

5A.4

Spatial and Temporal Object-based Evaluation of Numerical Precipitation Forecasts

Christopher A. Davis, Barbara G. Brown, and Randy Bullock
National Center for Atmospheric Research¹
Boulder, Colorado

1. Introduction

Several new methods of object-based verification have recently appeared in the literature and these are reviewed in Gilleland et al. (2009). The Method for Object Based Diagnostic Evaluation (MODE) is one of these methods (Davis et al. 2006a,b; 2009, henceforth D09). The method has been used to evaluate precipitation forecasts from relatively fine-resolution regional models. MODE first identifies distinct features in the forecast and observed rainfall fields by performing a convolution operation (which acts like a smoothing), followed by a thresholding to retain areas of greatest interest, and also to remove some kinds of errors in the data that contaminate lower rain rates.

Once an object has been identified, two-dimensional geometric properties of the object are computed. These include the centroid location, long and short axes, aspect ratio (the ratio of short to long axis lengths) and the orientation angle that the long axis makes with respect to the ‘x’ direction of the analysis grid.

Once attributes are defined for each object, attributes of objects in one field are compared, with attributes of all objects in the comparison field. The likelihood that a pair of objects, one in each field, match is computed using a fuzzy-logic algorithm based on a “total interest”, a summation of how closely objects in a pair resemble each other. Total interest for the j th object pair is defined:

$$I_j = \frac{\sum_{i=1}^M c_i w_i F_{i,j}}{\sum_{i=1}^M c_i w_i}. \quad (1)$$

Here F is the interest function that prescribes, on a scale from 0 to 1 with 1 being perfect, how closely a forecast attribute matches the observed attribute. The coefficient w is the weight assigned to that interest function and c is a function of attributes that describes the confidence in a partial interest value obtained from $w_i F_{i,j}$. Total interest comprises M interest functions that compare attributes of each object pair. Specific attributes are outlined in D09.

2. MODE in 3-D

The foregoing overview of the object identification and matching likelihood computation is valid for two-dimensional objects. However, precipitation systems are often characterized as much by their evolution in time as by their structure in two spatial dimensions. The timing of rainfall systems marks a major forecast challenge. Furthermore, treatment of time and space dimensions together gives information about the translation speeds of forecast and observed features, and can distinguish features that may be spatially proximate at a given time, but propagating in different directions from features that are further separated at a given time but propagating in similar directions at similar speeds. Such information, in turn, provides more diagnostic power for understanding model errors. Errors in physical representation of

¹The National Center for Atmospheric Research is sponsored by the National Science Foundation.

convection reveal themselves not so much by structural disparities with observations, but rather fundamentally different timing in the diurnal cycle and propagation characteristics.

The primary question we address in this paper is the extension of MODE to three dimensions (x,y,t) where t is time. Mathematically there little distinction between spatial and temporal dimensions. Thus one can view objects in 3-D as in Fig. 1, where the centroid is now defined in

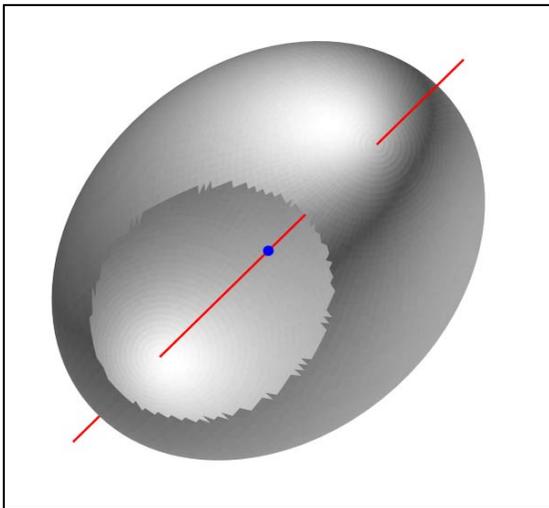


Figure 1. Hypothetical three-dimensional precipitation object in (x,y,t), where the major axis (red line) and centroid (blue dot) are shown. Time is the vertical direction.

three-dimensional space (x,y,t) and there are two axis angles that must be defined to specify the orientation of a feature. Furthermore, volume replaces area as a measure of object size.

In practice there are some subtleties, one being how to relate temporal and spatial displacements of objects. For a given object moving at velocity \vec{v} , $d\vec{x} = \vec{v}dt$ would be the obvious relation between spatial and temporal errors. In dimensional terms, this means that for an object moving eastward at 10 m/s, a displacement of 1 hour in time would be equivalent to a spatial displacement of 36 km.

However, the system-following relationship between space and time turns out to be not particularly relevant for assessing correspondence between simulated and observed precipitation features. First of all there is the problem presented by slow-moving systems, say, 3 m/s, in which a spatial error of only 100 km would be equivalent to a timing error of about 9 h. This could place “matching” systems in different phases of the diurnal cycle. Clearly if one adopts a constraint that the “environment” of the matching simulated and observed systems should be similar, we need to consider a severe penalty for temporal displacements greater than a few hours. The 2-D version of MODE severely penalizes spatial displacements of forecast and observed systems when the distance separating their edges becomes comparable to the scale of the systems. For instance, for two systems, each about 100 km in size, a displacement of 200 km or more would result in a relatively low likelihood of a match.

Thus, a practical relationship between space and time is the size of typical systems and a small fraction of the diurnal cycle. This implies a larger “effective speed” relating spatial and temporal errors. Herein we adopt $c=30$ m/s, meaning that a spatial displacement of 300 km is equivalent to a temporal displacement of about 3 h.

The above definition allows us to define a modified definition for the centroid separation,

$$d = \left((x_2 - x_1)^2 + (y_2 - y_1)^2 + c^2 (t_2 - t_1)^2 \right)^{\frac{1}{2}}.$$

Here, subscripts 1 and 2 refer to the forecast and observed objects, respectively, with centroids at $(x_{1,2}, y_{1,2}, t_{1,2})$. It is convenient to nondimensionalize the distance in the form of an interest map. This is simply a function of distance that varies from unity for $d=0$ and decays linearly to 0 at $d=1000$

km. The interest value represents the likelihood that the forecast object can be identified as the model's manifestation of the observed feature.

Each pair of forecast and observed objects will have an interest value between 0 and 1. A matrix of interest values is constructed, following D09, in which the interest value for each pair is one matrix element.

As a specific example, consider the WRF ARW forecast for June 15, 2002, during IHOP (Fig. 2). The hourly precipitation data are from the Stage IV analysis. The model precipitation is also an hourly accumulation. The ARW was initialized with the interpolated NCEP Eta analysis valid 00 UTC on the days under consideration. The model was integrated for 36 hours on a single domain with a grid

spacing of 3 km that covered a large fraction of the CONUS. Boundary conditions were obtained from the Eta forecasts. The cloud physics scheme was the Thompson microphysics scheme, and the boundary layer scheme was the Mellor-Yamada-Janjic TKE-predicting scheme.

A comparison of the hourly precipitation fields, displayed as objects in (x,y,t) space appears in Fig. 2. The cylindrical convolution and thresholding have been applied using two thresholds for hourly rainfall: 2 and 4 mm. The partial transparency of the 3-D objects makes visible the heavier rain region within the lighter-rain envelope. We also enforce a constraint of a minimum volume for a system to be included in the analysis, expressed as 300 grid volumes, with each grid volume 12-km in each spatial direction

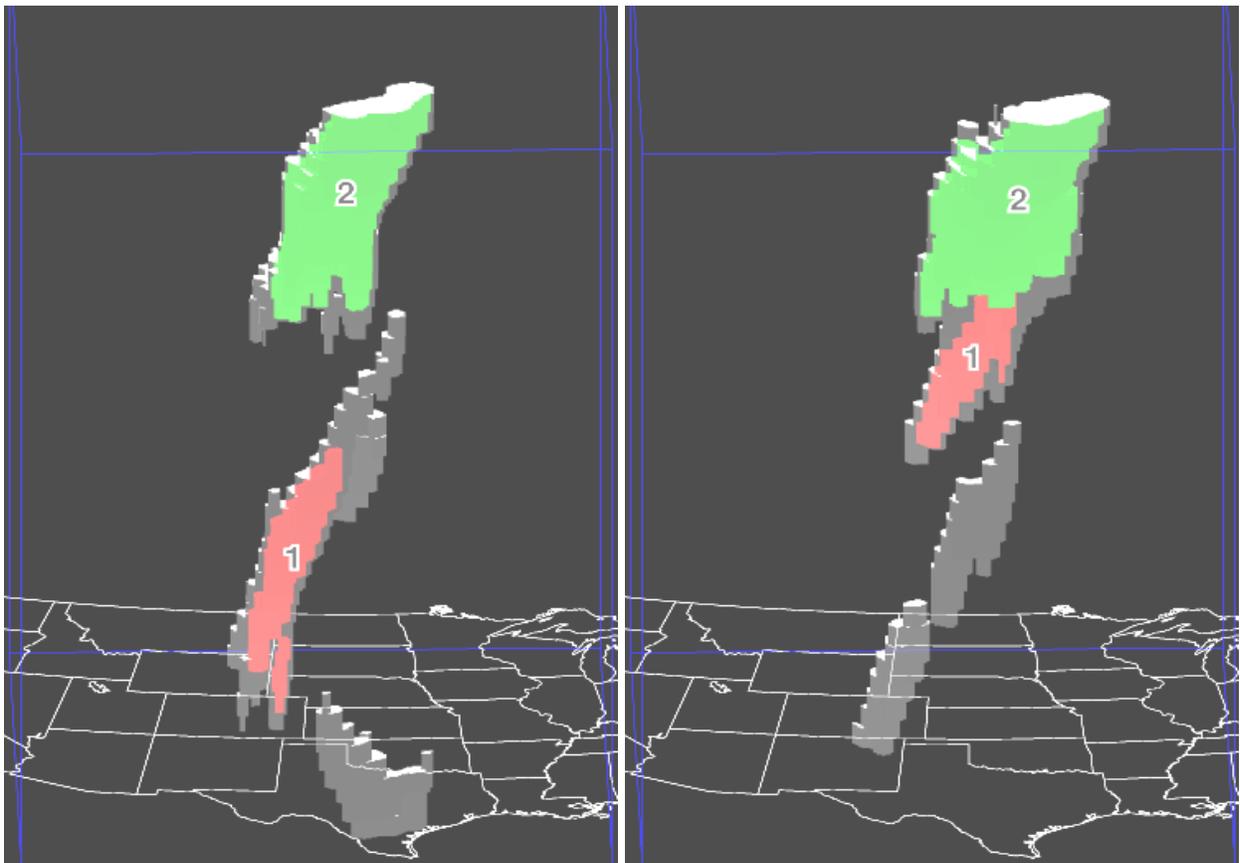


Figure 2. Space and time depictions of rain systems. Gray indicates the volume of systems exceeding 1 mm/h; colors (not indicative of matching between forecast and observations) show volume exceeding 4 mm/h.

and 1-h in time. A system 10x10x3 would be 120 km in each spatial dimension and 3 h in time.

		Observed Object		
		1	2	3
Forecast Object	1	0.06	0.00	0.00
	2	0.14	0.67	0.00
	3	0.00	0.27	0.00
	4	0.00	0.00	0.72

Table 1. Interest matrix for 15 June.

The interest matrix for this case is shown in Table 1. The values shown in red are notably larger than the other values, and indicate that forecast object 2 probably matches observed object 2, and forecast object 4 probably matches observed object 3.

Given the interest values, we define the “best match” for each forecast and observed object, that being the maximum across each row and column respectively. We then compute the median value of these maxima. For forecast objects, we compute the median of 0.06, 0.67, 0.27 and 0.72, which is 0.47. For the observed objects, the median of maximum interest is 0.67. To obtain an overall metric of forecast quality, we simply average the two median values and obtain 0.57, a metric known as the Median of Maximum Interest (MMI).

3. Results

Using the procedure outlined above we may compare forecast quality on different days, the result of which appears in Table 2. Note there is substantial variation among these days, especially at the lower threshold. The poorest forecast occurred on June 13, while June 12 was the best overall. This method can easily be extended to

comparing multiple models and to larger datasets. For 2-D objects this was demonstrated in D09.

Total interest can also be used to

Date of initialization (00 UTC)	MMI (T=2)	MMI (T=4)
10 June	0.56	0.75
11 June	0.42	0.59
12 June	0.75	0.70
13 June	0.34	0.56
15 June	0.57	0.50

Table 2. Median of Maximum Interest for five forecasts during IHOP, and two threshold rainfall values (2 and 4 mm). Forecasts initialized at 00 UTC on date shown and integrated 36 h.

assess which forecast systems likely correspond to observed rain systems, and from these pairs of objects, the joint distributions of other object attributes can be defined. An example, covering all pairs of systems during the 5 days shown in Table 2 whose total interest exceeded 0.5 is shown in Fig. 3. The scatterplot of system translation speed (derived from the slope of the major axis of each object relative to a “vertical” space=constant line (e.g. Fig. 2) suggests that there is no substantive bias in the forecast translation speeds, but there is not a high level of skill in the prediction of translation speed.

Another important attribute is system lifetime. The distribution of this parameter is truncated because of the minimum space-time volume constraint all systems must satisfy, and because simulations are of relatively short duration. However, from the scatterplot in Fig. 4, it is evident that the forecasts have some ability to simulate the lifetime of events, although there is a slight positive bias.

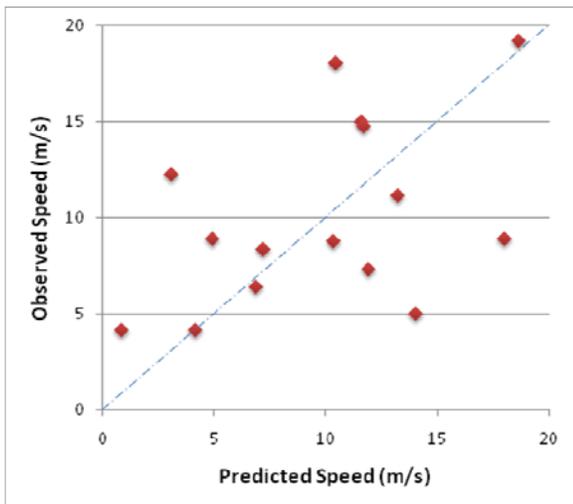


Figure 3. Scatterplot of predicted and observed translation speeds for pairs of systems with total interest exceeding 0.5. Dot-dashed line is 1:1.

4. Summary

The present paper has demonstrated an extension of the MODE method described in recent articles (e.g., D09) to include time evolution. This results in three-dimensional objects (x,y,t). Most of the concepts from MODE can be applied in a straightforward way to three dimensions provided the scaling between space and time can be established. The diurnal cycle drives

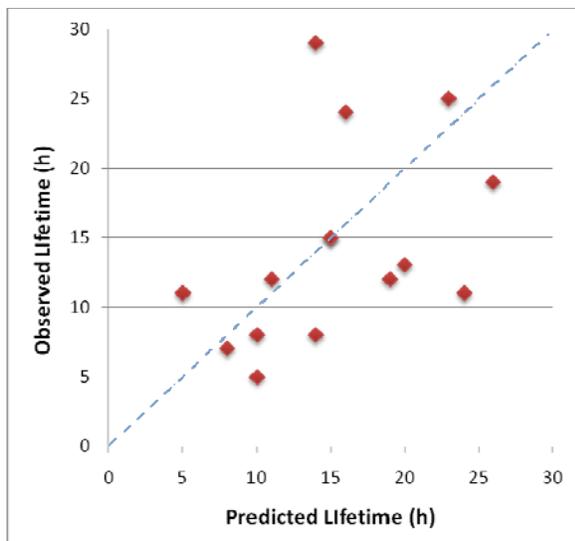


Figure 4. Scatterplot of forecast rain system lifetimes versus observed lifetimes (h).

a relatively rapid variation of environmental conditions compared to the time scale over which individual systems physically migrate into a different environment. Thus, the effective speed that relates space and time is larger than the speed of an individual rain system. We have chosen 30 ms^{-1} in this work; this is roughly double the typical speed of rain systems.

The 3-D MODE described herein results in a relatively small number of objects over a given time period compared with the number that result from the 2-D version of MODE, where time correlations are not considered. This has the advantage of reducing the complexity of rainfall forecasts to a few parameters that describe the geometric and physical properties of precipitation objects. Thus, the 3-D MODE can succinctly encapsulate forecast performance but provide diagnostic information at the same time that is not available from traditional verification metrics.

The next steps are to make the interest mapping more sophisticated, to extend the method to larger datasets and to compare forecasts from different models. We also wish to explore 3-D MODE applied to rainfall forecast problems with different space and time scales.

References

- Davis, C. A., B. G. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts, Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, 1782-1784.
- Davis, C. A., B. G. Brown, and R. Bullock, 2006b: Object-based verification of precipitation forecasts, Part II: Application to convective rain systems. *Mon. Wea. Rev.*, 1785-1795.
- Davis, C. A., B. G. Brown, R. Bullock, and J. Halley-Gotway, 2009: The method for object-based diagnostic evaluation (MODE) applied to WRF forecasts from the 2005 NSSL/SPC spring program. *Wea. Forecasting*, Accepted.
- Gilleland E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Spatial forecast verification methods inter-comparison. *Wea. Forecasting*. Submitted.