# The Gridding of MOS

BOB GLAHN, KATHRYN GILBERT, REBECCA COSGROVE, DAVID P. RUTH, AND KARI SHEETS

*Meteorological Development Laboratory, NOAA/National Weather Service, Silver Spring, Maryland*

(Manuscript received 5 October 2007, in final form 2 July 2008)

ABSTRACT

Model output statistics (MOS) guidance forecasts have been produced at stations and provided to National Weather Service forecasters and private entities for over three decades. As the numerical weather prediction models became more accurate, MOS followed that trend. Up until a few years ago, the MOS produced at observation locations met the basic need for guidance. With the advent of the Interactive Forecast Preparation System and the National Digital Forecast Database, gridded MOS forecasts became needed as guidance for forecasters. One method of providing such grids is to objectively analyze the MOS forecasts for points.

A basic successive correction method has been extended to analyze MOS forecasts and surface weather variables. This method is being applied to MOS forecasts to provide guidance for producing grids of sensible weather elements such as temperature, clouds, and snow amount. Guidance forecasts have been implemented for the conterminous United States for most weather elements contained in routine weather forecasts. This paper describes the method applied to daytime maximum temperature over the conterminous United States and gives example results.

## 1. Introduction

Model output statistics (MOS) guidance forecasts have been produced at points and provided to National Weather Service (NWS) forecasters and private entities for over three decades (Glahn and Lowry 1972; Carter et al. 1989). As the numerical weather prediction (NWP) models became more accurate, MOS followed that trend (WMO 1999, 249–251). Up until a few years ago, the MOS produced at observation locations met the basic need for guidance, although there have been many requests for guidance at other specific locations; for instance, an observation location recently established that had no observational record sufficient with which to develop so-called single-station (SS) MOS equations.

With the advent of the National Digital Forecast Database (NDFD; Glahn and Ruth 2003) and the method of producing fields for it—the Interactive Forecast Preparation System (IFPS; Ruth 2002)—guidance is needed on a grid (specific values at grid points) rather than, or in addition to, guidance at observation locations. Because MOS is usually developed for specific observational locations, the question arose as how to produce a grid of MOS forecasts, the grid points being far more dense than the observational locations.[1] One possibility is to combine stations into ''regions'' and to develop a MOS equation[2] so that it can be applied to any point, and specifically at grid points, within the region. This method (see Lowry and Glahn 1976), sometimes called regional operator (RO), is used when the predictand has a highly skewed distribution, and the rare categorical ''events'' are those of most interest. However, generally, the accuracy of temperature forecasts produced by this method is less than for SS forecasts.

Another difficulty in applying RO equations to grid points is that discontinuities appear at the boundaries between the regions used to produce and apply the

*Corresponding author address*: Bob Glahn, Meteorological Development Laboratory, 1325 East–West Highway, Silver Spring, MD 20910.
E-mail: harry.glahn@noaa.gov

DOI: 10.1175/2008WAF2007080.1

[1] The NDFD grid over the conterminous United States has a grid length of 5 km, and there are plans to go to 2.5 km.

[2] MOS is not limited to linear regression, but we use ''equation'' in this paper because operational MOS forecasts have been produced by that method. Predictors are many times derived from NWP variables so that they have a nearly linear relationship to the predictand. These ''linearized'' predictors in a linear statistical model compose a highly nonlinear process and have been preferred to the use of nonlinear statistical models for operational use.

equations. This is true especially for a fine grid; discontinuities are less noticeable when applied to random and/or less dense points. Boundary discontinuities between sets of RO equations can be mitigated by smoothing, but this is seldom satisfactory. An obvious solution is to use one equation set for the entire area [i.e., generalized operator (GO) equations], then there are no boundaries. However, such forecasts that compete favorably in accuracy with SS forecasts have heretofore been elusive. Another method of getting values at grid points representative of SS guidance is to apply an objective map analysis technique to the SS guidance.

One of the first operational analysis methods employed was basically that described by Bergthorssen and Doos (1955) and implemented by Cressman (1959; all reference to Cressman in this paper will be to this publication). It was first applied to geopotential heights at one or a few levels of the atmosphere on a grid with a spacing of 381 km and was based on radiosonde data. Wind observations were used through the geostrophic relationship to assist in the height analyses.[3]

As the science and computer capabilities advanced together, many other much more sophisticated schemes were developed and implemented, and the term ''objective map analysis'' gave way to ''data assimilation.'' These schemes have been designed largely to furnish initial conditions for NWP models and have characteristics the target model favors. Because this was essentially an assimilation of observations above the earth's surface (e.g., radiosonde, satellite), dynamic relationships could be used to produce analyses consistent in the horizontal and vertical dimensions and, eventually, also in time.

This paper describes the Bergthorssen–Doos–Cressman analysis technique and how it has been expanded and is being used to produce MOS daytime maximum temperature grids over the conterminous United States (CONUS). The same technique is being used for MOS forecasts of other weather elements, but space does not permit a full discussion of the element-specific software options necessary for the other elements. The basic method can also be used to analyze surface temperature and dewpoint observations.

## 2. The BCD analysis method

The basic scheme proposed by Bergthorssen and Doos (1955) and implemented by Cressman has in the past been called the BCD method (Glahn 1985) to identify it with the three individuals primarily responsible for bringing it into mainstream meteorology. The BCD method is one of successive correction and consists of making one or more passes over the data being analyzed, each pass correcting the previous analysis based on data in the vicinity of the grid point being modified. This requires an initial analysis or first guess to correct on the first pass.

The first guess, for many applications, does not have much effect, especially when the data density does not vary substantially over the analysis area. In fact, we have found that a constant for a first guess produces just as good a result, or even better in many cases, than some other possible choice, such as an analysis made for the same variable 12 h, or even 1 h, earlier.

For each pass, the BCD method consists of interpolating into the grid resulting from the previous pass (or the first guess for the first pass) to get the value implied by the analysis at each data point. Interpolation can be by an appropriate scheme, with a bilinear approach being, in general, probably as good as any other for this purpose and being computationally less intensive than, for instance, a biquadratic method. For each datum location, the difference between the interpolated value and the datum is found; that difference is applied to surrounding grid points and is usually weighted by the distance between the datum location and the grid point to be corrected. Those individual corrections are then averaged over all data points within a radius of influence $R$ around the grid point. The BCD method as implemented by Cressman and many others employs four passes over the data, each pass with $R$ decreasing from the previous pass in order to capture more detail, and has included a smoothing pass after one or more of the corrective passes.

## 3. Extension of the method to BCDG

Almost any application of the BCD scheme is specific to the situation and may need to be modified from the Cressman implementation. In developing gridded MOS, the changes and extensions were so extensive to justify adding a ''G'' to BCD to distinguish it from the basic technique. We describe here the application to MOS daytime maximum temperatures.

The area over which the analyses are made is the area of the NDFD, with a border of about 50 km.[4] One of the

[3] Some of the details of the evolution of this method used at the National Meteorological Center (NMC, now known as NCEP) are contained in Glahn and Hollenbaugh (1969) and the footnote reference in it.

[4] A narrow border is desirable for computational purposes, including smoothing. The relatively wide, 50-km border was requested by forecasters.

major challenges with objective map analysis is highly varying data densities. In our case, the vast difference in density is synonymous with whether the forecasts are over water or over land. To account for this difference, we used for a specific pass a different $R$ for land than for water forecast points. In addition to that, one can expect that the temperatures will differ considerably between water and land, even if there were not a difference in the data density. We have dealt with this by letting land forecast points influence only land grid points, ocean water forecast points influence only ocean water grid points, and inland water forecast points influence only inland water grid points. As a refinement, some forecast points on shorelines were designated as both inland water and land and could influence both types of grid points. This accommodation to water and land points generated, in effect, three analysis *systems* in one, but with a common grid and analysis.

Another major change to the basic BCD method was to introduce an elevation dependency over land. The change of an element's value with elevation normally varies with its location on the grid, time of day, the day of the year, and the synoptic situation. The method implemented was to let the data (MOS forecasts) tell us the vertical relationship, called here the vertical change with elevation (VCE).

Finally, a contour-following smoother was implemented to selectively smooth the field after one or more passes. Its purpose is to smooth when the elevation does not markedly vary, but still retain the high detail when warranted. The details of this smoother are described in the appendix.

In the following discussion, the word ''station'' is used instead of ''data point'' or some other designation to clearly differentiate a quasi-random MOS forecast from a grid point, even though this station may be only an observation site.

### *a. Determining the VCE*

The VCE plays a major role in how a datum affects a grid point. Consider a station $A$ with a specific temperature value TA at elevation EA, and another station $B$ a short distance away with a temperature value TB at elevation EB. To apply a correction based on station A to a grid point near station B, one needs to consider that

$$\mathrm{VCE} = (\mathrm{TB} - \mathrm{TA})/(\mathrm{EB} - \mathrm{EA}).$$

The VCE is computed for each station for each analysis. The specific value for a station $A$ is based on several stations $B_i$ that are close in horizontal distance and far apart in vertical distance, with the more stations the better,

$$\mathrm{VCE}(A) = \Sigma(\mathrm{TB}_i - \mathrm{TA})/\Sigma(\mathrm{EB}_i - \mathrm{EA}),$$

over all designated close stations $B_i$.[5]

Finding the stations $B_i$ is computationally expensive, so a list of ''pairs'' is found for each station A by preprocessing the metadata (locations and elevations). The preprocessing algorithm selects only stations within about 340 km of station $A$ with an elevation difference of at least 130 m; these specific values were found to produce lists of pairs appropriate for the analysis process.[6] The process is quite robust, computing only one parameter from several pieces of data.

### *b. The gridpoint correction algorithm*

Without an elevation correction, a station's contribution $D$ to the correction at a grid point *before distance weighting is applied* is simply the difference between the station value $S$ and the grid value interpolated to the station, BB:

$$D = S - \mathrm{BB}. \tag{1}$$

The elevation correction used is

$$D = (S - \mathrm{BB}) + \mathrm{VCE} \times \mathrm{ELEDIF}, \tag{2}$$

where VCE is the temperature change with elevation calculated for that station and ELEDIF is the elevation of the grid point minus the elevation of the station.

However, if an elevation correction were applied on successive passes without considering the lapse rate already existing in the analysis, it would greatly overcorrect. To consider the existing temperature change with elevation, the correction is

$$D = (S - \mathrm{BB}) + (\mathrm{VCE} - \text{existing change with elevation}) \times \mathrm{ELEDIF}. \tag{3}$$

The existing change with elevation between the grid point and the station, as determined from the analysis, is

$$(\mathrm{G} - BB)/\mathrm{ELEDIF},$$

where $G$ is the value at the grid point, so

$$\begin{aligned} D &= (S - \mathrm{BB}) + (\mathrm{VCE} \times \mathrm{ELEDIF}) \\ &\quad - [(\mathrm{G} - \mathrm{BB})/\mathrm{ELEDIF}] \times \mathrm{ELEDIF} \\ &= S - G + \mathrm{VCE} \times \mathrm{ELEDIF}. \end{aligned} \tag{4}$$

[5] For the improbable case when the denominator is zero, VCE is set to zero.

[6] Any number of algorithms could be derived to produce appropriate lists; space does not permit a full explanation of this particular preprocessing algorithm.

When the first guess is a constant, this is highly effective, because the existing change with elevation is zero, BB = $G$, and (4) devolves to (2).

We found that using only a partial elevation adjustment (ELCORR) as the radius was decreased on later passes gave better results than a full adjustment on all passes; this partial correction is

$$D = (S - \mathrm{BB}) + (\mathrm{BB} - G + \mathrm{VCE} \times \mathrm{ELEDIF}) \times \mathrm{ELCORR}, \quad (5)$$

where ELCORR $< 1$. When either VCE or ELEDIF = 0, (1) is used instead of (5).

Even though the existing VCE is removed (partially when ELCORR $< 1$) when the correction for a station is applied, the correction has to be made on an early pass because it may be that there will not be any stations close enough to the grid point to affect it on a later pass, and no elevation correction to that grid point would result. It is important to recognize that VCE is not a free-atmospheric lapse rate, but a value that is appropriate to BCDG in the *immediate vicinity* of the station for which it is calculated.

### c. Accommodation for land and water

Each grid point is designated as inland water, ocean, or land, and each station is designated as inland water, ocean, land, or both land and inland water. The latter designation (both land and inland water) is used rarely and only when a coastal forecast may be as representative of water as of land.

We had to make sure each water grid point had at least one water station within the largest $R$, or it would not change from its first-guess value. Because we have MOS forecasts for inland water for only the Great Lakes and the Great Salt Lake, those are the only inland water bodies we presently treat; otherwise, other inland water bodies are treated as land. An exception is that Lake Pontchartrain is treated as ocean water. We also had to make sure each land grid point had stations within the largest radius of influence in order to modify it from the first-guess value; this was not as great a problem because of the greater density of land stations. (An island with no station and no nearby coastal land station could be a problem, but has not been for our application so far.)

For maximum temperature, the ocean grid points could be assigned values from the forecasts at the few buoys by increasing the radius of influence $R$ to be large enough so that every grid point is modified by at least one point on at least the first pass. Given the consistency of the ocean forecasts up and down the coast, this is adequate.

Because of the land–water treatment, a specialized interpolation is used when the value at the station is to be inferred from the current analysis. This is bilinear when all four grid points surrounding the station are of the same type as the station (land or water) or can be used as the same type (both). When this condition is not met, the closest grid point of the four of the matching type is used.

## 4. Options in the BCDG software used for MOS maximum temperature analyses

Any full-bodied analysis system will have many parameters that can be used to tune for data density relative to gridpoint density, variation in data density over the grid, first-guess possibilities, error characteristics of the data, and smoothness versus detail desired in the analysis. In the BCDG software, the combination of possibilities is essentially limitless, and only a small fraction can be extensively tested. Previous experience was drawn upon to define reasonable choices for testing (Cressman 1959; Glahn et al. 1985; Glahn and Hollenbaugh 1969).

### a. First-guess option

As mentioned previously, a constant can be used as a first guess; we used the average of all maximum temperature values to be used in the analysis. We tried a first-guess grid composed of forecasts from GO equations, but the difference in the results was barely distinguishable, and the use of a constant is much simpler for operational implementation.

### b. Radii of influence and number of passes

Considerable experimentation confirmed previous experience (Cressman 1959; Glahn et al. 1985) that four passes were necessary and sufficient to capture the desired detail. The largest radius of influence was determined such that every grid point would have a correction made for it; the smallest must be such that the analyses are not unduly "spotty," showing more detail than a skilled meteorologist would accept as real. For maximum temperature over land, the values for passes 1–4 are 37.0, 27.5, 20.0, and 15.0 grid lengths, respectively, corresponding to 185, 137.5, 100.0, and 75.0 km on a 5-km grid. For water, these values were increased by a factor of 3.5 to accommodate the very sparse MOS forecasts.

### c. Quality control of the data

A central part of the analysis of the data is to not use a datum if it is obviously incorrect. This is a determination the software makes on each pass based on an acceptable difference (threshold) between the station value and the

value interpolated from the analysis. Obviously, this acceptable difference can be less for the later passes and depends, for the first pass at least, on the first guess used. This procedure can be quite effective for a rather smoothly varying field in space. However, for a field that has considerable finescale detail, incorrect data are hard to determine. Time-dependent error checking is not used in the analysis, but could be done in a preprocessing step. We determined the thresholds for the maximum temperature to be 60.0°, 25.0°, 21.0°, and 18.0°F for the four passes, respectively, based on extensive testing and meteorological judgment as to what were MOS values that contributed negatively to the quality of the analysis.

The procedure in BCDG is this. On each pass, before a correction is made based on a station's value, the difference between the datum and the interpolated value is found. If it exceeds 1.5 times the specified threshold, the datum is discarded for that pass; if it exceeds the threshold, but is less than 1.5 times the threshold, then the two closest neighbors are found. If either one of the two neighbors does not meet 0.6 times the threshold and the errors of both the station being checked and its neighbor are of the same sign, then the station is accepted. If this check does not allow the station to be accepted, a value is estimated from each neighbor by using that neighbor's datum and an average of the station's and neighbor's VCEs, as determined from the analysis, to estimate the station's value, and if either of these agree with the station within 0.6 times the threshold, then the datum is accepted. Otherwise, the datum is not used on this pass, but could be on subsequent passes. In addition, if the station is accepted, the station causing it to be accepted is also accepted on this pass if it meets 1.5 times the threshold. This computer-intensive closest-neighbor checking is rarely needed, but is highly effective when required and used. These values were determined by extensive testing and are in agreement with previous experience (Glahn 1985).

While it may seem unnecessary to quality control MOS forecasts, there are cases when, for some reason, a seemingly bad value is produced. This could occur in the early projections when an observation is used as a predictor and is in error. Or it could occur when the predictors for a particular situation lie outside the developmental sample space, and the equation does not give good results. Problems with data formatting and/or transmission can also cause unexpected errors.

### d. Type of correction

The type of correction $C$ can be specified by pass. There are three options:

$$C_1 = \frac{1}{n}\sum_{i=1}^{n} D_i, \text{ type1;}$$

$$C_2 = \frac{1}{n}\sum_{i=1}^{n} W_i D_i, \text{ type2; and}$$

$$C_3 = \frac{\frac{1}{n}\sum_{i=1}^{n} W_i D_i}{\frac{1}{n}\sum_{i=1}^{n} W_i}, \text{ type3,}$$

where

$$W_i = \frac{R^2 - d_i^2}{R^2 + d_i^2},$$

$d_i$ is the distance between the station and the grid point, $n$ is the number of stations affecting the grid point, and $D_i$ is the correction due to an individual station (see section 3b).

The first of these (type 1) weights each station's contribution to a change to the grid point equally. The second (type 2) weights the station's contribution according to the distance from the grid point; this emphasizes the closer stations but has the effect of decreasing the change made to the grid point because $W_i \leq 1$. The third (type 3) puts the sum of the weights in the denominator and increases the correction to the grid point over a type 2 correction. It was found that type 3 gave the best result, in agreement with previous work (Glahn 1985).

### e. Limiting positive VCE values

Usually, maximum temperature decreases with elevation. However, a significant number of the calculated VCE values are positive. These are legitimate and occur with inversions, and can occur along the western seacoast where the coastal temperatures are cooler than the temperatures in the nearby low mountains. To use a positive value in the immediate vicinity of the station is correct, but such a correction is not appropriate at greater distances. Therefore, positive VCEs are used, but are limited in horizontal extent to five grid lengths.

### f. Smoothing

The basic smoothing algorithm taken from Thomasell and Welsh (1962) is a generalization of the one used by Cressman:

$$\text{SG} = (G + bA)/(1 + b),$$

where SG is the smoothed value at a grid point, $G$ is the original value at that same grid point, $b$ is the smoothing parameter, and $A$ is the average of the four surrounding grid points.

The parameter $b$ can be specified by pass, and when $b = 0$, no smoothing is done. With $b = 1$, this is the same as what Cressman used; with $b = 2$, the average is weighted twice as much as the point being smoothed.[7] For the special smoothing along a contour, the same formula is used, but the average $A$ is computed from only the neighboring points along the ridge or valley. This algorithm can be applied to a nine-point stencil in a similar manner.

The contour-following smoother was found to be effective for not smoothing across ridges and valleys, and was implemented with $b = 4$ when a ridge point was higher than the two neighboring points perpendicular to the ridge by at least 100 m and when the valley point was lower than its two neighboring ''across valley'' points by at least 100 m. This smoother is applied once after the last pass. The 100-m value was determined by testing and is appropriate for not smoothing across valleys such as Death Valley on the 5-km grid we are using.

## 5. Assessing the quality

If one knew the correct value at each grid point, then some error measure, such as the mean absolute error (MAE), could be calculated and either displayed as a map or a summary statistic used to judge the correspondence between the analysis (grid) and the correct values. Of course, this is impossible, because the correct value is not known, which is the very reason an analysis is being done in the first place. Lacking this, there are two primary ways the quality of an analysis can be judged.

An objective way is to interpolate into the analysis to get a value at each data point used in the analysis and calculate a summary error measure. The problem is that almost any decent analysis procedure can fit the data very closely and still be poor where there are no data.

A viable option is to withhold a few data points at random as the analysis is being done, then calculate a summary error measure at just those points. The number of points withheld relative to the total number must be small enough to not materially affect the analysis. Replication can be used in which the withheld points are a different random set.

[7] It is noted that large values of b, indicating heavy smoothing, can change the phase of very small-scale variations.

Especially given that there is no good objective way to judge the quality at all grid points, another criterion is the extent to which a graphic produced from the grid ''looks meteorological.'' After all, among the many uses of a gridded product is to view it as a graphic.

Both subjective viewing and withheld data tests have been employed in assessing the efficacy of BCDG, as indicated below.

### *a. Meteorological appearance and consistency*

In most cases, the large-scale meteorological appearance of the graphics produced from individual grids was judged to be good by experienced meteorologists.[8] However, it was found by looping the graphics from the analyses produced from run to run (0000 and 1200 UTC cycles) that forecasts valid at the same time, though extremely similar in broad scale, ''pulsed''; that is, there were spots or small areas that were considerably different from run to run. It was also found that some small areas were consistently too cold or too warm.

The pulsing was caused by a combination of three factors:

1) There were MOS forecasts for some stations at one cycle but not the other; the analysis process would have data for a small area for one cycle, but not the other, so the same result could not be achieved at such points from run to run.
2) The MOS equations for the two cycles were not developed at exactly the same time and the predictors in the equations are not the same from run to run; therefore, the equations could give different answers for that reason alone.
3) The GFS model and its data assimilation on which the equations are based may perform slightly differently at different cycles, causing the MOS equations to give differing results.

A yo-yoing tendency in NWP guidance has been noticed by forecasters for years. As an attempt to correct this tendency, we merged two adjacent cycles. Specifically, the forecasts for the current cycle are merged with the forecasts from the previous cycle valid at the same time. That is, a 36-h forecast from the current cycle is analyzed together with the 48-h forecast from the previous cycle. This makes sense within an ensemble framework; we are in effect postprocessing an ensemble of two. This combining of cycles would have not been a good practice a decade

[8] MOS forecasts generally represent synoptic-scale conditions forecast by the driving numerical model, modified by seasonal and diurnal climatic factors specific to that location. The density of MOS forecasts does not allow a depiction of the sharp discontinuities such as fronts or drylines down to the 5-km scale.

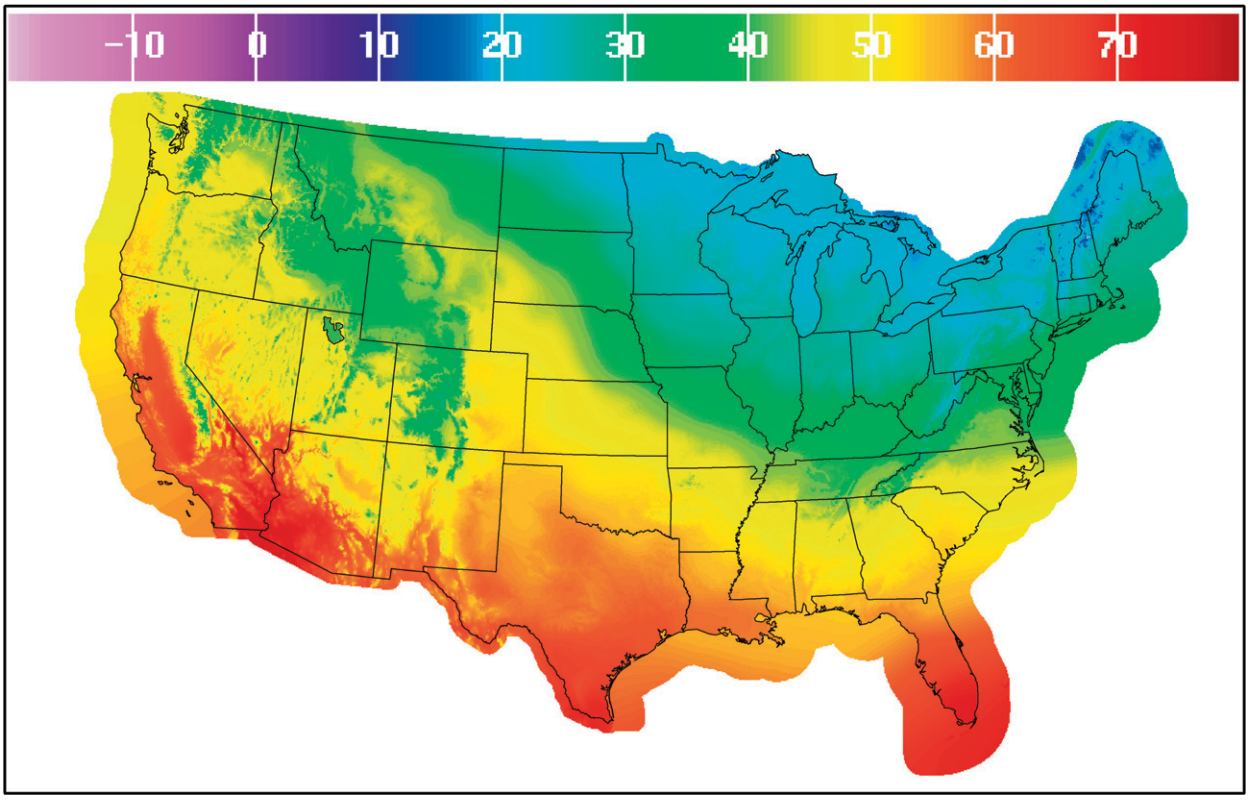


FIG. 1. Analysis of 48-h temperature MOS predictions (°F) based on 1200 UTC 26 Jun 2007 data.

ago, but NWP forecasts, and their associated MOS, have become of such quality that the differences between forecasts from cycles 12 h or less apart are small enough that the benefits outweigh the disadvantages. Withheld data tests have shown that doing so did not appreciably affect the accuracy of the forecasts, and the results were much more pleasing; the consistency from run to run was much improved. The tests reported in section 5b below were made with this average.

An example of a gridded maximum temperature MOS forecast is shown in Fig. 1 over the CONUS; Fig. 2 is an analysis of the same data without the land–water distinction and VCE adjustment. The differences in the two figures are striking. Both capture broad-scale features of the MOS forecasts, but Fig. 2 is ''blurry'' and does not account for the terrain features except in very broad terms. In contrast, the terrain is very well defined in Fig. 1. Some of the features in Fig. 1 that do not show up well in Fig. 2 are the Olympic Mountains, the Snake River valley in southern Idaho, the Bighorn Mountains in north-central Wyoming, Death Valley in southern California, the Grand Canon in northwestern Arizona, the ridge–valley pattern in Nevada, and the Columbia River and its tributaries from the north in Washington.

The water–land distinction is apparent in a few places in Fig. 1, but not in Fig. 2. For instance in Fig. 2, the ocean temperatures off the coast of southern California are reflective of the land, and the extension of land temperatures into the Gulf of Mexico from the coasts of Texas and Louisiana is noticeable. On the other hand, the temperatures at the buoys provide good ocean temperatures at those locations in Fig. 1 when a distinction is made between land and water. In Fig. 2, the Great Salt Lake is not defined, but in Fig. 1, it takes the temperatures of forecasts over the lake. There is also a distinction between southern Lake Michigan temperatures from the surrounding land in Fig. 1 but not in Fig. 2.

### *b. Withheld data tests*

A number of sets of MOS maximum temperature forecasts were analyzed, each with withheld data. The forecasts were made from the 1200 UTC run of the National Centers for Environmental Prediction's (NCEP) Global Spectral Model on 1 day each month from September 2005 through August 2006. The cases were chosen such that some portion of the grid might be a challenge for BCDG. For each of these 12 dates,

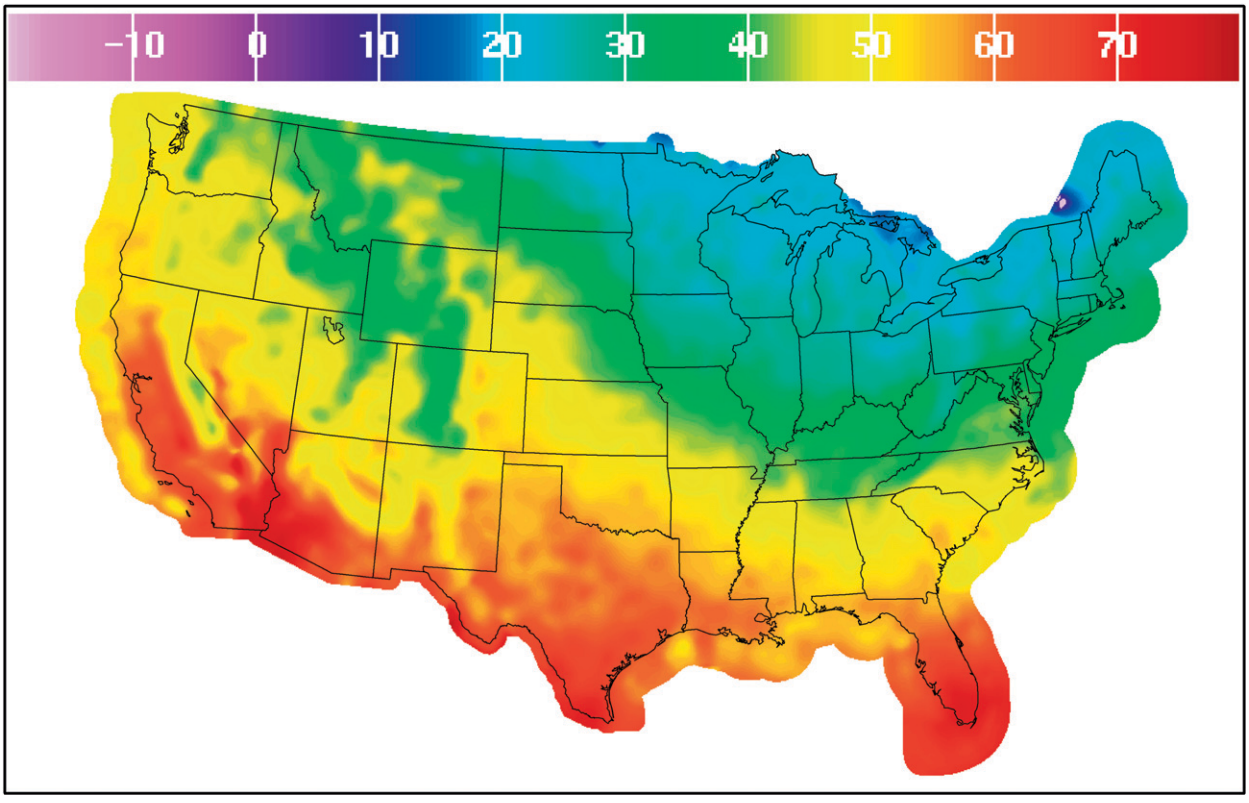


FIG. 2. Analysis of the same data as in Fig. 1, but here the land–water distinction and the VCE adjustment were not made.

maximum temperature forecasts were analyzed at projections every 24 h out to 186 h (7 days), for a total of seven projections. Analyses were made over the CONUS and also over only the West, with the West being defined as west of 105°W. For the CONUS runs, 20 stations were randomly withheld from each analysis; 10 stations were withheld for the runs over the West. The number of analyses made over the CONUS and over the West numbered 84. The total number of stations withheld over the CONUS was 1680 and over the West it was 840; this is on the order of 0.2% of the stations being analyzed, so it is likely the analyses were not affected except in the immediate vicinity of the withheld stations.

The randomization was done by randomly selecting the $X$ and $Y$ positions of the point to withhold on the grid (effectively randomly selecting the location on the grid), and then the closest land station to that point was withheld. This resulted in a reasonably uniform distribution of withheld points.[9]

After the analyses were completed, interpolations into the grids were made for both the stations used in the analyses and at the withheld station locations, and an MAE was calculated for each. The MAEs at the analyzed stations indicate the error if one were to attempt to recover the station's value from the grid. The MAEs at withheld stations indicate the average error one would get at random points on the grid. Both of these are measures of the quality of the analyses, especially the latter.

From this and other tests, it was determined that four passes were best in order to appropriately take account of the VCE and also fit the data adequately. This was determined by both analyzing errors and by viewing the resulting maps. Table 1 shows the results for pass 4 for both the withheld and nonwithheld stations over the CONUS and over the West.

From Table 1, it can be seen that

1) The elevation correction improved the analysis by 0.3°F or more at both withheld and nonwithheld (analyzed) stations over the CONUS and by about 1.0°F at the analyzed stations and by about 0.9°F at withheld stations over the more mountainous West.

[9] Withholding stations randomly selected from the list available produces biased results; too many stations are withheld in regions of dense data.

TABLE 1. The MAEs (°F) of MOS maximum temperature forecasts interpolated from the grid to stations used in the analysis and also to withheld stations with and without elevation correction for the CONUS and over only the mountainous West.

| Correction method | CONUS | | West | |
|---|---|---|---|---|
| | Analyzed | Withheld | Analyzed | Withheld |
| Elevation | 1.5 | 1.7 | 2.1 | 2.3 |
| No elevation | 2.0 | 2.0 | 3.1 | 3.2 |

2) The elevation correction afforded an average MAE of 1.7°F over the CONUS and 2.3°F over the West at withheld stations.

Bear in mind, the analyzed values were interpolated to calculate the MAEs, and this process also contains errors. Also, the calculated values may not be indicative of errors at very high elevations where there are few observations.

## 6. Current status of gridded MOS

The technique described here was made operational over the CONUS on a 5-km grid on 15 August 2006, not only for daytime maximum temperature and nighttime minimum temperature out to 7 days, but also for 2-m temperature and dewpoint, wind speed and direction, and relative humidity every 3 h out to 192 h; 3-h probability of thunderstorms out to 84 h; and 6- and 12-h probabilities of precipitation and 6- and 12-h probabilities of thunderstorms out to 192 h. Then on 5 June 2007, additional elements were added to the gridded MOS package: sky cover and wind gusts every 3 h out to 192 h, 24-h snow amount out to 132 h, and 6- and 12-h quantitative precipitation estimates out to 156 h.

The same basic method described here for maximum temperature is used for 3-h temperature and dewpoint, and minimum temperature; the other elements required some auxiliary procedures for which space here does not allow description. Relative humidity is calculated from the temperature and dewpoint, and a different technique is used to produce the probability of thunderstorms; these and all of the gridded MOS guidance products are available in graphical and binary formats from the gridded MOS information Web page (http://www.weather.gov/mdl/synop/gmos.php).

## 7. Future enhancements

Attention is now being placed on providing gridded guidance for Alaska, Puerto Rico, Hawaii, and Guam. The resolution will be at approximately 3.0 km. Each of these areas offers new challenges, and adaptations are being made to the technique. A method to recognize rain shadows in the lee of mountains will be developed and implemented, as will a method of modifying temperatures along coastlines to account for advection by MOS winds. The latter is needed because the density of the stations is not sufficient to account for the amelioration of contrasts between water and land under windy conditions. A radius of influence $R$ that varies by station will be incorporated. When the elements discussed above have been implemented for Alaska, Hawaii, Puerto Rico, and Guam, we will decrease the analysis grid length from 5.0 to 2.5 km for the CONUS.

The interelement consistency of the temperature variables is handled by 1) preprocessing to make the station values consistent and 2) a postprocessing step to make the 3-h temperature and dewpoint grids consistent. No other consistency checking is done. A more thorough consistency package is being built into BCDG that will consist of checking appropriate 3-h temperature values with daytime maximum and nighttime minimum temperatures at each grid point.

## 8. Conclusions

A viable way to analyze point data, including those points in rough terrain and in regions with high variability of data density, has been developed and described in this paper—specifically for maximum temperature. An average VCE is computed at each station by comparing neighboring values of the same weather element at different elevations, and those VCEs are applied in the analysis. Except for the lists of neighbors, the calculations are done on the fly, and they adapt well to changing synoptic situations, time of day, and day of year. This elevation correction is indispensable.

Gridded MOS is available online as part of the National Digital Guidance Database (NDGD; http://weather.noaa.gov/pub/SL.us008001/ST.opnl/DF.gr2/DC.ndgd/GT.mosgfs) for most of the elements in the NDFD and can be used as guidance by NWS forecasters in the IFPS process. These fields are spatially consistent and do not, in general, exhibit nonmeteorological discontinuities.

It was found that applying the analysis process to two consecutive 12-h cycles provided temporal continuity that analyzing only one cycle could not at this time provide. This ensemble approach was necessary because of the slightly different characteristics of the MOS station forecasts at the two cycles.

*Acknowledgments.* The provision of gridded MOS is totally dependent on the dedicated efforts of the persons in the Statistical Modeling Branch of MDL, led by Paul Dallavalle for many years, in acquiring and quality

controlling the voluminous data, developing the myriad equations, and making all software operational. Scripts provided by Judy Ghirardelli and Jerry Wiedenfeld for producing the many graphics for subjective quality determination were indispensable. We also thank Michael Schenk for performing the independent data verification. The lead author acknowledges the rekindling of his interest in on-the-fly calculations by Cliff Mass in a National Weather Service science and technology seminar he gave in Silver Spring, Maryland. This paper is the responsibility of the authors and does not necessarily represent a position of the National Weather Service or of any other governmental agency.

## APPENDIX

### The Contour-Following Smoother

In general, the smoother is applied on a five- or nine-point stencil centered on the grid point being smoothed. See section 4f for a definition of the ''normal'' smoother. The decisions in the smoother are made for each point and in the following order. Once a criterion is met, the following ones are not exercised.

1) When the point is a land point and is judged to be an island or spit of land, no smoothing is done.
2) When the grid points in the nine-point stencil are all water, smoothing is normal over the nine-point stencil (see section 4f).
3) When criterion 2 above is not met, then the grid points in the five-point stencil are all water, smoothing is normal over the five-point stencil.
4) When the grid points in the five-point stencil include both land and water, very light smoothing is done ($b = 0.25$). This allows some slight adjustment near the coast, and essentially lets the ocean temperatures (in the case of temperature) affect the near inland points, and vice versa.
5) When all points are land and the elevation of each point in the nine-point stencil does not vary by more than 200 m from the average (AVG) elevation of the points in the nine-point stencil, then the smoothing is normal over the nine-point stencil.
6) When all points are land and criterion 5 is not met, then the elevation of each of the points in the five-point stencil does not vary by more than 200 m from AVG, then the smoothing is normal over the five-point stencil.
7) When the center point is higher in elevation than all of the other four points in the five-point stencil, it is not smoothed.
8) When the center point is lower in elevation than all of the other four points in the five-point stencil, it is not smoothed.
9) When the elevation of neither of the two neighboring points in the $x$ direction from the point being considered does not differ by more than 200 m from that point, this is decreed to be a ridge, valley, or contour on a slope and smoothing is only over the three points involved. This is essentially smoothing along contours—hence the term ''contour following.''
10) The same process as in step 9 is applied except along the $y$ axis.

11 and 12) The same process as in steps 9 and 10 except in each diagonal direction. For the diagonal, the 200-m constant is increased by a factor of 1.414 to account for the greater distance between the points in those directions. If both criteria 11 and 12 indicate that smoothing can be done, all five points are used in the smoothing.

13) If none of the above tests is met, the contour through the point is judged to be strongly curved and is not smoothed.

## REFERENCES

Bergthorssen, P., and B. R. Doos, 1955: Numerical weather map analysis. *Tellus,* **7,** 329–340.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Wea. Forecasting,* **4,** 401–412.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374.

Glahn, H. R., and G. H. Hollenbaugh, 1969: An operationally-oriented small scale 500 mb height analysis program. ESSA Tech. Memo. WBTM TDL 19, U.S. Department of Commerce, 17 pp.

——, and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

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

——, T.L. Chambers, W.S. Richardson, and H.P. Perrotti, 1985: Objective map analysis for the Local AFOS MOS program. NOAA Tech. Memo. NWS TDL 75, 34 pp.

Lowry, D. A., and H. R. Glahn, 1976: An operational model for forecasting probability of precipitation—PEATMOS PoP. *Mon. Wea. Rev.,* **104,** 221–232.

Ruth, D. P., 2002: Interactive forecast preparation–The future has come. Preprints, *Interactive Symp. on the Advanced Weather Interactive Processing System (AWIPS)*, Orlando, FL, Amer. Meteor. Soc., 3.1. [Available online at http://ams.confex.com/ams/pdfpapers/28371.pdf.]

Thomasell, A., Jr., and J.G. Welsh, 1962: Objective analysis of sea level pressure and surface temperature, dewpoint and wind. The Travelers Research Center Tech. Pub. 19, Contract FAA/BRD-363, Hartford, CN, 87 pp.

WMO, 1999: *Numerical Weather Prediction Progress Report for 1998.* Tech. Doc. WMO/TD 968, 261 pp.