Final Report FAA Task 10.5.24— Develop, Test, and Improve the 3-km WRF-based High-Resolution Rapid Refresh

Jimy Dudhia (NCAR, NESL/MMM)

1. Introduction

This report is aimed at some aspects of the evaluation of the high-resolution 3-km simulations carried out at NCAR as part of the annual Spring Forecast Experiment (SFE) forecasts of severe storms in the eastern and central US. These forecasts used the 13-km RUC DDFI analyses for initial conditions at 12Z and 00Z. 48-hour forecasts were run with outputs placed on the Web in real time and were disseminated to the Storm Prediction Center (SPC) for their use as part of a set of forecasts that they evaluated in terms of potential usefulness to forecasters.

This report follows from a 2009 FAA report (Dudhia 2009) that focused primarily on the rainfall spin-up issue, given that in 2009 a significant number of forecasts were run with RUC analyses and cold-started with GFS analyses. Despite the lack of a diabatic term in the RUC DDFI at that time, these initial conditions clearly helped the spin-up of rainfall in the first six hours for the 00Z initializations. For the 12Z initializations, the RUC analyses prevented the 00Z peak from being over-amplified in terms of area-averaged rainfall rates.

In this 2010 report, focus is given to microphysics, which became a priority in the 2010 forecasts owing to a deficiency that was noticed in 2010 when compared to the forecasts run in 2009. Morris Weisman (NCAR/MMM) documented these problems at the WRF Users' Workshop, (Weisman, 2009 WRF Annual Users' Workshop, Boulder, CO), and this was isolated with some work during the summer. I became involved in this work after it became apparent from previous tests conducted by Wei Wang (NCAR/MMM) and Kevin Manning (NCAR/MMM) in consultation with Weisman that changes to the Thompson microphysics scheme between the 2009 forecast system and the 2010 forecast system were largely responsible for the differences. My role, working also with Greg Thompson (NCAR/RAL), was to determine specific changes within the microphysics scheme that may have led to the new behavior in 2010.

2. The Problem

During the 2010 forecast program, it became evident to Weisman and others following the forecasts that reflectivity features were not as well defined as had been seen previously, especially at the leading edge of convective systems with linear organization and propagation. This lack of reflectivity lines was also seen to be deficient in comparison to observed reflectivities. Figure 1 shows the 00 UTC June 2, 2010 case at the 6-hour forecast time. The microphysical scheme was immediately considered a potential source of the problem, particularly since in 2010 a newer version of the Thompson microphysics scheme was adopted in Version 3.2 of WRF, while in 2009 the older Thompson07 option, retained from Version 3.0 of WRF, was used.

Several sensitivity studies by Wang and Manning for about six selected cases compared switching between these two Thompson microphysics versions, and the results confirmed that the older scheme was capable of producing the linear features consistently with observations. Figure 2 shows the comparison for the above-mentioned case between the old and new Thompson versions. Apart from reflectivity structures, dynamical effects of reduced downdraft and cold-pool activity were evident from the equivalent potential temperature and the surface wind speeds (not shown). The older scheme was able to produce a larger area of low equivalent potential temperature and stronger gust front wind gusts, providing evidence of stronger downdraft circulations in the vicinity of the more-focused reflectivity lines. This is all consistent with a lack of intense rainfall associated with the newer scheme. This is further borne out by the overall season statistics on 12h–36h bias for 2010 compared to 2009 (Figure 3) that show that 2010 underestimated the area of precipitation at thresholds greater than about 25 mm per 24 hours, while 2009 actually overestimated these amounts. The biases also reveal that the new scheme overestimates the area of precipitation for all light and moderate thresholds down to 0.1 mm per 24 hours, while in 2009 there was almost no bias below 2 mm per 24 hours.

From these results it is apparent that the more densely-rimed particles, namely graupel, have some deficiency. Normally these particles would be responsible for focusing the rainfall and helping to drive the localized downdrafts, but in the situation of underestimated graupel, the precipitation may mostly come via snow that is carried back through the system (due to its fall speeds being up to three times slower than graupel's) and producing lighter stratiform rainfall further back.

3. Testing

It was at this point that the task became to isolate which code changes led to the specific behavior differences between the old and new Thompson schemes. In fact, numerous changes had been made between these two versions requiring a large series of tests to narrow down the main causes. The initial code tests were determined based on the most likely causes of difference in graupel behavior and in consultation with Greg Thompson on his changes. After several of these tests, however, it became clear that other changes that had not been immediately suspected were responsible, so the later testing was in terms of objectively dividing code changes into smaller and smaller groups until the main cause of the difference was found.

A representative case was chosen for testing, the 00Z June 2, 2010 run. This, as shown in Figures 1 and 2, had already developed distinct differences in behavior between versions in the first 6 hours of the simulation, making it ideal for extensive testing that only required short simulations and diverged quickly from identical initial conditions. A total of 15 tests were carried out with variants of the code mostly between Versions 3.1.1 and 3.2. All the tests will not be detailed here, just the main results.

a. Comparison of V3.0, V3.1, and V3.2 results

WRF V3.0 was the version used in the 2009 series of forecasts, and it is still an option in V3.2 using *mp_physics=98*. The main changes between V3.0 and V3.1 were the replacement of the drizzle parameterization with double-moment rain and a minor modification to the graupel intercept parameter function.

Results of the 6-hour simulation from June 2, 2010 are shown in Figure 4, which displays a mean profile of hydrometeor species in the 750-km square that includes the squall-line system. From these profiles, the primary difference is in the lack of graupel in V3.2 compared to the other two versions. The rain amount shows reductions between successive versions, too. This, however, is likely to be from changing the parameterization from drizzle to double-moment between V3.0 and V3.1 and from the reduction in graupel in V3.2, both of which tended to decrease rain amounts in the atmosphere.

Figure 5 shows frequencies of points in precipitation bins for the same area on a log scale. Each bin is twice the previous one, and the value of the bin is log2 (mm*10000), where "mm" is the 6-hour precipitation in mm. Bin 20 is about 100 mm, and bin 10 is about 0.1 mm. In these comparisons, the noticeable change between V3.0 and V3.1 is that V3.0 (solid) has significantly more light precipitation area than V3.1 (dashed). But, both have comparable contributions at the most intense bins, indicative of V3.1 preserving the leading-edge intense precipitation that leads to higher local accumulations. Comparing V3.2 with V3.1, it is seen that V3.2 (solid) loses the most intense bin, and also adds a peak at bin 14 (about 1.6 mm). V3.2 also generally increases light rainfall area over V3.1, that was already increased over V3.0. Comparing V3.2 with V3.0 emphasizes these two points again.

From these results, it is concluded that the light precipitation has increased significantly between each successive version, probably contributing to the increase in seasonal bias for light amounts seen in Figure 3. Meanwhile the loss of heavy precipitation amounts also seen in Figure 3 may have resulted from the differences seen between Version 3.1 and 3.2. This latter effect also had other impacts on the simulated reflectivity, and it was considered important to pursue with more tests to distinguish the changes between these latter two versions that led to the most difference.

b. Testing changes between V3.1 and V3.2

The differences between the V3.1 and V3.2 releases of the Thompson microphysics scheme (*mp_physics=8*) are summarized in several categories.

(i) The function for graupel intercept parameter was changed to depend on cloud and rain amount rather than graupel amount.
(ii) The look-up table was extended for rain/graupel collection term.
(iii) All look-up tables were made allocatable to save compile-time and memory.
(iv) Ice number concentration limits and the initiation function constant were changed.

(v) The rain self-collection and breakup were modified.
(vi) Limits of melting process of rain-drop formation were added.
(vii) Modifications from Peter Blossey (U. Washington) for conservation in terms involving graupel/rain collection and snow/rain collection were added.
(viii) Sedimentation processes were separated into different sub-steps for efficiency.
(ix) Changes to melting and freezing of small amounts of cloud between water and ice were made.
(x) A change to the rain-drop fall-speed as function of diameter was made.
(xi) A change to limits of rain-snow accretion was made.

Initial tests (e.g., t3 (test 3)), again using the 6-hour forecast from 00Z 2 June 2010, reversed the code changes related to graupel size distribution and rain/graupel accretion. These showed very little difference from V3.2 in either the mean hydrometeor profiles at 6 hours or in the surface rainfall bin distribution.

Later tests were able to rule out the rain-related changes (t5) and the look-up table related changes (t6). The next tests (t8 and t9) narrowed down the cause of the main differences to the ice-related, the conservation-, sedimentation-, or melting-related changes. The next two (t10 and t11) confirmed it was either the snow-melt (vi) or the Blossey conservation (vii) change. With the final tests (t12 and t13), the conservation change was found to have caused the main part of the difference between V3.1 and V3.2 results, both in the mean graupel profile and in the surface rainfall bins, where it was responsible for the peak at bin 14 (1.6 mm). However, it was found that the snow-melt change produced the other histogram peak around bin 9 (0.05 mm). The conservation change had no effect on the light-precipitation bins below bin 10, and the snow-melt change had no effect on the higher-precipitation bins above bin 11. But, these changes accounted for much of the increase of non-intense precipitation.

The “conservation” change was in the term whereby rain collects snow and is a source for graupel. The change turned out not to account for the fact that the graupel source was opposite in sign to the sink of rain and snow, so it was clearly a bug that has been uncovered by this evaluation process. By effectively adding a depletion term for graupel, the “conservation” change had an impact on convective structure and rainfall patterns that tend to diminish the focusing effect of graupel on intensity.

The minimum code change in module_mp_thompson.F that was required to correct the bug is as follows (< new, > V3.2).

```
At line 2024
<       if (temp(k).gt.T_0) then
---
>       if (temp(k).lt.T_0) then
>         prg_rcs(k) = prs_rcs(k) + prr_rcs(k)
>       else
```

Figure 6 shows the effect of this code change that removes the peak at bin 14, and adds a higher-intensity bin similar to V3.1 in Figure 5. Comparing Figure 7 with Figure 4 shows that the mean profile of graupel is the only one significantly changed, and now there is a comparable amount of graupel with Versions 3.0 and 3.1.

4. Acknowledgements

This work was supported by the FAA. I benefited greatly from discussions and initial analysis of the problem by Morris Weisman. Code explanations and discussions with Greg Thompson were responsible for a quick resolution of the central issues. The test simulations carried out by Wei Wang and Kevin Manning were vital to this project.

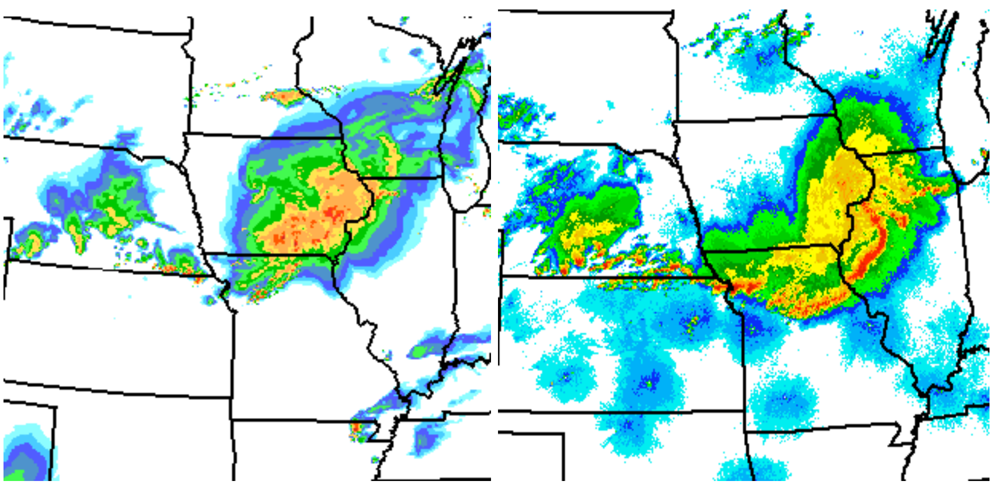

Figure 1.(left) 6-hour forecast reflectivity at 06Z June 2nd 2010, (right) observed reflectivity.

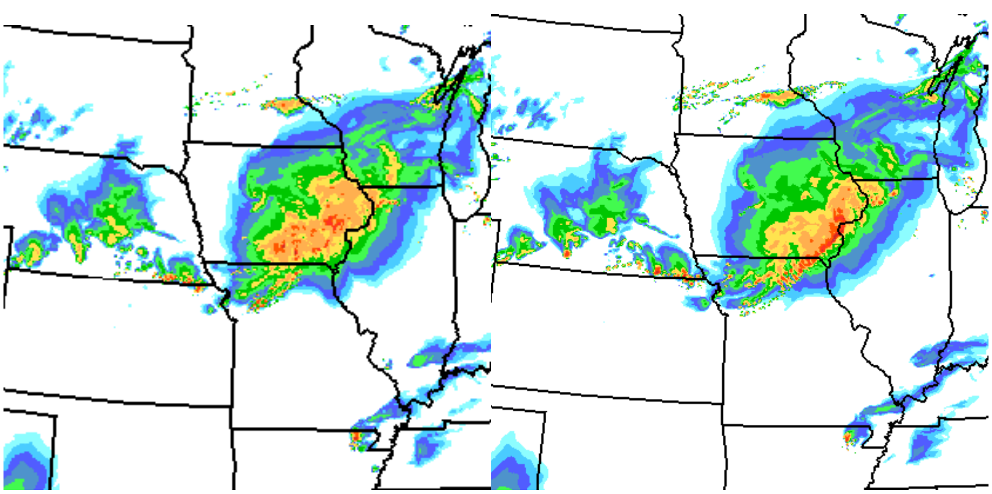

Figure 2. (left) same as Figure 1a, (right) same but with old Thompson scheme.

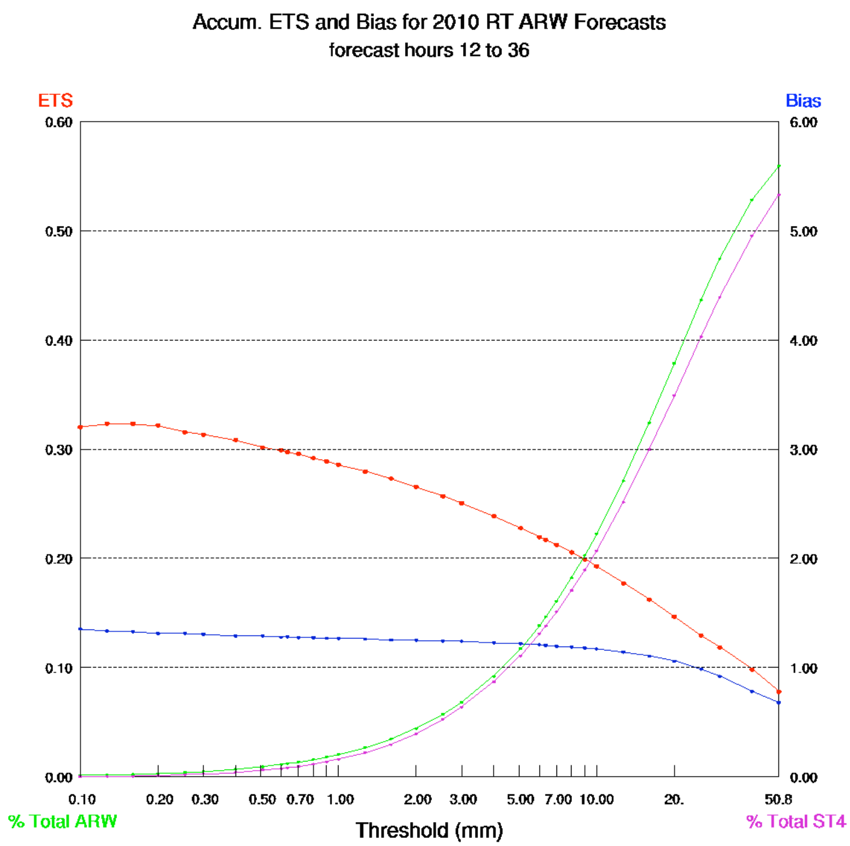

Figure 3a. Statistics for 2010 season versus Stage 4 precipitation, (blue) bias, (red) equitable threat score.

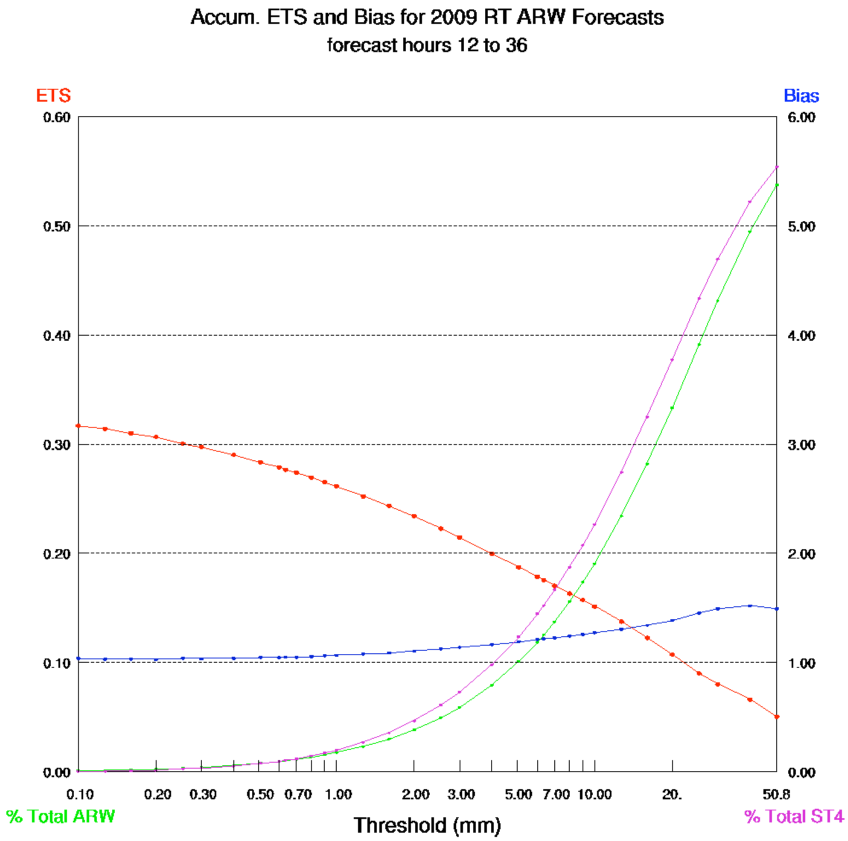

Figure 3b. Statistics for 2009 season versus Stage 4 precipitation, (blue) bias, (red) equitable threat score.

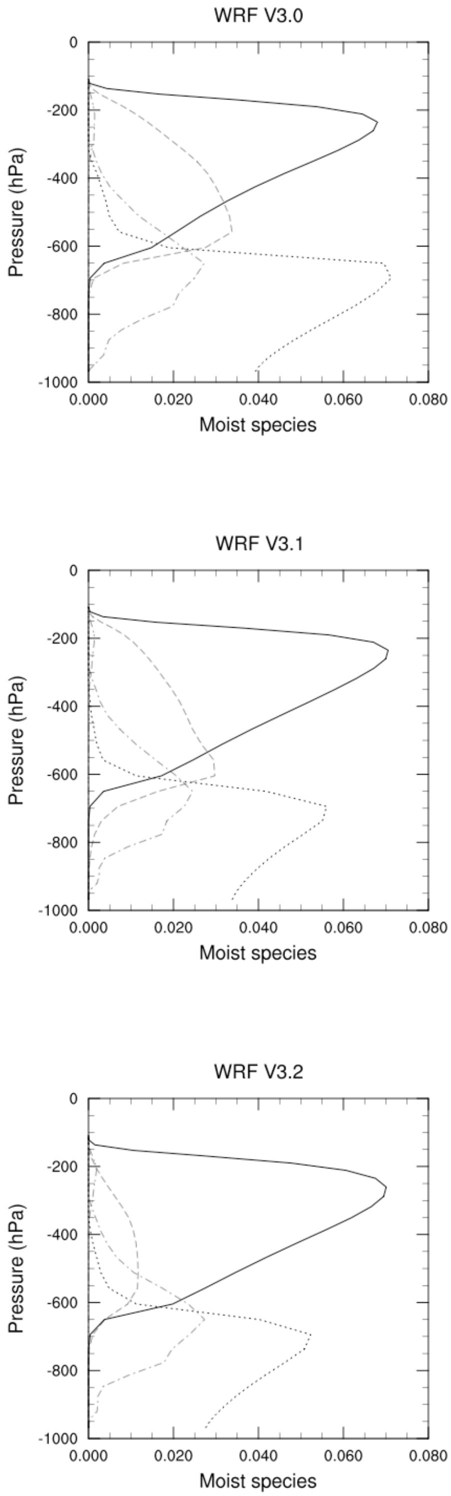

*Figure 4. Mean hydrometeor profiles for different versions. (solid) snow*0.1, (dash) graupel, (dot) rain, (dot-dash) cloud water, (dot-dot-dash) cloud ice.*

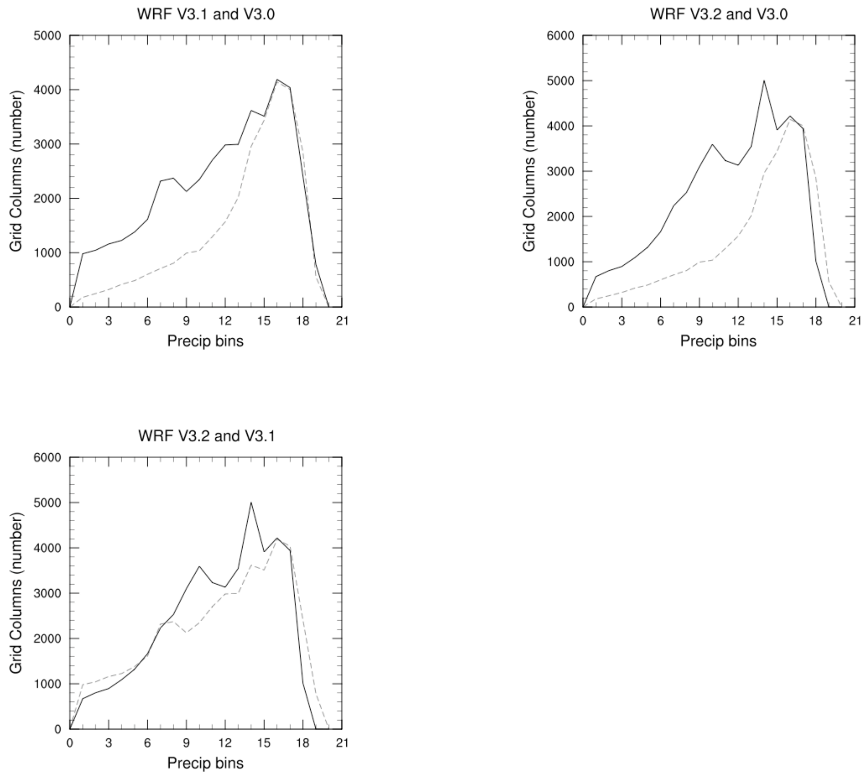

Figure 5. Six-hour rainfall accumulation bins comparing different pairs of versions (solid) newer version, (dash) older version, in each comparison.

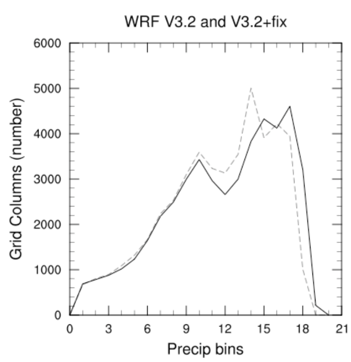

Figure 6. As Figure 5, but comparing code fix (solid) with V3.2.

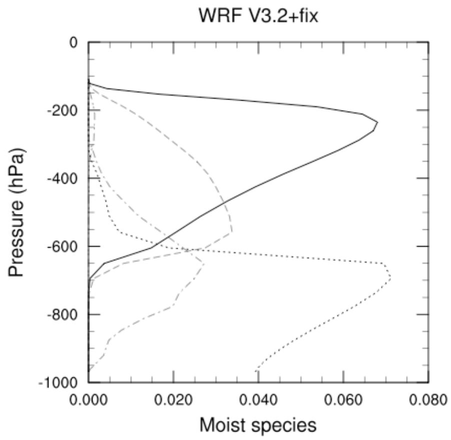

Figure 7. As Figure 4, but for code fix.