

### 9.3 APPLYING A PROCRUSTES SHAPE ANALYSIS VERIFICATION SCHEME TO NOWCAST ENSEMBLE FORECAST MEMBERS TO DETERMINE ACCURACY OF DIAGNOSED CONVECTIVE MODE

Neil I. Fox<sup>1</sup> and Steven A. Lack<sup>2</sup>

<sup>1</sup>University of Missouri-Columbia, Columbia, Missouri

<sup>2</sup>NOAA Earth Systems Research Laboratory, Boulder, CO

Collaboration with the Cooperative Institute for Research in Environmental Sciences (CIRES), Boulder, CO

## 1. INTRODUCTION

As part of a nowcast development project a new verification scheme has been developed that is designed to be more informative with regard to the nature of errors found in an object-oriented method. This methodology assesses an overall penalty function that is composed of errors due to differences in size, shape, location, intensity and orientation of precipitation objects. The concept behind this approach is to provide detailed error assessments that give forecasters indications of the nature of problems with forecasts that can inform decision making and development.

The original version of the verification scheme is described in detail in Micheas et al. (2007). The Procrustes verification technique described originally in Micheas et al. (2007) was slightly modified for use with meteorological precipitation fields. The adjustments made from the original version of the Procrustes scheme are minor and reflect changes in the interpretation of errors only and not with the original methodology. The original penalty took into account only shape and intensity sum of squares errors. The new penalty function is a cell-by-cell-based function which includes shape, intensity, dilation, rotation, and translation, and attempts to bring each error component penalty to similar orders of magnitude. Once penalties for each matched set are compiled a mean squared error is assessed for all matches in a forecast domain. In the current framework if there are more cells in the forecast an additional penalty is assessed for false alarms.

The first half of this paper uses some artificial geometrically shaped pseudo-reflectivity images to illustrate the operation of the new version of the verification scheme. The second portion shows the scheme as applied to an ensemble of nowcasts of a single case of severe weather in the St. Louis, MO area.

## 2. METHODOLOGY

Simple nowcast ensembles were generated by varying the reflectivity band selection in the K-Means cluster advection scheme in WDSS –II (Lakshmanan et al. 2007). This has the effect of forecasting the motion of storm cells based on the prior motion of different size clusters defined by the reflectivity threshold. As there is a correlation between cell size and reflectivity this is akin to performing an object-based spatial cascade. When performed there are clear differences between the motion and shape of the forecast storms.

The first stage of the verification involves the identification of cells in both the observed and forecast fields. There are two options in the original scheme for matching objects: matching based on minimizing the differences in shape between objects and matching based on minimizing the distance between centroids. As the two methods are not yet combined, the centroid difference was chosen as it matches closely with other object-oriented approaches. The identification of cells is based on a user-defined minimum intensity threshold and minimum size of object threshold. Once cells have been identified in both domains matching commences based on minimizing the centroid distances between matched pairs. Each observed cell is matched to exactly one forecast cell. Matched objects are then transformed into a similar coordinate system to assess penalties based on rotation ( $SS_T$ ), dilation ( $SS_D$ ), translation ( $SS_T$ ), and overall shape ( $RSE$ ). Maximum ( $SS_{max}$ ), minimum ( $SS_{min}$ ), and average intensity ( $SS_{avg}$ ) penalties are also assessed. The combined penalty,  $D$ , for one matched pair of cells is given by (1).

$$D = RSS^{0.5} + SS_{min} + SS_{max} + SS_{avg} + (1 - SS_D) * 100 + SS_R * 100 + SS_T^{0.5} \quad (1)$$

An overall penalty (1) for the matched pair of cells is calculated by summing the squared error of the individual error components. Once the penalty per each matched cell is assessed they are combined in a mean squared error penalty (based on the number of observed cells ( $N_{obs}$ )) for all matched cells in a given domain,  $D_{domain}$ , (2).

---

\*Corresponding author address: Neil I. Fox, 332 ABNR Building, Univ. of Missouri, Dept. of Soil, Environmental, and Atmospheric Science, Columbia, MO, 65211; email: foxn@missouri.edu

$$D_{domain} = \frac{\sum D}{N_{obs}} \quad (2)$$

Finally, if there are more forecast objects than observed objects an additional penalty is assessed for false alarms based on the ratio of the amount of forecast cells to observed cells. The adjusted cell-based total penalty given each idealized case now makes more intuitive sense and using a cell-based penalty function allows the user to examine error characteristics of a particular cell of interest in the domain.

The above equation (1) attempts to bring all of the error components measured by the Procrustes shape analysis verification tool to a common order of magnitude. For example, the maximum rotation error for an individual cell is  $\pi/2$  (1.57) for a maximum squared error of 2.47. For an object that needs no resizing (dilation = 1), the  $SE_D$  term in (1) becomes 0. A large squared error for translation may be on the order of 10000 (100 km) outlining the need for adjustments to the cell-based error to get everything close to the same order of magnitude; hence multiplication factors on dilation and rotation and the square roots of some other squared error components in (1). A perfect forecast results in  $D=0$  in (1) for an individual cell,  $D$  may be considered a squared error term for a matched pair of cells; for more than one cell in a domain the total domain penalty would be a mean squared error.

It is now evident that objects that must be rotated and translated to a large extent have the highest total penalty, as well as those cases that differ in intensities.

### 3. GEOMETRIC CASES

Five geometric cases were used in this study with known error characteristics. Each case involves known displacement errors and include different sized objects and rotation issues. The output from the Procrustes verification tool will shown for each case including graphics produced with a table showing the breakdown of error components for each case (Table 1). Figure 1 shows the observed object (Geom000) that all subsequent objects are matched to. Two images are produced including the original image complete with intensities and an image showing the cell identification with zero intensity. In this experiment just one cell being identified for each case. The rest of the sample geometric cases will not be displayed herein.

Figure 2 shows a display of the Procrustes fit. The observed object's (Geom000) outline is in black while the forecasted object (Geom001) is in red. The overlay

(fit) is shown in blue and used to assess the residual (shape) error. It visually gives the user the goodness of fit. In Figure 2 the fit is nearly perfect.

The numerical output for the second case can be seen in Table 1. This case is similar to the first except that there is an additional translation error. The fits for the remaining three cases (Geom003, Geom004, and Geom005) are shown in Figures 3, 4 and 5. Of note is the third case (Geom003) as the fit is not perfect as an ellipse is being matched to a perfectly circular object. This results in a large penalty for the residual (shape) error and is shown in Table 1.

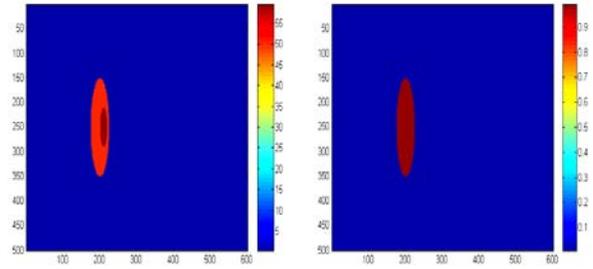


Figure 1: The observed object (Geom000) with intensity shown on the left and the object identification shown on the right.

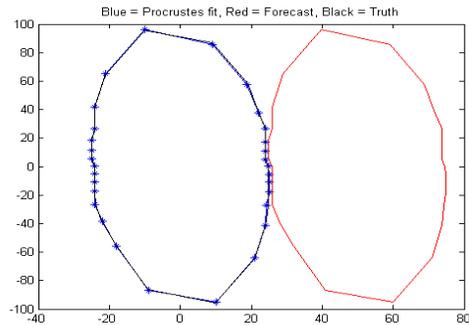
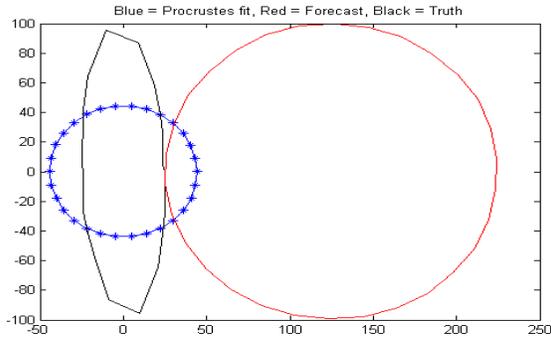
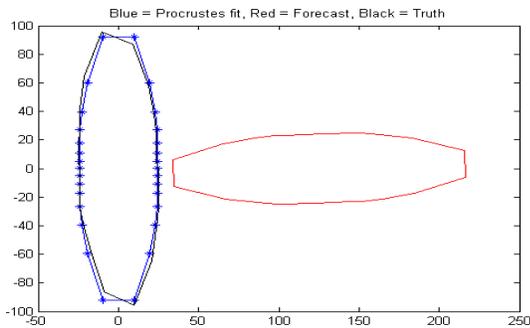


Figure 2: The Procrustes fit (blue) for Geom000 (black) and Geom001 (red). The user can see the goodness of fit as well as the translation error of the matched pair.

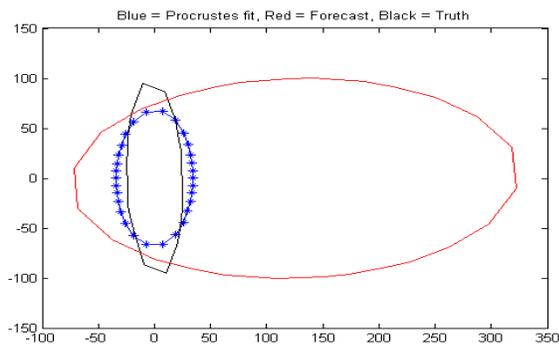


**Figure 3: The Procrustes fit for Geom000 to Geom003.**

Of note in Figures 4 and 5 is the 90° rotation present in Geom004 and Geom005. This is reflected in the rotation squared error in Table 1. The values are near 2.3 which corresponds to approximately around 1.5 radians. This represents the maximum attainable rotation error. This value is multiplied by 100 for the assessment of the final penalty for the matched pair.



**Figure 4: The Procrustes fit for Geom000 to Geom004.**



**Figure 5: The Procrustes fit for Geom000 to Geom005.**

Truth=Geom000	Geom001	Geom002	Geom003	Geom004	Geom005
Truth ID	1.000	1.000	1.000	1.000	1.000
Fore ID	1.000	1.000	1.000	1.000	1.000
Max Int Truth	100.000	100.000	100.000	100.000	100.000
Min Int SE	0.001	0.000	0.005	0.000	0.003
Max Int SE	0.000	0.000	0.000	0.000	0.000
Mean Int SE	0.000	0.000	1.113	2.791	1.503
Dilation SE	1.001	1.000	0.196	1.024	0.116
Rotation SE	0.000	0.000	0.015	2.359	2.306
Translation SE	2489.600	39932.000	6884.600	15783.000	5304.400
Residual SE (fit)	3.352	22.714	15197.000	190.060	4328.900
Cell Adj Error	51.787	204.630	289.270	380.410	459.090
Tot Error	3.353	22.714	15198.000	192.850	4330.400
Mean Adj. Error	51.787	204.630	289.270	380.410	459.090
Abs location errors					
Magnitude	50	200	125	125	125
Direction	90	90	90	90	90

**Table 1: The individual components with their associated squared error for each geometric object matched to a single truth (Geom000). Absolute location errors are also given which match to the description of the geometric cases.**

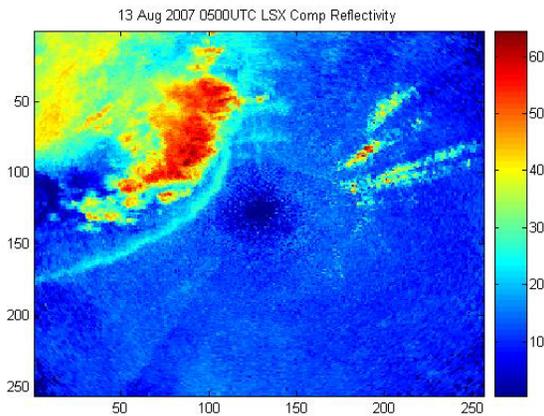
The errors presented in Table 1 readily show both the comparative skill of each “forecast”, but also how the error in each case is composed of different components. The interpretation of the overall errors are straightforward in these cases as there is only one cell being matched each time, and this means that the mean and cell adjusted errors are equal. One can clearly see how the error in shape (as opposed to translation or area that are typically the only errors quantified) can dominate, as shown in Geom003 and Geom005. This could be interpreted as a model poorly forecasting the mode or type of precipitation event thereby providing valuable information to the forecaster or developer. Similarly in Geom004 and Geom005 there are significant contributions to the overall error from the alignment of the cell and this is being quantified by the rotation error. Again this could signify some kind of error in storm morphology.

Overall, the usefulness of this particular verification scheme is readily apparent. It allows for the breakdown of error into components and the penalty function can be user-defined so that important errors become the dominant player in the end result. It also allows an assessment of the meaning of the magnitudes of errors produced in cases where the verification scheme uses an open-ended error scale.

### 3. NOWCAST ENSEMBLE CASE

The methodology for an alternate application for the WDSS-II nowcasting scheme is quite simple. The user can select different ranges of reflectivity on which to calculate the storm motion to generate the nowcast. The default storm motion calculation within the segmotion algorithm utilizes a range from 20 to 60 dBZ, which then gets divided into bins for storm motion calculation. By simply changing the lower and upper end of the range, the user can alter the bins, thus changing how the storm motion is calculated on different spatial scales. Systematically altering the range of reflectivity and running the nowcaster multiple times can generate an ensemble product by taking the mean of the solutions.

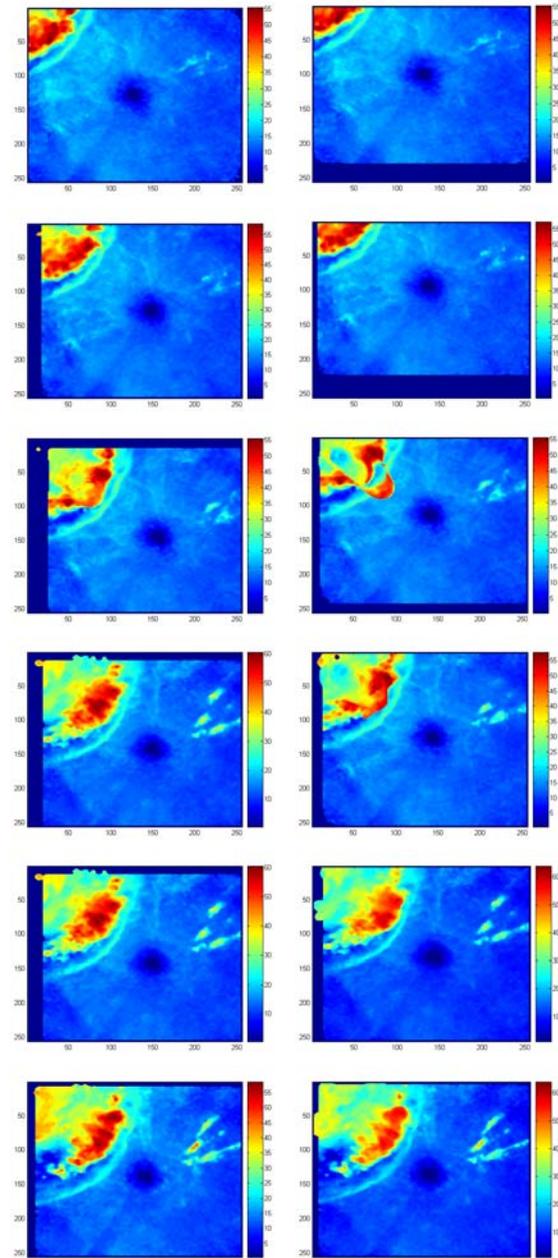
A sample case for this alternative WDSS-II K-means nowcast methodology is shown using data from 13 August 2007 centered over the St. Louis, MO (KLSX) radar. The case involved a severe linear convective system bow-echo event that passed through the St. Louis metropolitan area causing significant damage. For demonstration purposes this paper concentrates on a single time 0500 UTC and the nowcasts for that time for periods of up to 60 minutes prior to then. The actual radar reflectivity from that time is shown in figure 6.



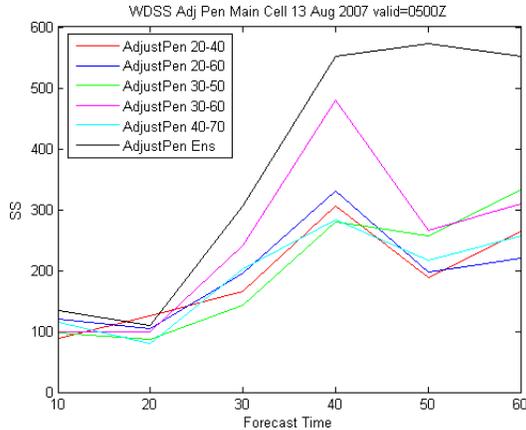
**Figure 6: Actual radar reflectivity image from 0500 UTC on 13 August 2007 from the St. Louis (LSX) radar.**

Forecasts for this time were made using segmotion ranges of 20-40 dBZ, 20-50 dBZ, 30-50 dBZ, 30-60 dBZ, and 40-70 dBZ, as well as an ensemble constructed as a mean of the five fields at each lead time. Two example sequences of nowcast reflectivity fields are shown in figure XX. These are from the 30-60 dBZ and the 40-70 dBZ ranges and show contrasting levels of success. While the 40-70 dBZ displays a smooth and consistent nowcast sequence, the 30-60 dBZ range has a period of excessive bowing in the

central section of the line in the middle of the forecast period. In this case the nowcast scheme must be separating the line into three clusters and advecting the central cluster faster than it should for a period before the entire pattern returns to a more realistic outlook.



**Figure 7: Nowcast sequences for the 30-60 dBZ reflectivity band (left hand column) and the 40-70 dBZ band (right hand column).**

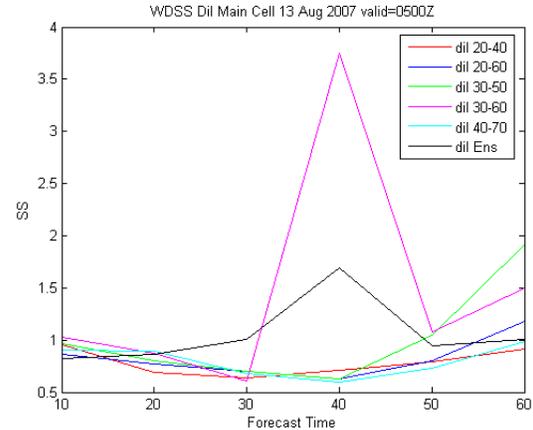


**Figure 8: The total error diagnosed for the main cell detected in the nowcast field.**

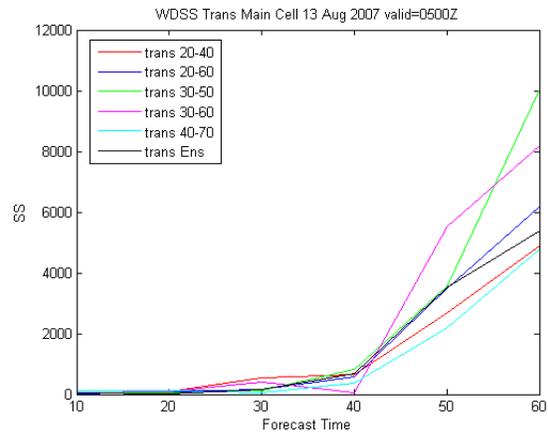
This inaccuracy is reflected in the Procrustes verification output. Figure 8 shows the total penalty recorded for the main cell diagnosed, and the large error found for the 30-60 dBZ nowcast for the 40-minute lead time compared to the other forecasts in the ensemble. It is interesting that all the individual ensemble members have their largest error at this lead time, indicating that the behavior of the storm at this time was less predictable. However, with the Procrustes scheme it is possible to separate the different error components to determine what aspect of the nowcast is performing badly.

In this case a large proportion of the error comes from the dilation factor. As shown in figure 9, the dilation error for the main cell in the 30-60 dBZ field spikes at this time as a result of the main cell appearing as three discrete cells at this time. One of the three is recognized as the main cell and matched to the single large cell in the actual image, but the Procrustes fit requires that the single forecast cell matched is dilated greatly, resulting in a large penalty. On the other hand, the other forecast runs actually have a minimal dilation penalty at this time as they produce a single cell that matches the actual cell size well.

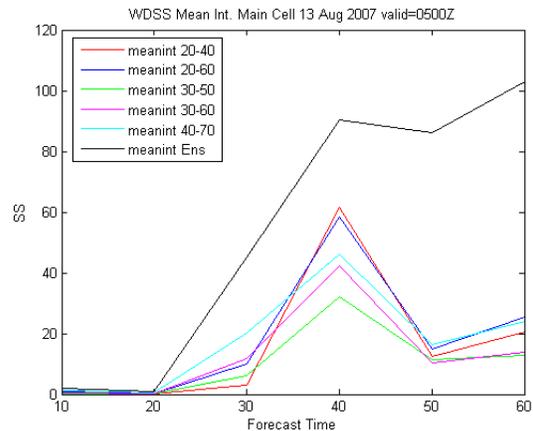
The separation of the line into three segments in this forecast also produces a greater translation error than in the other forecasts, and this is seen in figure 10.



**Figure 9: Main cell dilation penalty.**



**Figure 10: Main cell translation penalty.**



**Figure 11: Main cell intensity error.**

#### 4. DISCUSSION

Overall the total adjusted penalty reveals a problem with the 30-60 dBZ segmentation forecast. The separation of the penalty function into its components shows that this forecast generates a storm that is the wrong size and has a poorer location forecast than in the other nowcast runs. The total penalty, in contrast to the individual cell penalty, also includes an additional error due to the mismatch in the number of detected cells between the forecast and the actual. Therefore, one could determine that this particular forecast field was splitting the single line into more than one segment and advecting part of the line at the wrong speed, while retaining a good representation of the shape.

The use of geometric cases allows one to assess the performance and behavior of the verification scheme. In particular, it permits examination of the various error components both in combination and individually, such that fundamental problems with forecasts can be rapidly identified. It also allows one to assess weighting schemes for the different components of the penalty function so that they are comparable, and do not skew performance assessment in the direction of one type of error. For instance, we plan to use the minimization of the penalty function to provide a robust cell matching system that does not rely solely on the distance between or the characteristics of the cells to be matched. In order for this to succeed it is necessary that the weighting of the components of the penalty function are set in a manner that the cell matching proceeds in a realistic way.

#### ACKNOWLEDGEMENTS

This research was funded by the National Science Foundation Award # ATM-0434213. We would like to thank Sakis Micheas for his help and advice, and George Limpert for his assistance.

#### 5. REFERENCES

Lack, S.A., 2007: Cell identification, verification, and classification using shape analysis techniques. Ph.D. dissertation, Dept. of Soil, Environmental, and Atmospheric Sciences, University of Missouri, 147 pp.

Lakshmanan, V., T. Smith, G. J. Stumpf, and K. Hondl, 2007: The warning decision support system - integrated information (WDSS-II). *Wea. Forecasting*, **22**, 596-612.

Micheas, A., N.I. Fox, S.A. Lack, and C.K. Wikle, 2007: Cell identification and verification of QPF ensembles using shape analysis techniques. *J. Hydrology*, **343**, 105-116.