6.3

INSURING TEMPORAL AND SPATIAL CONSISTENCY IN SHORT RANGE STATISTICAL WEATHER FORECASTS

by

Bob Glahn* and Jerry Wiedenfeld
Meteorological Development Laboratory
Office of Science and Technology
National Weather Service, NOAA
Silver Spring, Maryland

1. INTRODUCTION

Statistical forecasts have been produced for many years by a number of organizations and with a variety of techniques. Purely Markov models have been used for very short range forecasts. Postprocessing of Numerical Weather Prediction model data has been done to produce interpretative guidance out to several days in advance.

Many times stepwise regression is employed where a plethora of potential predictors are "screened" to produce regression equations. Usually, the selection process is based on minimizing the mean square error of a predictand [equivalently, maximizing the Reduction of Variance (RV)], and a suitable stopping procedure is used to decide on the number of predictors to include in the equations. With sufficient data samples, equations can be produced for one or more variables, for many locations, and for several projections in time. Consistency of the forecasts produced by these equations is of concern--consistency (1) among variables (e.g., temperature and dew point; cloud amount and precipitation), (2) among spatial locations, (3) among projections made from one start time, and (4) among forecasts valid at the same time and place, but made at different times. Certainly, meteorological conditions may warrant significant and rapid temporal and/or spatial changes, but many times the conditions are such to suggest the guidance should exhibit consistency.

* Corresponding author address: Bob Glahn, National Weather Service, Room 10214, 1325 East West Highway, Silver Spring, MD 20910; email: harry.glahn@noaa.gov

2. CONSISTENCY AMONG VARIABLES

There are certain consistency relationships that should obtain across variables. Perhaps the simplest one is that the dewpoint temperature cannot exceed the air temperature, and the forecasts should obey that relationship. In the Meteorological Development Laboratory (MDL), we have mitigated the possibility of inconsistencies by developing temperature and dew point equations "simultaneously," meaning that predictors are selected based on the RV to either, and then those exact predictors are put into each equation--one for temperature and one for dewpoint. This does not *guarantee* the least squares solution will produce consistent results, but is better than giving free reign to both variables. When inconsistent forecasts *do* occur, the forecasts are averaged and the average is used for both elements.

In addition, the temperature and dew point are developed in projection "blocks" that cover the periods of "daytime" maximum (max) temperature and "nighttime" minimum (min) temperature. The temperatures (and dewpoints) in the daytime block are developed simultaneously with the max temperature; the temperatures (and dewpoints) in the nighttime block are developed simultaneously with the min temperature. This helps to keep the predicted 3-hourly temperatures from exceeding the max and being lower than the min. Again, this does not guarantee consistency, and a postprocessing step is used to make them consistent. In cases of inconsistency, the max is adjusted to the maximum 3-hourly temperatures within the daytime period and the min is adjusted to the minimum of the 3-hourly temperatures in the nighttime period. Early testing determined that this procedure actually slightly increased the accuracy of the max and min overall (see Dallavalle et al. 1980). That is, in the cases of inconsistency, the max was, on average, too low, and the min was too high.

Equations for wind speed and u- and v-components are developed simultaneously. Since the speed is not computed from the components, this procedure is probably most important for direction forecasts. Even though the components can be used to calculate speed, such speeds will have a low bias (Glahn 1970).

Other possible inconsistencies also need to be addressed--for instance, total cloud amount and ceiling height; visibility and obstructions to vision; precipitation probability, quantitative precipitation, and precipitation type; and wind speed and wind gusts.

3. TEMPORAL CONSISTENCY

By temporal consistency, we mean the consistency among forecast projections of a weather variable made at the same time, e.g., temperature forecasts made for projections 6, 9, ..., 72 hours after a numerical model run postprocessed statistically from numerical model output. Because of high redundancy of model forecasts across elements and time, and because of the natural diurnal variability of some variables, such as temperature, temporal inconsistencies have not played a dominant role in the evaluation of such forecasts. However, when forecasts are made at hourly intervals, the expected diurnal change from one forecast to the next will not necessarily overshadow such inconsistencies, and some specific consideration is necessary to enhance consistency.

If each forecast projection is dealt with independently in the statistical process, different predictors may be picked by screening regression for the different projections. Many different combinations of predictors will yield nearly the same RV over the developmental sample, but the forecasts made from the resulting equations may exhibit different degrees of consistency/inconsistency from hour to hour depending on the predictors selected. The forecasts from the Localized Aviation MOS Program (LAMP; Glahn and Ghirardelli 2004; Ghirardelli 2005) are at hourly intervals.

Three methods of screening regression were tested in the LAMP wind direction and speed development to see whether they would produce significantly different degrees of consistency/inconsistency:

1) Each hourly projection was treated independently in the selection process and predictors could be selected from projections other than the predictand projection time,

2) each hourly projection was treated independently in the selection process, but predictors could be selected from only the predictand projection time, and

3) constraints imposed by the LAMP software were used in predictor selection.

Regression equations for three predictands were developed simultaneously for each projection--the u- and v-wind components and wind speed. The wind components were used to compute the wind direction. Potential predictors included the u- and v-wind components and speed of the initial observation and of the MOS synoptic scale forecasts, as well as other variables (see Wiedenfeld 2005). The MOS forecasts were included because LAMP is an update system, and it is desired that the forecasts at the later projections be quite consistent with MOS forecasts valid at the same time.

Due to LAMP's short range--hourly projections of 1 through 25 hours--the initial observation was chosen in all three selection methods. The synoptic scale MOS predictors were also chosen. The restrictions imposed by the LAMP software are essentially as follows:

1) A predictor will be selected on the basis of the highest RV contribution to any one of the predictands being dealt with simultaneously and to any one of the 25 hourly projections (1 through 25).

2) A predictor selected will be in the equations for all predictands (i.e., simultaneous development).

3) A predictor selected will be in the equation for each projection, except the projection of the predictor will be the same as the predictand (the predictors "march" with the predictand). The initial observation, cannot, of course, march--it remains the same in all equations.

4) A predictor selected for one projection may not contribute much to other projections (even with the marching option) and the coefficients would go to zero based on user controlled minimum RV and collinearity thresholds. When that happens, constraints are imposed such that:

a) A persistence predictor (initial observation), once it is eliminated from consideration for lack of sufficient RV for projection X will not be used for any projection greater than X. This assumes the initial observation is of most use in the early projections.

b) A MOS predictor, once it is eliminated from consideration for lack of sufficient RV for projection X will not be used for any projection less than X. This assumes the MOS predictors are of most use in the later projections.

c) For any other predictor, once it is eliminated from consideration for lack of sufficient RV for projection X, will not be used for any projection less than X or for greater than X, depending on the lesser number of projections for which it is eliminated. For instance, a predictor that has a zero coefficient for a predictand at projection 5 (20) will not be used for projections of 1 through 4 (20 through 25).

The result of 4) is that a predictor cannot "come in" and "go out" in a random fashion depending on very slight differences in additional RV.

A number of cases of wind speed forecasts were analyzed at to find temporal inconsistencies. Overall, all three methods were remarkably consistent, but a few cases of inconsistency were found. One such event was for Richmond, Virginia, on January 25, 2000 (see Fig. 1). This was a situation with strong northerly flow at the first projection (1000 UTC) tapering off to light northwest winds 24 hours later. All methods produced good direction forecasts (not shown). The speed forecasts were fairly good, starting off a little high. Free choice of all predictors [Method 1) above] gave the most non-verified variability, and the LAMP software [Method 3) above] the most consistent. The LAMP forecasts were also the most accurate during the projections of high variability. Despite finding only slight differences of temporal inconsistency in the wind direction and speed forecasts for these three methods, we believe the added complexity of the selection process in LAMP is worth the effort if only a small percentage of the forecasts are improved.

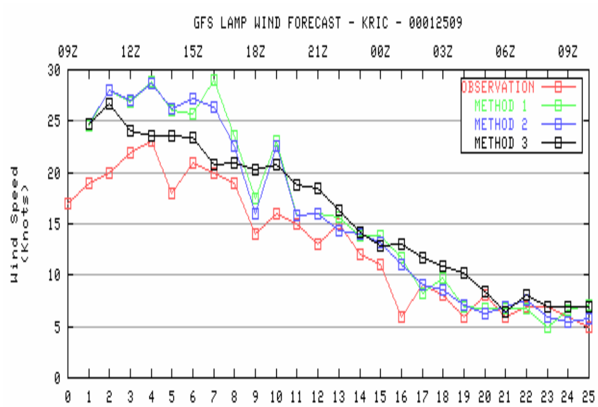

Figure 1. LAMP wind speed comparison for Richmond VA, January 25, 2000, for the three methods.

4. CONSISTENCY OF FORECASTS MADE AT DIFFERENT TIMES

It has been noted, especially by field forecasters, that the numerical models are not as consistent from run to run as desirable. Postprocessing statistically can remove the overall biases, but there may still be situation-dependent biases that remain. It is possible this can be successfully treated by a weighted average of MOS forecasts made for the same time from successive model runs, either by applying the weighting to the postprocessed forecasts, or by including predictors from successive model runs in the equations. MDL will be looking into these possibilities, now that the MOS forecasts (see Section 5) may play a more direct role in the total forecast production paradigm.

5. CONSISTENCY AMONG LOCATIONS

Spatial consistency is not usually a problem in any competent statistical system when forecasts are made at stations (observing sites). The stations are of such low density that any differences in the forecasts can be due to their separation. However, when forecasts are made for high density gridpoints, say ≤ 5 km spacing, then inconsistencies can easily show as not meaningful meteorologically.

With the advent of the National Digital Forecast Database (Glahn and Ruth 2003), guidance is needed for grids, not just for observation locations. Developing for a gridded system has its challenges (Dallavalle and Glahn 2005; Glahn and Dallavalle 2006). There is, in general, no predictand data set for gridpoints, and even if a field exists that can be used as a surrogate of the desired predictand (like radar data for precipitation), the non-normal distribution of the variable and the concentration on rare events almost dictate the approach be regional to insure a sufficient sample size for the method to be stable. By this, we mean all data within a region are combined and one equation set derived that can be applied to any and all points within the region. This essentially insures spatial consistency within the region, but unacceptable inconsistencies can occur at the regional boundaries. Two such examples are shown in Figs. 2 and 3.

Fig. 2 shows the probability of a thunderstorm in a 20-km square box for 19 hours after model initialization of 0900 UTC, July 27, 1997. While there is general agreement there will likely be thunderstorms in the Nebraska area, the boundary between two regions is very apparent and not suitable for providing to customers without some sort of postprocessing.

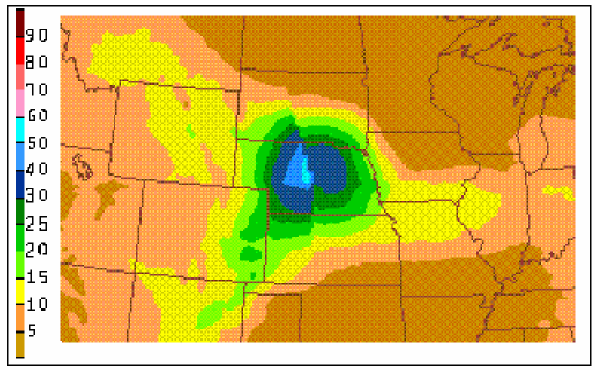

Figure 2. Probability of a thunderstorm in a 20-km box over a 2-h period ending 19 hours after 0900 UTC July 27, 1997.

Fig. 3 shows the probability of measurable precipitation (PoP) for a 36-h projection from January 15, 2005. The bands of probability do not come together pleasingly at the boundaries between the regions. It is obvious one boundary is along the eastern slopes of the Cascade and Sierra Mountains, and although it is reasonable the probability is lower on the eastern slopes, one might expect a transition zone parallel to the mountains would show up on a 5-km grid, rather than an abrupt change. Another boundary, oriented in an east-west direction, is in northern California with no terrain justification. Again, postprocessing is needed if these equations are to be implemented on a grid. The placing and viewing of MOS forecasts on a grid will undoubtedly play a major role in how our MOS development is done in the future.

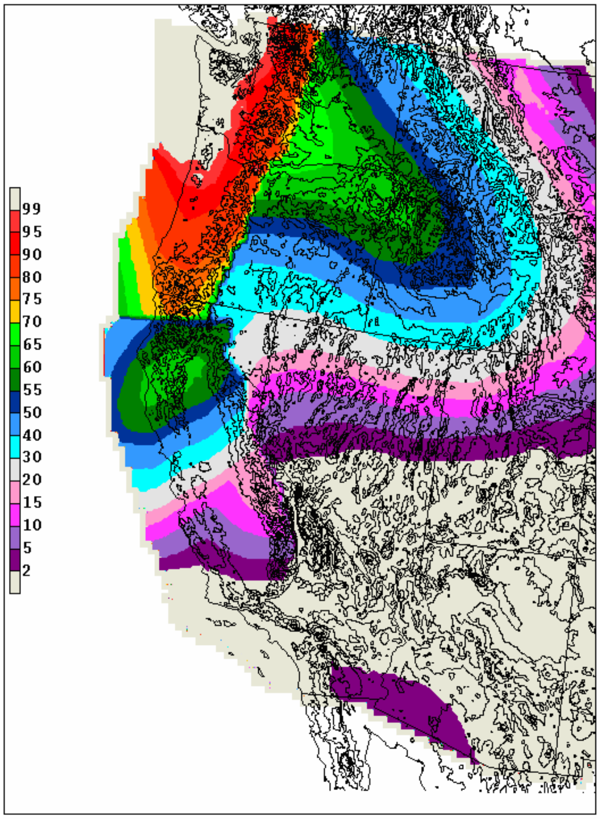

Figure 3. Probablity of Precipitation for a 12-h period ending 36 hours after January 15, 2005

6. FORCING OF PREDICTORS

Much of the development is done by screening predictors, as indicted in the Introduction. However, a useful technique is to "force" certain predictors that are deemed desirable to be in the equations. For instance, it is known that both the initial ceiling observation and the MOS forecasts of ceiling are going to be important to LAMP in predicting ceiling, so it may be advantageous to require them to be in the equation before the screening of additional predictors is started. Doing so in each region of a regional development will insure to some measure the same predictors in each set of equations, and may reduce discontinuities across regional boundaries. The forced predictors should still be subject to variance and collinearity constraints to help control instability of the equations.

In fact, the final LAMP wind development was done by forcing the u- and v-components and speed of each of the initial observation, MOS, and a LAMP geostrophic wind (Wiedenfeld 2004). This seemed to produce the best results.

7. SUMMARY

When developing and implementing a large statistical system--one that provides guidance for many weather elements, many projections, at several times per day--great care must be taken to insure that the guidance is not blatantly inconsistent. Methods will vary by weather element and will depend on whether the guidance is for relatively sparse locations or for a fairly dense grid.

Methods must be developed to make the forecasts consistent throughout a contiguous area (e.g., the conterminous United States or Alaska). One method is to use only one region; then there will be no boundaries. However, such an approach usually gives less than the possible accuracy or skill. Another approach is to somehow smooth the boundaries, possibly through an objective analysis procedure (Glahn and Dallavalle 2006). The latter, especially, is new ground, and MDL is working on solutions.

8. ACKNOWLEDGEMENTS

Many persons have contributed to the MOS and LAMP efforts on which this study is based. We thank those persons, and especially Scott Scallion for helping to examine many of the wind forecast case studies.

9. REFERENCES

Dallavalle, J. P., J. S. Jensenius, Jr., and W. H. Klein, 1980: Improved surface temperature guidance from the limited-area fine mesh model. Preprints *Eighth Conference on Weather Forecasting and Analysis*, Denver, CO,. Amer. Meteor. Soc., 1-8.

___, M. C. Erickson, and J. C. Maloney, III, 2004: Model Output Statistics (MOS) guidance for short-range projections. Preprints, *20th Conference on Weather Analysis and Forecasting*, Seattle, WA, Amer. Meteor. Soc., CD-ROM, 6.1.

___, and B. Glahn, 2005: Toward a gridded MOS system. Preprints, *21st Conference on Weather Analysis and Forecasting*, Silver Spring, MD, Amer. Meteor. Soc., 13B.2.

Ghirardelli, J. E., 2005: An overview of the redeveloped Localized Aviation MOS Program (LAMP) for short-range forecasting. Preprints *21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., 13B.5.

Glahn, H. R., 1970: A method for predicting surface winds. *ESSA Technical Memorandum* WBTM TDL 29, Environmental Science Services Administration, U.S. Department of Commerce, 18 pp.

___, B., and J. P. Dallavalle, 2006: Gridded MOS--Techniques, status, and plans. Preprints, *18th Conference on Probability and Statistics in the Atmospheric Sciences*, Atlanta, GA, Amer. Meteor. Soc., 2.1.

___, and J. E. Ghirardelli, 2004: The new and improved Localized Aviation MOS Program (LAMP) analysis and prediction system. Preprints, *20th Conference on Weather Analysis and Forecasting*, Seattle, WA, Amer. Meteor. Soc., J12.3.

___, and D. P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.*, **84**, 195-201.

Wiedenfeld, J. 2005: Localized Aviation MOS Program (LAMP): Statistical guidance of wind speed, direction, and gusts for aviation weather. Preprints *21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., P1.48.