P1.89 COMPARISON OF IMPACTS OF WRF DYNAMIC CORE, PHYSICS PACKAGE, AND INITIAL CONDITIONS ON WARM SEASON RAINFALL FORECASTS

William A. Gallus, Jr.
Iowa State University, Ames, Iowa

1. INTRODUCTION

A series of simulations were performed for 15 warm season convective system cases occurring during August 2002 using several different WRF (Weather Research and Forecasting) model configurations to compare the sensitivity of the rainfall forecasts to changes in dynamic core, physics package, and initial conditions. Most simulations used 8 km grid spacing, but a few 10 km simulations available from the WRF Developmental TestBed Center (DTC) for these cases (Seaman et al. 2004; Bernardet et al. 2004) were used to determine sensitivity to initial conditions and the small change in grid spacing. The tests were motivated in part by recent findings that variations in model physics can lead to more diversity among ensemble members than changes in initial conditions (e.g. Stensrud et al. 2000; Gallus and Segal 2001), and that members tend to cluster first by model, next by physics, and lastly by initial conditions in a mixed model, physics and initial condition ensemble (Alhamed et al. 2002). The present study seeks to examine if changes in dynamic core result in similar impacts to changes in model, and quantify the impacts of changes in these three different components of model configurations. In addition, 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. METHODOLOGY

The WRF model was run with 8 km grid spacing for 15 events occurring during August 2002. All of the 8 km runs were integrated for 48 hours, with initial and lateral boundary condition data provided from 40 km Eta GRIB output. The domain covered most of the central United States and can be seen in Fig. 1. Four different versions of the WRF model were run at 8 km grid spacing, with two different dynamic cores and two different physics packages. The dynamic cores included the ARW (Advanced Research WRF) and NMM (Nonhydrostatic Mesoscale Model). One physics package, denoted NCEP, used the Betts-Miller-Janjic (BMJ; Betts 1986, Betts and Miller 1986, Janjic 1994) convective parameterization, Miller-Yamada-Janjic planetary boundary layer (PBL; Janjic 1994) scheme, and GFDL radiation package. The other, denoted NCAR, used the Kain-Fritsch (KF; Kain and Fritsch 1992) convective scheme, YSU PBL scheme, and Dudhia/RRTM radiation package. Both physics packages used the Ferrier et al. (2002) microphysics scheme and NOAH land surface scheme. The four configurations resulting from the use of the different dynamic cores and physics packages will help to show the sensitivity of rain forecasts to the choice of dynamics or physics.

In addition to these 4 configurations, two other WRF versions will be used to determine sensitivity to initial conditions and small changes in grid spacing. These versions, run by the WRF-DTC (Bernardet et al. 2004), used the ARW dynamic core but were initialized with RUC output and run with a 10 km grid. Rainfall from these 10 km runs was remapped to the standard 8 km grid using procedures typically used at NCEP.

To determine the sensitivity of rainfall forecasts to changes in physics, dynamics, and initial conditions, peak rain rates, domain total rain volume, and correspondence ratios between pairs of configurations will be used. Correspondence ratio (CR; Stensrud and Wandishin 2000) is the ratio of the number of grid points in a set of model runs (2 in the present study) that all show precipitation above a given threshold to the number of grid points where at least one run shows precipitation above that threshold (intersection / union). Correlation coefficients were also computed but were found to be so strongly influenced by the fineness of precipitation structures that the parameter did not provide useful information about sensitivity. As will be shown later, the use of NCAR physics resulted in far more detailed precipitation patterns. Any pair of runs that included at least one run using NCAR physics had extremely low correlation coefficients, whereas comparisons between two runs where both used the NCEP physics, which produced much smoother rainfall fields, resulted in much higher correlation coefficients.

3. RESULTS

The 15 events chosen from August 2002 (Table 1) all included substantial areas of

---

Corresponding author address: William A. Gallus, Jr., ISU, Agronomy Hall 3025, Ames, IA. Email: wgallus@iastate.edu.

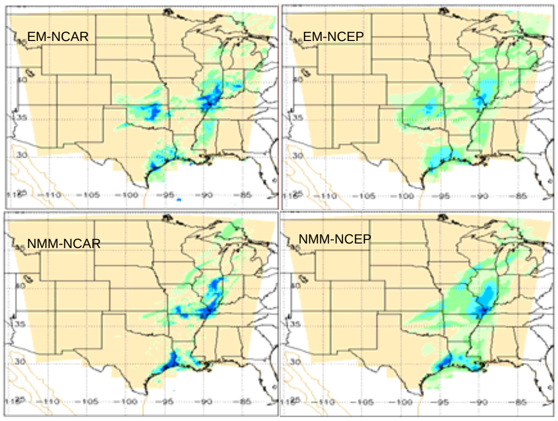

Figure 1: Rainfall in the first 6 forecast hours of a case initialized 12 UTC 28 August 2002. Runs using ARW (also known as EM) dynamic core shown at top, NMM at bottom. NCAR physics used in left panels; NCEP used at right.

convective rainfall within the 48 hour forecasts. Peak observed 6 hour rainfall totals within the first 24 hour period of each forecast generally exceeded 3 inches (Table 1), with rain volume in the model domain region ranging from 13.4 km^3 on April 9 to 34.7 km^3 on August 13 (Table 2). Observations are taken from 4 km gridded Stage IV multi-sensor data (Baldwin and Mitchell 1987) remapped to the model's 8 km grid. In 8 of the 15 cases, the observed rainfall intensity was largest in the 12-18 h forecast period (00-06 UTC), and in only one event was it largest in the first 6 hours (not shown). Observed domain rain volume behaved similarly with no cases having the largest 6-hourly volume during the first 6 hours of a forecast.

Case	Obs	A-R	A-P	N-R	N-P
8/3	3.24	4.60	2.22	5.49	1.70
8/4	4.38	7.34	2.77	4.56	1.23
8/5	4.74	5.45	1.85	4.63	1.38
8/8	4.43	4.57	2.78	5.85	1.43
8/9	2.57	3.69	2.83	4.65	1.68
8/11	2.84	3.69	2.23	3.11	2.10
8/12	3.56	3.27	2.42	6.69	1.69
8/13	4.80	4.20	3.49	2.80	1.30
8/16	3.16	4.09	2.77	5.15	2.03
8/18	4.48	2.70	3.29	4.91	1.33
8/20	3.53	4.33	4.87	5.09	1.48
8/21	4.11	7.25	2.41	4.69	2.08
8/22	3.05	6.74	3.42	4.85	1.18
8/26	3.30	5.09	2.72	4.99	0.78
8/28	2.81	2.84	2.94	7.37	1.23
AVE	**3.67**	**4.65**	**2.87**	**4.99**	**1.51**

Table 1: Peak 6-hour rainfall amounts (inches) within first 24 hours of each forecast for all 15 cases. Second column from left shows Stage IV observations averaged to common 8 km grid. Bottom row is the average for all 15 cases. Notation A is used for ARW core, N for NMM, R for NCAR physics, P for NCEP physics.

Peak 6-hourly rainfall amounts (Table 1) and total domain rain volume (Table 2) in the first 24 hours of the forecast period varied substantially among the four different WRF configurations using different dynamic cores and physics packages. Peak rain intensities were overestimated in most events when the NCAR physics package was used with both the ARW (12 of 15 cases) and NMM (13 of 15 cases) 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). 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 (Table 2). Despite the tendency for the NCAR physics to produce much greater peak rain intensities than the NCEP physics (on average, 60% larger rates with the ARW core and 230% larger with the NMM core), total rain volume was much more comparable between the two physics packages, with both ARW runs wetter than both NMM runs. It thus appears that peak rain rates are much more sensitive to the physics package used than the dynamic core, but total domain rain volume is more sensitive to the dynamic core than the physics used.

Case	Obs	A-R	A-P	N-R	N-P
8/3	24.5	27.8	28.8	26.9	23.4
8/4	18.7	23.2	25.0	19.1	19.4
8/5	18.3	24.7	28.8	19.2	21.4
8/8	20.0	21.7	25.1	19.8	21.9
8/9	13.4	18.1	19.0	19.4	15.8
8/11	18.8	22.2	18.0	20.7	15.5
8/12	26.3	34.7	34.1	33.7	28.4
8/13	34.7	50.4	50.7	43.3	36.2
8/16	31.6	38.3	40.9	35.7	32.9
8/18	18.0	21.1	17.5	16.8	16.1
8/20	23.1	33.6	31.6	27.4	30.8
8/21	31.4	43.3	41.9	40.1	36.4
8/22	19.6	31.0	32.9	25.6	26.5
8/26	19.0	20.1	15.2	18.9	12.7
8/28	17.4	23.3	20.3	20.8	14.7
AVE	**22.3**	**28.9**	**28.7**	**25.8**	**23.5**

Table 2: Domain total rain volume (km^3) in the first 24 hours of each forecast. Second column from left based on Stage IV observations averaged to common 8 km grid. Bottom row is the average for all 15 cases. Notation as in Table 1.

Figure 1 shows rainfall forecast variability among the four 8 km WRF runs using different dynamic cores and physics packages for the first 6 hours of a forecast initialized at 12 UTC 28 August. Although this figure shows only one 6 hour period from a total of 120 available, 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 two panels of Fig. 1) compared to the NCEP physics (right two panels). This difference is most likely related to the different convective parameterizations used. The NCAR physics used the KF scheme while the NCEP package used the BMJ scheme. Numerous studies (e.g., Gallus 1999; Grams et al. 2005) have shown that the KF scheme permits more grid-resolved precipitation to occur and results in both isolated heavier amounts and finer scale structure than the BMJ scheme. A more detailed look at the figure shows, however, that although the structure of the rainfall regions is strongly influenced by the physics package used, the general locations of rainfall may be more influenced by the dynamic core used.

Note that both of the ARW (EM) runs concentrate precipitation in three regions: (i) Oklahoma and southeastern Kansas, (ii) Texas Gulf coast, and (iii) near the Mississippi River. The WRF runs using the NMM core, however, show precipitation in slightly different areas. The rainfall in Oklahoma and Kansas is less concentrated and instead a more linear feature is hinted at from parts of Oklahoma northeastward into eastern Iowa and Wisconsin. The rainfall near the Gulf coast of Texas is more concentrated near the border with Louisiana. The rainfall area near the Mississippi River may be least changed from the ARW runs, although both NMM runs show somewhat less rainfall in Mississippi. 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.

To perform a more thorough analysis of sensitivity to changes in dynamic core, physics package, and initialization, CRs between any two model runs were computed for all 6 hour time periods. Table 3 shows average CRs for 2 rainfall thresholds (.01 and .5 inch) for all 6 hour periods within the first 24 hours of the forecast for 6 couplets reflecting a change in one model component alone (dynamics, physics, initial conditions) and 5 other couplets reflecting changes in multiple components. The model runs compared in Table 3 are listed from smallest to largest CR, or from greatest impact on the rainfall forecast to 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 8km 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.

Rank	CR (.01)	Changes	CR (.5)	Changes
1	.311	All (APr-NRe)	.066	All (ARr-NPe)
2	.330	D+IG (NCAR)	.092	All (APr-NRe)
3	**.347**	**P (NMM)**	**.105**	**P (NMM)**
4	.356	D+P (AP-NR)	.113	D+P (AR-NP)
5	.374	All (ARr-NPe)	.113	D+IG (NCAR)
6	.401	D+P (AR-NP)	.142	D+P (AP-NR)
7	**.424**	**D (NCAR)**	**.146**	**IG (NCEP)**
8	**.457**	**P (ARW)**	**.152**	**IG (NCAR)**
9	**.490**	**IG (NCAR)**	**.157**	**P (ARW)**
10	**.503**	**IG (NCEP)**	**.240**	**D (NCEP)**
11	**.565**	**D (NCEP)**	**.244**	**D (NCAR)**

Table 3: CRs ranked from most sensitivity to least for comparison of component (D = dynamic core, P = physics package, IG = initial conditions and grid spacing) changes at rainfall thresholds of .01 and .5 inch. Changes involving one component alone are boldfaced. Parenthetical expressions show (i) runs compared when multiple components are changed (notation as in Table 1 with r indicating RUC initial conditions and 10 km grid, and e indicating Eta initial conditions and 8 km grid), and (ii) dynamic core or physics package held constant when only one component is changed.

Examining just those couplets where one component alone was changed (boldfaced in Table 3), it can be seen that a change in 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 and grid size were changed. Rigorous hypothesis testing following Hamill's (1999) resampling technique showed that the 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 IG(NCEP) where confidence was 90%]. For all other couplets where only one component was changed, the CRs were larger than those when multiple components were changed.

The temporal evolution of CRs over the full 48 hours of the forecast is shown for the .01 inch threshold in Fig. 2, and the .50 inch threshold in Fig. 3. For the light threshold, except in the first 12 hours, the same general ranking holds at all times (Fig. 2). 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 .1 higher for a physics package change when the ARW core is used compared to the NMM core. If one assumes that roughly 10% of the model domain was

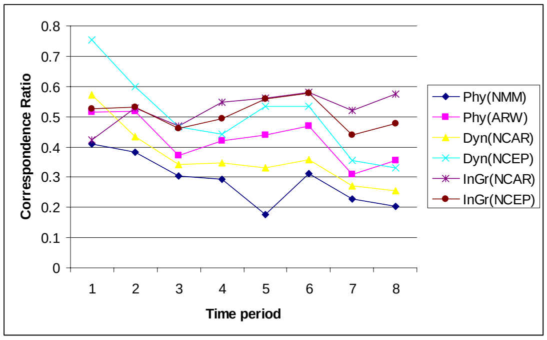

Figure 2: Temporal evolution of CRs for .01 inch rain threshold in 6 WRF configuration comparisons. Time periods 1-8 correspond with 0-6, 6-12, 12-18, 18-24, 24-30, 30-36, 36-42, and 42-48 h forecast periods.

forecasted to experience at least .01 inch of rainfall (roughly 10,000 points), this difference in CRs reflects about a 1000 grid point decrease in points (roughly 250 x 250 km area) where both model runs predicted rainfall above the threshold, and a roughly 1500 grid point increase 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 fine-scale 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 6-12 hours of the forecast.

For the light threshold, during the first 6 hours of the forecast, sensitivity to initialization data set 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 hours, the sensitivity to initialization changes generally lessens with time over the first 24-36 hours. Thus, by hour 18, both couplets reflecting the impact of changes in initialization data set have higher CRs than any other couplet. For most couplets, the decline in CR levels off or switches to an increase after the first 18-24 hours. A local maximum is present around hour 36 implying the forecasts become somewhat more similar at this time, corresponding to 18-00 UTC in the day 2 forecast period. This is typically the period when the troposphere is most convectively unstable. CRs drop quickly after this time with most of the dynamic core and physics changes showing their lowest values 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 hour period to determine the statistical significance of differences. In general, standard deviations were roughly .05 for each test shown in Fig. 2, and differences in the curves were significant with 95% confidence if CRs differed by approximately this amount or more. Thus, 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, which are restricted to much smaller areas of the model domain, some small differences can be seen in the behavior of CRs (Fig. 3) over time. 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 hours of the forecast when the greatest sensitivity occurs due to a change in initialization data set and grid spacing. As with the lighter threshold, sensitivity to initial condition changes becomes relatively less pronounced with time, with both of these couplets having larger CRs than in the other tests after the 24-30 hour period. Also similar to the trends present at the lighter threshold, the least sensitivity in the first 6 hours arises from a change in dynamic core when NCEP physics are used. However, the sensitivity increases greatly 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 .01 inch threshold, for a large portion of the forecast period (all times after 6 hours except the 24-30 h period), the sensitivity is greater for changes in dynamic core when the NCEP physics are used

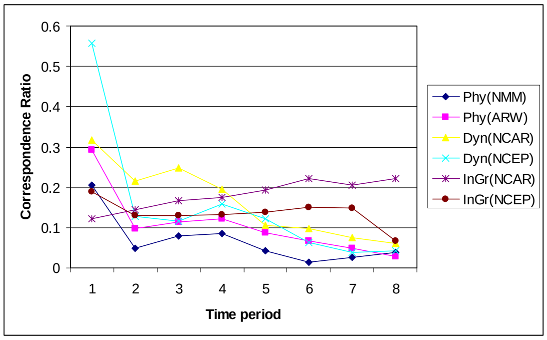

Figure 3: As in Figure 2 except for the .5 inch rain threshold.

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 .01 inch threshold. Apparently at the heavier rainfall thresholds the areas of heavy rainfall are relatively small in runs using the NCEP physics, and changes in dynamic core result in more variation in the forecast than when NCAR physics are used. Bias scores (not shown) support this conclusion, with much smaller values at the .5 inch threshold in runs using NCEP physics than in runs using NCAR physics.

Resampling techniques applied to the .5 inch threshold showed 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 differences in CRs in the tests shown in Fig. 3 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. One, denoted NCAR, used the KF convective parameterization, YSU PBL scheme and Dudhia/RRTM radiation, while the other, denoted NCEP, used the Betts-Miller-Janjic convective scheme, Miller-Yamada-Janjic PBL scheme and 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 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.

It was found that sensitivity to any one component was often influenced by other components. The greatest sensitivity resulted from changes in the physics package when the NMM dynamic core was used. This sensitivity was found to be statistically significantly larger than that valid for most other component changes at most times, especially for lighter rainfall amounts. 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 (statistically significant) 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 was generally smaller than that of changes in dynamics or physics, except in the first 6-12 hours of the forecast.

Acknowledgments

The author would like to thank Dr. Robert Gall for the invitation to visit the WRF-DTC during summer 2004, and for the assistance of several WRF-DTC scientists in accessing retrospective WRF runs and in running the new 8 km simulations. These scientists included Jim Bresch, Wei Wang, and Louisa Nance at NCAR, and Ligia Bernardet, Andrew Loughe, and Linda Wharton at NOAA's Forecast Systems Laboratory. The study was supported in part by NSF grant ATM-0226059

5. REFERENCES

Alhamed, A., S. Lakshmivarahan, 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 multi-sensor U.S. precipitation analysis for operations and GCIP 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, *5th WRF/14th MM5 User's Workshop*, Boulder, CO, 22-25 June, 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.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column test using GATE wave, BOMEX, and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693-709.

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: Impact of resolution and choice of convective parameterization. *Wea. Forecasting*, **14**, 405-426.

Gallus, W. A., Jr., 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.

Grams, J. S., W. A. Gallus, Jr., L. S. Wharton, S. E. Koch, A. Loughe, and E. E. Ebert, 2005: The use of a modified Ebert-McBride technique to evaluate mesoscale model QPF as a function of covective system morphology during IHOP 2002. *Wea. Forecasting*, (submitted).

Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer and turbulence closureschemes. *Mon. Wea. Rev*., **122**, 928-945.

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.

Seaman, N., R. Gall, L. Nance, S. Koch, L. Bernardet, G. DiMego, J. Powers, and F. Olsen, 2004: The WRF process: Streamlining the transition of new science from research into operations. Preprints, *5th WRF/14th MM5 Users' Workshop*, Boulder, CO, 22-25 June, 173-176.

Stensrud, D. J., and M. S. Wandishin, 2000: The correspondence ratios in forecast evaluation. *Wea. Forecasting*. **15**, 593-602.

Stensrud, D. J., 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.