Using MODE with CloudSat Data

Randy Bullock

National Center for Atmospheric Research, Boulder, Colorado

1 INTRODUCTION

This paper discusses a preliminary version of an extension of MODE (the Method for Object-Based Diagnostic Evaluation) for use with satellite data. MODE is one of the tools in the MET (Model Evaluation Tools) verification software package. Many users of MET have expressed a desire to be able to use satellite data in verification. This project is intended as a first step in that direction.

The main purpose of this effort is to resolve the data in the vertical curtain into *objects*, in an effort to create a version of MODE that is "vertical" rather than "horizontal", which is the traditional version of MODE. Object-based verification has become widely used in the last decade, and extending the basic approach to other verification scenarios (such as satellite data) is something that would be very useful to both forecast creators and end-users.

2 DATA

For our observations, we used reflectivity data from the CloudSat 2B-GEOPROF files. Our forecast data was reflectivity derived from model data—specifically, NOAA's Rapid Refresh model.

3 METHODS

Some adaptations were needed in order to make MODE work with the vertical "curtain" data that the NASA A-trains satellite constellation produced. This section gives an overview of those adaptations and also a discussion of objects in general.

3.1 *Satellite Ground Track*

The path of the satellite ground track was used to cut a (curved) vertical slice out of a 3D model reflectivity field to compare with the CloudSat observations. The bearing angle (eastward from north) could be used to determine the direction orthogonal to the path at each point. If β is the bearing angle, then $\tan\beta = -(dL/d\phi)\cos\phi$ where the minus sign shows up because we are considering west longitude to be positive. Note that β cannot be assumed constant along the track. Adding or subtracting 90° to this value will give the orthogonal direction, but we used a vector approach instead. If we denote the tangent vector to the track as **T**, then **T** $= \sin\beta$ **E** $+ \cos\beta$ **N** where **E** is the unit vector in the local east direction, and **N** is the unit vector in the local north direction. See Figure 1. Then (assuming a spherical Earth with unit radius), at position **P** on the ground track, the orthogonal direction (call it **V**) is given by the cross product **V** $=$ **T** × **P**. The vectors **T** and **V** give an orthonormal basis for the tangent plane to the Earth at **P**. The vector **V** was used to search orthogonal to the track. Figure 2 shows a comparison of the CloudSat observation field with the averaged forecast field obtained in this way.

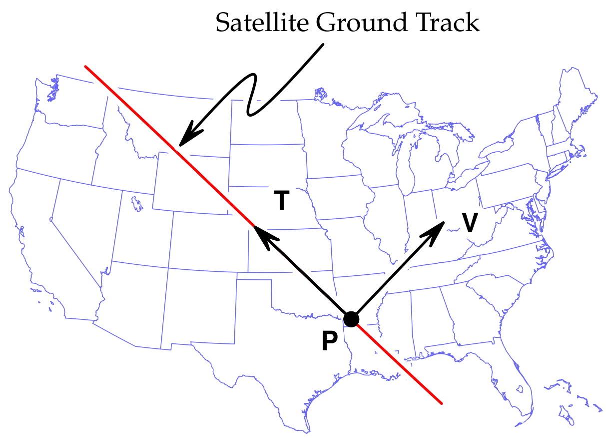

Figure 1

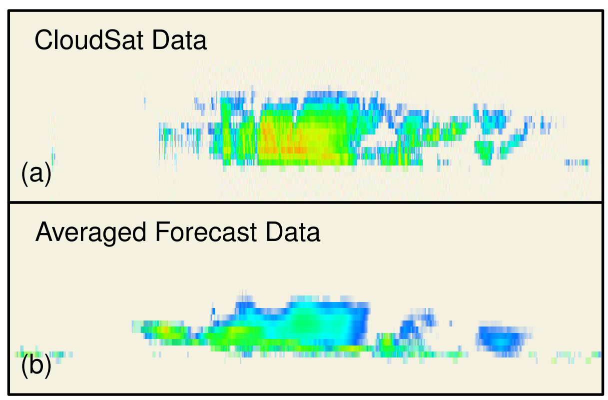

Figure 2

3.2 *Objects*

What are objects in this context? Objects are just regions of interest. For example, in a precipitation forecast

Corresponding author address: Randy Bullock, P. O. Box 3000 Boulder CO, 80307. email: bullock@ucar.edu

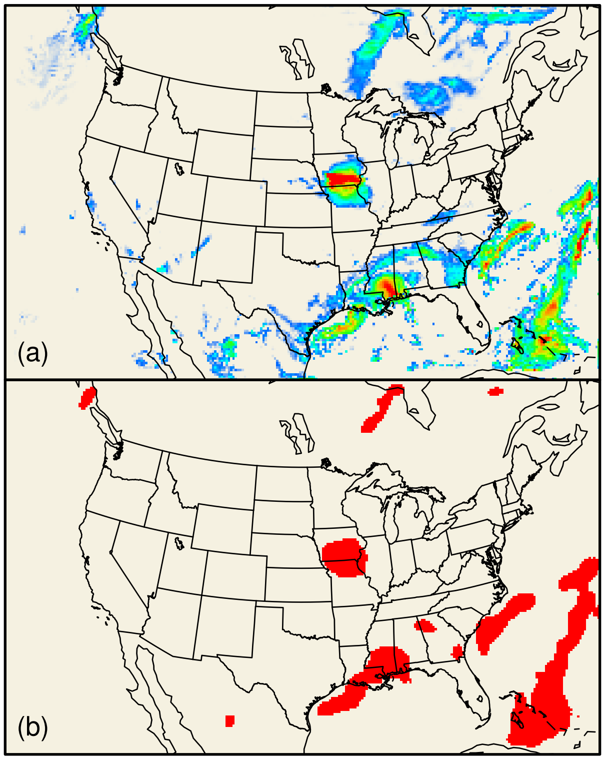

Figure 3

(see Figure 3a), the objects would be high rainfall areas. Resolved objects are shown in Figure 3b. Examples of objects in other contexts might be clouds, hurricanes, or regions of high wind divergence.

Resolving the raw data field into objects is a four step process. (See Figure 4.) Starting with the raw data (Figure 4a), we apply a convolution filter. This is basically a spatial smoothing operation. The results can be seen in Figure 4b. A threshold is then applied to this smoothed field, giving an on/off mask field (Figure 4c). This determines the object boundaries. Finally the raw data is restored to the interior of the objects (Figure 4d), allowing one to examine distributions of raw data values inside the objects.

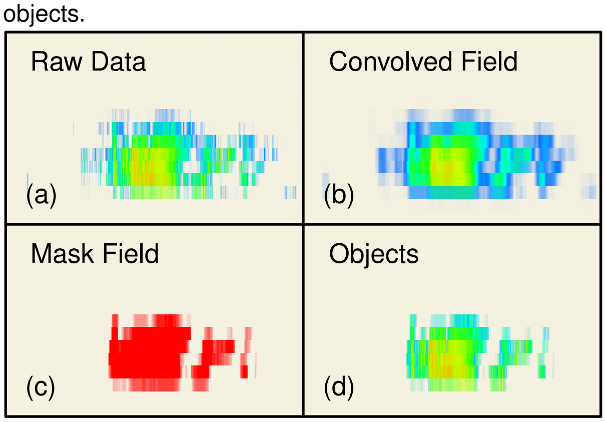

Figure 4

Once the objects have been resolved in the forecast and observed fields, one can examine spatial errors, intensity biases, alignment and other measures of forecast quality that correspond more closely to one's intuition than traditional verification measures such as probability of detection (POD) or root mean square error (RMSE).

3.3 *Attributes*

Once the objects are in hand, comparison of forecast and observed objects is done by means of various calculated object *attributes*. Object attributes come in two varieties: those for single objects, and those for object pairs, one object from the pair coming from the forecast field, and one from the observed field.

Example attributes for single objects are area, centroid location and orientation angle. As an illustration of some simple object attributes, see Figure 5. If we consider the blue shape as an object, then we can calculate attributes for that object. Figure 5a shows the *centroid* of the object. It is basically the geometric center of the object. Figure 5b shows the *axis* of the object. It gives information on the spatial orientation of the object. We need the calculated axis to be invariant with respect to rotations of the object, so a simple least squares line fit was ruled out. The dotted line in Figure 5c illustrates the *convex hull* of the object. It is used in the calculation of several other attributes.

For object pairs, example attributes are intersection (or overlap) area, union area and symmetric difference. These are by no means complete lists of the attributes that are calculated by MODE, but they will hopefully give the reader some idea of what attributes are, and their relation to the objects they describe.

Essentially, attributes allow one to quantify and compare the similarity (or lack thereof) of objects to each other. What would be essential ingredients of a good forecast? Clearly, most people would respond that a good forecast is one that gets the overall intensity, spatial location and orientation, and size (among other things) of the features roughly correct. But these are just the object attributes as calculated by MODE. These measured attributes thus correspond very closely to user's intuitive ideas of forecast quality, and so the results of doing verification with an object-based methodology are much more easily interpreted by both experts and everyday users of the forecasts.

3.4 *Matching and Merging*

Attributes are used to compare objects with a view to performing two operations, both of which involve associating some objects with others.

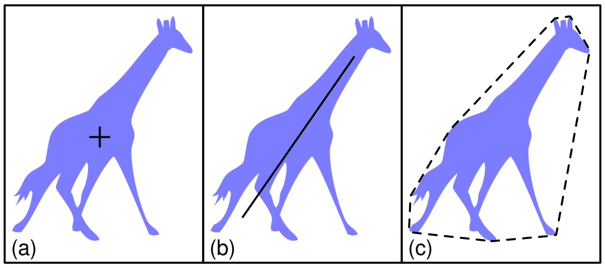

Figure 5

One operation is *merging*. This is the association of objects in the same field with each other. For example, one might like to consider a collection of two or more forecast objects as being one composite object. Similarly, merging can be done in the observed field to produce composite objects there. Once the merging has been done, one can calculate new object attributes for these composites.

The other operation is *matching*. This is the association of objects in the forecast field with objects in the observed field. This is where the comparison of forecast with observations actually happens. Often, it is composite objects rather than simple objects that are matched, since clusters of objects in one field will often have counterparts in the other field.

Not all objects end up participating in these two processes. There will be objects in one field that have no counterpart in the other field. Thus we often have objects that represent either misses or false alarms.

3.5 *Fuzzy Logic*

Matching and merging are done with a fuzzy logic engine. The attributes are combined using interest maps, weights, *etc.*, to produce a single quantity, the *total interest*, that is thresholded and which determines which matches and merges are actually performed. There are many tunable options that guide this process in traditional MODE, although this preliminary extension of MODE used with CloudSat data uses a simplified version of the fuzzy logic engine. Eventually, this will be expanded to include more attributes and interest maps, but for this prototype version it was decided to keep things simple.

4 CONCLUSIONS & FUTURE WORK

This has been an initial foray into the world of satellite data for MODE in particular, and also MET in general. Verification using satellite data has been a long-standing omission in the MET verification software package. While a robust satellite data verification system will likely be the work of several years, this may help the process get started.

Eventually, we would like to extend MET's use of satellite data to other areas, such as cloud forecasts, space weather, and more spatial dimensions. The MODE object-based verification technique should have many applications in these areas.

5 ACKNOWLEDGEMENTS

NCAR is sponsored by the National Science Foundation. This research is in response to requirements and funding by the National National Aeronautics and Space Administration (NASA). The views expressed are those of the author and do not necessarily represent the official policy of NASA or NCAR.

The author, along with the rest of the MET Development Team, would like to thank our co-workers on this project, especially the folks at the CloudSat Data Processing Center at Colorado State University.

In addition, we would like to thank our sponsor, NASA, for making this work possible.

6 REFERENCES

Information on CloudSat data is available at the CloudSat Data Processing Center website:

www.cloudsat.cira.colostate.edu/index.php

Information on the MET verification software package and on MODE can be obtained from:

www.dtcenter.org/met/users