

5.3 STATUS OF NOWCASTING THUNDERSTORM INITIATION: WHERE DO WE GO FROM HERE?

Rita Roberts*, Eric Nelson, James Pinto, Tom Saxen, and Cody Phillips

National Center for Atmospheric Research, Boulder, Colorado

1. INTRODUCTION

Arguably, the greatest challenge in providing accurate spatial and temporal nowcasts of thunderstorms is in forecasting storm initiation. Operational 10 cm wavelength radars can be used to analyze, monitor and track existing thunderstorms (>35 dBZ) but is of limited use in monitoring the rapid growth of nascent- (>10 dBZ) and non-precipitating cumulus clouds (< 10 dBZ) beyond ~50 km in range. Other data sources must be used to capture when and where new storms will form. Satellite (Roberts and Rutledge, 2003) and surface station data, numerical model output, radar-detected surface convergence boundaries, stability profiles, terrain, storm climatology (Wilson et al. 2007) and forecaster input are all ingredients that are currently being factored into the production of 0-2 hr storm initiation nowcasts (Saxen et al, 2006).

In this paper, we evaluate the current state-of-the-art in producing storm initiation nowcasts using statistical results from two demonstrations of the NCAR thunderstorm nowcasting (Auto-nowcaster) system conducted during the summer of 2006. However, standard methodologies computing PODs, FARs, and CSIs over very large domains fail to illustrate the true skill of most nowcasting systems because initiation nowcasts represent a significantly smaller fraction of events when compared to the large number of storm extrapolation nowcasts that typically dominate the statistics. Thus, two new approaches are presented here for computing statistical performance of storm initiation nowcasts that serve to highlight the value of forecaster input and improved accuracy in storm initiation nowcasts. Following these results, we explore if further improvements in accuracy can be achieved.

2. AUTO-NOWCASTER DEMONSTRATIONS

The NCAR Auto-nowcaster system is a unique automated nowcasting system designed to

combine radar, satellite, surface station, sounding and numerical model data not only for the detection and extrapolation of existing storms, but to provide 0-1hr nowcasts of storm initiation, growth and decay as well (Mueller et al, 2003). Several years have been spent refining this system to obtain accurate storm initiation nowcasts using a fuzzy logic-based engine to automatically combine predictor fields retrieved from observational datasets. More recently, environmental predictor fields representing the stability of the atmosphere are obtained from operational numerical model derived products. Derived fields such as CAPE, CIN, relative humidity, equivalent potential temperature, vorticity and frontal location are included in the suite of fields used to assess the likelihood for the environment to support new convection.

A key ingredient in accurately predicting storm initiation is the detection of low-level convergence boundaries which have been shown to play an important role in triggering new convection. Initially automated algorithms have been used to detect convergence boundaries in radar data. However these algorithms have difficulty detecting all portions of convergence boundaries that often extend between radars spaced 300 km or more apart.

Previous demonstrations of the Auto-nowcaster (or ANC) for Federal Aviation Administration and NWS-funded activities have shown that forecaster input into the ANC process adds consistency, reliability and accuracy to the 0-1 hr short term, time and location specific thunderstorm nowcasts (Roberts et al, 2005, Nelson et al 2006). This performance was achieved when forecasters, using an interactive ANC display tool, entered the locations of surface convergence boundaries observed in radar, surface station and satellite data.

In 2006 NCAR collaborated with the National Weather Service (NWS) and Weather Services

*corresponding author address: Rita Roberts, NCAR, P.O. Box 3000, Boulder, CO 80307; email: rroberts@ucar.edu

Incorporated (WSI) Corporation to include forecaster-entered, low-level convergence boundaries into the Auto-nowcaster in real-time to provide more complete representation of surface convergence boundaries and ideally, more accurate thunderstorm initiation nowcasts. This collaboration was in the form of two Auto-nowcaster demonstrations conducted at the NWS Forecast Office in Ft. Worth Texas for the Forecaster-Over-the-Loop project and with WSI over the central and northeastern sections of the U.S. during the spring and summer of 2006.

3. NWS FORECASTER-OVER-THE-LOOP

Workstations were installed at the Dallas/Ft. Worth NWS Forecast Office (FO) (see Fig. 1 that displayed operational datasets, ANC-related fields and the GUI controls associated with the interactive forecaster tools. NWS forecasters were trained on system and entered boundaries once every couple hours as needed. They could select the dominant synoptic regime of the day and also adjust the storm initiation likelihood field as needed when it was clear that the numerical model CAPE and CIN output fields were running too high or low on a given day.

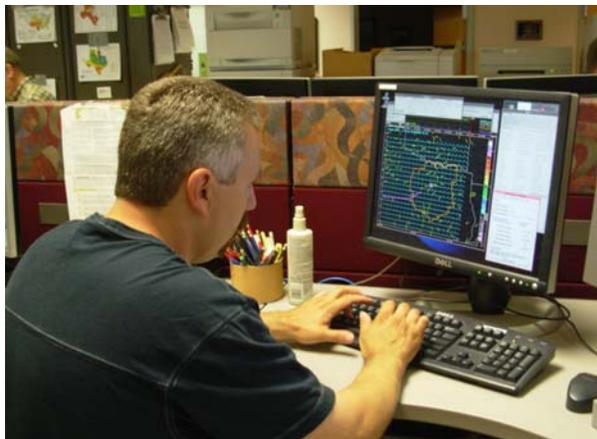


Figure 1. NWS forecaster entering boundaries on the ANC system.

The Auto-nowcaster was run over the NWS Dallas/Ft. Worth County Warning Area (CWA), the domain of responsibility for forecasters warning of hazardous weather (see Fig. 2). A mosaic of 7 radars was used in the validation process. Examples of ANC storm initiation nowcasts and the forecaster-entered convergence boundary associated with the passage of a cold front through the CWA on 30 March 2006 are shown at forecast time in Fig. 2a. Validation of the nowcasts

is presented in Fig. 2b. Storm initiation was moving through the CWA and all new storm initiation was triggered by cold frontal forcing. This relatively clear-cut event is presented here to facilitate the discussion of the different statistical approaches used in evaluation of the ANC demonstrations. Forecast and validation plots from later in the event are shown in Fig. 3 on the sub-grid scale.

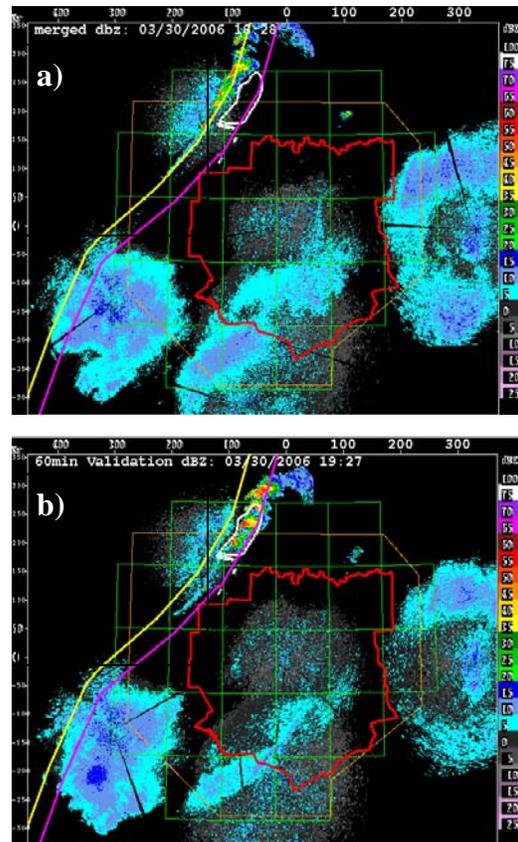


Figure 2. Forecast domain and validation grids for the NWS Dallas/Ft. Worth Forecaster-Over-the-Loop Demonstration on 30 March 2006. Red polygon bounds the County Warning Area (CWA); orange polygon bounds the full-domain used for computing the real-time validations statistics; the full-domain is divided into 1 deg latitude/longitude (green) boxes for computing sub-domain statistics. Yellow polyline is the forecaster-entered convergence boundary; magenta polyline is the 60 min extrapolated boundary position produced by ANC system. a) ANC 60 min storm initiation nowcasts (white polygons) overlaid onto the low-level radar reflectivity mosaic image at 18:28 UTC, forecast time. b) ANC 60 min storm initiation nowcast from (a) overlaid onto the radar reflectivity mosaic at 19:27 UTC, forecast valid time.

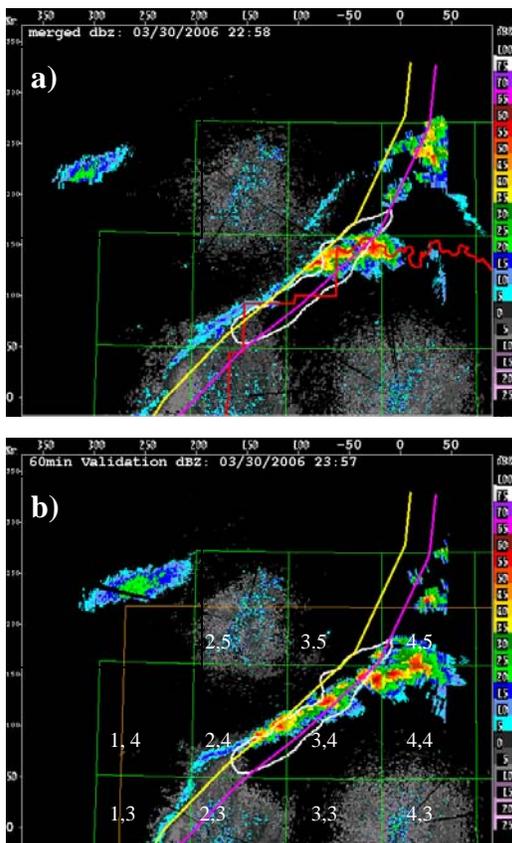


Figure 3. Same as Fig.2, except on the sub-grid scale on 30 March 2006. a) Storm initiation nowcasts overlaid onto reflectivity at 22:58 UTC; b) nowcasts overlaid onto reflectivity at 23:57 UTC validation time. Sub-grid box numbers listed in white.

3.1 Methodology for Statistical Evaluation

The overall performance of the ANC 60 min storm initiation and extrapolation nowcasts on 30 March is shown in Fig. 4 using standard statistical calculations. The yellow curve is the performance of storm nowcasts based on extrapolation of existing storms. Forecaster-entered boundaries are included in the set of predictor fields used to produce these particular nowcasts. This curve is a benchmark for performance; CSI scores higher than this curve represent greater skill and accuracy. The blue curve represents the performance of both *storm initiation* and the growth and decay nowcasts combined with forecaster input into the process (i.e., forecaster-entered boundaries). The magenta curve is performance of storm initiation, growth and decay nowcasts without forecaster-entered boundaries included. Thus, differences observed between the blue and magenta curves is an indication of the impact that forecaster-entered boundaries have on

the accuracy and skill of the storm initiation portion of the nowcast. When computing statistics over the full validation domain, it is possible to see the impact and improvement in the storm initiation nowcasts when forecaster-entered convergence boundaries have been included in the nowcast process. This is evident when comparing the blue and magenta curves in Figure 4b, particularly early in the event

(between ~1645-2100 UTC) when storm initiation is occurring frequently during that 5 hr period. Later in the day, after 2100, it is difficult to separate out the CSI skill associated with the storm initiation nowcasts, as nowcasts of extrapolated storms generally dominate the statistics over the very large domain where many storms are occurring simultaneously. Yet one need only look at the radar imagery at 2357 UTC in Fig. 3b to know that in addition to the persistence of existing storms along the NE portion of the cold front, new storm initiation is occurring along the SW portion of the cold front. Figure 3a shows that quite reasonable storm initiation nowcasts were forecast for this area 60 min in advance.

For this reason, a new approach was taken to perform the statistical evaluation of storm initiation nowcasts. The approach was simply to compute the statistics over smaller domain sizes that are more relevant to the scale on which the many different discrete areas of convection are occurring. This is illustrated in Fig. 3b, where the full validation domain (gold outline in Fig. 2) was sub-divided into regularly-spaced 1 deg latitude by 1 deg longitude boxes. Validation statistics are still computed on a 1km grid-to-grid comparison, as is done over the full domain. We have found that this approach provides us a more detailed and informative look at nowcast performance (see Figs. 5-8) and a better scientific understanding of the factors that increase or decrease the accuracy of the nowcast. For example, statistics shown in Figs. 5 and 8 illustrate ANC storm initiation, growth and decay nowcasts with (blue) and without (magenta) forecaster-input were mostly similar in the regions of these two sub-grid domains. However, at discrete time periods, validation of nowcasts-with-human do show increase in CSI accuracy of 0.2, at around 1853 UTC (Fig. 5). An increase in CSI of this magnitude is not evident in the full domain CSI statistics for the same time period (see Fig. 4b). Examination of the statistics shown in Figs. 6 and 7 illustrate this point even more dramatically, in the time periods centered around 00-01 UTC, when compared with the

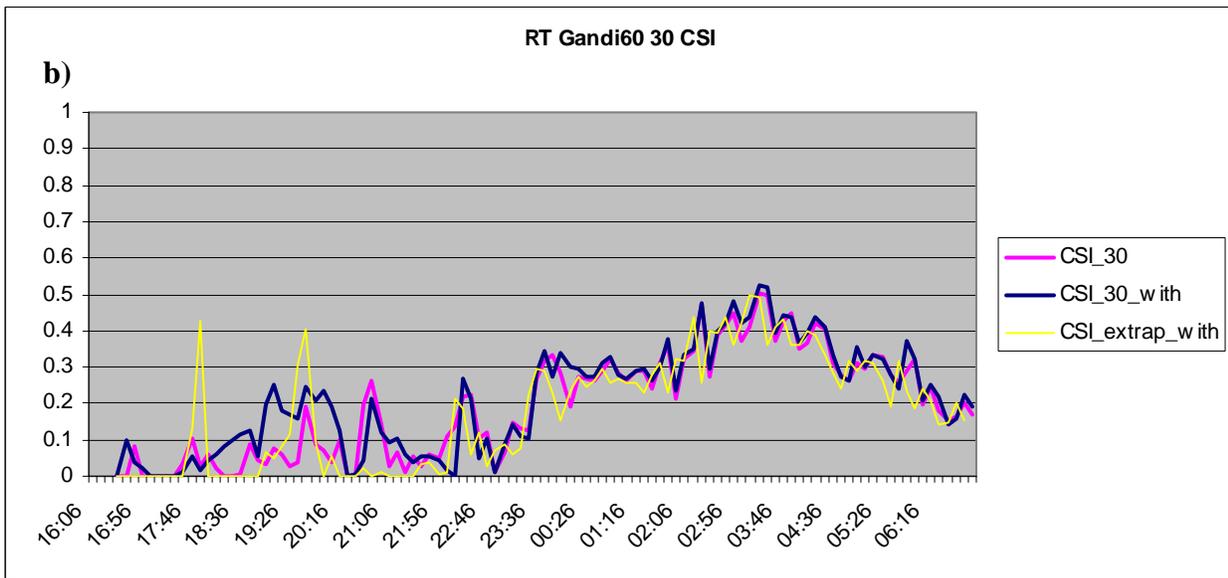
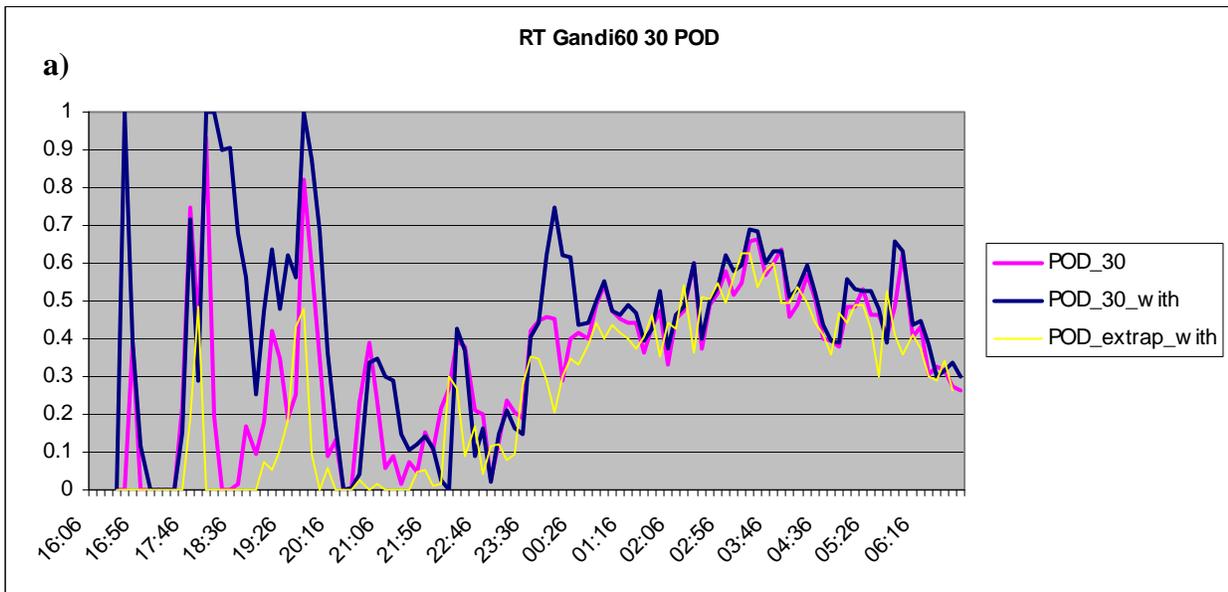


Figure 4. Standard statistics computed over the full validation domain and for the full period of convective weather on 30 March 2006. a) Probability of Detection (POD) scores; b) Critical Success Index (CSI) scores. See text for explanation of the different curves.

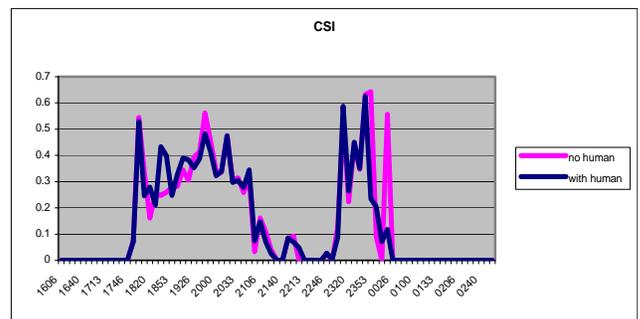
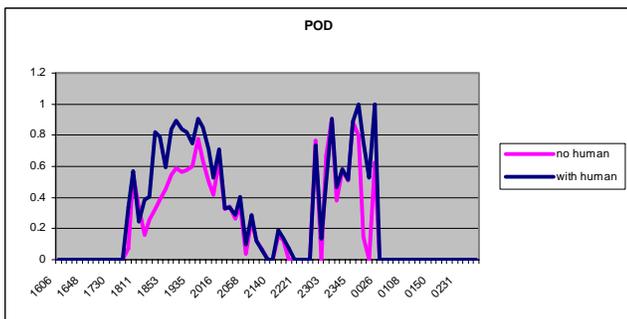


Figure 5. POD and CSI scores for sub-grid Box 3,5 (see Fig. 2b for box location).

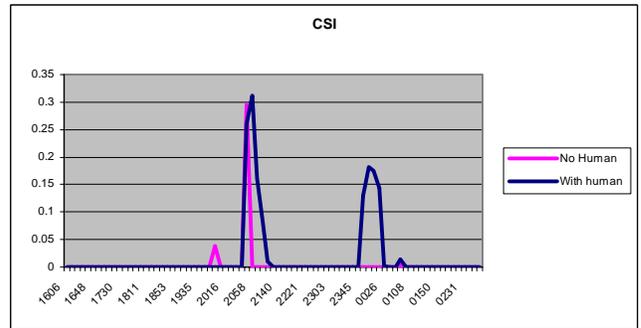
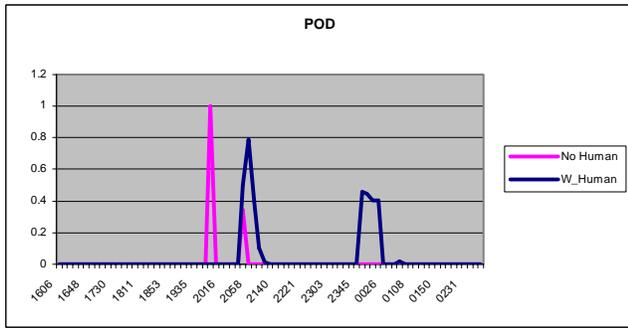


Figure 6. Same as Fig. 5 but for Box 2,4.

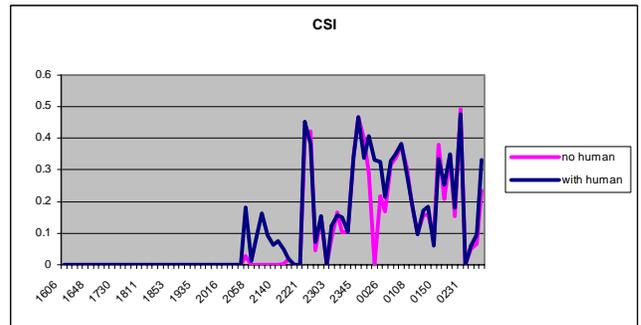
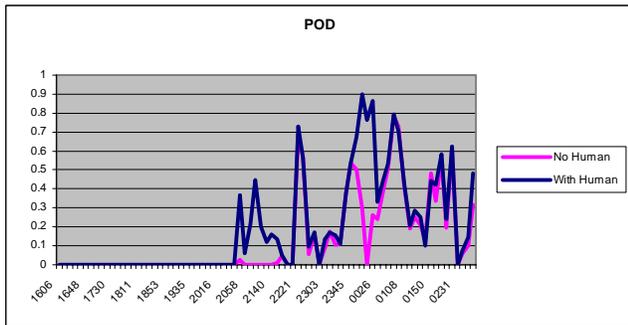


Figure 7. Same as Fig. 5 but for Box 3,4.

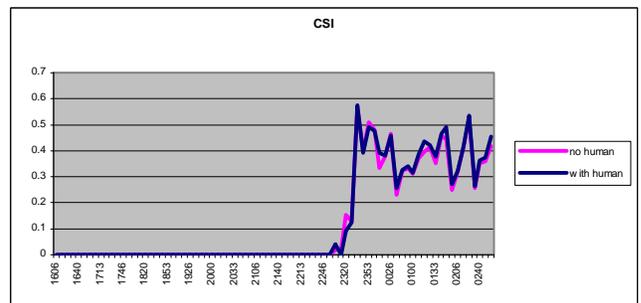
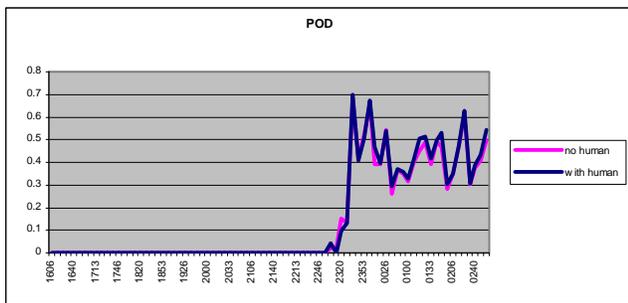


Figure 8. Same as Fig. 5 but for Box 4,4.

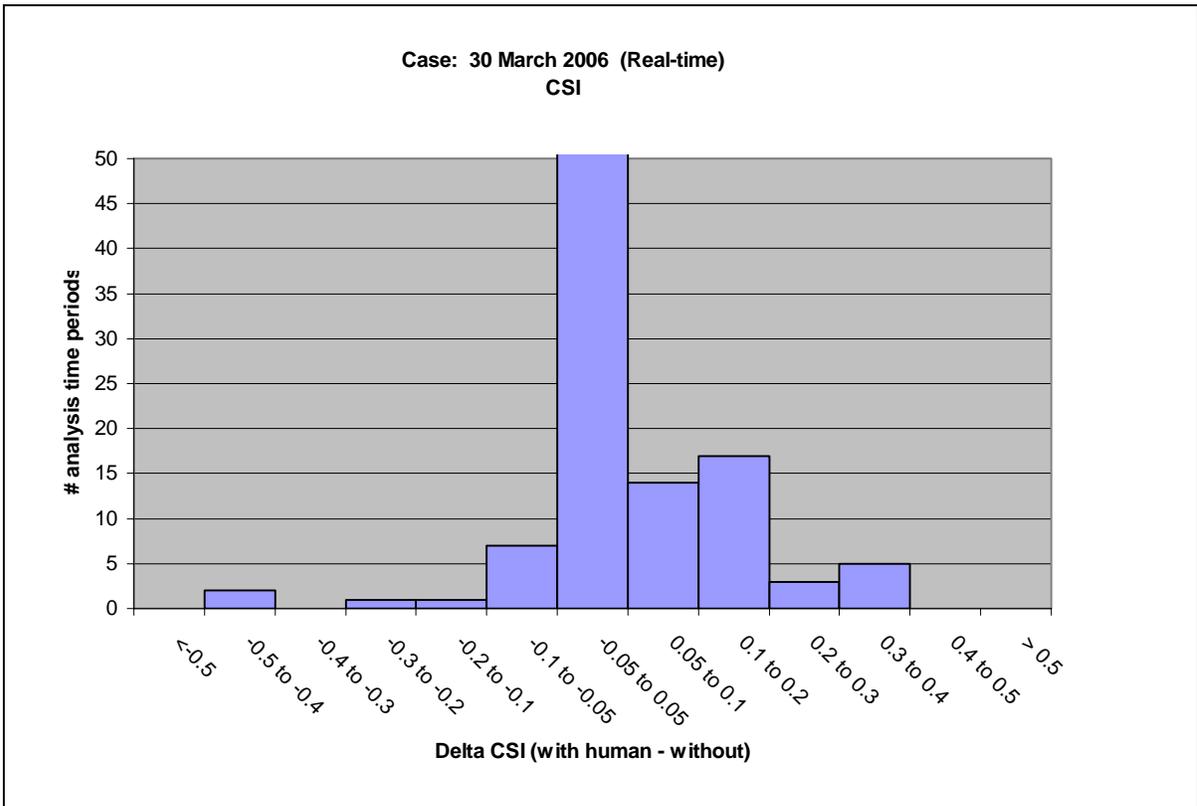
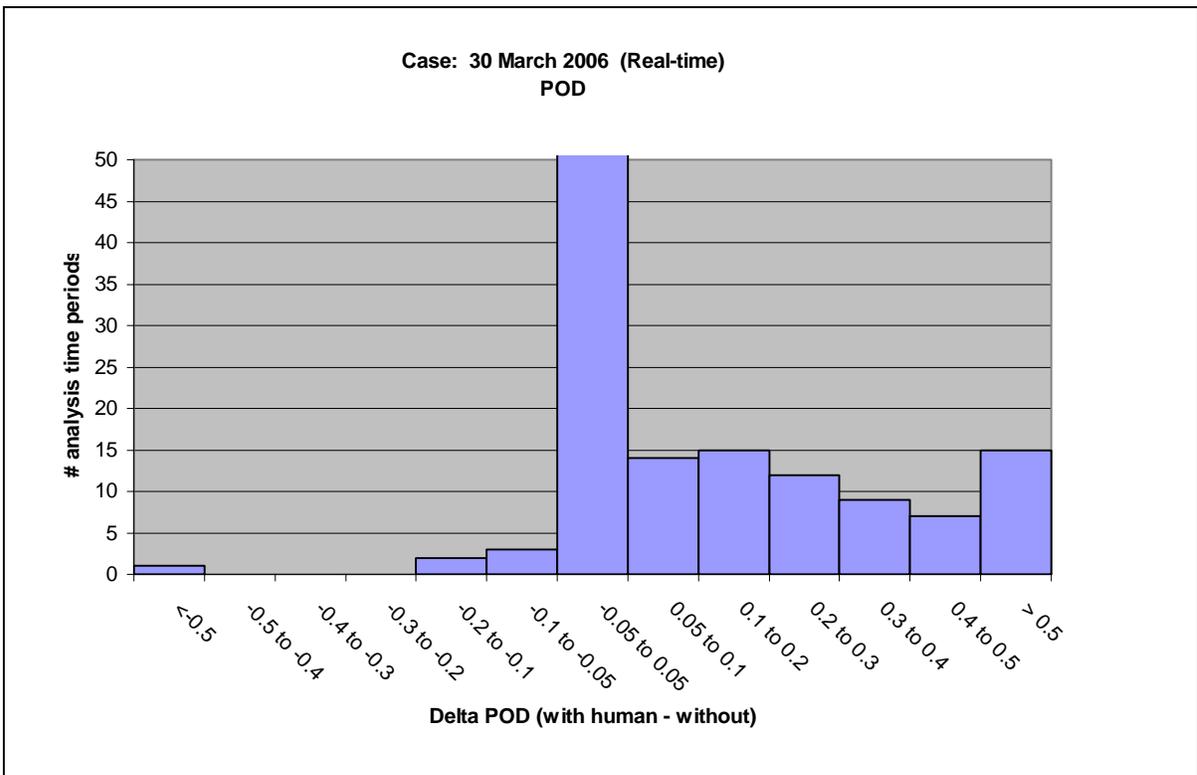


Figure 9. Summary histogram plots of the change in POD and CSI storm initiation scores when comparing nowcasts made with and without human input into the ANC system. Histograms are combined results of the change in POD and CSI scores computed over all the sub-grid boxes.

statistics in Fig. 4 for the same time period that show very little evidence of the increased accuracy of the storm initiation nowcasts when forecaster-entered boundaries are included in the ANC system. Expanding upon these results an additional step was taken to combine the statistics from the individual sub-grid boxes over the duration of the event (14 hrs of data for the 30 March case) to evaluate the *overall* performance in storm initiation nowcasting when forecaster-entered boundaries are used in determining storm initiation. Changes in POD and CSI are plotted in Figure 9. Positive change represents increase in accuracy of nowcasts when forecaster input is included. Negative changes indicate the forecaster-input has decreased the accuracy of storm initiation nowcasts. *Figure 9 graphically illustrates that the use of forecaster-entered boundaries in the forecast process leads to increased POD of storm initiation and increase in overall accuracy of CSI.* Equally important is the large portion of statistics clustered at zero change in the histogram plot. These statistics represent to a large extent all the “no” categorical skill scores; i.e., the nowcasts for “no” storm initiation that validate. The high proportion of values clustered near zero, along with the very low percentage of negative change also reinforce that the role of the human in the nowcast process does little to no harm.

Although the 30 March event is only one event, analyses of sub-grid box statistics for several other days and events from the Forecaster-Over-the-Loop demonstration show similar results.

4. WSI DEMONSTRATION

The WSI forecaster interacted with the ANC system in an operational environment between 0900 and 1800 EDT, Monday through Friday. While NWS forecasters interacted with the ANC system in addition to their many other operational duties, the forecasters at WSI were dedicated solely to supporting the ANC demonstration during the summer of 2006. Forecaster interactions included entry of significant boundaries (large outflows, lake breeze fronts, cold/warm/stationary fronts) and a capability to “nudge” interest or likelihood values when forecaster intuition indicated that the ANC nowcasts were running too “warm” or too “cold” (i.e., over- or under-alerting throughout the domain). The domain for this demonstration is shown as a red outline in Fig. 10 and similar to the NWS demonstration, storm initiation nowcast performance was difficult to

evaluate over the large domain. Thus, a second approach for validating the accuracy storm initiation nowcasts with and without human-input was explored.

4.1 Methodology for Statistical Evaluation

The storm forecast results (which include initiation and extrapolation) were evaluated using a boundary-relative verification methodology. That is, only regions within approximately +/- 100 km of a human-entered boundary were verified against the mosaic of WSR-88 reflectivity (see white polygon region for the example shown in Fig. 10). Thus, analogous to the approach used for the NWS, a sub-domain is employed to compute validation statistics, but this sub-domain encloses only the area of the forecast domain impacted by a forecaster-entered boundary over the duration of the event. Validation statistics are therefore computed over the sub-domains that bound the track of each boundary.

The verification statistics were computed for storm nowcasts without (Fig. 10a) and with (Fig. 10b) human-input. In this example, storm initiation nowcast products are rendered in shades of blue rather than bounded regions defined by white polygons (as in Figs. 2 and 3), but the process for producing the nowcasts is identical. The forecaster-entered boundaries (yellow polylines) are overlaid in Fig. 10b. The difference in storm initiation nowcasts without and with forecaster input is evident by comparing the blue shaded regions in Fig. 10a with Fig. 10b. In particular, the new convection that occurred in the southeastern portion of the domain was well forecast due to the forecaster-entered boundary that contributed to the storm initiation nowcast in that area.

4.2 Results from the WSI demonstration

The storm initiation field in Fig. 10 is verified by comparing the 35 dBZ threshold in the WSR-88D data to three likelihood levels (given by 25, 30 and 32 dBZ) which indicate slight, moderate, and high risk of storm initiation in 60 min. Figure 11 gives the cumulative distribution of differences in CSI skill scores between human and no-human interaction for 3 randomly selected boundaries on 7 June 2006. The different colors indicate the three storm likelihood levels. The first thing of note is that the skill is basically unchanged (most of the differences in CSI occur around zero) throughout most of the verification domain shown in Fig. 10. There is also a significant number of grid points where the skill has changed with improvements in

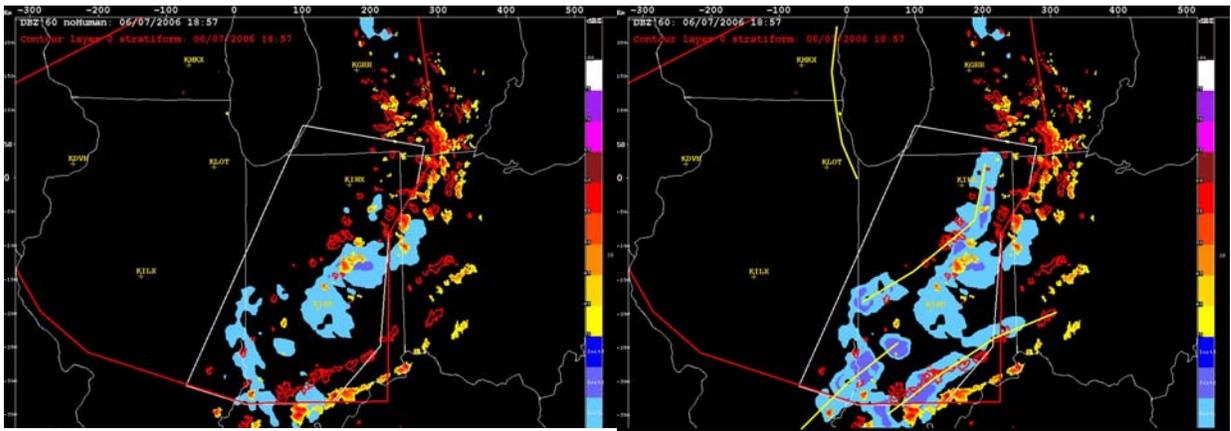


Figure 10. The 60 min storm initiation nowcasts without (left panel) and with (right panel) forecaster-entered boundaries. Blue shades represent likelihood for storm initiation (light blue – initiation possible, dark blue – initiation likely). Orange-red shades indicate extrapolated positions of existing storms. The red contours depict the verifying positions of the storms. Yellow polylines are forecaster-entered boundaries. White polygon represents sub-domain used for validation statistics.

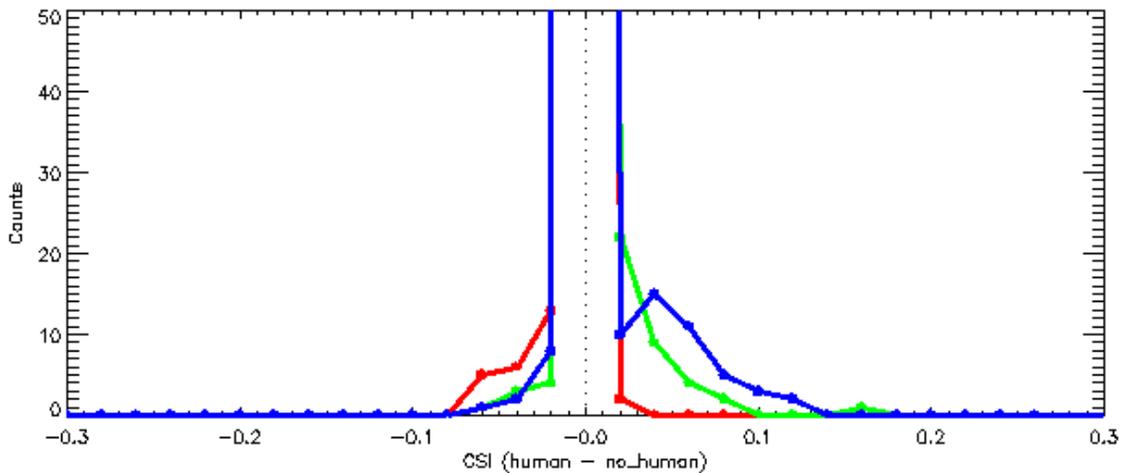


Figure 11. Distribution of CSI differences indicating the change in CSI skill score when comparing forecasts with human-entered boundaries with those without boundaries. Initiation likelihood levels of slight (red), moderate (blue), great (green) were verified.

CSI skill scores or forecast accuracy exceeding degradation of forecasts, indicating better forecasts overall. In this example, nearly all the improvement occurs at the higher likelihood thresholds (green and blue curves) with slightly reduced skill at the lowest likelihood threshold (red curve). The reduced skill at the lowest storm initiation level is the result of increased false alarms (not shown) which is expected in the prediction of thunderstorm initiation events that are so highly intermittent in space and time.

5. CONCLUDING REMARKS

Two new approaches to statistical validation of storm initiation nowcasts have been demonstrated on data collected during two Auto-nowcaster demonstrations conducted last summer. Storm initiation nowcasts have been made by incorporating a variety of observational and numerical model output fields into a fuzzy logic system to produce 60 min nowcast fields. The most recent improvement to storm initiation nowcast accuracy, as demonstrated with the

statistical results here, has been the inclusion of forecaster input (convergence boundaries) into the process. However, based on the generally low CSI values (typically < 0.5) shown here and in other studies, it is clear that there is plenty of room for additional improvement. The question is, whether there are current forecast parameters or observations now available that can lead to further improvement in short term nowcasting of thunderstorms? We explore a few of these possibilities below.

Most short term nowcasting systems have been designed to forecast surface-triggered convection. As mentioned above, surface convergence boundaries are a critical component in thunderstorm formation. However, several studies have shown that elevated convection accounts for ~50% of convective storms that occur, with approximately 60% of the storms that occur at night being triggered by elevated processes. One step toward improvement of the short term nowcasts will be to include additional fields and fuzzy logic for nowcasting convection that is not surface-based (Cai et al, 2007).

Radar climatology studies that illustrate preferred regions for storm initiation (Saxen et al, 2007; Wilson et al 2007) and the pre-disposition for storms to form at selected times during the diurnal cycle (Pinto et al, 2005) are factors that are known to be important but are difficult to incorporate with high reliability into short term, grid-specific nowcasting systems. A recent concept is to make better use of climatology information of preferred storm initiation locations and frequency, terrain impacts, preferred stability profiles and low-level wind direction over areas of heightened interest, rather than over a full, grid-to-grid specific domain.

Making use of new or enhanced observational datasets should likely contribute to improvement in thunderstorm nowcasts as well. This includes 1) enhancements of our national network of surface mesonetworks, 2) use of microphysical and storm evolution information obtained from satellite-derived products made by EUMETSAT and from the future GOES-R series of IR channels, 3) use of TDWR, X-band and CloudSat radar data for extension of coverage of clear air, cumulus cloud, and storm information obtained from the WSR-88D network of radars, and 4) incorporation of low-level moisture (refractivity) fields that can be retrieved from WSR-88D radars (Roberts et al., 2006).

Perhaps the promising prospects for improving short term thunderstorm nowcasts may be realized from numerical modeling efforts, as grid resolutions continue to increase to handle convective scale processes, as explicit rather than parameterized cumulus convection is employed, and as 3D- and 4D-VAR systems assimilate not only reflectivity and velocity, but moisture fields as well, to provide improved convergence and vertical velocity detection and forecast fields.

As the performance of the numerical models improve, it is anticipated that the blending of observations with model output (Wilson et Xu, 2006) will be one of the tools not only for improving nowcast accuracy but extending the prediction of storm initiation out beyond 1-2 hr.

6. REFERENCES

Cai, H., R. Roberts, D. Megenhardt, E., Nelson, and M. Steiner, 2007: Tackling the challenge of nowcasting elevated convection. *Preprints, 33rd Conf. on Radar Meteor.*, AMS, Cairns, Australia.

Nelson, E., R. Roberts, T. Saxen, and H. Cai, 2006: The NWS/NCAR "Forecaster-Over-The-Loop" Fort Worth Operational Demonstration: Human enhancement of a thunderstorm nowcasting system. *5th International Conference on Mesoscale Meteorology and Typhoon*, Boulder, CO,

Mueller, C., T. Saxen, R. Roberts, J. Wilson, T. Betancourt, S. Dettling, N. Oien, and J. Yee, 2003: NCAR Auto-Nowcast System. *Wea. and Forecasting*, **18**, 545-561.

Roberts, R., and S. Rutledge, 2003: Nowcasting storm initiation and growth using GOES-8 and WSR-88D data. *Wea. and Forecasting*, **18**, 562-584.

Roberts, R., and co-authors, 2005: The NWS/NCAR Man-in-the-Loop (MITL) nowcasting demonstration: forecaster input into a thunderstorm nowcasting system. *Preprints, WWRP Symposium on Nowcasting and Very Short Range Forecasting*, Toulouse, France.

Roberts, R., F. Fabry, N. Rehak, P. Kennedy, J. Fritz, D. Brunkow, J. Wilson, L. Mooney, T. Crum and E. Nelson 2006: The Colorado REFRACTT demonstrations. *Preprints, 4th European Conf. on*

Radar in Meteor. And Hydrology, Barcelona, Spain.

Saxen, T., R. Roberts, H. Cai, E. Nelson, and D. Breed, 2006: Recent and planned to the NCAR Auto-Nowcast System. *Preprints, 5th International Conference on Mesoscale Meteorology and Typhoon*. Boulder, CO.

Saxen, T., and co-authors, 2007: The operational mesogamma-scale analysis and forecast system of the U. S. Army Test and Evaluation Command. *Submitted to J. of Applied Meteor.*

Wilson J., and M. Xu, 2006: Experiments in blending radar echo extrapolations and NWP for nowcasting convection. *Preprints, 4th European Conf. on Radar in Meteor. And Hydrology*, Barcelona, Spain.

Wilson, J., M-X Chen, Y-C Wang, L. Wang, 2007: Nowcasting thunderstorms for the 2008 summer Olympics. *Preprints, 33rd Conf. on Radar Meteorology*, AMS, Cairns, Australia.