378 EVALUATION OF GFS WATER VAPOR FORECASTI ERRORS DURING THE 2009-2010 WEST COAST COOL SEASON USING THE MET/MODE OBJECT ANALYSES PACKAGE

Wallace Clark[1,2,6], Huiling Yuan[1,8,3], Tara Jensen[1,4,5], Gary Wick[1,2], Ed Tollerud[1,3,5], Randy Bullock[1,4,5], Ellen Sukovich[1,2,7], John Halley Gotway[1,4,5], Paul Oldenburg[1,4,5], Isidora Jankov[1,3,9]

[1]HMT-DTC Collaboration Project, [2]NOAA/ESRL/PSD, Boulder, CO, [3]NOAA/ESRL/GSD, Boulder, CO, [4]NCAR/RAL, Boulder, Colorado, [5]DTC (Developmental Testbed Center), Boulder, CO, [6]Science and Technology Corporation, Boulder CO, [7]CIRES (Cooperative Institute for Research in the Environmental Sciences), University of Colorado, Boulder CO, [8]School of Atmospheric Sciences and Key Laboratory of Mesoscale Severe Weather, Ministry of Education, Nanjing University, Nanjing, Jiangsu, China, [9]CIRA (Cooperative Institute for Research in the Atmosphere), Colorado State University, Fort Collins, CO.

INTRODUCTION

Continuous advancement in the development and use of numerical weather prediction models, coupled with tremendous advances in remote sensing capabilities, particularly via satellite over the oceans, of the critical meteorological parameters needed to initialize these models, has powered several decades of steady improvement in weather forecasting. This is particularly true for the domain of this study, which focuses on forecasts of atmospheric river (AR) induced extreme precipitation events occurring along the west coast of the United States. The occurrence of such an event is generally now predicted two days or more in advance. But the uncertainties with respect to location and intensity remain sufficiently large that the forecast can miss the exact river basin and significantly over or underestimate the amount of the resulting precipitation, which depends in a complex way on the underlying coastal topography and the speed and direction of flow of the stream of moisture. Because of the West Coast's vulnerability to sometimes catastrophic flooding along this coastal domain, it is important to reduce these forecast uncertainties until both short term (2 day) and longer term (greater than 5 day) mitigation can be based on information of sufficient fidelity that rational and optimal decision making is possible.

That improvements are still needed is reflected by the observation that due to the remaining current uncertainty in forecasts of AR track and forecasts of the intensity of the rainout along the track, traditional site based skill scores (e.g., POD, FAR, and CSI), when applied to extreme precipitation events, tend to imply almost no forecasting skill, despite the fact that the forecasters now know with some certainty that an event of some significance will occur within a larger time-space domain containing the site. Thus, although these traditional skill scores sound an alarm that forecast improvement is needed, they provide little information on how to make the improvements. This suggests an additional approach needs to be added to the verification toolbox. One such approach is the object based approach.

Corresponding author address: Wallace Clark, NOAA \ESRL\PSD2, 325 Broadway, Boulder, CO 80503; Wallace.l.clark@noaa.gov.

In this study, conducted as part of the HMT-DTC collaboration project, a joint effort of the Hydrometeorological Testbed and the Developmental Testbed Center, we examine the use of an object based approach to quantify the uncertainties remaining in forecasts of extreme precipitation events along the United States West Coast. We do this through application of DTC's MET/MODE utility, a software tool that allows custom definition and creation of objects from input forecast and observation data fields. For the study described in this paper the data field is a gridded two dimensional (latitude, longitude in ½ degree steps) array of Integrated Vapor Transport IVT values (i.e., vapor flux). The study is based upon the real time GFS model output over a selected domain in the Northeast Pacific covering the 2009-2010 West Coast cool season.

Since observed directional wind fields at altitude over the Northeast Pacific (NEP) needed for calculating IVT were not readily available, in this study we work entirely with model data, letting the GFS analysis fields serve as the observations against which 24, 48, 72, and 96 hour lead-time GFS forecasts are compared. Selected MODE object attributes are used to create metrics meant to quantify the degree of agreement between analysis and forecast objects with respect to object location, size, shape and intensity. As detailed in the discussion, straight forward application of the MET/MODE object attributes, while intuitively simple, resulted in certain complex biases that make interpretation of some of the new metrics difficult. The elimination of these biases will need further development.

WHAT IS MODE

MODE stands for Method for Object-based Diagnostic Evaluation. It is an object-based verification software tool provided in the MET (Model Evaluation Tools) package developed and supported by the Developmental Testbed Center (DTC). This package of tools is readily available and intended to provide the research community with a common software package incorporating the latest advances in forecast verification. Figure 1 sketches the steps used by the MET/MODE software to find and define objects within an input data field. Figure 2 sketches two objects on overlaid data fields, one forecast and one observed, being compared, and illustrates the concepts of object centroid, centroid distance, and object intersection (i.e., overlap).

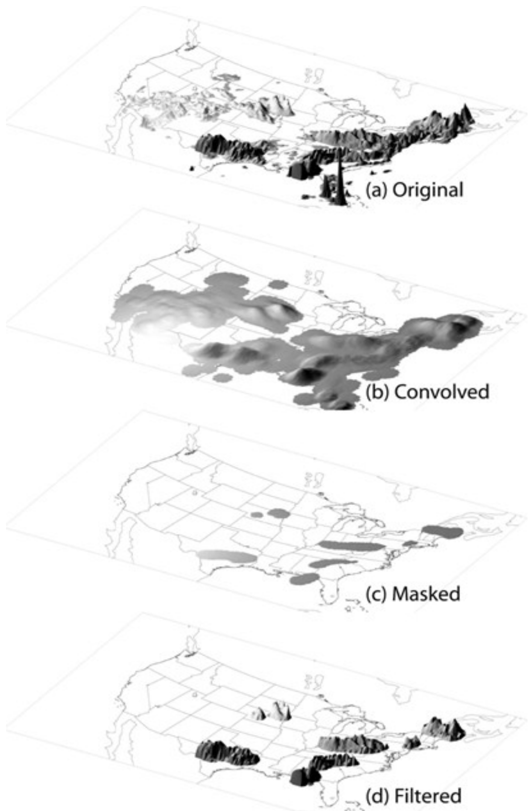


Figure 1: The four steps MODE uses to define an object: 1) A gridded data set is created. 2) This data set is smoothed using a customizable algorithm so that in the next step objects of the desired size and smoothness will be created. 3) A threshold or logical criterion is applied. Pixels that satisfy the criterion are associated with a bit map value of 1, otherwise 0. This bit map space is used to calculate many of the MODE object attributes. Objects are defined on the bit map as the sets of adjacent pixels with values of one. 4) A new gridded data set is created by filling in the object pixels with their original values. Some MODE object attributes are based on this data set.

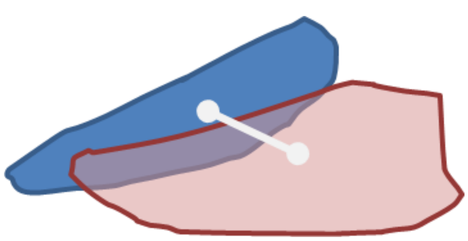


Figure 2: A schematic of a forecast (blue) object and an observation (pink) object. Each object obviously has an area. The area where the objects intersect is called the overlap area in this paper. The white dots represent the respective centroids (center of gravity in bit map space).

Figure 3 shows an example of the objects determined by MODE for a particular case.

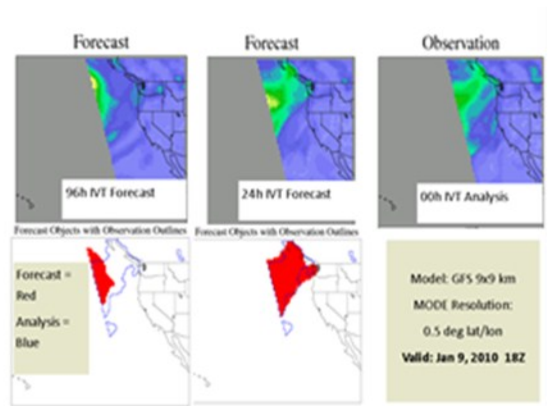
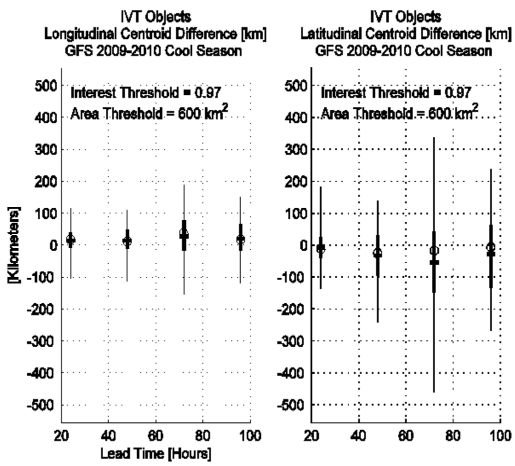


Figure 3: The panels in Figure 3 were built from the graphical output of MODE. Along the top row from right to left it displays the 96 h and 24 h GFS forecast of Integrated Vapor Transport, to be compared with the third panel, the GFS analysis. The leftmost lower panels show the MODE determined IVT objects, where the forecast objects are in solid red and the Analysis objects ('Observations') are outlined in blue. As might be expected, throughout the data set the 24 h forecast objects resemble the analysis much more closely than the 96 h. The panels represent the full data set domain and the western boundary applied to the data set before analysis to keep the study near the west coast and to reduce the IVT object size is evident.

USE OF MODE OBJECT ATTRIBUTES AS SKILL METRICS

3.1 CENTROID DISPLACEMENT

As a preliminary example, the increase in uncertainty of location as lead time increases is suggested in the following preliminary statistics of centroid displacement of IVT objects (contiguous pixels of smoothed IVT greater than 25 cm m/s).
It is plausible that the centroid displacements summarized in Figure 2 resemble the locational uncertainty a forecaster faces in predicting the location of AR associated extreme precipitation events, although there are several caveats, or points that need to be addressed in further studies. First, the exact nature of this connection remains to be demonstrated. Second, in this analysis the size of the displacements are sometimes biased small due to truncation of objects that haven't fully crossed the arbitrary western boundary we used to restrict the study to objects near the coast. And finally, there may be a few cases where the object shapes found have created unrealistic matches. Nonetheless the fact that for 50 percent of the object comparisons the centroid locations differ by more than plus or minus 100 km suggests that placement error is a significant factor in keeping skill scores down.



Legend

Median = o, Mean = +

Percentile Ranges: 2.5—97.5 = Thin line, 25—75 = Thick line

Figure 3: A statistical summary of centroid distance (forecast—analysis), summarizing longitudinal (left panel) and latitudinal displacements observed over the whole cool season. That there appears to be a southerly forecast bias is apparent and examination of the thick vertical bars that represent the middle half of the distribution shows that the uncertainty in location tends to grow lager with forecast lead time, at least out to 72 hours. The significance of this bias is questionable because of the object truncation effects of the western boundary to the data set. However, the uncertainties expressed by the percentile ranges serve as a lower bound to the difference in location of the two objects.

3.2 UNCERTAINTY IN IVT OBJECT AREA

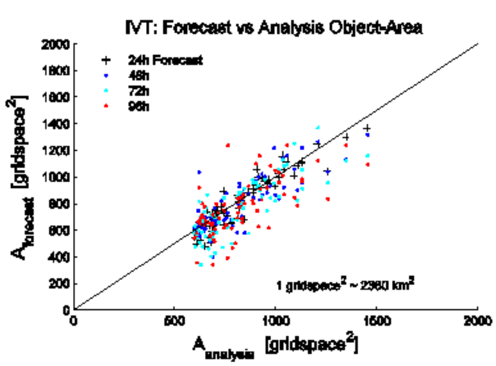


Figure 4: This scatterplot plots the forecast versus analysis object area. The wider spread of the red dots (96 h lead times) compared to the black '+' symbols (24 h) are consistent with the expectation that the uncertainty grows with forecast lead time. Since this uncertainty does not appear to change with object area, the fractional error will decreases with area. Figure 4 below details the observed increase in uncertainty with forecast lead time.

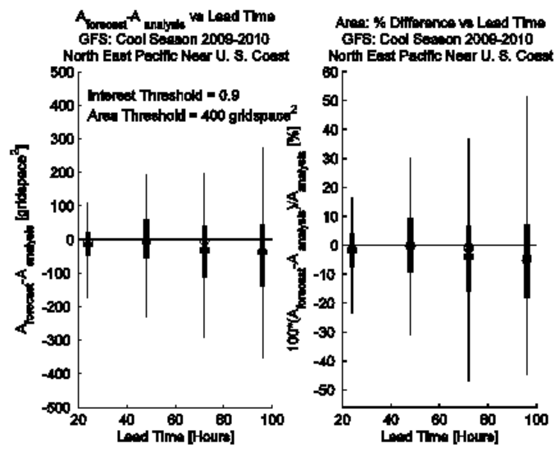


Figure 4: Boxplots of the difference between the forecast and analysis areas sketch the statistical distributions of the area differences (forecast minus analysis area) , broken out by forecast lead time, on the left and as a percent of the analysis object area on the right. As before, the circles locate the medians and the plus symbol, visible as a short horizontal bar, locates the mean. The thick vertical bars, which lengthen with lead time, depict the interquartile range, while the thinner vertical line depicts the 2.5 to 97.5 percentile range. The biases shown, especially at the larger lead times, may be largely due to the effect of the western boundary line imposed upon the data to restrict the analysis to near shore. For example, when the analysis object is farther advanced toward the coast than the forecast object its area may appear to be larger than the artificially truncated area of the forecast.

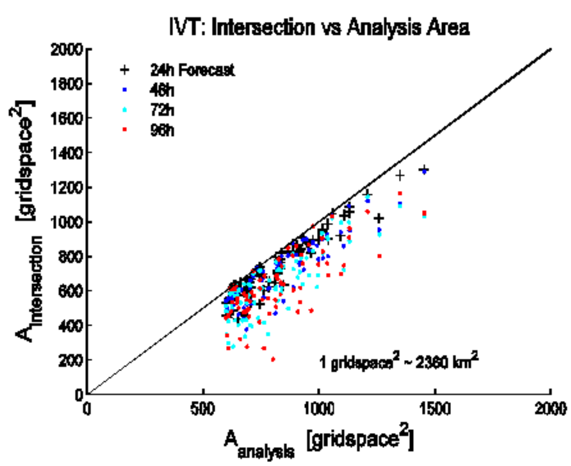


Figure 5: This figure plots the area common to both the two objects versus the area of the analysis. For a perfect forecast the intersection area would equal the analysis area. The smaller the intersection area is

relative to the analysis object area, the more misses there will be in a grid point skill score analysis. Clearly the 24 h forecast (black '+' symbols) are more skillful than the 96 h (red dots).

3.3 UNCERTAINTY IN IVT OBJECT INTENSITY

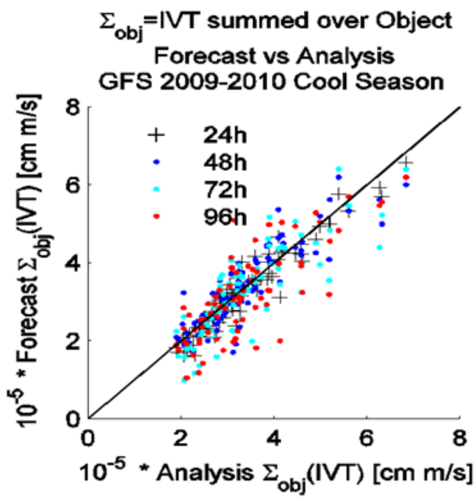


Figure 6: Similarly to the area analysis, the uncertainty in total IVT—the sum of the IVT pixel values contained in the object, which is physically related to the potential rainout of an event—does not appear to change with the magnitude of IVT, but does increase with forecast lead time.

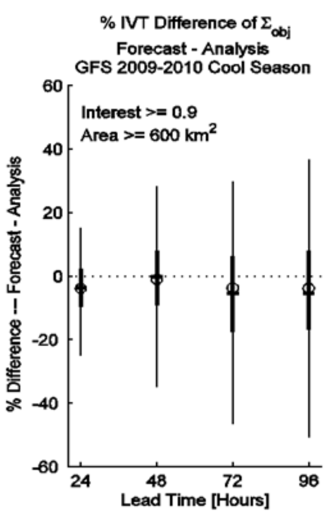


Figure 7: The boxplots symbols are here defined in the same way as in the previous figures. For brevity only the % difference statistics are shown. Both the interquartile and 95th percentile ranges tend to increase with forecast lead time. Again, the significance of the apparent biases is difficult to determine because of the object truncation effects caused by the application of the western boundary.

4.0 DISCUSSION

An exploratory study of the use of MET/ MODE object attributes as a basis for development of diagnostic verification metrics that can provide quantitative measures of the uncertainties in West Coast extreme event precipitation forecasts. IVT, or height integrated vapor flux, which is related to the source of moisture feeding extreme precipitation events rather than being the event itself, was selected for this study because this data field allows the precursors of precipitation events to be identified well out to sea and tracked as they come ashore. It is assumed that the uncertainty in forecasts of the location and intensity of these precursor (IVT) objects is closely related to and diagnostic of the causes of uncertainty in the forecast precipitation field itself.

The study utilizes 2009-2010 cool season GFS 6Z and 18Z GFS valid time model output. The GFS analysis data are utilized as MODE's observational field, against which the 24h, 48h, 72h, and 96h GFS forecasts are compared.

Out of many possible MODE attributes four are tried here as verification metrics: 1) centroid distance; 2) object area; 3) fractional forecast/ analysis object overlap; and 4) total IVT summed over the object. The results are promising, but more work is needed to make them more interpretable and relevant to forecast verification. A major difficulty in this study resulted from an effort to make the centroid difference easier to interpret physically by imposing a western boundary on the data field. As was the intention, this kept the analysis close to the coast and reduced the size of the objects, which otherwise could, and sometimes were, nearly as large as the whole Northeast Pacific. The application of this western boundary did indeed help to make the centroid distance more relevant and interpretable, but with respect to the other attributes, it also produced complicated biases by truncating the forecast and/or analysis objects when they were not completely within the domain.

Some uncertainty characteristics were common to all four attributes. As expected, the uncertainties increased significantly with forecast lead time. The magnitude of the uncertainty appeared to be independent of the magnitude of the attribute, meaning the fractional error of these attributes decreases with magnitude.

Of the four attributes the centroid distance seems to be the easiest to interpret in this study because the effect of the western boundary truncation is to consistently reduce its size. This allows the found separations to be interpreted as lower bounds to the actual object separations. It seems likely that these centroid distances are closely related to the spatial error in forecasts of extreme precipitation. If this proves to be true, then this uncertainty is bounded on the low end by the latitudinal and longitudinal distributions of centroid difference shown in Figure 3. The validity of the southern bias shown in Figure 3 is less certain due to the distortions caused by the western boundary. If the comparison had been made versus actual observations instead of the analysis, the differences should be somewhat larger, if for no other reason than that the observational data set will have its own uncertainties which should add to the uncertainties found here.

Since in any particular cast it may be either the forecast or analysis object that is more greatly truncated, the interpretation of the area, ratio of area overlap, and total object IVT metrics are more difficult to interpret. From a physical viewpoint these attributes seem intuitively appropriate to diagnostic verification of AR sourced precipitation event precursors. However, their routine application will depend on techniques being found that obviate the biasing effects of domain boundaries.

These results are consistent with the hypothesis that spatial uncertainties (location, shape, and timing) are significant sources of error that keep traditional skill scores low. However, more development is needed to make these and other MET/Mode attributes the quantitative metrics of diagnostic forecast skill desired.

BASIC REFERENCES

The following two references are provided as entry points to MET and MET/MODE and the literature on atmospheric rivers.

1) MODE: www.dtcenter.org/met/users/support/online_tutorial/METv2.0/mode/index.php

2) Atmospheric Rivers: Neiman, P. J., et al., 2008, J. Hydrometeorology, Vol. 9, pg. 22.