LESSONS LEARNED ON SHORT-TERM EXPLICIT PREDICTION OF STORMS THROUGH A JOINT IHOP_2002 RETROSPECTIVE STUDY

Juanzhen Sun*, Morris Weisman, Fei Chen, Chris Davis, Kyoko Ikeda, Yubao Liu, Kevin Manning, Roy Rasmusson, Greg Thompson, Stan Trier, Qingnong Xiao, Wei Yu
National Center for Atmospheric Research, Boulder, Colorado
Istvan Geresdi
University of Pecs, Pecs, Hungary

1. INTRODUCTION

An across-laboratory research program, named STEP (Short Term Explicit Prediction), was established at the National Center for Atmospheric Research (NCAR) in 2006 with the objective of improving the short term explicit (convection permitting) forecasting of high impact weather. The STEP program enables the coordination of NCAR's exiting wide-range of activities and expertise across the institution with emphasized themes to achieve this objective. These activities range from basic understanding of high impact weather to development and real-time demonstration of data assimilation and forecasting systems. Last year, STEP organized a joint retrospective study using IHOP_2002 data with the participation of eight projects covering topics of basic understanding of deep convective systems, high-resolution data assimilation and numerical forecasting, physical parameterization, nowcasting system, and advanced verification technique. The main objective of the joint retrospective study was to evaluate the current capabilities, which are developed at NCAR with the goal of supporting the research and operational community, for short-term (0-12 hours) explicit forecasting of storms.

The objective of this paper is to report what we have done and learned on the modeling aspect of the joint study, with the emphasis on model initialization and sensitivity of physical parameterizations. The period of June 10-16, 2002 was selected for the common focus of the study. This period was chosen because of the high frequency of convection occurrence and the large variety of convective systems. Wilson and Roberts (2005) documented Twenty-two initiation episodes for the seven days. The statistical analysis of observed rainfall revealed that the regime of convection for the one-week period corresponded well to translating synoptic cold front pattern with significant southeastward propagation. The synoptic regimes and resulting convections for this period along with a description of baseline experiment will be described in section 2. The data assimilation experiments and their results will be described in section 3. Section 4 is devoted to the sensitivity of physical parameterization schemes to short-term forecasting. The role of environmental condition on deep convection is examined in section 5. The summary and concluding remarks are given in section 6.

2. DESCRIPTION OF THE WEATHER REGIMES AND RESULTING CONVECTION

The overall weather pattern (see Fig. 1) started with a strong upper-level trough in the Pacific Northwest on the 10th, which progressed steadily eastward across the northern plains of the US, ending up over the Great Lakes by the 15th. This resulted in weak southwesterly upper level flow over the domain of interest on the 10th, which gradually shifted to westerly and then moderate northwesterly upper level flow by the 15th, with occasional short-wave features propagating through this overall upper-level regime. Key surface features at the beginning of the period included
generally moist, unstable conditions supportive of convection through most of the region, with a dry line in the western portion of the domain. Convection remained mostly scattered on the 10th, with a few stronger cells along the dry line in eastern Texas and west-central Kansas. A more organized convective system developed in central and southeastern Kansas late afternoon of the 11th, ahead of a cold front that was slowly progressing southeastward from Wyoming and Nebraska. This front was located in south-central Kansas by the afternoon of the 12th, triggering a new squall line that propagated through most of Oklahoma overnight. Another significant convective system associated with the cold front propagated through Oklahoma and northern Texas on the morning of the 13th, stabilizing most of the region through the 14th, except for some scattered convection limited to the high plains of eastern Colorado and eastern New Mexico. Moist, unstable southerly flow returned to the western portion of the domain on the 15th, which, along with an upper-level disturbance in the northwesterly flow, resulted in severe convection being triggered in northwest Kansas by mid-day.

Fig. 1. 300 mb Eta analysis of geopotential height (contour), wind speed (color), and wind vector for June 10 – 15.

This convection organized into a strong, bow-shaped system as it propagated southeastward through Kansas during the afternoon, continuing to propagate southward through Oklahoma and northern Texas during the evening.

A baseline simulation was produced for this period over a domain of 800 x 750 km using
the WRF-ARW model at 3 km horizontal grid resolution, initialized both at 00 UTC for a 36h forecast, and 12 UTC for a 24h forecast. No assimilation procedures were used for this run (e.g., cold start), so that the impact of such techniques could be more easily judged. This baseline simulation uses Thompson microphysics, MYJ PBL scheme, and Eta analysis for initialization. It serves as a benchmark for data assimilation and physical parameterization sensitivity studies that will be discussed in the following sections.

3. HIGH-RESOLUTION DATA ASSIMILATION AND FORECASTING EXPERIMENTS

Three data assimilation systems have been developed for WRF, including WRF-3DVAR and 4DVAR, nudging-based RTFDDA (Real Time Four Dimensional Data Assimilation), and WRF Ensemble Kalman Filter (EnKF). These data assimilation systems were applied to high-resolution data in the past including radar observations. In this study, the WRF 3DVAR and RTFDDA were run for the one-week period with 3km and 3.3 km resolutions respectively, while 4DVAR and EnKF only for individual cases due to computation limitations and the inadequate level of system readiness. In this section, we will only describe the multi-day experiments and results of RTFDDA and 3DVAR. The 4DVAR and EnKF case studies will be described in other papers. Because the emphasis of the Joint IHOP retrospective study is on the 0-12 hour explicit forecasting of convective systems, all assimilation systems included radar data in their experiments.

The Doppler radar data assimilation in WRF 3DVAR was described by Xiao et al. (2005, 2007, 2008). These studies demonstrated positive impact of Doppler radar on precipitation forecasts on a model grid of 10 km. The higher resolution of 3 km used in this study is a good test for the 3DVAR’s ability in high-resolution data assimilation. Radial velocity and reflectivity from 35 NEXRADs on the domain of the baseline run covering the states of Oklahoma, Kansas, and part of the surrounding states. These radar data were used in 3DVAR, RTFDDA, and 4DVAR after pre-processing and quality control. The WRF 3D-Var background error statistics was performed using the WRF forecasts from 14 April to 15 May 2007 and interpolated to the IHOP domain. Two seven-day (June 10-16, 2002) experiments, without and with radar data respectively, were conducted. Figure 2 shows the rainfall verification of the two experimental results against NCEP Stage IV precipitation data, along with the skill score of the baseline forecasts described in the last section. The verification results indicate that WRF 3D-Var with only conventional GTS data assimilation degrades the Eta analysis and results in inferior rainfall forecasts. With Doppler radar data assimilation, the rainfall verification ETS score is increased and the rainfall forecasts are improved up to 9 hours. Comparison between the experiments with/without radar data assimilation clearly show the benefits of radar data assimilation in rainfall forecast. However, the
A troublesome result is that the WRF 3D-Var with the 3-km grid spacing degraded the baseline forecasts when only conventional data are assimilated. We are currently investigating the reasons for the degradation.

One of the possible reasons for the degradation is related to the background error statistics. To investigate the sensitivity of the forecast with respect to background error statistics, an experiment is conducted using a different background error statistics. This background error statistics was obtained using WRF 10 km forecasts from 05 to 19 June 2002. The length scale and variance were reduced empirically for the 3km experiments. The squall-line case from 1200 UTC 15 to 0000 UTC 16 June 2002 is selected to perform this sensitivity experiment. We carried out the WRF 3D-Var at 1200 UTC 15 and conducted 12-h WRF forecast. The result clearly indicated that the forecasts with the statistics that was generated using the period that covers the IHOP retrospective study period had improved performance (figure not shown). More work on the proper specification of the background error statistics is being conducted.

Another possible reason for the 3DVAR forecast degradation in the no radar experiment is that the experiment was conducted without any continuous cycling. An experiment with the cycling is being conducted to verify the hypothesis. Due to the balance relations used in 3DVAR do not represent the convective-scale dynamics, a spin-up is still inevitable in 3DVAR forecast. In this regard, the nudging-based RTFDDA may have an advantage over the 3DVAR because of its 4-D nature of assimilation.

The RTFDDA system (Liu et al. 2008a,b) is an operational WRF-based multi-scale, rapid cycling mesoscale analysis and forecasting system. The main data assimilation algorithm of RTFDDA is “observation-nudging”, which is formulated to effectively combine diverse observations, including all standard WMO upper-air and surface observations and non-conventional data such as satellite winds, aircraft ACARS, radar VAD wind profiles, QuikScat sea surface winds, and high-density surface observations from various mesonets, with the full-physics WRF-ARW model. RTFDDA generates dynamically and adiabatically consistent meso- and small scale analyses, and thus is able to produce dynamically and cloud-precipitation “spun-up” nowcasts and short-term forecasts. This ability is particularly important for summer-convection prediction because a model forecast from “cold-start” initial conditions which include no or little convection information can be seriously impaired by convective “spinning-up” processes for the first a few to more than ten hours of forecasts, which can be observed in majority of operational NWP forecasts. A limitation of “nudging-based” FDDA is that it can only assimilate direct prognostic variables, e.g. temperatures, winds and moisture. A new hybrid-scheme has been developed to incorporate radar reflectivity and radial wind measurements. In the hybrid scheme, WRF 3DVAR is firstly employed to assimilate radar data and then the 3DVAR analyses are combined into the WRF-RTFDDA analyses through “grid-nudging” approach.

Two experiments, without and with radar data respectively, were run using RTFDDA on the IHOP data of June 10-16, 2002. The hybrid scheme is only used in the experiment with the radar data. Since these two experiments used WRF version 3 instead of version 2 as in previous experiments, a baseline experiment starting from GFS analysis was rerun using the new version. These experiments covers an inner domain of 733 km x 733 km with a horizontal resolution of 3.3 km, nested within a 10 km
domain. Figure 3 shows the ETS score and bias of 6-hour accumulated precipitation with respect to precipitation amount over the 7 days. It clearly shows that the data assimilation experiments via RTFDDA without or with radar improve the precipitation skill measured by ETS over the GFS cold start experiment. In addition, the assimilation of radar observations results in significant improvement over the experiment with RTFDDA without the use of radar data, which is especially evident through the reduction of the BIAS at the higher precipitation amount. Statistics of wind, temperature, and humidity revealed that the bias and the RMS error of these fields were all reduced with RTFDDA conventional data data assimilation and further reduced with radar data assimilation (figure not shown).

Figure 4 compares forecasts of 0-6h rainfall of cold-started WRF, WRF-RTFDDA, and WRF-RTFDDA-hybrid for 13 June 2002 squall line case with the NCEP Stage IV precipitation analysis. It can be seen that cold-started WRF produces too strong, too compact convection systems and with rain cores dislocated from the observed. WRF-RTFDDA greatly corrects these problems and the hybrid method appears to forecast the best morphologies and intensity of the observed system.

Lessons learned from the initialization experiments are:

a) WRF 3DVAR cold-start without radar observations could result in poorer precipitation forecasts than a cold-start that are solely based on interpolation of large-scale, coarser-resolution analysis. However, when radar data are included in the 3DVAR assimilation, improved precipitation forecasts are obtained.

b) Background error covariance is an important component in 3D-Var data assimilation. Further studies on the proper generation and tuning of background error statistics are needed for high-resolution data assimilation. Our experiment using background error covariance from ensemble forecasts with the same case (Xiao and Sun 2007) and the experiment carried out in Fig. 3c using the background error covariance from WRF forecasts in the same IHOP period implies that the flow-dependent background error covariance may play a vital role in high-resolution data assimilation.

c) The nudging-based RTFDDA initialization resulted in improved short-term forecasts over the GFS cold start. The hybrid technique that nudges the WRF 3DVAR analysis that includes radar radial velocity and reflectivity was proved to be a feasible methodology for assimilating radar observations in RTFDDA. The success may be attributed to the improved dynamical balance through the gradual nudging of the 3DVAR increments to the model.

4. SENSITIVITY EXPERIMENTS WITH RESPECT TO MODEL PHYSICS

Basic sensitivity testing with respect to physics considered variations to PBL schemes (e.g., YSU versus MYJ), strength of the land-surface coupling, and microphysics schemes (WSM6 versus Thompson). Generally, more forecast sensitivity was found by varying the initial and boundary conditions as oppose to varying the physics, which is in agreement with Weisman et al. (2008).

In order to gain an in-depth understanding of the performance of the commonly used microphysical schemes in WRF, simulations of an idealized two-dimensional squall line
were performed using a retrospective sounding from the squall line case observed on 12 June 2002 during the IHOP retrospective study period. Previous studies have shown sensitivity of squall line dynamics to microphysics parameterization (e.g., Yang et al. 1995). A unique aspect of this study is the inclusion of the detailed microphysics schemes of Geresdi et al. (2005) in addition to three more convectional mixed-phase bulk microphysics schemes [Lin, Farley, and Orville (1983); WSM 6-class (Hong et al. 2004); and Thompson et al. (2004)] that are commonly used in the WRF model. The model domain had a horizontal dimension of 501 km with a 1-km grid spacing and 81 vertical levels. The convection was initialized with a 1.5°C warm bubble and the simulation time was five hours. All of the four microphysics schemes have five hydrometeor categories: cloud water, rain, cloud ice, snow, and graupel.

Figure 5 shows Hovmuller diagrams of rainfall simulated with the four schemes. Using the Lin and WSM schemes, the initial convection was short-lived and did not propagate with the leading edge of the precipitation system. In comparison, the Detailed and Thompson schemes produced a relatively more organized system. The simulations also produced a stronger and deeper cold pool in the Lin and WSM schemes consistent with the faster propagating system. Detailed analysis showed that the sensitivity is related to bulk versus bin-wise treatment of particle fall speed and assumed rain and graupel size distributions of the bulk schemes versus variable size distribution in the Detailed scheme.

The Thompson scheme showed a similar cold pool evolution as in the Detailed scheme, because in the Thompson scheme, y-intercept parameter in rain and graupel particle size distributions varies depending on the mixing ratio and allows size-sorting to occur as in the Detailed model. The superior performance of the Thompson scheme is also observed in the real data forecasts. A statistical verification is being performed to confirm the visual observation.

Another in-depth study that is related to the sensitivity of the physics parameterization is on the land-atmosphere coupling. We have investigated the sensitivity of convection initiation (CI) and quantitative precipitation forecasts (QPFs) to the strength of the land-atmosphere coupling in the model. The strength of the land-atmosphere coupling is strongly influenced by the surface exchange coefficient, \( c_H \), which controls the total energy transferred from the land surface to the atmosphere. \( c_H \) is in turn affected by the treatment of surface layer physics and roughness length for heat and moisture, \( z_0 \). In the coupled WRF/Noah land surface modeling system, \( z_0 \) is related to the roughness lengths for momentum,
\[ z_{0m} = z_{0m} \exp\left(-kC_{zil} \sqrt{Re}\right) \]

where \( k = 0.4 \) is the von Kármán constant, \( R_e \) is the roughness Reynolds number and \( C_{zil} \) is an unknown empirical constant.

Chen and Zhang (2009), however, note that the appropriate value for \( C_{zil} \) required to yield reasonable values of \( C_{zil} \) may vary substantially depending on the height of the vegetation canopy. Since the strength of the land-atmosphere coupling influences the depth and thermodynamic properties of the atmospheric boundary layer (e.g., temperature, water vapor mixing ratio), it, in turn, likely influences the timing and intensity of convective precipitation in the model. To test the role of land-atmospheric coupling on CI and QPFs we varied the value of \( C_{zil} \) from 0.01 to 1.0 for 24-h forecasts initiated at 12 UTC for each of the six days that compose the IHOP retrospective period. This range for \( C_{zil} \) is consistent with that found empirically by Chen and Zhang (2009), with WRF currently having a default \( C_{zil} \) of 0.1. Here, a \( C_{zil} \) value of 0.01 (1.0) is indicative of larger (smaller) values of \( C_{zil} \), and hence

The six-day average dependency of QPFs on \( C_{zil} \) over the 800 \( \times \) 750 km baseline run domain is displayed in Fig. 6.

Notice that both the strong and weak coupling runs under-predict precipitation amounts for the first several hours due to model spin-up issues associated with a cold start. However, the area-averaged early evening maximum simulated precipitation values for the \( C_{zil} = 0.01 \) strong coupling runs precede those of the \( C_{zil} = 1.0 \) weak coupling runs by \( \sim 2 \) h and are on average almost twice as large. The maximum amounts in the weak coupling runs are more consistent with the Stage-IV radar-derived rainfall analyses (Fulton et al. 1998). The observed timing of the rainfall maximum falls between those of the strong and weak coupling runs. This significant QPF sensitivity is comparable, if not greater, than that associated with differences in land-surface parameters, such as soil moisture, over model forecasts of similar duration (e.g., Trier et al. 2004). Some caution, however, needs to be exercised in generalizing the current results since constant \( C_{zil} \) values were used in the current simulations and empirically-based observational studies (Chen and Zhang 2009) suggest strong dependence on vegetation type. We are currently analyzing simulations in which \( C_{zil} \) varies with vegetation category across the model domain and will report results in a future study.

The major lessons learned through the physics sensitivity study is summarized below:

a) Generally, more forecast sensitivity was found by varying the initial and boundary conditions as oppose to varying the physics.

b) The Thompson microphysical scheme produced superior short-term forecasts than the Lin and WSM6 schemes based the results of the simulated squall-line study. The real data run also seems to indicate that the Thompson scheme produced better forecasts than WSM6. A statistical verification is being performed to verify this conclusion.

c) The short-term forecasting of convection is sensitive to the treatment of land-atmospheric coupling. For short term
forecasting of convection, it is possibly more important or at least equally important to correctly model the land-atmospheric coupling than to accurately specify the vegetation and soil conditions.

5. KEY PHYSICAL PROCESSES CONTRIBUTING TO SHORT TERM FORECASTING SUCCESS OR FAILURE

For the forecast period of June 10-16 2002, we found that the forecast skill not only depended on the initialization method and the physical parameterization scheme but also varied considerably from case to case. As an example, the 12 UTC forecast for 12 June reasonably forecast the convective system in northern Oklahoma, although there was an over prediction of convection in the Texas panhandle. The 24h forecast valid the same time, however, erroneously forecast a convective system to be in southern Oklahoma instead. In contrast, the strong, bow-shaped convective system observed at 00 UTC on June 16 was very well forecast 12h in advance, and also reasonably forecast 24h in advance. The success and robustness of the forecast of this system may be attributed to the fact that it was triggered in response to a well defined shortwave in the northwesterly flow that was well represented in the larger scale analyses. We therefore ask the question: what are the key processes contributing to short term forecasting success and failure? To answer this question diagnostic studies are being conducted. These studies investigate the mesoscale environmental conditions, such as thermodynamical stability and horizontal wind shear in the vertical, on convective initiation and convective-scale processes that are inherent to convective cells, such as cold pool.

Because explicit models of convection do not use a cumulus parameterization, the triggering of convection must be treated on the grid scale. On the one hand, this implies that understanding how convection is initiated and maintained in explicit models is potentially useful for understanding the processes in the real atmosphere. However, such models still retain significant error in the placement and timing of convective systems. It is not clear whether this error resides more in the evolution of the so-called convective environment or is inherent in the small-scale processes that lead to deep convective cells to form. A diagnostic has been developed that examines in detail the evolution of the inhibition to convection. Here we express it as the maximum negative buoyancy (Bmin), or the maximum negative difference between the temperature of a lifted parcel and the environment at the same level (Fig. 7).

The evolution of this parameter can be evaluated by noting the change in the initial temperature and pressure of the parcel, and the change in the environmental virtual temperature at the level where the buoyancy is most negative. While traditional convective inhibition (CIN) is an integral measure of negative buoyancy, one does not obtain a continuous field of CIN because it is dependent on CAPE being positive. This restriction does not occur for Bmin. Furthermore, high-time-resolution output from explicit models allows us to see the evolution of Bmin and relate it to processes on various scales.

Preliminary analysis revealed that convective initiation has a reasonably good correlation with the value of Bmin. The overlapping of the small negative Bmin zone and the horizontal convergence band is found to be critical for convective initiation.

![Fig. 7. Schematic illustrating the Bmin parameter which is defined for the most unstable air parcel. The vertical location for Bmin is depicted by the horizontal red line and its value is the indicated Tparcel – Tenvironment. The remainder of this figure is adapted from Crook (1996).](image-url)
It was also found that once an organized system (e.g., a squall line) is formed, convection is able to maintain itself for several hours despite increasingly negative values of $B_{min}$ in the environment ahead it. This is consistent with the simulation results of Parker (2008) that concluded that intense localized lifting at the leading edge of mature convective systems can often support organized convection despite increasing negative buoyancy.

Further diagnostic studies are being conducted to examine the different forcing mechanisms for daytime and nighttime convection. The role of cold pool and low-level wind shear in convective initiation and development is also being studied using high-resolution analysis that assimilates Doppler radar wind and reflectivity. Examination of the key processes contributing to the success and failure of model forecast will be examined using the diagnosis, analysis, and observations.

6. SUMMARY AND CONCLUDING REMARKS

In this paper, we have described a joint study on short-term explicit prediction of deep convection using IHOP data of June 10-16 2002. Two of the current WRF high-resolution data assimilation systems, WRF 3DVAR and RTFDDA, were evaluated. WRF model sensitivities with respect to physical parameterization schemes pertaining to the short term forecast were investigated. In each of the sections, we have discussed major lessons learned through this study. The results shown here are preliminary because the work is still in progress.

7. REFERENCES


