Study of the impacts of grid spacing and physical parameterizations on WRF simulations of convective system rainfall and morphology

Report on WRF-DTC Visit of W. Gallus, I. Jankov and E. Aligo – summer 2005

Introduction

This project was originally designed to build upon work accomplished in a WRF-DTC visit during 2004 by W. Gallus and I. Jankov (some of these results will be discussed later). The project would have explored the sensitivity of simulations of warm season convective systems to grid size and physical parameterizations using grid spacings ranging from roughly 4 to 12 km. However, the original goals and tasks of the project were adjusted prior to the visit based on discussions with DTC personnel. It was determined that a significant need existed for research using the Rapid Refresh WRF (WRF-RR). Because this version of the WRF may replace the RUC model and is intended to run on a very large domain (432 x 338 points on 35 levels), such fine grid spacings cannot be used. At the time of the visit, the WRF-RUC code being run at FSL was thought to be the best candidate to become the WRF-RR, and this 13 km WRF-ARW version was thus the one that we decided to use for our research. The code was provided to us by Tanya Smirnova at FSL. Our use of this version of the WRF meant that we could no longer explore issues related to a range of grid spacings, although we did perform some runs at 4 km grid spacing to allow comparisons with the 13 km results.

Tasks/Goals

The research effort consisted of three tasks and goals:

1) perform controlled experiments (may have been the first such controlled experiments performed for a sizable sample of cases) to examine the differences resulting from use of the MYJ and YSU PBL schemes, with a goal to try to correct problems that had been identified in the MYJ scheme (too cool and moist with too shallow a boundary layer).

2) compare simulations of mesoscale circulations in convective systems in the WRF-RUC with detailed observations with a goal of determining where strengths and weaknesses lie in the simulation of these circulations.

3) examine the impacts on precipitation forecasts and simulations of circulations from changes in the Grell-Devenyi convective scheme used in the model, with a goal of improving the rainfall forecasts and the depiction of low-level cold pools, which was believed to be too weak.

Methodology

Most of the simulations performed used the 13km WRF-ARW model run on the CONUS domain. This version of the WRF used the Grell-Devenyi (GD) convective parameterization, WSM-5 microphysics, MYJ PBL scheme, Dudhia/RRTM radiation schemes, and the RUC land-surface scheme. Because archived data to run the model for past cases was somewhat limited, we were encouraged to simulate events from 2005. Seven cases from June 2005 (June 4, 5, 7, 8, 9, 10, 12) were simulated, with most runs initialized at 12 UTC and integrated for 24 hours, except for the June 10 run which was run for 30 hours. For two dates, June 5 and 7, 00z runs were also performed, making a total of 9 "cases" from 2005. Most of these events included active and substantial convective systems, although two cases were more quiescent to allow better examination of PBL effects. For all of these cases except June 10 and June 12, both the MYJ (control) and YSU PBL schemes were used to permit a study of differences in the simulations related to the two PBL schemes.

For several cases with active convection near profiler sites, 4 km runs were also performed over a smaller domain (roughly the same number of grid points as in the CONUS 13km runs). These cases included June 4, 8, 9, 10 and 12. These 4 km runs were performed without the use of a convective scheme. To permit detailed comparison of simulated mesoscale circulations with observations, one event from 2003 for which BAMEX data were available was also simulated. For the June 4 and June 9 cases, the GD scheme was run in the 4 km runs to allow a comparison with 4 km runs not using GD. In addition, for all convective cases except June 8, 13 km runs without a convective scheme were run to help assist in the understanding of the positive and negative impacts of the GD scheme.

To permit objective verification of weather parameters over a large number of points, the WRF post-processor was used on all 13 km model output. To get the post-processor to work with the data on ijet took a significant amount of effort and the assistance of several DTC personnel. Post-processed output was sent to Andrew Loughe at FSL who then incorporated it into databases allowing for computation of many verification parameters and nice display via a web site (http://www-ad.fsl.noaa.gov/users/loughe/projects/wrf/PBL_expt_summer_2005_RT/)
.

Results from summer 2005 DTC visit

TASK 1: Investigation of PBL schemes

The analysis of the 7 cases where both the MYJ and YSU schemes were used in 13 km runs included both a subjective and an objective component. Subjective analysis of soundings was performed at selected points. The analysis of these soundings supported earlier studies that suggested the MYJ scheme was often too cool and moist within a too shallow PBL while the YSU scheme was too warm and dry in a too deep PBL. Figure 1 shows an example from a 6 hr forecast (valid 18 UTC) over Iowa in a region of clear skies. Soundings were grouped by whether conditions were clear or cloudy (using a 700 mb threshold of 70% RH to define the cloud regions). The differences in PBL schemes were more pronounced in clear conditions, implying a strong relationship to radiative heating. It was also found that the forecasts verifying at 12z did not demonstrate the large differences found for forecasts verifying at 00z, again supporting the idea that any problems in the schemes are related to daytime solar heating.

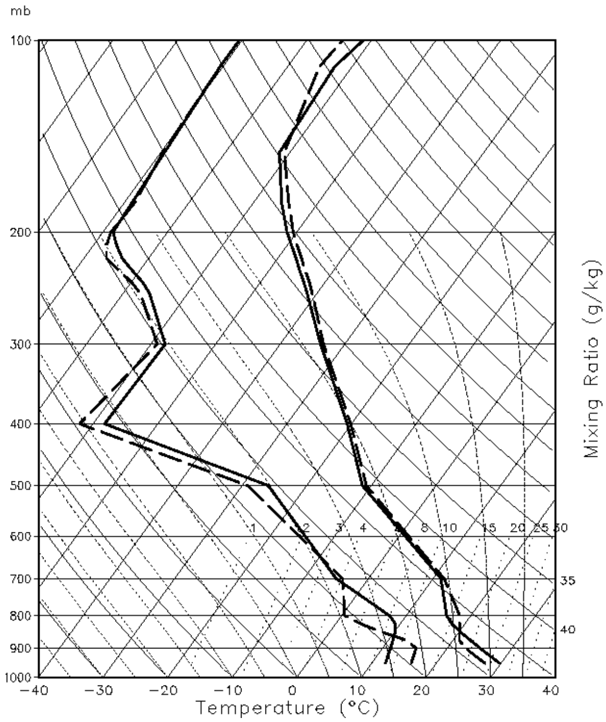

Figure 1: Temperature and dew point profiles from WRF runs using the YSU (solid) and MYJ (dashed) PBL schemes, valid at the 6h forecast time of 18 UTC.

Further analysis of areally averaged sensible and latent heat fluxes over the Central Plains under clear sky conditions and for the two different PBL schemes (YSU and MYJ) indicated that during the day time (especially around noon) the MYJ scheme tends to have a noticeably lower sensible heat flux compared with the YSU scheme. A reversed pattern was found for the latent heat flux; it was usually higher in the case of the MYJ scheme, but not significantly.

Lower sensible heat flux values and lack of entrainment at the top of the CBL support the fact that the MYJ scheme often produces shallower and cooler CBLs. Shallower CBLs produced by the MYJ scheme along with higher values of the latent heat flux were moister. Thus, the different partitioning between sensible and latent heating (i. e. Bowen ratio) and the fact that the schemes handle mixing differently may explain the causes of differences in performance between the schemes.

Specific tasks undertaken are listed below:

1) Scaling based on CBL moisture difference

Based on typical differences in specific humidity and potential temperature in the lower atmosphere, scaling has been performed to evaluate if these differences were forced by the difference in partitioning between the sensible and latent heat only.

The difference in specific humidity (Dq) between the two PBL schemes due to differences in latent heat flux (DHl) was scaled as:

$Dq=(DHl*t)/(L*h*Rho)$ (1)

Assuming the average of $DHl=50\ W/m^2$ during a daytime of duration $t=40000$ s, with $L=2.5\times10^6$ J/kg, CBL depth $h=1000$m and $Rho=1.2$, then late in the afternoon $Dq=.66$ g/kg. Assuming $h=1500$m then, $Dq=.44$g/kg. Assuming $DHl<50W/m^2$, which was more typical for our simulations, then Dq should be less than .44g/kg.

Differences in q between runs using both schemes are accompanied by a difference in CBL depth (Dh), and thus additional drying (Dqi) occurs because the CBL penetrates deeper (by Dh) into the drier layer aloft:

$Dqi=(Dh*Delq)/h$ (2)

where Delq is the averaged initial (i.e morning) decrease of q within the layer Dh. For $Dh=300$ m, $Delq=3$ g/kg and $h=1000$ m, the scaled value for Dqi is about 1g/kg.

It appears that in the afternoon the Dqi would be larger than Dq for the cases that we examined. Namely, drying of the YSU is likely to be more due to its deeper CBL and thus additional elevated dry air mixing, rather than due to its lower Hl.

2) Scaling based on CBL theta difference

With regard to differences in theta, DTH, between the two schemes; in analogy to (1):

$DTH=(DHs*t)/(h*Cp*Rho)$ (3)

Based on our simulations we use $dHs=100\ W/m^2$, which results in DTH=3.3K.

If the changes in the CBL theta and q are only due to surface fluxes, then $DTH/Dq\sim L/Cp\sim2.5$. However, the scaling above suggests that the mixing effect of q as the CBL deepens should have an important role in generating the CBL differences in q between the two schemes (YSU and MYJ).

In analogy to (2) scaling can be done for DTHi.

In any case better agreement between the PBL surface schemes for Hs and Hl should reduce the CBL characteristic differences between both schemes.

3) Some sensitivity simulations

The first experiment was based on the fact that the entrainment thermal flux at the top of the CBL is typically considered to be ~-.2*Hs. Assuming a lack of an entrainment effect with the MYJ scheme, our test run used 1.2*Hs for the sensible heat flux at the surface while solving the vertical eddy diffusion for Theta (scaling approach used in a simple convective mixed layer, less popular for multi-level models). Areally averaged values indicated a slight warming of CBL Theta (~1K) and an increase in the CBL depth of about 120m. The surface sensible heat flux actually decreased by several W/m^2 and the latent heat increased for the same amount compared with the original MYJ simulation.

4) Further testing

Further tests were based on the fact that the two schemes within their surface layer formulations use different surface roughness lengths for z_{0t} and z_{0q}. This may have a large impact on surface flux partitioning.

Typically it is assumed that z_{0q}=z_{0t}, with z_0>z_{0t}, z_{0q} as is the case in the MYJ surface layer. In reality z_{0q} is somewhat smaller than z_{0t}. By modifying the MYJ code to have z_{0q}<z_{0t} it is likely to change the Bowen ratio to enhance drying.

For our first test, we performed a run with z_0=z_{0t}=z_{0q}. It resulted in a further increase of latent heat, a slight warming of the CBL (verified by both arealy averaging 2-meter temperature as well as by using Iowa Environmental Mesonet archived surface temperature data) and an increase in CBL height of 50m on average. However it is possible that by redefining z_{0q} in the original code, to be less then z_{0t} the latent heat flux will decline and the sensible heat flux will increase.

On the other hand, within a surface layer subroutine used by the YSU scheme the effective z_{0q} is significantly lower than z_0, while z_{0t}=z_0. This relation between the roughness lengths is likely to yield overpredicted sensible heat flux and underpredicted latent heat flux. In this case, the test will be to examine the Bowen ratio response to a 'reasonable' prescription of z_{0q}=z_{0t} or z_{0q}<z_{0t}, where z_0>z_{0t}, z_{0q}. The first test with z_{0t} and z_{0q} equalized to z_0 resulted in a decrease in sensible heat flux of about 100 W/m2 on average, lowering CLB height and its moistening. The same test was repeated but with more realistic prescriptions of z_{0t} and $z_{0q.}$ These were set to be roughly 0.1* $z_{0.}$ The results showed the same trends in latent heat flux, sensible heat flux, CBL height and moisture, but with lower magnitudes.

It was also noted that the YSU scheme tends to produce 'smoother' profiles as compared to profiles produced by the MYJ scheme. Some related preliminary tests have been performed but further work is needed to investigate in more detail. Multiple profiles produced by the two schemes were compared at lower levels at different times. Preliminary results indicated that at lower levels the profiles simulated by the run that used the MYJ scheme are sometimes smoother compared to YSU profiles, which is in contrast to what was earlier noted. Vertical cross sections of TKE indicated a drop in the TKE magnitudes at the top of the CBL but the values were still relatively high. This may explain the appearance of smoothing in the MYJ profiles.

It was also documented in some cases that the tropopause was slightly less well-defined when the YSU scheme was used compared to runs using the MYJ scheme. This result may suggest that the mixing used in the YSU scheme, determined from boundary layer characteristics, has a potentially undesirable impact throughout the atmosphere.

Objective verification using large numbers of model grid points was performed using the NCEP verification system. However, the value of this system was limited because it only performs the verification at those levels where mandatory upper air data are available. Thus, quantitative error measures were restricted to the 1000, 850, and 700 mb levels near/within the boundary layer, and to 00 and 12 UTC times. One recommendation we would have for the DTC is to make changes in this system to allow verification at a wider range of levels, even though this might require vertical interpolation of the observations and/or model output.

Finally, it was determined that the differences in fluxes between the two schemes are sensitive to the vegetation type. For individual points in the Corn Belt, the MYJ scheme generally had less sensible and more latent heat flux than the YSU scheme. In wooded areas, the results were opposite. Also of note, no matter what the surface fluxes were, the PBL depth was greater in the YSU scheme. This result suggests that the different ways each scheme handles vertical mixing may play an important role in explaining differences. Further work to explore some of these findings is anticipated, and it is hoped that a paper will be submitted later in 2006 with I. Jankov as main author.

TASK 2: Comparison of simulated mesoscale circulations with observations

Five cases were simulated where active convective systems occurred in the central United States. Analyses were performed to compare simulated mesoscale circulations with those observed in one case, June 10-11. Observational data from 2005 were obtained although the fine-scale nature of the circulations makes a comparison difficult, since there were no special observations available in 2005, as were available in 2002 and 2003 with the IHOP and BAMEX projects. To facilitate somewhat greater comparison with observation, a few 2003 BAMEX cases were also simulated. For these active convective events, simulations were run with both 13 and 4 km grid spacing.

As of this point, extensive analysis has only been performed for the June 10-11, 2005 case. For this event time-height cross sections of winds from 5 profiler stations in KS and OK were examined. Each profiler had witnessed the passage of a squall line(s) during the evening and early morning hours of 11 June. Winds ahead and behind the squall line were commonly 20 m/s or higher at around 1 km AGL and below. Winds of 25 m/s behind the squall line were not uncommon, which could be indicative of a rear-inflow jet. Vertical cross-sections have been constructed after using time-space conversion on profiler data. These data have been combined with velocity data from radar to assist in pinpointing the various circulations associated with the observed squall lines.

One of the comparisons performed explored differences between two 4 km runs, one using Grell-Devenyi (GD) and the other fully explicit. Both runs missed some of the areas of rainfall, but both also developed some convection in western Oklahoma that eventually evolved into one or more squall lines. The run without the convective scheme developed its squall line too early and too far west in OK. Both runs overestimated precipitation amounts and moved their squall lines too slowly.

Regarding winds, observations suggested winds of 18-25 m/s in the 1-2 km layer, with even greater speeds (up to 35 m/s) closer to the ground. The run using GD was roughly 5 m/s weaker with winds in this layer than the run without GD. The weaker winds agreed well with observations, although the location of the wind maximum was better in the stronger no-GD case. Comparisons of surface winds and temperature were complicated by differences in the evolution of the squall lines in the runs compared to observations. It did appear that both runs overestimated horizontal temperature gradients by 2-3 C per 60 km. This research, begun during the DTC visit, will be incorporated into the ongoing PhD dissertation work of E. Aligo. The anticipated date of completion for this Ph.D. is 2008.

TASK 3: Examination of sensitivity of convective system forecasts to convective scheme

For most of the same cases used in Task 2, an additional component of our work was to explore how runs using the GD scheme differed in their rainfall and circulation depictions compared to runs not using GD (and observations). This task involved running the WRF-RUC version of GD in the 13 km version of the model, as well as the RUC version of GD, and a modified version of the WRF-RUC GD. In addition, some simulations were run at 13 km without the convective scheme, and with the Kain-Fritsch scheme substituted. Additionally, 4 km runs were performed without any convective scheme, and in a few cases with the GD scheme used.

For the June 12 case, the GD scheme was modified in consultation with George Grell, to reduce the impact of low-level drying somewhat (by increasing the parameter edtmax from .8 to .9, and making several other changes). The impact in that case appeared to be very slight, with usually no more than a 1 mm impact in hourly rainfall accumulation at any hour at any point (compare Figs. 2 and 3). In addition, for that case and 3 others, the version of the GD scheme used in the RUC model was also tested (the code supplied by G. Grell). That version had a much more noticeable impact on the forecast, and seemed to result in more intense precipitation maxima, and possibly more similar forecasts to runs using no convective scheme. The 24 hour precipitation for the June 12 case with the RUC version of GD is shown in Fig. 4. The precipitation in the same period from the fully explicit 13 km run is shown in Fig. 5, and from the KF run in Fig. 6. Note for this case that the RUC version of the GD produces results most similar to the fully explicit run, whereas all of the other configurations are more similar. These results suggest that it is difficult to obtain meaningful changes in the forecast from rather modest adjustments of parameters internal to the GD scheme, but much more significant impacts can be made with careful tuning (as shown in the RUC GD results). Such tuning was beyond the scope of this project, but is recommended for future study if the GD scheme is used in the WRF-RR.

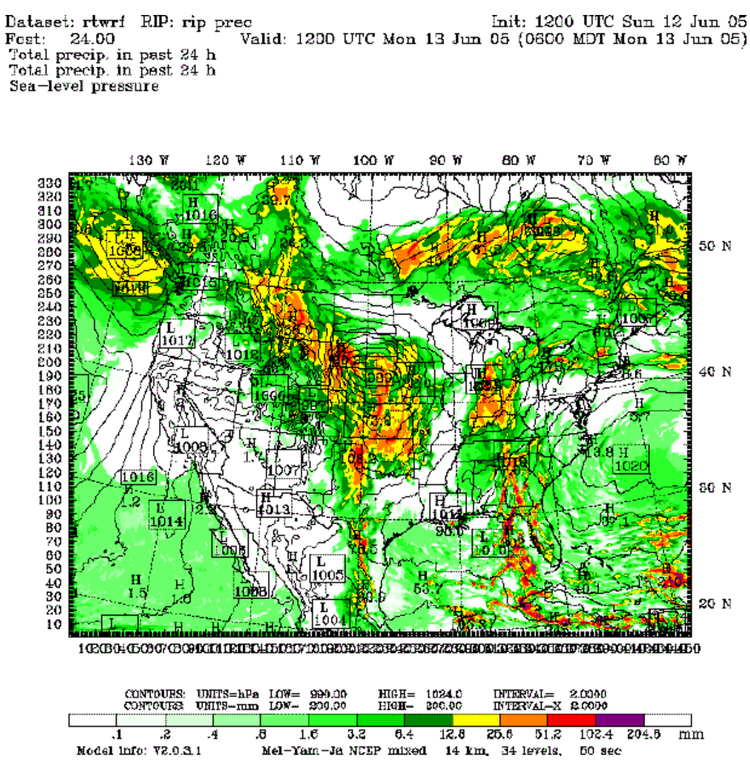

Figure 2: 24 hour accumulated precipitation during 12-12 UTC period June 12-13 2005 from the WRF run using standard Grell-Devenyi scheme

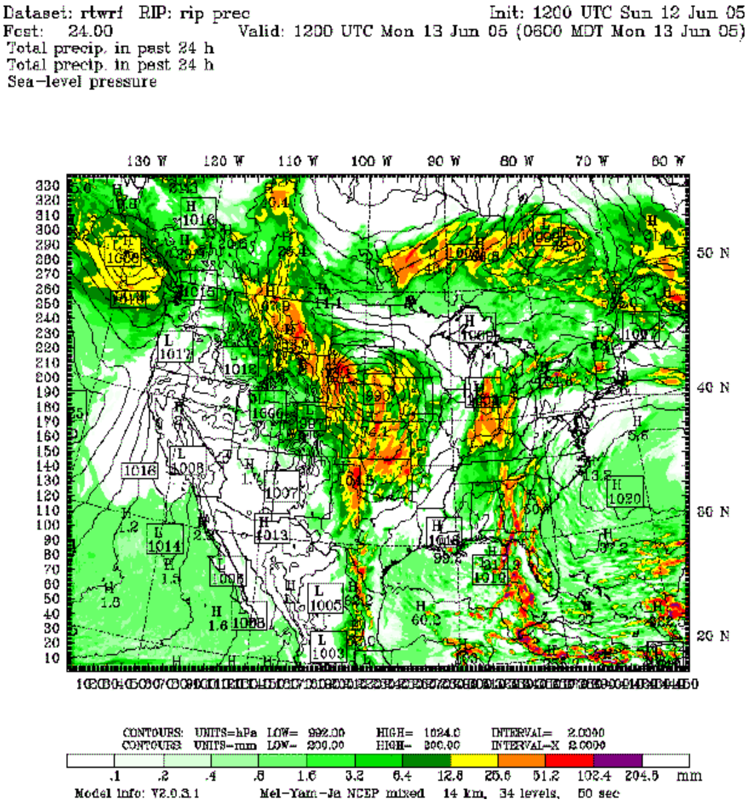

Figure 3: As in Figure 2 but for modified Grell-Devenyi scheme with decreased drying below cloud base.

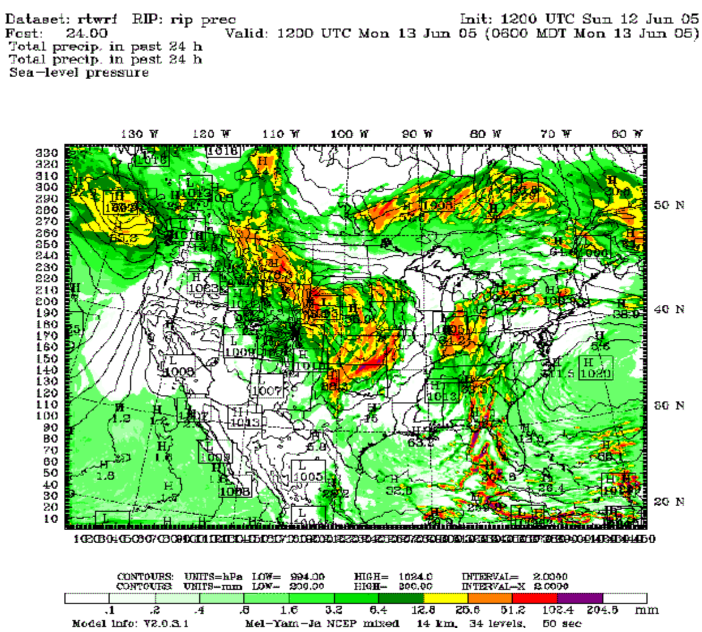

Figure 4: Same as in Fig. 2 except for run using RUC version of Grell-Devenyi scheme.

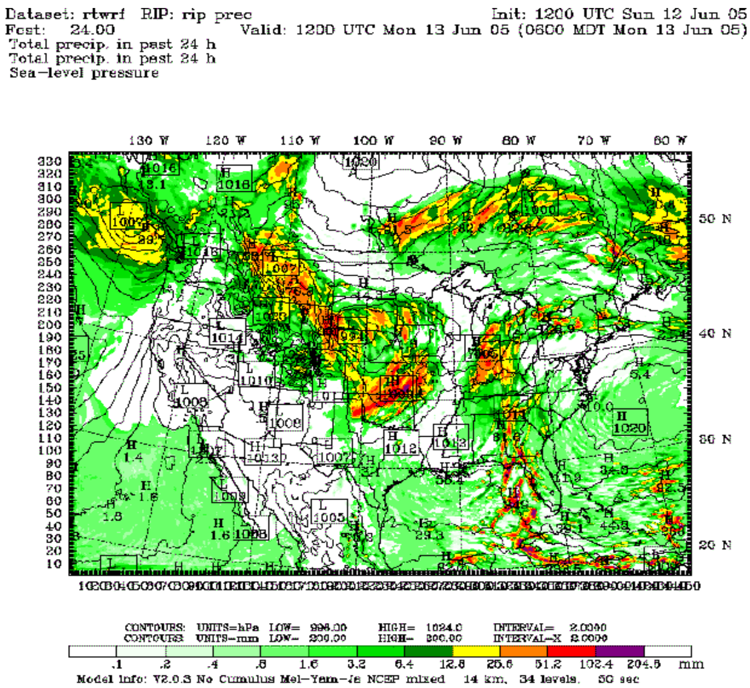

Figure 5: Same as in Fig. 2 except for fully explicit run.

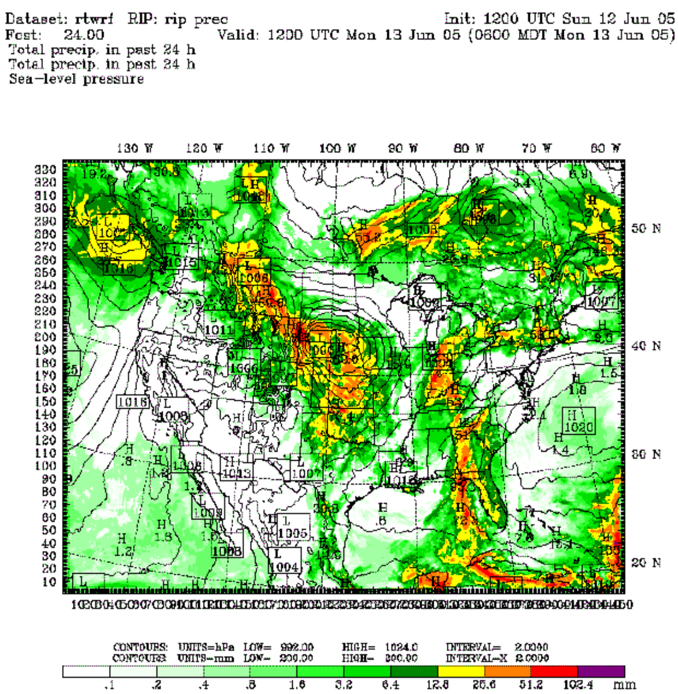

Figure 6: As in Fig. 2 except for run using Kain-Fritsch scheme.

An attempt was made in this project to explore the errors associated with each closure in the GD scheme. If some closures seem more prone to errors than others, different weighting might be assigned to improve the forecasts. We were able to obtain some plots for each of the closures in one case, but there were problems in interpretation of the results, and consultation with G. Grell did not result in a solution. It would appear that this particular investigation would require a large investment of time.

Objective verification of the precipitation forecasts from these various model runs was performed using the WRF post-processor and precipitation extraction routines made available at the DTC, and a database system constructed by Andrew Loughe. Figure 7 shows Equitable Threat Scores and Bias scores for the full sample of cases for 24 hour accumulated rainfall. For this sample of cases, the best ETSs for a 24 hour period at most thresholds occurred in the WRF run that used the standard GD scheme but substituted the YSU PBL scheme instead of the MYJ. Differences were fairly small among most model runs, although the run not using a convective scheme earned the lowest ETSs at all thresholds above .1 inch. The standard WRF run, based on the WRF-RUC configuration used by Smirnova, tended to have the worst problem with a high bias at most thresholds. Despite earning relatively lower ETSs, the fully explicit run had good bias scores for most light to moderate rainfall thresholds. For thresholds of .75 inch or more, all configurations except the KF runs had a serious problem with high bias.

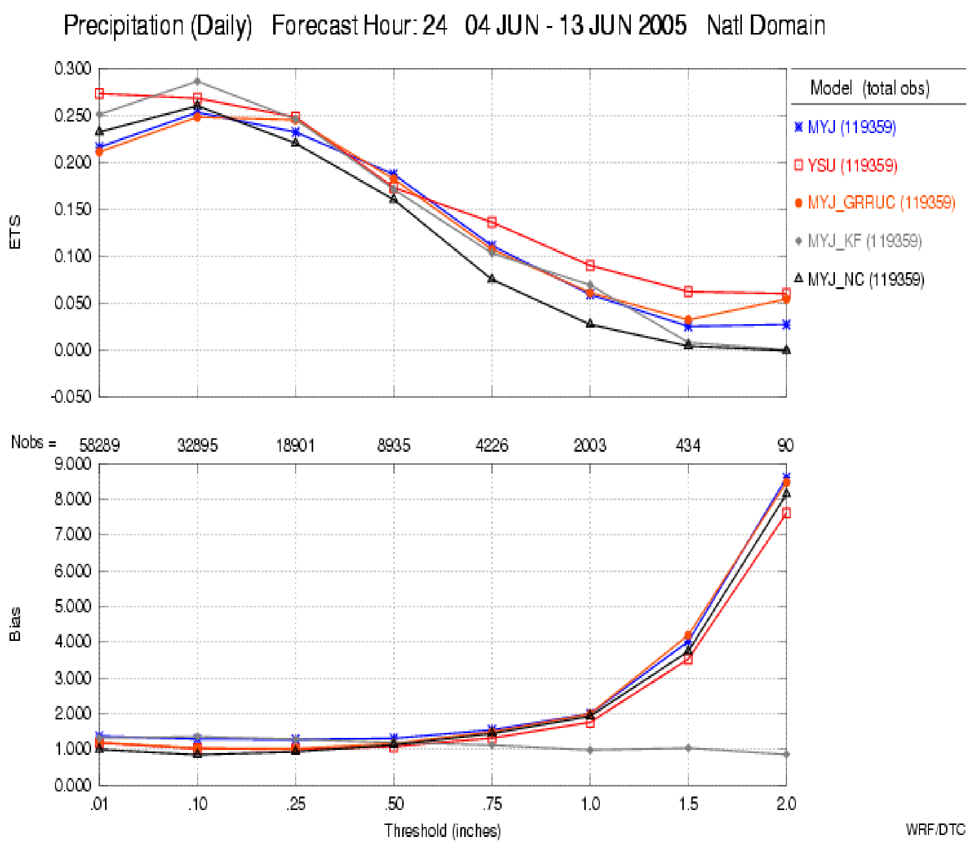

Figure 7: ETS and Bias scores for the full sample of cases from the various configurations of WRF tested.

Results from 3 hourly verifications (figures not shown) were different from those for the 24 hour period. An examination of every 3 hours showed that the fully explicit runs performed best in terms of ETS for the light precipitation amounts at all hours through the 15 h forecast. For most thresholds, however, at most times, no particular configuration was consistently better with ETSs or bias scores. Despite the high bias that most models showed in the 24 h precipitation, all versions showed a low bias in 3 hour accumulations during the 00-09 UTC period, indicating possible problems during the period when convective systems are typically growing upscale.

Simulations performed with 4 km grid spacing showed that the use of the GD scheme does result in noticeable impacts on the forecast. No objective verification was performed with the 4 km runs. The GD scheme resulted in cold pools that were too strong compared to those in the fully explicit runs.

Results from summer 2004 visit

As mentioned in the introduction, the original focus of our 2005-06 DTC project was to expand upon findings obtained during a visit to the DTC during summer 2004. One key component of the research conducted during that summer was an investigation of the impacts of using different physics, dynamic cores, and initial conditions on warm season rainfall forecasts. Because this research has resulted in one publication in Monthly Weather Review (Gallus and Bresch, 2006), a brief overview of that project is being included.

A series of WRF-ARW simulations for 15 events occurring during August 2002 were performed over a domain matching the central United States domain used during the DTC retrospective run project. The goal of this work was to allow a comparison of the sensitivity of warm season rainfall forecasts to changes in model physics, dynamics, and initial conditions. The WRF-ARW simulations were run with 8 km grid spacing. These runs were compared to the WRF-NMM retrospective runs, and the WRF-ARW retrospective runs that used a slightly different grid spacing and a different initialization (RUC instead of Eta). In all of the runs, two different physics packages 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 YSU PBL scheme and Dudhia/RRTM radiation. Other physical schemes were the same (e.g., NOAH land surface model, Ferrier et al. microphysics) in all runs.

Simulations suggested that the sensitivity of the model to changes in physics is a function of which dynamic core is used, and the sensitivity to 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 (see Figs. 8 and 9). 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 that present 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. 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 opposite is true with a greater sensitivity occuring when NCEP physics is used. 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.

Sensitivity to initial conditions (Eta versus RUC 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 hours of the forecast when it is comparable. As might be expected for warm season rainfall, the fine-scale 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.

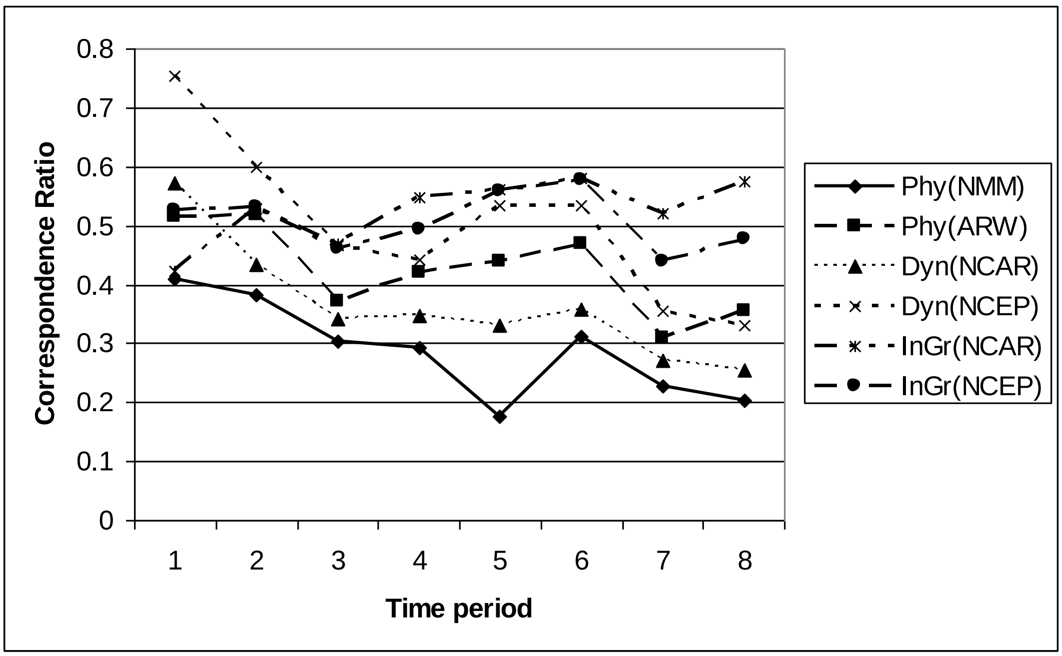

Figure 8: Temporal evolution of CRs for .01 inch rain threshold in 6 WRF configuration comparisons (dynamic core or physics package held constant indicated in parentheses). Time periods 1-8 correspond to 0-6, 6-12, 12-18, 18-24, 24-30, 30-36, 36-42 and 42-48 hour forecasts.

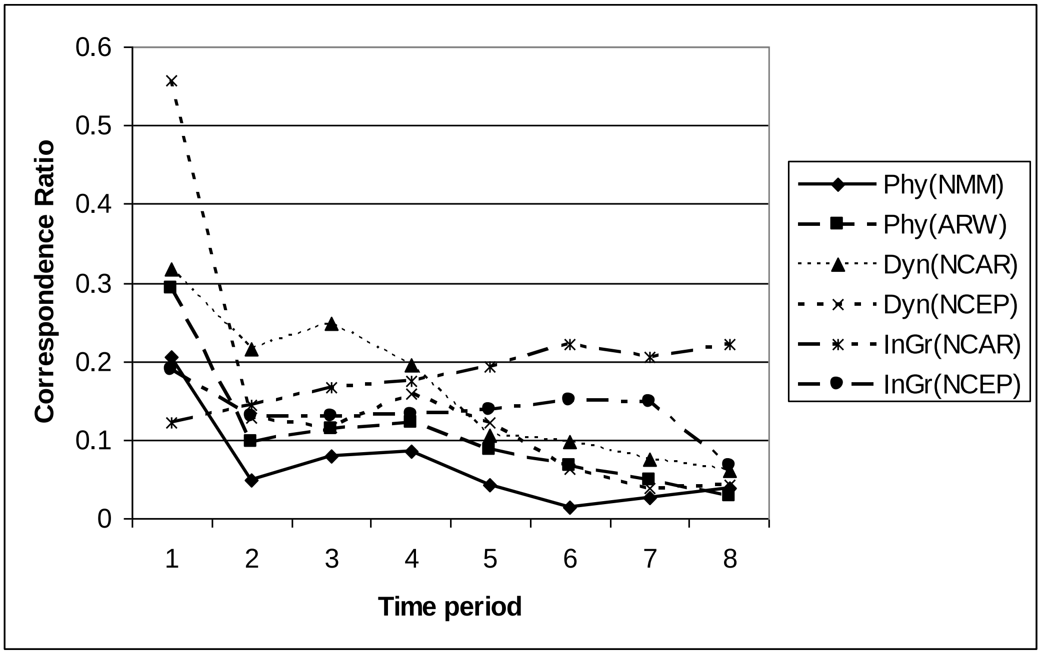

Figure 9: As in Fig. 8 except for .50 inch rainfall threshold.

Components of Project that Could Become Part of Software Suite Offered by DTC to WRF users

The research undertaken explored impacts of existing WRF codes and thus no new software was developed by the research team. However, the 2004 research greatly benefited from the use of the Ebert Mc-Bride code which was adapted to handle the retrospective run data by Gallus with much help from L. Wharton at FSL. The Ebert-Mc Bride verification code should be made available to WRF users since it provides a different way of performing mesoscale verification. In addition, the 2005 research benefited from the work of A. Loughe at FSL who set up the objective verification program with web access. If this process could be streamlined, it would be another great benefit for WRF users.

Impact of Results on Operational Forecasting

Further work may need to be performed before the results of our WRF DTC visits can improve operational forecasting, particularly our 2005 visit results. The 2004 study suggests that even for warm season convective system precipitation, where it might be assumed that mixed physics creates the most spread in an ensemble system (due mostly to the use of convective schemes), the use of different dynamic cores (with the appropriate physics package) can also result in substantial spread, and may be a valuable component to an ensemble system. The results also suggest there is little value in using mixed initial conditions. However, it is likely that mixed lateral boundary conditions may have more value. The impacts of varied initial conditions tended to become small after 6-12 hours.

The 2005 results suggest that there are improvements that could be made to the PBL schemes to reduce temperature and moisture errors. It was the research team's understanding that at least one change was being tested at NCEP. The team had hoped to test this change as well, but the code was not made available from NCEP in time to work with it. It is more difficult to state than any changes were found that could be made in the Grell-Devenyi scheme to improve forecasts. The research shows that it is possible to make changes in that scheme that do substantially impact the forecasts of rainfall, but it appears no particular change, or even the use of a completely different convective treatment, greatly changes ETSs. Bias scores for heavier precipitation could be improved, though, by switching to the KF scheme.

Publications resulting from the DTC projects

Gallus, W. A., Jr., 2005: A comparison of impacts from dynamic core, physics, and initialization dataset in WRF simulations of warm season convection. 17th Conf. on Numerical Weather Prediction. Washington, D.C., Aug. 1-5, 2005, P1.89.

Gallus, W. A., Jr., and J. F. Bresch, 2006: Comparison of impacts of WRF dynamic core, physics package, and initial conditions on warm season rainfall forecasts. Mon. Wea. Rev. (in press).

Jankov, I., and W. A. Gallus, Jr., 2006 (proposed): Differences in performance between various PBL schemes for warm season clear-sky events. (submission planned for Wea. Forecasting).