USE OF A MODIFIED EBERT-MCBRIDE TECHNIQUE TO VERIFY IHOP QPF AS A FUNCTION OF CONVECTIVE SYSTEM MORPHOLOGY

Jeremy S. Grams, William A. Gallus, Jr., Linda S. Wharton, Steven Koch, Elizabeth E. Ebert, and Andrew Loughe+

[⁺]NOAA - Forecast Systems Laboratory, Boulder, Colorado [#]Bureau of Meteorology Research Center, Melbourne, Victoria, Australia

1. INTRODUCTION

J13.4

Summertime convective systems are among the most difficult weather events for operational meteorologists and numerical models to predict. The International H₂O Project (IHOP) that took place from 09 May to 26 June 2002 was designed to help improve the understanding and prediction of this convection. Highresolution model datasets produced for this project offered the opportunity to investigate precipitation forecast accuracy as a function of convective system The object-oriented Ebert-McBride morphology. Technique (Ebert and McBride, 2000) for verification of gridded quantitative precipitation forecasts (QPFs) was used as the principal driver for this study.

The analysis of Ebert-McBride Technique (EMT) objective verification measures in combination with a morphological classification reveals whether certain types of systems, such as linear convective systems with trailing stratiform rain areas, were better predicted than other types within the models. The overriding goal of this study was to use these objective measures to improve our understanding of the relationship between convective system morphology and precipitation forecast accuracy.

2. BACKGROUND

The EMT was used to evaluate the performance of the 12-km Eta model and a 12-km diabatically initialized ("hot start") version of the MM5 model run during IHOP. A detailed radar-based morphological analysis of observed systems was performed for all contiguous rain areas (CRAs) identified in the IHOP domain. This was done for the first 6 hours of each model run available during the IHOP period.

The first stage of classification involved using the observed system highlighted in the stage IV 6-h accumulated precipitation grid (Fig. 1). Next, the observed CRA system was cross-referenced with an observed system indicated in radar observations.

The radar-based system classification was based on mesoscale convective system (MCS) characteristics. The radar based morphology used 2 km NIDS composite base reflectivity radar images with a temporal resolution of 30 minutes.

In order to classify each observed system identified by the EMT, objective criteria for MCS types had to be defined. The criteria for the MCS types were based mainly on previous literature dealing with MCS classification (e.g., Bluestein and Jain 1985; Parker and Johnson 2000).

In our survey, we defined a radar-based MCS as a convective system containing continuous or discontinuous convective echoes that propagated and/or organized in nearly the same manner as other convective echoes within the system. We required the MCS to have at least 30 dBZ of base radar reflectivity over at least a 100 by 100 km area and at least 40 dBZ in a 50 by 50 km area. Both dBZ conditions had to exhibit temporal continuity of at least three hours. Using the Z-R relationship of Z = $300 \times R^{1.4}$, approximate rain rates can be calculated for these thresholds. In the case of 30 dBZ, a rain rate of around 0.10 in/hr is assumed. For 40 dBZ, the rate is around 0.50 in/hr.

The following section describes the definitions used for the observed system classification.

2.1 Linear vs. Non-Linear Classification

The first series of classifications began with a distinction between linear and non-linear systems for those meeting the MCS criteria. Since not all observed systems identified by the EMT met our radar-based MCS criteria, separate categories had to be made for these "marginal" systems. The classification scheme included seven general types of systems:

a. Continuous Linear (CL)

A nearly continuous major axis of at least 40 dBZ convective echoes, of at least 100 km length, which share a nearly common leading edge and move approximately in tandem. In addition, the major axis must be at least 3 times as long as the minor axis.

b. Continuous Linear Bowing (CLB)

In addition to the linear criteria above, a linear bowing system must contain a bulging, convex shape of continuous convective cells with a tight reflectivity

^{*} Corresponding author address: Jeremy S. Grams, lowa State University, Dept. of Geological and Atmospheric Sciences, Ames, IA 50011; e-mail: jgrams@iastate.edu



FIG. 1. Example of CRA output from the Ebert-McBride Technique. Forecast of rain above 0.25 inch threshold is outlined in purple in upper left. Displacement vector (in red) shifts forecast area to the northeast. Observed rain above 0.25 inch threshold is outlined in purple in lower left, with the shifted forecast overlaid in magenta. Outer purple box shows the area over which CRA statistics (shown to the right) were calculated.

gradient on the front edge of the convective region. This shape must be identified for at least 1.5 hours on radar.

c. Continuous Non-Linear (CNL)

If the criteria for a CL or CLB are not met, but the system contains a nearly contiguous region of echoes that satisfy the minimum size criteria for an MCS, then the CNL classification is given.

d. Discontinuous Areal (DA)

If the above minimum size MCS requirements are not met in a continuous area but are met in an area of discrete convective elements, in which no element is separated by more than 200 km from another, then the DA classification is given.

e. Isolated Cells (IC)

If discrete cells are too small, isolated, or lack temporal continuity to meet the above classification, but have at least 40 dBZ in a 20 by 20 km area and at least 30 dBZ in a 40 by 40 km region, then the IC classification is given. It is well-understood that a 12 km model cannot fully resolve isolated cell events.

f. Nothing Produced (NP)

If none of the above criteria are met, then the observed system in the CRA is classified as nothing produced.

g. Orographically Fixed (OF)

The IHOP domain encompassed the eastern slopes of the Rocky Mountains on its western boundary, and the Black Hills on its northern boundary. Thus, a separate category was made for those systems that remained nearly stationary over an orographic feature.

In the 6-h CRA window, multiple radar-based systems might be observed within one larger CRA, since the EM technique uses the union of accumulated observed precipitation and forecasted precipitation. When this was the case, the system with the greater temporal, spatial, and/or rain volume was classified as the dominant type. Likewise, an evolution of types can exist within one system. The type that occurred the majority of the time was classified as the dominant one.

2.2 Additional Linear Classifications

For every linear type system (either CL or CLB), additional sets of classifications were performed using the taxonomy proposed by Parker and Johnson (2000) and Bluestein and Jain (1985). First, the arrangement of stratiform rainfall with respect to the convective region was classified using the definitions given by Parker and Johnson (2000). Second, a classification was made based on Bluestein and Jain's (1985) four definitions for squall line development.

a. Stratiform Classification

Parker and Johnson (2000) defined three areas where stratiform precipitation is present with respect to convective precipitation in an MCS. Their three categories: trailing (TS), leading (LS), and parallel (PS), were used in this study. Combinations of these types were noted, when both were seen for at least 1.5 hours.

b. Development Classification

Bluestein and Jain (1985) defined four types of development for a squall line MCS. Their four categories: broken areal (BA), broken line (BL), back building (BB), and embedded areal (EA) were used in this study.

3. MODIFIED EBERT-MCBRIDE PARAMETERS

The main purpose of this study was to find out how well the MM5 and Eta predicted systems as a function of morphology. This was accomplished by taking the averages of statistical parameters produced by the EMT. A brief description of the modified EMT is provided below.

3.1 Overview of Ebert-McBride Technique

The Ebert-McBride Technique (EMT) uses contiguous rainfall areas (CRAs) as a way to determine error statistics. These CRAs are made up of the union of observed and forecasted rainfall areas, which exceed a user-specified threshold amount. A displacement vector is then determined by shifting the forecast area to either maximize correlation coefficient or minimize root mean square error. The forecast is permitted to shift within an expanded CRA box whose size is controlled by the user. After the displacement is computed, various measures of error can be determined.

3.2 Improvements to the Ebert-McBride Technique

Since the EMT was originally used in Australia with rainfall systems evaluated on a time scale of one day, the technique needed to be modified for use on the mesoscale convective systems observed over the IHOP region, as described below.

a) The percentage of grid points allowed to shift off the gridded domain was reduced from 50% to 0.1% for displacement calculations involving root mean square error (RMSE) minimization. This fixed a problem where some forecasted rain areas were shifted off the grid, away from the observed area, in an attempt to minimize RMSE. For our study, we also found that maximizing correlation coefficient to determine where the forecast should shift, produced more reasonable results than minimization of RMSE (Fig. 2). For this reason, we permitted a portion of the forecast precipitation area (up to 25%) to shift off the grid in some cases where the forecasted rainfall was near the edge of the grid. This resulted in displacements that better agreed with subjective impressions.

b) When using correlation coefficient maximization instead of RMSE minimization, we found that negative displacement errors could occur with the original EMT error decomposition, which compared original and shifted RMSE values. In these cases, pattern errors also exceeded 100%. We developed a new error decomposition process using correlation coefficient terms, based on Murphy (1995). This process replaced the three-term EMT technique, in which the total error is decomposed into displacement, volume, and pattern errors. The "conditional bias" quantity appearing in Fig. 1 is an additional term derived from Murphy's scheme.

c) The critical mass threshold defines a minimum volume of rainfall necessary for a system to be identified by the EMT. This threshold was originally used for 24-h periods. Since our study focused on the first six hours of a model forecast, we reduced the critical mass threshold by a factor of four. This corresponds to a combined forecast and observed system producing at least 0.25" of rain in six hours over a 200 by 200 km area. In the complete absence of a forecasted system, this threshold will allow the EMT to identify observed systems roughly on the same order as our minimum MCS criteria.

d) Since mesoscale models at best only resolve 6 delta-x features, a low-pass filter was applied to the stage IV observed rainfall accumulation grid prior to application of the EMT. The filtering created patterns that were more similar to those produced by the 12 km grid spacing models.

e) The only change made to the code for statistical measures was the determination of volume and rain rate over just the forecast or observed portion of the CRA. The rest of the CRA statistics were computed over the union of observed, forecasted, and shifted forecast of rain, as done in the original EMT.

f) The CRA critical threshold of 0.25" occasionally resulted in multiple observed systems being identified as one CRA. This problem occurred more frequently in the MM5 compared to the Eta. Likewise, point-to-point verification statistics showed that the MM5 had a much higher wet bias at the 0.25" threshold than the Eta. Therefore, for the MM5 forecasts, the threshold to identify systems was raised to 0.50". The change resulted in a more reasonable division of CRAs in each time period. For application of the EMT with both the Eta and MM5 runs, a rainfall threshold of 0.25" was used to define observed systems.



FIG. 2. Example of the CRA displacement vector calculation for either minimizing RMSE (left image) or maximizing correlation coefficient (right image). Note how the forecast shifted realistically to the observed area when maximizing correlation coefficient.

In addition, a bug in the MM5 Hot Start procedure during the early portion of the IHOP period resulted in very heavy amounts of QPF produced by the model relative to observations. A change to the MM5 code was implemented on 25 May 2002, which attempted to correct this bug. This change subsequently resulted in more reasonable forecasts. As a result, the MM5 statistics discussed below were divided into a pre-MM5 period (prior to 25 May), and a post-MM5 period (25 May to the end of the project).

4. OBSERVED MCS TYPE DISTRIBUTIONS

A total of 234 CRAs were identified in the evaluation of the MM5 model during the entire IHOP period, compared to 209 in the Eta. For all of these CRAs identified by the EMT, the observed system within that region was classified as described earlier using 2 km NIDS radar data. The focus of the examination to follow is on model performance as a function of the observed system morphology.

Of the CRAs identified by the EMT, 27 (12%) of the MM5 CRAs did not satisfy any observed system criteria and were classified as nothing produced. 16 (8%) of the Eta CRAs were classified as nothing produced. Nothing produced is equivalent to a "false alarm" model forecast, with little, if any, observed rainfall. The MM5 had slightly more instances of this happening compared to the Eta. A small amount of the observed systems, 9 (4%) in the MM5 CRAs and 7 (3%) in the Eta CRAs were classified as orographically fixed. These systems were excluded from further analysis in this study.

Next, looking at just the systems where our radarbased criteria were met over the Plains, a division can be made between linear and non-linear systems. Linear encompassed both continuous linear and continuous linear bowing systems. Non-linear included continuous non-linear, discontinuous areal, and isolated cells systems. In the MM5 CRAs, 67 (34%) observed cases were classified in the linear division, with 131 (66%) as non-linear. The Eta CRAs also had 34% of the observed cases as linear, and 66% as non-linear. Thus, both models' CRAs had approximately twice as many observed non-linear systems as linear systems.



FIG. 3. Histogram of observed systems (general types, stratiform types, and development types) for the combined model (MM5 and Eta) count of CRAs identified.

Figure 3 shows the combined model (MM5 and Eta) CRA totals for all observed systems. 111 (85%) were continuous linear, with far fewer, 19 (15%), as continuous linear bowing. Non-linear counts were led by the continuous non-linear category with 121 (48%), followed by discontinuous areal with 96 (38%). Isolated cells only had 37 (14%).

For the stratiform types, trailing stratiform dominated with 89 (68%). The trailing stratiform/parallel stratiform type (generally large systems since stratiform rain occurred in both regions) garnered the second highest total with 20 (15%). The categories of leading stratiform, parallel stratiform, and leading stratiform/parallel stratiform each had less than 10 occurrences. The small numbers in these classifications would likely prevent statistical significance of results, and thus we discuss them below as a combined nontrailing stratiform category with 21 (16%).

Among the development types, broken areal was the most common with 61 (47%) counts. Broken line had 38 (29%), with back building at 28 (22%). Embedded areal had only 3 (2%) and was therefore excluded from further study.

5. CRA STATISTICAL ANALYSIS

Statistics were calculated for the following parameters: rain volume, rain rate, maximum rainfall,

correlation coefficient, and mean square error decomposition. This analysis was performed for all observed systems over the Plains, where at least the isolated cells criteria were met. Model errors are then presented as a function of the observed morphology classification system derived above.

5.1 Rain Volume

Even with its model verification at the 0.50" threshold and an observed threshold of 0.25", the pre-MM5 was extremely wet. For all CRAs, precipitation forecasts were ~400% wetter than the observed systems. The bias was greater for non-linear systems than linear ones, with forecasts ~600% wetter than the observed. The post-MM5 forecasts were much closer to the observed volume. Its average rain volume for all systems, 0.765 km³, was slightly less than the 0.832 km³ for all observed systems. However, this "dry bias" would not have existed if the MM5 forecasts were verified on the 0.25" threshold. This should be kept in mind when interpreting the results below.

For the Eta, a small wet bias existed for all systems, with 1.116 km³ forecasted compared to 0.980 km³ observed. More importantly, both the Eta and post-MM5 showed a dry bias with linear systems and a wet bias with non-linear ones, as shown in Figure 4. It is possible that the dry bias with linear systems reflects the lack of transport of condensate away from more intense convective cells (not included in the convective schemes), a process known to be very important in the upscale growth of organized linear systems (e.g., Rutledge 1986). The biggest discrepancy came in the isolated cell category, where around a ten to one wet bias exists in the Eta; with around a five to one wet bias in the post-MM5. Note also that the sample size differed in the evaluation of each model, since the post-MM5 was evaluated from May 25th to the end of the IHOP period, whereas the Eta was evaluated over the entire IHOP period. This results in different amounts for the observed category shown in each graph. The large wet biases may imply the models are too eager to develop convection. In some ways, the isolated cell events are not too different from the nothing produced events large false alarm rates are present.

Stratiform types also show the dry bias in linear systems (Fig. 5). Both models are too dry for trailing stratiform. They are generally dry biased for the combination categories involving leading stratiform and parallel stratiform.



FIG. 4. Average rain volume (km³) for general types for the post-MM5 forecast (green) and observed (blue); Eta forecast (purple) and observed (grey).



FIG. 5. Average rain volume (km³) for stratiform types for the post-MM5 forecast (green) and observed (blue); Eta forecast (purple) and observed (grey).

Both models gave a mixed bag of results in development types (Fig. 6). The Eta did best with broken areal, but it had much more significant biases for the broken line and back building types. Meanwhile, the post-MM5 did best with back building, followed by modest dry biases for broken line and broken areal. The



FIG. 6. Average rain volume (km³) for development types for the post-MM5 forecast (green) and observed (blue); Eta forecast (purple) and observed (grey).

most obvious trend for development types was that both models have dry biases for linear systems forming as broken lines.

5.2 Rain Rate

The MM5 showed the ability to produce different average rates of rainfall for various systems (Fig. 7a). The post-MM5, when adjusted downward by 0.25" to account for its original verification on the 0.50" threshold, showed a slightly higher rain rate for continuous linear bowing than observed. It showed a fairly high rain rate for isolated cells systems, which had the lowest observed average rain rate.



FIG. 7a. Average rain rate (in/6-hr) for general types for the post-MM5 (green) and observed (blue).

Unlike the MM5, the Eta's forecasted average rain rate (for all CRA points above the 0.25" threshold), was lower than observed for every classification category, except isolated cells (Fig. 7b). It also showed consistency with producing nearly the same rain rate for all general types. This may imply the model doesn't have the capability to differentiate its rate of rainfall for highly efficient precipitation systems, like bowing lines, versus those with low efficiency, such as isolated cells. Gallus (1999) showed that the Eta with the Betts-Miller-Janjic (BMJ) convective scheme (as in the present study) was also fairly insensitive to changes in horizontal resolution. The BMJ scheme seems is so aggressive at drying the atmosphere that small-scale structures more likely to be produced in the gridresolved component of the rainfall are often eliminated. Operational forecasters have long noted that the rainfall forecasts from the Eta appear to be overly smooth and lacking of fine-scale structure. The current analysis agrees with those observations.

Stratiform types and development types were also fairly homogeneous for rain rates in the Eta. The post-MM5 exhibited some subtle differences in these divisions. For stratiform types, the post-MM5 slightly underforecasted rain rate for trailing stratiform and trailing stratiform/parallel stratiform. It had a slight wet bias for the leading stratiform, parallel stratiform, and leading stratiform/parallel stratiform combination group. In development types, it underforecasted both broken areal and broken line development; while it overforecasted the back building type.



FIG. 7b. Average rain rate (in/6-hr) for general types for the Eta (purple) and observed (grey).

5.3 Maximum Rainfall

Maximum rainfall was the highest observed amount of precipitation in the models' 12km grid and in the filtered stage IV observed accumulation grid.

The post-MM5 overpredicted rainfall maxima for all general, squall, and development types, except for continuous non-linear in the general category (Fig. 8a). This overprediction was clearly seen in both the continuous linear bowing and isolated cells. Here predicted rainfall maximum were ~200% and ~300% greater than observed, respectively.

Conversely, the Eta underpredicted rainfall maxima for all general, squall, and development types, except for isolated cells in the general category (Fig, 8b). In isolated cells, its prediction of the maximum was very close to the observed. Just like in the average rain rate category, the Eta was very homogeneous in its distribution of maximum rain rate for each category. This led to the observed maximum rain rate being ~300% greater than forecasted in the continuous linear bowing type. Again, the tendency of the Eta to have far smaller maximum rain rates than observed agrees with Gallus's (1999) paper which showed that the use of the BMJ scheme prevented large rainfall amounts from occurring with fine grid resolution. When the Kain-Fritsch scheme was used instead, much larger rain rates were noted. In Gallus (1999), the maximum rain rates occurred in regions with large grid-resolved rainfall components. An analysis of convective and grid-resolved rainfall components with the IHOP cases has not yet been done.



FIG. 8a. Average maximum rainfall (in) for general types for the post-MM5 forecast (green) and observed (blue).



FIG. 8b. Average maximum rainfall (in) for general types for the Eta forecast (purple) and observed (grey).

5.4 Correlation Coefficient

Both models showed a better average correlation coefficient for linear systems than non-linear ones. Correlation coefficient is sensitive to accurate depiction of substantial rainfall gradients, which generally occupy more of a system's area in nonlinear events. The Eta was just slightly higher than the post-MM5.

Out of the general types, both models were significantly worse with correlation in the isolated cells category (Fig. 9). This could be expected based on the overforecast of rain volumes, rain rates, and maximum rain amounts.

When looking at the improvement of correlation from the original forecast to the shifted forecast, for CRAs in which a displacement occurred, one can see common themes in both models (Fig. 10). Both have larger improvements for non-linear systems, than linear ones. In the general types, both models show the least improvement in the continuous linear-bowing. For the Eta, the improvement in the isolated cells category is particularly dominating. For the development types, both have their greatest improvement in back building systems, especially the post-MM5. Both models also have the least improvement in broken line systems, especially in the Eta.



FIG. 9. Average correlation coefficient (R^2) for general types for the post-MM5 (green) and Eta (purple).



FIG. 10. Average improvement of correlation coefficient (R^2) for general types for the post-MM5 (green) and Eta (purple) shifted CRAs.

5.5 Mean Square Error Decomposition

The total mean square error (MSE) was nearly twice as big in the post-MM5 (0.947 in^2) than the Eta (0.509 in^2) . The post-MM5 had greater errors than the Eta for all error decomposition terms (displacement, volume, and pattern) in both linear and non-linear divisions. More importantly, both models produced their greatest errors in linear systems, versus non-linear ones.

For general types, both models had their greatest total MSE in continuous linear bowing systems (Fig. 11). The isolated cells category was nearly equal in the amount of error in the post-MM5. However, the isolated cells category provided the least amount of total MSE in the Eta.

For stratiform types, both models have their greatest error in the leading stratiform, parallel stratiform, and leading stratiform/parallel stratiform

combination category (Fig. 12). The graphs suggest that when leading stratiform precipitation is present, the Eta had much higher MSE for that category compared to trailing or parallel stratiform.



FIG. 11. Average total MSE (in²) for general types for the post-MM5 (green) and Eta (purple).



FIG. 12. Average total MSE (in^2) for stratiform types for the post-MM5 (green) and Eta (purple).

6. CONCLUSIONS

The EMT was modified and applied to forecasts of convective system rainfall from the 12km Eta and MM5 models during IHOP. Several objective measures of skill and error were averaged for different morphological types. The EMT suggests that the Eta performs better than the post-MM5 (generally has better averages of error measures). However, systematic deficiencies appear in both models.

The peak amounts in the Eta were too light for every observed category except isolated cells. The post-MM5 was too heavy with the peak amount of rain forecasted in most categories. Average rain rates in the Eta were insensitive to system type, a result not agreeing with observations. Average rain rates were consistently too low. The deficiencies in the Eta model generally agree with routine observations from operational forecasters, and are probably more related to the use of the BMJ convective scheme than any deficiency in the model numerics.

In general, both models appeared to do the best with continuous non-linear and continuous linear systems. Part of this may be explained by the fact that these two groupings have the highest representation versus the other general types in our study. Both models did worse with isolated cells and continuous linear bowing systems. These represent opposite ends of the spectrum of MCSs and were represented the least amount in our data. These results confirm expectations that models have a tougher time with the systems that tend to be the biggest and smallest with respect to size and precipitation amount. Also of note, both models tended to be too dry with linear systems and too wet with nonlinear ones.

For stratiform types, the trailing stratiform/parallel stratiform combination appeared to be the best overall forecasted in the Eta. Trailing stratiform and the trailing stratiform/parallel stratiform combination appeared to be the best two categories in the post-MM5. This implied that the catch-all category of leading stratiform, parallel stratiform, and leading stratiform/parallel stratiform combination, was likely the worst forecasted type by the models. However, with such a small dataset to study, the statistical significance of leading stratiform and parallel stratiform as separate categories was not met.

The three development types showed mixed results with no clear cut conclusions. It did seem that back building is the one type that is somewhat harder for the models to forecast.

Future work with this dataset will include exploration of "why" these results were obtained. The statistical data should allow us to garner some fundamental understanding of the behavior of the Eta and MM5's representation of moisture processes versus the observed system morphology.

7. ACKNOWLEDGMENTS

This research was supported in part by NSF Grant 0226059 and by a grant to the Forecast Systems Laboratory from the U.S. Weather Research Program.

8. REFERENCES

- Bluestein, H. B., and M. H. Jain, 1985: Formation of mesoscale lines of precipitation: severe squall lines in Oklahoma during the spring. *J. Atmos. Sci.*, **42**, 1711-1732.
- Ebert, E. E., and J. L. McBride, 2000: Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrology*, **239**, 179-202.
- Gallus, W. A., Jr., 1999: Eta simulations of three extreme precipitation events: Impact of resolution and choice of convective parameterization. *Wea. Forecasting*, **14**, 405-426.
- Murphy, A. H., 1995: The coefficients of correlation and determination as measures of performance in forecast verification. *Wea. Forecasting*, **10**, 681-688.
- Parker, M. D., and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 3413-3436.
- Rutledge, S. A, 1986: A diagnostic numerical study of the stratiform region associated with a tropical squall line. *J. Atmos. Sci.*, **43**, 1337-1358.