Evaluation of WRF-ARW High-Resolution Tropical Storm Forecasts in 2005 Season

Wei Wang[1], Chris Davis[1], Joe Klemp[1], Greg Holland[1], and Mark DeMaria[2]

[1]National Center for Atmospheric Research[*], P.O. Box 3000, Boulder CO 80307

[2] NOAA/NESDIS

1. Introduction

The 2005 hurricane season turned out to be a record year for the Atlantic / Gulf of Mexico. It had 30 named storms, breaking the previous record of 21 named storms in the region. The widespread disruption caused by these hurricanes has highlighted the need for further improvements in hurricane forecasting. This season also provided a good opportunity to evaluate the capabilities of the newly developed Weather Research and Forecasting (WRF) Model for high-resolution hurricane forecasting.

The Advanced Research WRF (ARW) Model (Skamarock et al. 2005) was run in test mode to forecast hurricane track and intensity in real-time during August, September and October of the 2005 hurricane season. The forecast experiments were set up in part to support the RAINEX field campaign (*http://orca.rsmas.miami.edu/rainex*). In this paper, the ARW forecasts are verified and compared to a number of operational forecast products for the intensity and track forecasts.

2. Model Configurations

The ARW Version 2.1.0 was used in these experiments. No tuning was performed for the hurricane forecasts. Two configurations were used. A single domain, 12 km grid was run to forecast 5-day tracks. This domain covers an area of approximately 5500 by 4200 square kilometers. The physics used for this domain includes the Kain-Fritsch cumulus scheme, Yonsei University planetary boundary layer scheme, five-layer soil diffusion model, RRTM long-wave radiation, and MM5-Dudhia short-wave radiation scheme. The second configuration comprised two domains, 12 and 4 km, two-way nested grids, with the 4 km grid automatically relocated to keep the hurricane in the center of the domain (Tenerelli and Chen, 2001; Michalakes et al., 2005). The physics used for this configuration is the same as those for the 12 km single grid, except that no cumulus scheme is used on the 4 km grid. The sea-surface temperature was prescribed from the NCEP daily analysis and kept constant during the model integration. The model was initialized with GFDL initial condition, which contained a bogused vortex, whenever the data were available. When GFDL was not available, the GFS analysis was used. GFS forecast was used for specifying the lateral boundary conditions. The model configurations were run once (0000 UTC) or twice a day (0000 and 12000 UTC) depending on the availability of the computer time and severity of the storm.

3. Verification Methods

The track and intensity data from the ARW forecasts were verified again several operational forecast products. The track is defined as the location of the minimum sea-level pressure, and intensity of the storm is measured by 10 m winds directly from the model forecasts. For the 12 km forecasts, thirty-four forecasts from five storms (Katrina, Maria, Ophelia, Rita and Wilma) were verified. For the 4 km forecasts, thirty-five forecasts were verified for four storms (Katrina, Ophelia, Rita and Wilma). For intensity forecasts, five operational products were also verified: the official NHC forecast (OFCL) and forecasts from two statistical schemes (SHP5 and DSHP), GFDL, and the Florida State University Super Ensemble (FSSE). For track forecasts, the comparison included: the official NHC forecast, climatological persistence forecast (CLP5), and forecasts from GFDL, GFS, NOGAPS, FSSE and the UK Met Office. The comparison is homogeneous in that all forecasts correspond to the same time periods.

4. Results

a. track forecasts

The verification results for the ARW 12 km track forecasts are shown in Figure 1. The ARW track errors are comparable to the other operational products between 12 to 48 forecast hours, and grew from 34 nm at hour 12 to 72 nm at hour 48.

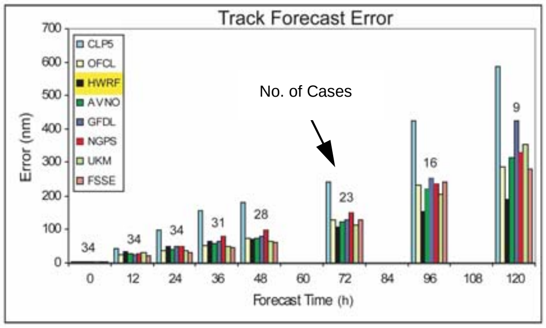

Figure 1: ARW 12 km track errors (in black, labeled HWRF) as compared with other operational products. The number of cases verified is indicated above the bars.

For the day 3 to day 5 track forecasts, the ARW track errors are smaller than those from other operational products by an average of 20 nm at day 3, and 141 nm by day 5 (CLP5 excluded, since it is not a model-based product). Like all models used in the verification, the track errors grew with the length of the forecast, but the errors from the ARW forecasts grew at a noticeably slower rate. For example, the track errors from the official forecasts grew from 23 nm at hour 12 to 233 nm at hour 120, while the ARW forecasts grew from 33 nm at hour 12 to 190 nm at hour 120.

Figure 2 shows the ARW track forecast errors from the 4 km runs compared with other available operational products. The results are generally similar to those from the 12 km forecasts, except for the error at 96 hours which is considerably larger than the results from the 12 km runs. It is noted that the two cases that are verified for the 96 hour forecasts were from two early Wilma forecasts which had relatively large timing errors.

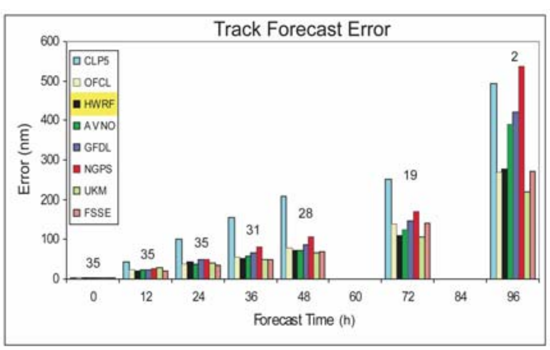

Figure 2: ARW 4 km track errors (in black) as compared with other operational products. The number of cases verified is indicated above the bars.

b. Intensity Forecasts

While it is encouraging to see ARW's track forecasts compare well with other operational products, especially in the day 3 to day 5 forecasts, it is more challenging to see if a model like ARW can add any value to the intensity forecasts, especially given the historical lack of skill by models in this area.

The intensity of the ARW model forecasts were measured by the 10 m winds directly output from the model. Figure 3 shows the comparison of the 12 km ARW with available operational products. The ARW intensity error was worse at hour 12 (15.6 kts), and became similar to others at hours 24 and 36 (16.9 and 17.4 kts, respectively). As the forecast time increases, the ARW intensity errors improved in comparison to the other forecasts. By forecast hour 120, the ARW had the smallest error of 13.9 kts. The ARW also had the smallest bias errors at hours 96 and 120 and the second smallest at hour 72 (not shown).

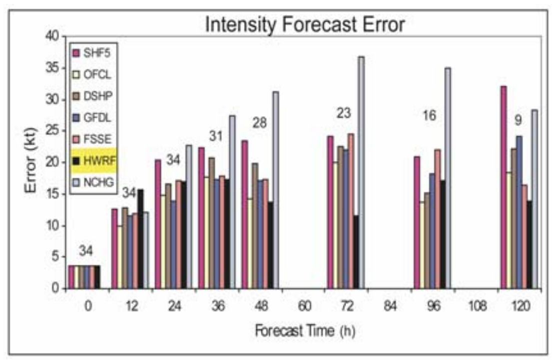

Figure 3: ARW 12 km intensity errors (in black) as compared with other operational products. The number of cases verified is indicated above the bars.

The intensity error verification for the 4 km ARW is presented in Fig. 4. The errors are comparable to other operational products from hour 12 to 48, but the forecasts were better than most other products at hour 72 by an average of 7.2 kts, and at hour 96 by an average of 11.6 kts.

It is clear from Figures 3 and 4 that both ARW's 12 km and 4 km forecasts outperformed most of the operational products at day 3 to day 5 forecasts for the storms verified here. When comparing ARW's 12 km and 4 km forecasts, the 4 km forecasts gave slightly better intensity forecasts at hours 12 and 24, but slightly worse

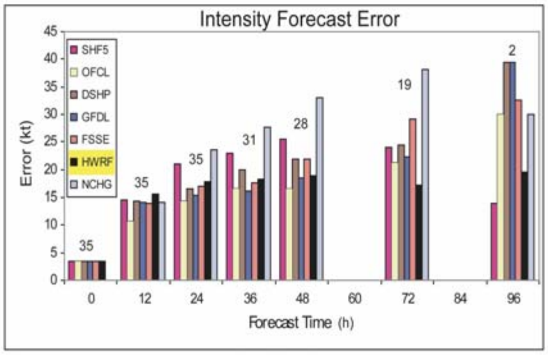

Figure 4: ARW 4 km intensity errors (in black) as compared with other operational products. The number of cases verified is indicated above the bars.

results at later hours (not shown). This is still under investigation.

The 4 km configuration, however, provided some aspects of the forecasts that were not possible with a coarser grid, or a grid that uses a cumulus parameterization scheme. Figure 5 shows an example of such a forecast. It shows the model simulated radar reflectivity at 62 forecast hours (left panel) for Katrina at landfall. When it is compared with the observed reflectivity (right panel), it shows that the size of the hurricane eye, the structure and strength of the rainbands, especially to the northeast of the storm center, are represented reasonably well by the model.

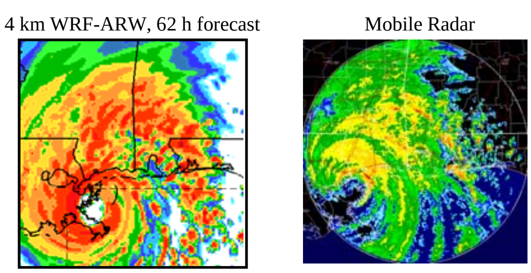

Figure 5: 4 km, 62-hour WRF-ARW forecast of maximum reflectivity for Katrina, initialized at 0000 UTC August 27 and valid at 1400 UTC August 29 (left), and Mobile radar reflectivity at the same time (right).

5. Discussion and Summary

We consider that the lack of skill in the early time periods arises from the relatively poor initial conditions obtained by simple interpolation of the bogused GFDL initial condition or the coarse grid GFS analysis. An analysis of the model data indicates that the model can take up to 24 hours to properly adjust. This points to the substantial benefit that can be obtained from the use of a proper data assimilation approach to specifying the initial condition.

Despite of the shortcomings in the model initialization, our comparison of the un-tuned ARW model with operational forecast products for a part of the active 2005 hurricane season is very encouraging. The ARW is very competitive at longer time scales and we consider that the use of a proper data assimilation approach will improve the shorter term forecasts substantially. A limited test using very high resolution nest also indicates the potential for substantial improvement in intensity forecasts. These results have encouraged us to continue further developments aimed at improved modeling of track, intensity and storm structure.

References:

Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, 2005: The Weather Research and Forecasting Model: Software architecture and performance. *Proceedings of the Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology*. Eds. W. Zwiefhofer and G. Mozdzynski, World Scientific, 2005, pp 156 – 168.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: *A description of the Advanced Research WRF Version 2*. NCAR Tech Note TN-468+STR. (Available at *www.wrf-model.org*.)

Terenelli, J. E., and S. S. Chen, 2001: High-resolution simulation of Hurricane Floyd (1999) using MM5 with a vortex-following mesh refinement. *Preprints, 18th Conference on Weather Analysis and Forecasting / 14th Conference on Numerical Weather Prediction*, 30 July – 2 August, 2001, Ft. Lauderdale, Florida, AMS, J54 - J56.

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.