia rapid radiative transfer model (RRTM) radiation. The National Centers for Environmental Prediction (NCEP) used the Betts–Miller–Janjic (BMJ; Betts 1986; Betts and Miller 1986; Janjic 1994) convective scheme with the Mellor–Yamada–Janjic (Janjic 1994) PBL scheme and the Geophysical Fluid Dynamics Laboratory (GFDL) radiation package. Both physics packages used Ferrier et al. (2002) microphysics and the Noah land surface scheme (Ek et al. 2003). These four simulations permit a comparison of the sensitivity of rainfall forecasts to changes in the dynamic core and to changes in the physics packages. It should be noted that differences in the numerics of the dynamic cores do result in some unavoidable small differences in the initializations (e.g., ARW runs directly use temperature data whereas NMM runs infer temperature hydrostatically using initial geopotential heights). These differences are small and should not greatly impact the forecasts. It also should be noted that 8-km grid spacing is likely finer than the range for which these convective parameterizations were originally designed, and the physical justification for use of such schemes may be questionable (e.g., Molinari and Dudek 1992). However, Kain and Fritsch (1998) have shown that adequate simulation of precipitation may require the use of such parameterizations at grid spacings as fine as 5 km, and these schemes have been used in recent years with similar grid spacings quasi operationally at NCEP.

In addition, WRF runs using the ARW dynamic core and initialized with the rapid update cycle (RUC) output instead of Eta output were performed by the DTC group (Bernardet et al. 2004) and are used to determine the sensitivity of the WRF-ARW runs to initial conditions. Some small differences are present in the grid design of these RUC-initialized runs with a horizontal grid spacing of 10 instead of 8 km, and 51 vertical layers instead of 60. The slightly coarser vertical and horizontal grid spacings were necessary due to computational constraints at the time the DTC group performed these simulations. Simulations of a few other cases with WRF

suggested that the small changes in grid spacing and number of vertical levels present in this comparison likely have a much smaller impact on the rainfall simulations than the differences in initial conditions. Rainfall from both the 10-km runs and all 8-km runs was remapped to an independent 8-km grid using procedures typically used at NCEP to allow a comparison of the output.

As a measure of forecast accuracy, the equitable threat score (ETS; Schaefer 1990) and bias were used, where

$$\text{ETS} = \frac{(\text{CFA} - \text{CHA})}{(F + O - \text{CFA} - \text{CHA})}, \quad (1)$$

$$\text{CHA} = O \frac{F}{V}, \quad (2)$$

and

$$\text{BIAS} = \frac{F}{O}. \quad (3)$$

In the above equations, each variable indicates the number of grid points at which (i) rainfall was correctly forecasted to exceed the specified threshold (CFA), (ii) rainfall was forecasted to exceed the threshold (F), (iii) rainfall was observed to exceed the threshold (O), and (iv) a correct forecast would occur by chance (CHA), where V is the total number of evaluated grid points. An ETS of 1 would occur with a perfect forecast, with lower values showing a less accurate forecast. Values of BIAS significantly higher (lower) than 1 indicate that the model notably overpredicted (underpredicted) areal coverage.

To determine the sensitivity of rainfall forecasts to changes in physics, dynamics, and initial conditions, the correspondence ratio (CR; Stensrud and Wandishin 2000) was used, where

$$\text{CR} = \frac{I}{U}, \quad (4)$$

and I is the number of grid points where all evaluated model runs (two if pairs are compared as in the present study) show rainfall exceeding a specified threshold (intersection of rainfall areas), and U is the number of grid points where at least one run shows rainfall above the threshold (union of rainfall areas). Greater diversity in solutions occurs when the CR is small. Although the spread ratio, the inverse of the CR, has an advantage over the CR in that large values reflect large spread, it becomes extremely large in those events where the intersection is small, and it can be unbounded in some situations. Therefore the CR has been used in this study. The correlation coefficient was also computed for all simulations, but because the rainfall structure

TABLE 1. Peak 6-hourly rainfall rates (in.) during the first 24 h of each forecast and total domain rain volume (km^3 , shown in parentheses) for 15 events during August 2002. Observed values taken from stage-IV multisensor data are also indicated, and averages for the full sample are boldfaced and shown in last row.

Case	Observed	ARW–NCAR	ARW–NCEP	NMM–NCAR	NMM–NCEP
3 Aug	3.24 (24.5)	4.60 (27.8)	2.22 (28.8)	5.49 (26.9)	1.70 (23.4)
4 Aug	4.38 (18.7)	7.34 (23.2)	2.77 (25.0)	4.56 (19.1)	1.23 (19.4)
5 Aug	4.74 (18.3)	5.45 (24.7)	1.85 (28.8)	4.63 (19.2)	1.38 (21.4)
8 Aug	4.43 (20.0)	4.57 (21.7)	2.78 (25.1)	5.85 (19.8)	1.43 (21.9)
9 Aug	2.57 (13.4)	4.20 (18.1)	2.83 (19.0)	4.65 (19.4)	1.68 (15.8)
11 Aug	2.84 (18.8)	3.69 (22.2)	2.23 (18.0)	3.11 (20.7)	2.10 (15.5)
12 Aug	3.56 (26.3)	3.27 (34.7)	2.42 (34.1)	6.69 (33.7)	1.69 (28.4)
13 Aug	4.80 (34.7)	4.20 (50.4)	3.49 (50.7)	2.80 (43.3)	1.30 (36.2)
16 Aug	3.16 (31.6)	4.09 (38.3)	2.77 (40.9)	5.15 (35.7)	2.03 (32.9)
18 Aug	4.48 (18.0)	2.70 (21.1)	3.29 (17.5)	4.91 (16.8)	1.33 (16.1)
20 Aug	3.53 (23.1)	4.33 (33.6)	4.87 (31.6)	5.09 (27.4)	1.48 (30.8)
21 Aug	4.11 (31.4)	7.25 (43.3)	2.41 (41.9)	4.69 (40.1)	2.08 (36.4)
22 Aug	3.05 (19.6)	6.74 (31.0)	3.42 (32.9)	4.85 (25.6)	1.18 (26.5)
26 Aug	3.30 (19.0)	5.09 (20.1)	2.72 (15.2)	4.99 (18.9)	0.78 (12.7)
28 Aug	2.81 (17.4)	2.84 (23.3)	2.94 (20.3)	7.37 (20.8)	1.23 (14.7)
Avg	3.67 (22.3)	4.65 (28.9)	2.87 (28.7)	4.99 (25.8)	1.51 (23.5)

shows so much more finescale detail when the NCAR physics are used, correlation coefficients were almost always far smaller in any comparison involving a run with these physics than in comparisons of runs using NCEP physics. Correlation coefficients therefore did not appear to add meaningful insight into the sensitivity of forecasts and are not shown.

3. Results

The 15 events chosen from August 2002 (see Table 1) all included substantial areas of convective precipitation within the 48-h forecasts. Peak observed 6-h rainfall totals at any grid point within the first 24 h of each forecast period generally exceeded 3 in., with rain volume in the model domain ranging from 13.4 km^3 on 9 August to 34.7 km^3 on 13 August. Observations taken from 4-km gridded stage-IV multisensor data (Baldwin and Mitchell 1997) were remapped to the 8-km grid used for verification. It should be noted that Schwartz and Benjamin (2000) have found the stage-IV multisensor data to be wetter for rainfall amounts under 0.5 in. in 24 h than gauge-only data, and drier for heavier amounts. In 8 of the 15 cases, the observed rainfall rates were largest between 0000–0600 UTC (corresponding to the 12–18-h model forecast period), and in only one event was it largest in the first 6 h (not shown). Observed domain rain volume behaved similarly with no cases having the largest 6-hourly volume during the first 6 h of a forecast. These results suggest that any model spinup delaying rainfall production should not seriously impact the simulation of the most active periods of convective system rainfall.

Peak 6-hourly rainfall amounts and total domain rain volume in the first 24 h of the forecast period varied substantially among the four different WRF configurations using different dynamic cores and physics packages (Table 1). Peak rain rates were overestimated in most events when the NCAR physics package was used with both the ARW (12 out of 15 cases) and NMM (13 out of 15 cases) dynamic cores. Peak rain rates were underestimated often when the NCEP physics package was used in both dynamic cores (11 times in ARW, all 15 times in NMM). These differences in behavior of peak rain rates as a function of physics package are most likely related to the use of the BMJ convective scheme in the NCEP package, and the KF scheme in the NCAR package (e.g., Gallus 1999).

Overestimates of rain volume were common in all four configurations, with every case overestimated by ARW–NCAR, 12 overestimated by NMM–NCAR, 11 by ARW–NCEP, and 10 by NMM–NCEP. Overestimates were particularly large in the first 6 h (not shown), ranging from a 40% overestimate in NMM–NCAR to a 205% overestimate in ARW–NCEP. Both runs using NCEP physics had particularly large overestimates in the first 6 h, potentially evidence of model spinup effects. In addition, overestimates of rain volume in all configurations may be at least partly a consequence of the use of the Ferrier et al. microphysics scheme. Jankov et al. (2005) found in a study of a different set of warm season events that the microphysical scheme choice could have a significant impact on the total domain rain volume. The Ferrier et al. scheme was substantially wetter than the NCEP-5 class scheme (Hong et al. 1998), although it was not as wet as the Lin

TABLE 2. ETSs and BIAS values for rainfall thresholds of 0.01, 0.25, and 0.5 in. (indicated in parentheses) averaged for 6-h accumulation periods during day 1 for the 15 events shown in Table 1.

	ARW-NCAR	ARW-NCEP	NMM-NCAR	NMM-NCEP	ARW-NCAR (RUC-init)	ARW-NCEP (RUC-init)
ETS (0.01)	0.304	0.293	0.304	0.274	0.288	0.288
BIAS (0.01)	1.410	1.871	1.034	1.810	1.500	1.862
ETS (0.25)	0.145	0.147	0.159	0.119	0.145	0.138
BIAS (0.25)	1.284	1.225	1.101	0.924	1.352	1.280
ETS (0.5)	0.092	0.085	0.103	0.058	0.094	0.078
BIAS (0.5)	1.103	0.886	1.167	0.525	1.160	0.960

et al. (1983). Despite the tendency for the NCAR physics to produce much greater peak rain rates than the NCEP physics (on average 60% larger with ARW and 230% larger with NMM), total rain volume was much more similar between the two physics packages, with both ARW runs wetter than both NMM runs. Even if the first 6 forecast hours are excluded because of the excessive rain volumes in the runs using NCEP physics, the volumes differ less as physics are changed than the rain rates did. Thus, *peak rain rates are much more sensitive to the physics package used than the dynamic core used, but total domain rain volume may be more sensitive to the dynamic core choice than the physics used.*

Table 2 shows ETS and bias values averaged over all 6-hourly periods during the first 24 h of the forecasts of the 15 events. The skill of the forecasts is comparable in five of the six configurations, with the only exception being in the NMM run using NCEP physics, where ETSs are noticeably lower at each threshold. As is typical for warm season rainfall (e.g., Gallus and Segal 2001), skill decreases rapidly for heavier amounts, and little skill is present for rainfall amounts of 0.5 in. or more. All configurations except NMM running with NCAR physics tend to have too high of a bias for light rain. The high bias is especially pronounced in runs using the NCEP physics. At the 0.5-in. threshold, the runs using NCEP physics evidence a bias less than 1, or a tendency to underpredict areal coverage. The bias trends (too high for light rain, too low for heavy rain) in the runs using NCEP physics agree with trends found by Gallus and Segal (2001) for the NCEP Eta model when the BMJ convective scheme was used. Overall, no particular configuration stands out as being much better than any other, although there is a suggestion that the NMM core using NCEP physics did not perform as well as other configurations (although recall that this configuration in Table 1 performed best in terms of domain rain volume).

Figure 1 shows rainfall forecast variability among the four 8-km WRF runs using different dynamic cores and

physics packages for the 12–18-h forecast period for a run initialized at 1200 UTC 13 August. Although this period is only one from a total of 120 available 6-h periods, it demonstrates the typical variability seen when these changes were made in the model. Among the more obvious differences is the finer-scale structure and greater intensities of rainfall occurring when the NCAR physics package is used (left panels in Fig. 1) compared to the NCEP physics (right panels). This difference is most likely related to the different convective parameterizations used. Past studies (e.g., Gallus 1999) have shown that the KF scheme permits more grid-resolved precipitation to occur and results in both isolated heavier amounts and more finescale structure than the BMJ scheme. It should be noted, however, that although the structure of the rainfall regions is strongly influenced by the physics package used, the intensity and spatial coverage of rainfall is also influenced noticeably by the dynamic core used. There is some indication in Fig. 1 of drier solutions in both NMM runs than in the ARW runs, particularly when NCEP physics were used, as was evident in Table 1. These results subjectively suggest that rainfall forecasts are sensitive to both changes in dynamic core and physics package, although the impacts from changes in these routines are manifest in different ways.

Observed precipitation is plotted in Fig. 2. A comparison of Figs. 1 and 2 suggests that all four simulations differ substantially from the observations. All model versions produced too much precipitation in Illinois and areas farther north, although the NMM run using NCEP physics was so much drier than the other versions everywhere that this problem was not as pronounced. All versions also failed to depict a region of rainfall in northern Mississippi and western Tennessee out ahead of the main line. Farther south, the model versions agreed better with observations, although they depicted a more linear character to the rainfall than was observed. These results suggest that insufficient spread would be present in an ensemble using these four configurations.

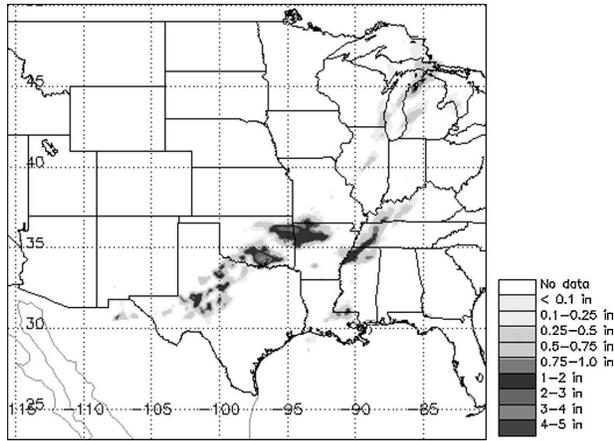


FIG. 2. Observed rainfall (from stage-IV multisensor analysis) in the 0000–0600 UTC period on 14 Aug 2002 corresponding to the 12–18-h forecast period in the runs shown in Fig. 1.

To perform a more thorough analysis of sensitivity to changes in dynamic core, physics package, and initialization, CRs for pairs of model runs (or couplets) were computed for all 6-h time periods. Table 3 shows CRs for two rainfall thresholds (0.01 and 0.50 in.) averaged for all 6-h periods within the first 24 h of the forecast for six couplets reflecting a change in one model component alone (dynamics, physics, and initial conditions) and five other couplets reflecting changes in multiple components. The model runs compared in Table 3 are listed from smallest to largest CR, or from the greatest impact on the rainfall forecast to the least impact.

As might be expected, the largest impacts (smallest CRs) for both thresholds occurred when all three components were changed, although for the lighter threshold, there was a large difference in ranking between the case when the 10-km ARW core (RUC initialization) running with NCEP physics was contrasted with the 8-km NMM (Eta initialization) using NCAR physics, and the case when the 10-km ARW core (RUC initialization) running with NCAR physics was compared to the 8-km NMM (Eta initialization) using NCEP physics. As will be evidenced by many of the other couplets, the sensitivity to any one component is a function of the other components.

Examining just those couplets where one component alone was changed (boldface in Table 3), it can be seen that a change in the physics package alone while using the NMM core resulted in a bigger impact on the forecast than in several couplets where the dynamic core and either the physics package or the initialization data (with inherent change in grid size and vertical layers) were altered. Rigorous hypothesis testing following Hamill’s (1999) resampling technique showed that the

TABLE 3. The CRs ranked from most sensitivity to least for comparison of component changes [dynamic core (D), physics package (P), initial conditions and grid spacing (IG)] at rainfall thresholds of 0.01 and 0.50 in. Changes involving one component alone are boldfaced. Parenthetical expressions show (i) runs compared when multiple components have been changed (notation A is used for ARW core, N for NMM, R for NCAR physics, and P for NCEP physics, with r indicating RUC initial conditions and a 10-km grid, and e indicating Eta initial conditions and an 8-km grid), and (ii) dynamic core or physics package held constant when only one component has been changed.

CR (0.01)	Components changed	CR (0.5)	Components changed
0.311	All (APr-NRe)	0.066	All (ARr-NPe)
0.330	D + IG (NCAR)	0.092	All (APr-NRe)
0.347	P (with NMM)	0.105	P (with NMM)
0.356	D + P (AP-NR)	0.113	D + P (AR-NP)
0.374	All (ARr-NPe)	0.113	D + IG (NCAR)
0.401	D + P (AR-NP)	0.142	D + P (AP-NR)
0.424	D (with NCAR)	0.146	IG (with NCEP)
0.457	P (with ARW)	0.152	IG (with NCAR)
0.490	IG (with NCAR)	0.157	P (with ARW)
0.503	IG (with NCEP)	0.240	D (with NCEP)
0.565	D (with NCEP)	0.244	D (with NCAR)

sensitivity to physics while using the NMM core was statistically significantly larger than the sensitivity found for all other couplets shown in boldface in Table 3 [with 95% confidence in all cases except for initial conditions and grid spacing (with NCEP) where confidence was 90%]. For all other couplets where only one component was changed, the CRs were larger (less spread) than those when multiple components were changed.

The temporal evolution of CRs over the full 48 h of the forecast is shown for the 0.01-in. threshold in Fig. 3, and the 0.50-in. threshold in Fig. 4. For the light threshold, except in the first 12 h, the same general ranking holds at all times (Fig. 3). The greatest sensitivity (lowest CR) is present when the physics package is changed in WRF runs using the NMM dynamic core. The next greatest sensitivity occurs when the dynamic core is changed while using the NCAR physics package. Interestingly, the sensitivity to this dynamic core change is greater than that for a change in physics package when the ARW dynamic core is running. At most times, CRs are at least 0.1 higher for a physics package change when the ARW core is used compared to when the physics are changed with the NMM core. If one assumes that roughly 10% of the model domain was forecasted to experience at least 0.01 in. of rainfall (roughly 10 000 points), this difference in CRs reflects about a 1000 gridpoint decrease in points (roughly 250 × 250 km area) where both model runs predicted rainfall above the threshold, and a roughly 1500 gridpoint in-

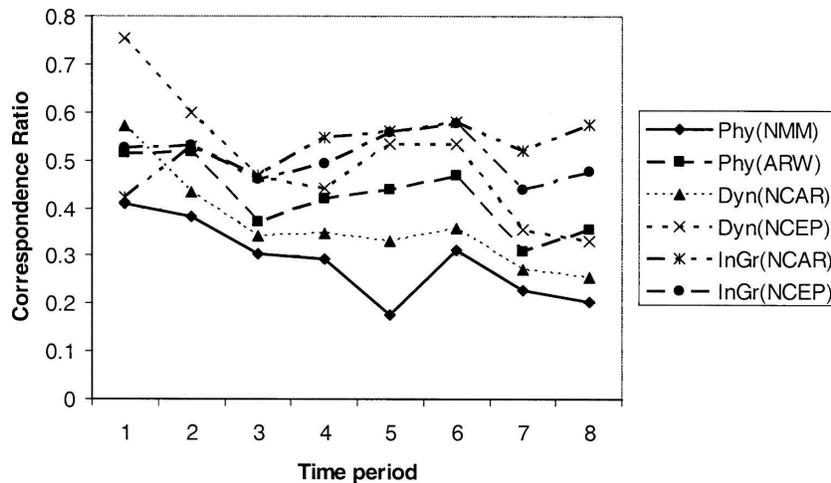


FIG. 3. Temporal evolution of CRs for 0.01-in. rain threshold in six WRF configuration comparisons (dynamic core or physics package held constant indicated in parentheses). Time periods 1–8 correspond to 0–6-, 6–12-, 12–18-, 18–24-, 24–30-, 30–36-, 36–42-, and 42–48-h forecasts.

crease in the number of points of disagreement where only one of the runs predicts rainfall above the threshold.

The CRs reflecting sensitivity to a change in dynamic core are likewise much higher (less sensitivity) when the NCEP physics package is used than when the NCAR package is used. This result is understandable since the broad precipitation regions created by the BMJ convective scheme in the NCEP package likely minimize changes in CR when the dynamic core is changed. Small changes in location of rainfall areas are more likely to influence CR when those rainfall areas are small with substantial finescale structure, as occurs with the KF scheme in the NCAR physics package. The

lack of sensitivity to changes in dynamic core when the NCEP physics are used is especially pronounced in the first 12 h of the forecast.

For the light threshold, during the first 6 h of the forecast, the sensitivity to initialization dataset is substantial. The CR when NCAR physics are used is nearly as small as the lowest value, which was associated with a change in physics package. However, whereas the sensitivity to changes in dynamic core or physics increases in most cases through the first 24 h, the sensitivity to initialization generally lessens with time over the first 24–36 h. Thus, by hour 18, both couplets reflecting the impact of changes in initialization have higher CRs than any other couplet. For most couplets,

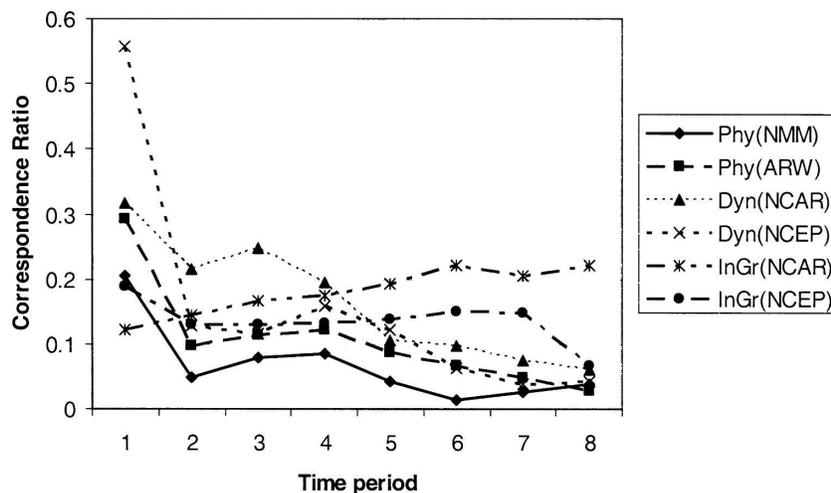


FIG. 4. Same as in Fig. 3 but for 0.50-in. rainfall threshold.

the decline in CR levels off or switches to an increase after the first 18–24 h. A local maximum is present around hour 36 implying the forecasts become somewhat more similar at this time, corresponding to 1800–0000 UTC in the day-2 forecast period. This is typically the period when the troposphere is most convectively unstable, and areal coverage of light rain in the models (and observations) is maximized. The CRs drop quickly after this time with most of the dynamic core and physics changes showing their lowest CRs in the 42–48-h period, a time when nocturnal MCSs are often at their mature stages within this model domain.

Resampling techniques were applied to the data in each 6-h period to determine the statistical significance of differences (*p* values not shown). In general, standard deviations were roughly 0.05 for each test shown in Fig. 3, and differences in the curves were significant with 95% confidence if CRs differed by approximately this amount or more. At all times, the sensitivity to physics was significantly larger when using the NMM dynamic core than when using the ARW core. Likewise, sensitivity to dynamic core choice was significantly larger when NCAR physics were used than when NCEP physics were used. At most times, the sensitivity to physics with the NMM core was significantly larger than that for any other model component examined.

For heavier rainfall amounts (threshold of 0.50 in.), which are restricted to much smaller areas of the model domain, some small differences can be seen in the temporal behavior of CRs (Fig. 4). Once again, the biggest sensitivity at most times is associated with a change in physics package in runs using the NMM dynamic core. One exception to this general trend is present during the first 6 h of the forecast when the greatest sensitivity occurs due to a change in initialization dataset. As with the lighter threshold, sensitivity to initial conditions becomes relatively less pronounced with time, such that both of these couplets have the largest CRs after the 18–24-h period. Also similar to the trends present at the lighter threshold, the least sensitivity in the first 6 h arises from a change in dynamic core when NCEP physics are used. However, the sensitivity increases greatly over time for this heavier threshold, and at most other times is similar to that for changes in physics when the ARW core is used, and changes in the dynamic core when NCAR physics are used. Unlike the trends present at the 0.01-in. threshold, for a large portion of the forecast period (all times after 6 h except the 24–30-h period), the sensitivity is greater for changes in dynamic core when the NCEP physics are used than it is when the NCAR physics are used. It was pointed out that the broad regions of relatively light rainfall produced by the NCEP physics result in high CRs for the

0.01-in. threshold. Because areas of heavy rainfall are relatively smaller in runs using the NCEP physics than in those using NCAR, changes in dynamic core result in more variation in the forecast at the 0.5-in. threshold for the NCEP runs than for the NCAR ones. Bias scores (not shown) support this conclusion, with much smaller values at the 0.50-in. threshold in runs using NCEP physics than in runs using NCAR physics.

Resampling techniques applied to the 0.50-in. threshold (*p* values not shown) indicated fewer cases where differences were statistically significant (with 95% confidence). Sensitivity to physics changes while using the NMM core were still significantly larger than those of most other component changes except when compared to physics changes while using the ARW core, and dynamic core changes while using NCEP physics. Apparently the small areas of heavier rainfall are influenced enough by changes in most model parameters that the differences in CRs shown in Fig. 4 are not statistically significant.

4. Conclusions

A series of tests were performed at 8- and 10-km grid spacing with the WRF model to compare the sensitivity of rainfall forecasts to changes in model physics, dynamics, and initial conditions/grid spacing. Fifteen warm season rainfall events from August 2002 were examined. Both the ARW and NMM dynamic cores were used, along with two physics packages: NCAR used the KF convective parameterization, YSU PBL scheme, and Dudhia/RRTM radiation, while NCEP used the BMJ convective scheme, Mellor–Yamada–Janjic PBL scheme, and the GFDL radiation package. Other physical schemes were the same (e.g., Noah land surface model, Ferrier et al. microphysics) in all runs. All four of the model configurations using these dynamic cores and physics packages were initialized using Eta output. The ARW dynamic core runs also were compared with 10-km grid spacing WRF–ARW runs performed by the WRF–DTC (Bernardet et al. 2004) for these cases using RUC output for initialization to determine sensitivity to initial condition dataset and the small change in grid spacing (and number of vertical layers).

It was found that sensitivity to any one component was influenced by other components. The greatest sensitivity resulted from changes in the physics package when the NMM dynamic core was used. For light rainfall amounts, the next strongest sensitivity was from a change in dynamic core while NCAR physics were used. The use of NCEP physics had a much smaller impact for light rainfall, likely due to the large and

smooth rainfall regions produced by the BMJ convective scheme in that package. For heavier rainfall, the ranking of sensitivity to changes in specific components varied much more over time. Because the NCEP physics package led to a much smaller bias at the heavier amounts than the NCAR physics package, runs were generally more sensitive to a dynamic core change under the NCEP physics than under the NCAR physics, unlike the behavior noted for lighter rainfall. For both thresholds evaluated, the impact of initial condition changes (with small changes in grid spacing and vertical layers) was generally smaller than that of changes in dynamics or physics, except during the first 6–12 h of the forecast.

Acknowledgments. We thank Dr. Robert Gall for the invitation for W. Gallus to visit the WRF-DTC during the summer of 2004, and for the assistance of WRF-DTC staff in accessing retrospective WRF runs, analyzing the precipitation output, and in running the new 8-km simulations with the same version of WRF on the WRF-DTC computers. These people included Wei Wang, Louisa Nance, and Dave Gill at NCAR, and Andrew Loughe, Ligia Bernardet, and Linda Wharton at NOAA's Forecast Systems Laboratory. We would also like to acknowledge the helpful comments of two anonymous reviewers. This study was supported in part by NSF Grant ATM-0226059.

REFERENCES

- Alhamed, A., S. Lakshmiarahan, and D. J. Stensrud, 2002: Cluster analysis of multimodel ensemble data from SAMEX. *Mon. Wea. Rev.*, **130**, 226–256.
- Baldwin, M. E., and K. E. Mitchell, 1997: The NCEP hourly multisensor U.S. precipitation analysis for operations and GCIIP research. Preprints, *13th Conf. on Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 54–55.
- Bernardet, L. R., L. Nance, H.-Y. Chuang, A. Loughe, and S. E. Koch, 2004: Verification statistics for the NCEP WRF pre-implementation test. Part I: Deterministic verification of ensemble members. Preprints, *Fifth WRF/14th MM5 Users' Workshop*, Boulder, CO, National Center for Atmospheric Research, 229–232.
- Betts, A. K., 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc.*, **112**, 677–692.
- , and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709.
- Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, **52**, 1434–1456.
- , M. J. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.
- Ek, M. B., K. E. Mitchell, Y. Lin, P. Grunmann, E. Rogers, G. Gayno, and V. Koren, 2003: Implementation of the upgraded Noah land-surface model in the NCEP operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.
- Ferrier, B. S., Y. Jin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in NCEP Eta model. Preprints, *15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., 280–283.
- Gallus, W. A., Jr., 1999: Eta simulations of three extreme precipitation events: Sensitivity to resolution and convective parameterization. *Wea. Forecasting*, **14**, 405–426.
- , and M. Segal, 2001: Impact of improved initialization of mesoscale features on convective system rainfall in 10-km Eta simulations. *Wea. Forecasting*, **16**, 680–696.
- , J. C. Correia, and I. Jankov, 2005: The 4 June 1999 derecho event: A particularly difficult challenge for numerical weather prediction. *Wea. Forecasting*, **20**, 705–728.
- Hamill, T. M., 1999: Hypothesis test for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167.
- Hong, S.-Y., H.-M. H. Juang, and Q. Zhao, 1998: Implementation of prognostic cloud scheme for a regional spectral model. *Mon. Wea. Rev.*, **126**, 2621–2639.
- Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.*, **123**, 2181–2196.
- Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 928–945.
- Jankov, I., W. A. Gallus Jr., M. Segal, B. Shaw, and S. E. Koch, 2005: The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall. *Wea. Forecasting*, **20**, 1048–1060.
- Kain, J. S., and J. M. Fritsch, 1992: The role of convective “trigger function” in numerical forecasts of mesoscale convective systems. *Meteor. Atmos. Phys.*, **49**, 93–106.
- , and —, 1998: Multiscale convective overturning in mesoscale convective systems: Reconciling observations, simulations, and theory. *Mon. Wea. Rev.*, **126**, 2254–2273.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk scheme of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.
- McBride, J. L., and E. E. Ebert, 2000: Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. *Wea. Forecasting*, **15**, 103–121.
- Molinari, J., and M. Dudek, 1992: Parameterization of convective precipitation in mesoscale numerical models: A critical review. *Mon. Wea. Rev.*, **120**, 326–344.
- Noh, Y., W. G. Cheon, S.-Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.*, **107**, 401–427.
- Olson, D. A., N. W. Junker, and B. Korty, 1995: Evaluation of 33 years of quantitative precipitation forecasting at the NMC. *Wea. Forecasting*, **10**, 498–511.
- Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575.
- Schwartz, B. E., and S. G. Benjamin, 2000: Verification of RUC-2 precipitation forecasts using the NCEP multisensor analysis. Preprints, *Fourth Symp. on Integrated Observing Systems*, Long Beach, CA, Amer. Meteor. Soc., 182–185.
- Seaman, N., R. Gall, L. Nance, S. Koch, L. Bernardet, G. DiMego, J. Powers, and F. Olsen, 2004: The WRF process: Streamlin-

- ing the transition of new science from research into operations. Preprints, *Fifth WRF/14th MM5 Users' Workshop*, Boulder, CO, National Center for Atmospheric Research, 173–176.
- Stensrud, D. J., and M. S. Wandishin, 2000: The correspondence ratios in forecast evaluation. *Wea. Forecasting*, **15**, 593–602.
- , J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.
- Wandishin, M. S., S. L. Mullen, D. J. Stensrud, and H. E. Brooks, 2001: Evaluation of a short-range multimodel ensemble system. *Mon. Wea. Rev.*, **129**, 729–747.