RC23057 (W0210-188) October 29, 2002
Other

IBM Research Report

Case Studies of an Operational Mesoscale Modelling System in the Northeast United States

Anthony P. Praino, Lloyd A. Treinish,
Zaphiris D. Christidis, Annettte Samuelsen
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

3.10

CASE STUDIES OF AN OPERATIONAL MESOSCALE MODELLING SYSTEM IN THE NORTHEAST UNITED STATES

Anthony P. Praino*, Lloyd A. Treinish, Zaphiris D. Christidis and Annette Samuelsen
IBM Thomas J. Watson Research Center, Yorktown Heights, NY

1. INTRODUCTION

The *Deep Thunder* mesoscale weather modelling and visualization system has been run semi-operationally since January 2001 at the IBM Thomas J. Watson Research in Yorktown Heights, NY focusing on the New York City metropolitan area. The model used in this work is derived from earlier work supporting the 1996 Centennial Olympic Games in Atlanta (Snook et al, 1998). This is a highly modified version of the modelling system described by Pielke et al, 1992. The model is configured in a triply-nested domain structure. The outermost domain covers much of the northeast United States at 16 km. The middle domain includes the greater New York City tri-state area at 4 km and the innermost domain covers metropolitan New York City at 1 km. In addition, all nests are configured at 31 vertical levels. Figure 1 shows the domain configuration for the 4 km and 1 km nests, with the boundary of the latter marked in red on this terrain map. Locations of National Weather Service metar reporting stations used in this study are shown in white. Selected airport (IATA) locations and municipal centers are indicated in black. Currently, two 24-hour forecast runs are produced each day. The details of the system architecture, modelling and implementation are discussed in Treinish et al, 2003. We present herein the preliminary results of a portion of our continuing work in overall system implementation and benchmarking.

2. DATASETS AND METHODS

Currently, the data for both initial and boundary conditions, the latter nudged every three hours, for each

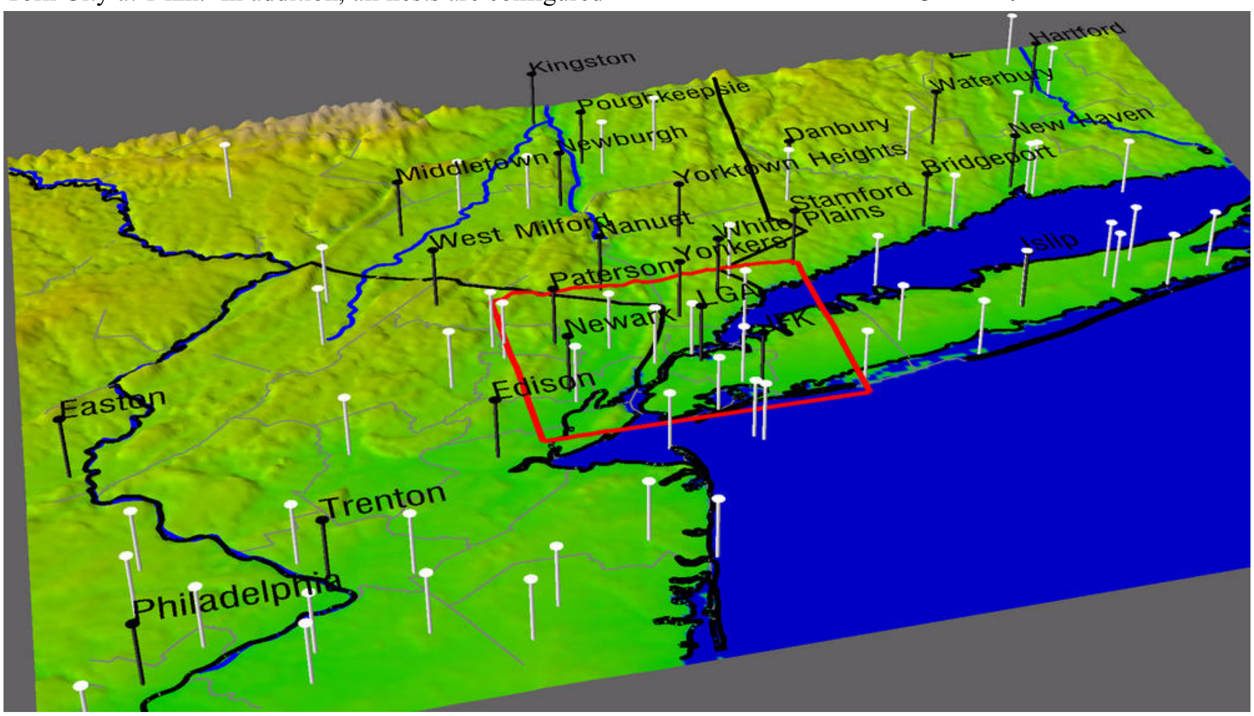

Figure 1. Inner Model Nests and Metar Reporting Stations.

model execution are derived from the Eta synoptic-scale model operated by the National Centers for Environmental Prediction.These data are available after sampling to 40 km resolution every three hours on the AWIPS 212 grid and interpolated to 27 isobaric levels. Forecast model runs are now fully automated and initiated at 0 UTC and 12 UTC daily. Each model execution requires roughly two hours to compute on the available IBM parallel computing system (Treinish et al, 2003). Post-processing of model results creates static and animated visualization products as well as diagnostic and statistical analyses used for system administration and model verification.

After each model run, the results of all three nests combined in a multi-resolution structure (Treinish, 2000) are bilinearly interpolated to the locations of the National Weather Service metar stations at 10 minute steps of forecast time. After the observations corresponding to each model run become available, a verification process is initiated in which these spatially interpolated results are statistically analyzed and com-

*Corresponding author address: Anthony P. Praino, IBM T. J. Watson Research Center, P. O. Box 218, Yorktown Heights, NY 10598, `apraino@us.ibm.com`

pared to parsed and quality-checked surface observations. Given the roughly hourly sampling of the metar data, the two *Deep Thunder* data values nearest in time to the observation time are linearly interpolated to the closest temporal match. An analogous process is applied to each Eta-212 grid as part of the automated pre-processing. However, the observational data are temporally interpolated to the nearest Eta time (every three hours) for the variable of interest. This yields a set of evaluation tables and visualizations for each model run as well as the aggregation of all model runs during the previous week. In addition, we have used this information to begin to perform some long-term verification, which are discussed below.

This process is not as straightforward as it should be due to a variety of inconsistent samplings in space, time and observables (e.g., precision and error). There are also occasional problems with the quality and available of the observations, as well as noise due to the measurement process, which impact the results. Although simple quality control (e.g., range checking) is used to eliminate measurements which are clearly out of range, that can be insufficient in practice. In addition, wind data for speeds below 3 mph as well as variable wind data and gusts are not included for statistical calculations.

Since the measurements are made above the surface (2 m for temperature, dew point, etc. and 10 m for wind), and the model topography is only an approximation of the actual station elevation, simple corrections are applied. Temperature data are adjusted for elevation differences between model and observation using the standard lapse rate of 6.5 degrees per km. Exponential atmospheric pressure change ($p(z) = p(0)e^{-z/z0}$) is used for correction of the difference in elevation between the observations and the model data pressure.

A variety of standard statistics are computed, for each model forecast (both *Deep Thunder* and Eta) in total, by time and by location. The statistical products which compare observations with model output include bias, mean absolute error, mean square error, root mean square error, skill score, etc. Only a handful of those statistics are discussed herein. In addition, contingency tables are produced to compare *Deep Thunder* precipitation results with rain gauge data at different thresholds. These results will be discussed in a future paper.

3. COMPARISON AND VERIFICATION OF *DEEP THUNDER* AND ETA FORECASTS

Although model runs as part of this project were first generated in early 2001, the creation of forecast products was part of on-going development, building of infrastructure and automation (Treinish et al, 2003). Regular twice-daily forecasts as outlined above have been in production since early 2002. Therefore, only a subset of the forecast results are being considered in this initial quantitative and qualitative evaluation.

3.1 Quantitative Analysis

As a first step, a relatively long time series (July 2001 through September 2002) of model results from both Eta and *Deep Thunder* vs. observation are considered. Usually one run per day was completed during the first portion of this period (i.e., initialized at 0, 6, 12 or 18 UTC) with two at 0 and 12 UTC typically each day in the latter half. The observations considered are only from the 55 metar locations inside the 4 km domain. The results of this initial analysis are shown in Figures 2 through 8. They each show either root mean square errors or biases as a function of forecast time for both *Deep Thunder* (blue) and Eta (red). In each case, the former are averaged to every three hours to enable direct comparison with the Eta statistics.

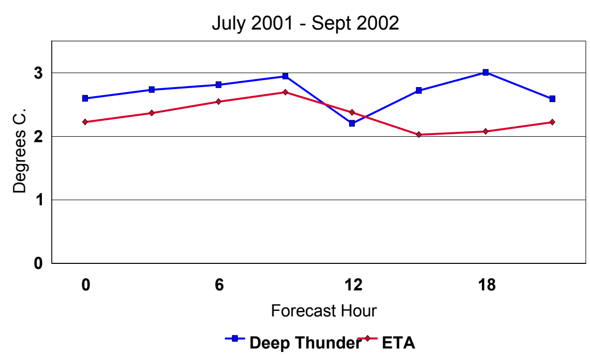

Figure 2. *Deep Thunder* and Eta Temperature Errors.

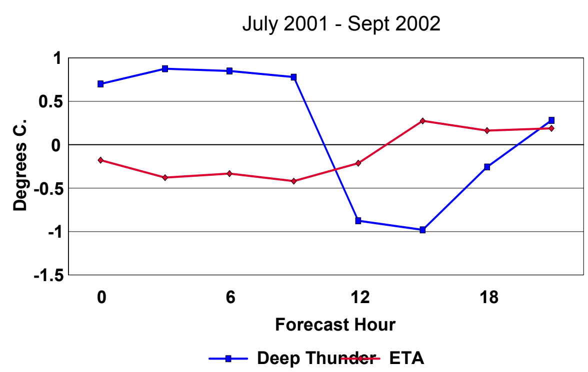

Figure 3. *Deep Thunder* and Eta Temperature Biases.

Figures 2 and 3 show the comparison of temperature data. They indicate that *Deep Thunder* is comparable to Eta with root mean square errors typically less than three degrees. Biases shown in Figure 3 are within plus or minus one degree.

Dew point root mean square error (Figure 4) comparisons show *Deep Thunder* within four degrees, while Eta is slightly better. However, the former has a positive one to two degree bias.

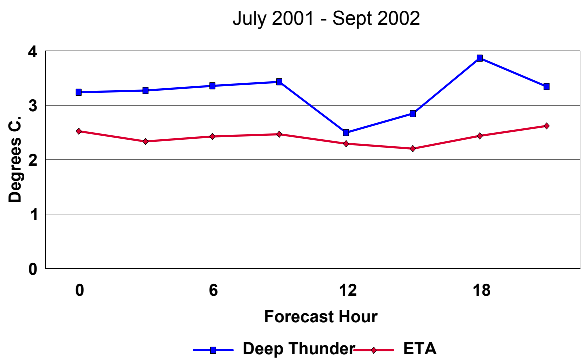

Figure 4. ***Deep Thunder*** **and Eta Dew Point Errors.**

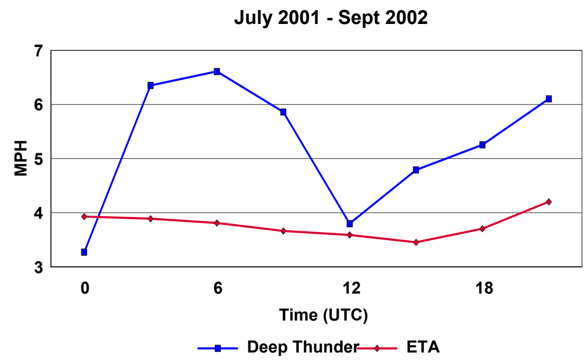

Figure 7. ***Deep Thunder*** **and Eta Wind Speed Errors.**

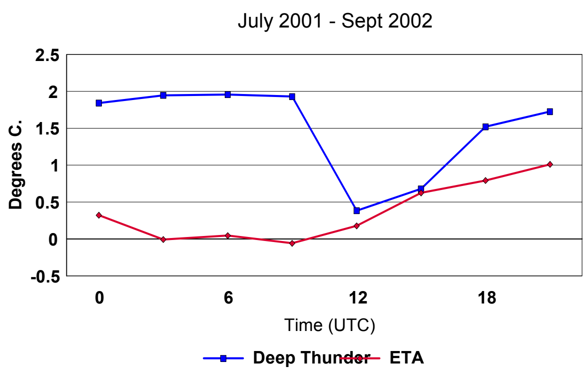

Figure 5. ***Deep Thunder*** **and Eta Dew Point Biases.**

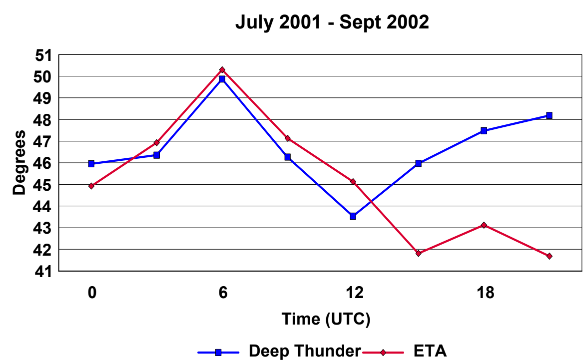

Figure 8. ***Deep Thunder*** **and Eta Wind Direction Errors.**

Deep Thunder is comparable to Eta in pressure (not shown) with a relatively constant root mean square error of about 0.5 inches of mercury. As shown in Figure 6, there is a significant low bias for both models.

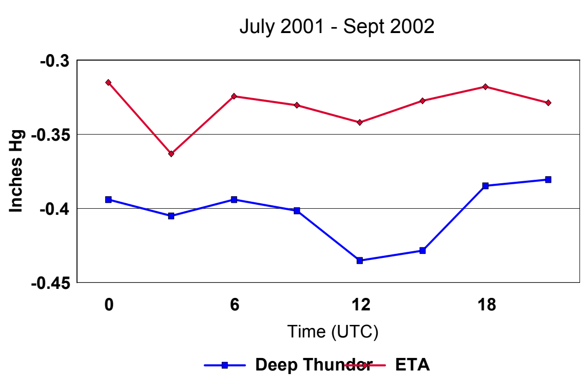

Figure 6. ***Deep Thunder*** **and Eta Pressure Biases.**

Wind speeds for *Deep Thunder* (Figure 7) show a much wider range and larger root mean square errors. In contrast, wind directions as shown in Figure 8 were similar to Eta. *Deep Thunder* has less of a bias for wind direction (not shown).

All of the aforementioned results for *Deep Thunder* show a consistent reduction in both errors and bias at 12 hours of forecast time. This is potentially due to both model correction and nudging of the boundary condition, but requires further study.

3.2 Qualitative Evaluation of Events

Complementary to the aforementioned statistical analyses, we have made qualitative comparisons of specific severe or unusual weather events within the forecasting domain. Obviously, this is problematic, given the lack of appropriate observations as well as numerical access to other model data of relevance. In general, we have observed good predictions of numerous convective events in terms of timing, location and intensity.

For example, consider a line of severe thunderstorms preceding a cold front in southeastern New York that occurred during the evening on May 31, 2002. A sample of the *Deep Thunder* results is illustrated in Figure 9 with a model run initiated at 12 UTC that day. Such products were available operationally at approximately 17 UTC (i.e., six to seven hours before the event occurred).

Figure 9 is a qualitative, yet comprehensive, three-dimensional visualization of the model results at 2340 UTC. It shows a terrain map, colored by a forecast of total precipitation, where darker shades of blue indicate heavier accumulations. The map is marked with the location of major cities or airports as well as river, coast-

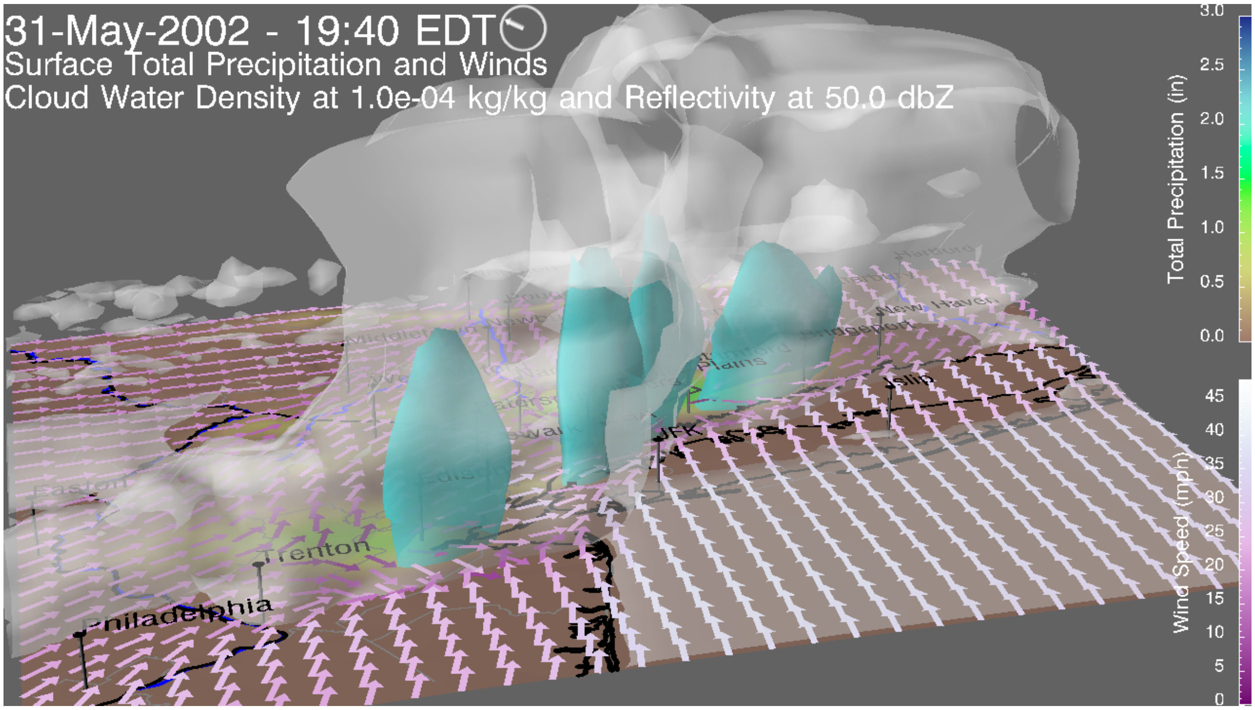

Figure 9. Model Forecast for 5/31/2002 Convective Event at 2340 UTC.

line and county boundaries within the 4 km nest. In addition, there are colored arrows indicating predicted winds, with the lighter color being faster winds and the arrow direction corresponding to the direction to which the wind is flowing. Above the terrain is a forecast of clouds, represented by a three-dimensional translucent white surface of total cloud water density (water and ice) at a threshold of 10^{-4} kg water/kg air. This approximation of a cloud boundary shows the typical *anvil* shaped structure of a cluster of thunderstorm cells. Within the cloud surface are translucent cyan surfaces of forecast reflectivities at a threshold of 50 dbZ, that correspond to rain shafts for individual convective cells. The frontal boundary and trailing areas of precipitation are quite clear.

Figure 10 illustrates NexRad radar observations from the nearby National Weather Service field office in Upton, NY at approximately the same time as in Figure 9 (2343 UTC). The left-hand panel shows composite reflectivities in short-range mode while the right-hand side are radar-derived storm total precipitation esti-

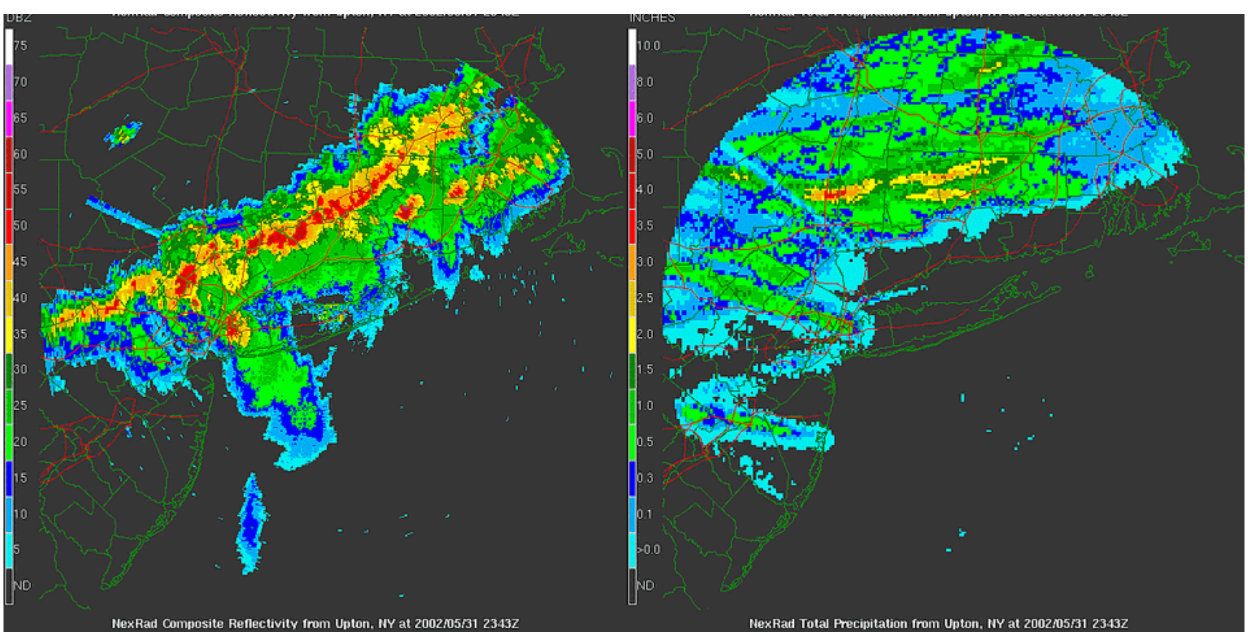

Figure 10. Radar Composite Reflectivities and Precipitation Estimates for 5/31/2002 2343 UTC.

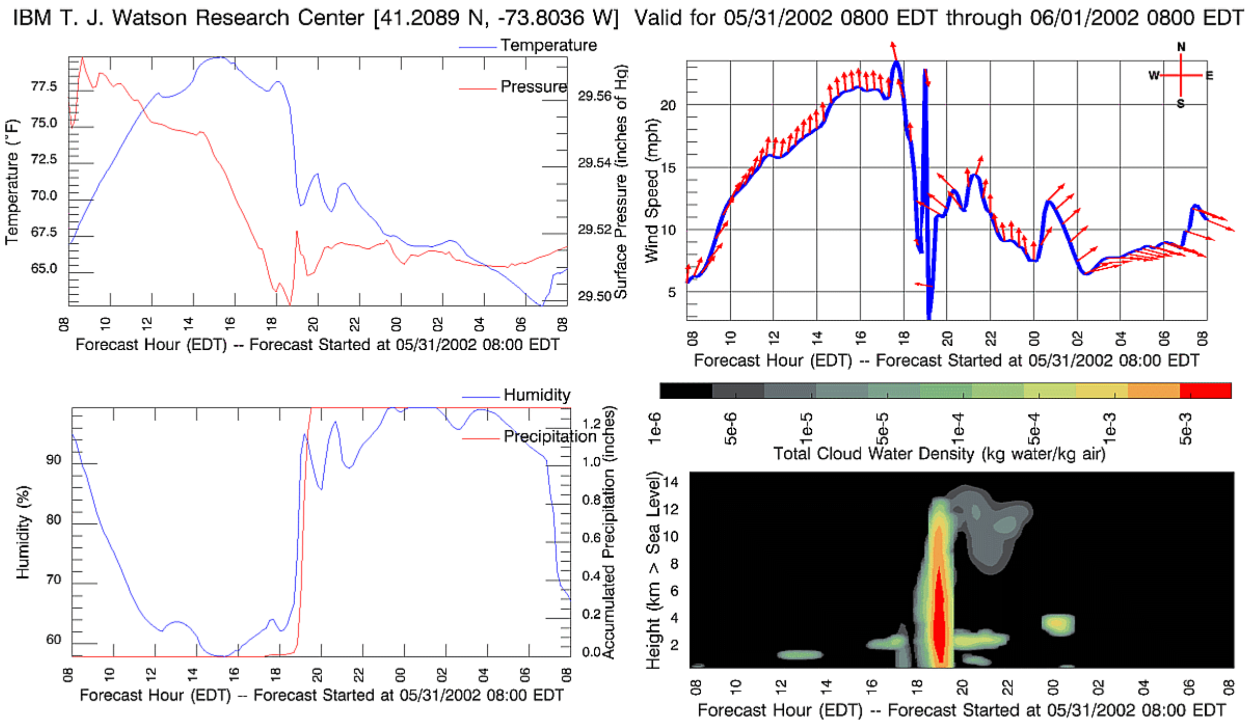

Figure 11. Model Forecast Meteogram for 5/31/2002 Convective Event (24-Hour Model Run Initiated at 12 UTC).

mates. Visual comparison of these images or more appropriately animation sequences of them through this period of time, illustrates good correspondence in the location and timing of the frontal boundary as well as intensity of radar echoes.

The model results can be further evaluated via Figure 11. This meteogram consists of four panels showing forecasted surface data. In all cases, the variables are shown as a function of forecast time bilinearly interpolated to a specific location (i.e., the authors' offices) from the data generated for the 4 km nest. They can be used to isolate the predicted time of the frontal and thunderstorm passage at this location. The plots on the left each show two variables while those on the right show one. The top left plot presents temperature (blue) and pressure (red). The middle left panel shows humidity (blue) and total precipitation (red). Since the precipitation is accumulated through the model run, the slope of the curve will be indicative of the predicted rate of precipitation. The top right plot illustrates forecasted winds -- speed (blue) and direction (red). The wind direction is shown via the arrows that are attached to the wind speed plot. The arrows indicate the predicted (compass) direction to which the wind is going. The bottom right plot is a colored contour map of forecasted total (water and ice) cloud water density as a function of elevation and time, following the color legend on the top of the panel. Portions of the plot in black imply time or elevations where there are little or no clouds. In combination, all of these panels show a sharp discontinuity at about 2300 UTC (1900 local time) indicative of the frontal and thunderstorm passage. The rapid movement of the front at this time is further shown by the forecasted wind speed at that time. Visual observations by the one of the authors in this vicinity indicate that the actual passage was at approximately 2330 UTC.

Therefore, the location and timing of the predicted event are quite good. However, a portion of the squall line that passed through southern New Jersey was not predicted by the model. This may be due to the fact that the line was approaching the boundary of the 4 km nest. In addition, forecasts like this tend to have a positive bias in precipitation amounts for convective events compared to rain gauge data. However, the correspondence to radar-based estimates as shown between Figures 9 and 10 is relatively good.

Another qualitative case to be considered is the unusual late spring snow storm that occurred in the Catskill Mountains of New York on the morning of May 18, 2002. The results of a *Deep Thunder* forecast initiated at 1200 UTC on May 17 are shown in Figure 12. It illustrates predicted accumulated snow in south-central New York State (within the 16 km nest) as a contour map, following the legend to the upper right. Products such as these were available operationally at approximately 17 UTC (i.e., almost 15 hours before significant snow fall began). As with the other example discussed above, location and timing of the event were good as corroborated via observations from snow spotters that were reported to local National Weather Service offices. In this case, model snowfall accumulations were somewhat higher (e.g., 10 vs. 8 inches) than the actual reports.

4. DISCUSSION AND FUTURE WORK

Overall, the results to date are mixed. *Deep Thunder* has shown very good skill at prediction of severe or unusual weather events, especially those involving sig-

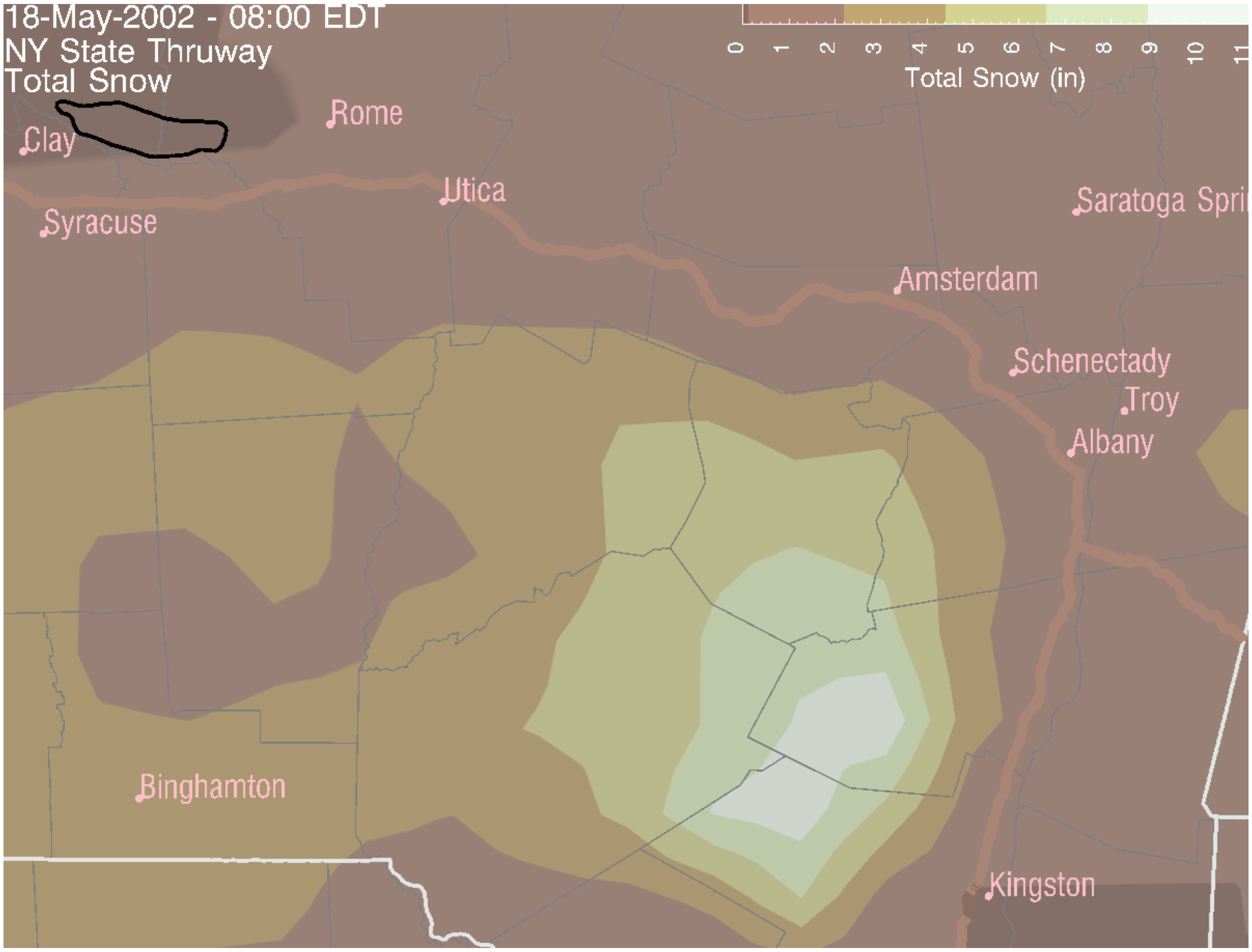

Figure 12. Model Forecast for Accumulated Snowfall at 5/18/2002 1200 UTC.

nificant convection. This could be viewed as an expected outcome given the considerable amount of computation used to make such high-resolution results and detailed microphysics practical. In contrast, the preliminary statistical analysis indicates skill that at best is comparable to that of the Eta results used to prepare initial and boundary conditions. Hence, there is some consistency with issues discussed recently about the utility of high-resolution, mesoscale modelling (e.g., Mass et al, 2002).

However, this is an on-going effort with considerable potential. For example, the capability to produce quantitative comparisons and automated analyses for model runs are a relatively recent addition to the overall system's capabilities. Hence, there has been no significant model tuning based upon these findings. Preliminary results that are at least favorable when compared with Eta and observation is encouraging. This suggests a number of next steps to refine the model operations and forecast quality.

The first effort will be additional analyses. This will include identification of potential seasonal bias as well as an isolation of geographic biases inherent in the *Deep Thunde*r results. This can be extended to consider temporal biases (i.e., forward or backward lags between model and observations). Although other statistics have been calculated, they and the results for precipitation have not been addressed herein, and require examination. Of course, this work will be extended to include more detailed examination of interesting weather events. The overall statistics will be improved by utilizing observations made by weather stations operated independently of the National Weather Service and more frequent model runs. Additional quality control techniques will be applied by examining neighboring observations for both temporal and spatial consistency. Similarity theory will be used to applied a better correction of the model results to the shelter height of all observations to enable a more accurate comparison.

To aid in the improvement of forecast quality, the ability to leverage the expected availability of full-resolution 12 km Eta results on the AWIPS 218 grid as well as some assimilation of observations will be implemented to enhance both initial and boundary conditions.

Since a focus of this project is to provide customized capabilities to assist in weather-sensitive business operations, efforts will also be addressed to determine and apply alternate metrics for measuring business value. These will serve to provide an evaluation of *Deep Thunder* that is complementary to the traditional meteorological verification.

5. ACKNOWLEDGEMENTS

One of the authors, Annette Samuelsen, is a Ph.D. candidate in oceanography at Florida State University. She contributed to this effort during her summer 2002 internship at the IBM Thomas J. Watson Research Center.

This work is supported by the Mathematical Sciences Department at the IBM Thomas J. Watson Research Center.

6. REFERENCES

Mass, C. F., D. Owens, K. Westrick, B. A. Colle. *Does Increasing Horizontal Resolution Produce More Skillful Forecasts*. **Bulletin of the American Meteorological Society**, **83**, no. 3, pp. 407-430, March 2002.

Pielke, R. A., W. R. Cotton, R. L. Walko, C. J. Tremback, W. A. Lyons, L. D. Grasso, M. E. Nicholls, M.-D. Moran, D. A. Wesley, T. J. Lee and J. H. Copeland. *A Comprehensive Meteorological Modeling System - RAMS*. **Meteorology and Atmospheric Physics**, **49**, 1992, pp. 69-91.

Snook, J. S., P. A. Stamus, J. Edwards, Z. Christidis, J. A. McGinley. *Local-Domain Mesoscale Analysis and Forecast Model Support for the 1996 Centennial Olympic Games*. **Weather and Forecasting**, **13**, no. 1, pp. 138–150, January 1998.

Treinish, L. A. *Multi-Resolution Visualization Techniques for Nested Weather Models*. **Proceedings of the IEEE Visualization 2000 Conference**, October 2000, Salt Lake City, UT, pp. 513-516, 602.

Treinish, L. A., A. P. Praino and Z. D. Christidis. *Implementation of Mesoscale Numerical Weather Prediction for Weather-Sensitive Business Operations*. To be published in **Proceedings of the Nineteenth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology**, February 2003, Long Beach, CA.