NOAA Technical Memorandum
NWS MDL 85

ERROR ESTIMATION OF OBJECTIVE ANALYSIS OF SURFACE OBSERVATIONS

Bob Glahn
Jung-Sun Im

Meteorological Development Laboratory
Silver Spring, MD
November 2011

UNITED STATES
Department of Commerce
John E. Bryson
Secretary

National Oceanic and
Atmospheric Administration
Jane Lubchenco
Under Secretary

National Weather Service
John L. Hayes
Assistant Administrator

Table of Contents

ERROR ESTIMATION OF OBJECTIVE ANALYSIS OF SURFACE OBSERVATIONS

Bob Glahn and Jung-Sun Im

ABSTRACT

As part of the Localized Aviation Model Output Statistics (MOS) Program (LAMP), the Meteorological Development Laboratory is analyzing surface data reports on an hourly basis. The Bergthorssen-Cressman-Doos-Glahn (BCDG) analysis program being used for gridding MOS forecasts has been tailored to analyze surface observations. It is desired to know the errors involved in these analyses. While the actual errors are unknowable for several reasons, they can be estimated. Withheld data tests can be made to get an estimate of overall errors, or more correctly put–differences between the data values and the analyses. On any given analysis, one would expect the errors at a specific location to be a function of some knowable parameters, such as distances between the reporting locations and the grid points, the terrain roughness, the density of reporting locations, and the variability of the data values, all in the immediate vicinity.

We have made analyses of surface temperature and dewpoint over the conterminous United States every fifth hour for one year. On each analysis, 20 land stations and one water station were randomly withheld from the available data. For each withheld datum, the analysis value at that site was estimated by bilinear interpolation. The differences between these interpolated values and the actual observations were related to knowable parameters through least squares regression, one relationship for land and another for water, for each variable. These regression equations were then applied, respectively, to each land and water grid point for a specific hour. This produced an estimate of the error of the analysis for each grid point for that specific time. This paper describes the process and shows results.

1. INTRODUCTION

Analyses of surface-based meteorological observations have many uses. Such analyses are part of the assessment of the synoptic situation necessary for weather forecasting; this is as true today as it was 60 years ago. The methods of analyses have, however, changed since then. In the mid 1950's, data were being plotted by hand on maps and the analyses performed by humans. Since then, various methods of automated analyses have been developed in concert with the development of the digital computer. Summarizing a mesoscale conference, Horell and Colman (2005) state, "Thus, the NWS has an immediate and critical need to produce real-time and retrospective analyses at both a high spatial and temporal resolution in order to create the NDFD forecasts as well as to verify their accuracy."

Objective (i.e., computer produced) analyses of meteorological data was being considered even before 1950. One of the first objective techniques to appear was the least squares fitting of a polynomial over a fairly large area to the data (Panofsky 1949). Although this was refined by Gilchrist and Cressman (1954) to fit the data over a small area and was actually used in early numerical weather prediction experiments, it never became widely used. George Cressman, the first director of the National Meteorological Center (now the National Centers for Environmental Prediction, NCEP) and then director of its forerunner the Joint Numerical Weather Prediction Unit of the Weather Bureau, recognized the potential for a technique developed by Bergthorsson and Doos (1955), and put a version of that into operation for analyzing upper air geopotential heights (Cressman 1959). This successive correction technique consisted of making multiple passes over the data, correcting each grid point on each pass by the data in the immediate vicinity, immediate vicinity being defined by a radius of influence which was constant over the analysis area, but was decreased for successive passes. A very similar technique, basically differing only in the distance weighting factor, was proposed by Barnes (1964) and has been used extensively. Barnes (1964) method was proposed to be used with one or two passes. He suggested "...direct application of the scheme to obtain maximum detail in regions wherein the data densities vary considerably is not recommended." Achtemeier (1989) suggested Barnes scheme be extended to three passes.

Other very sophisticated methods of analysis, now called data assimilation, have been developed and are in operation at national centers worldwide for providing initial conditions for numerical models. These latter methods employ relationships among free atmospheric variables that are not as effective for use with variables observed at the earth's surface, although they have been adapted for surface variables and are used for the Real-Time Mesoscale Analysis (RTMA; Pondeca and Manikin 2009; Manikin and Pondeca 2011) being run at NCEP.

The method as used by Cressman has been used extensively in the Local Aviation MOS Program (LAMP) and has been called the BCD method (see Glahn et al. 1985). As part of the Meteorological Development Laboratory's support to the aviation community, and the Next Generation Air Traffic Control System Program (NextGen) in particular (see Ghirardelli and Glahn 2010; 2011), we have further developed BCD, which is called now BCDG, the initials of the names of the primary developers (Bergthorssen, Cressman, Doos, Glahn). This method has been described by Glahn et al. (2009) for the analysis of MOS forecasts and has been adapted for analysis of surface observations. Although the basic BCDG method is described in Glahn et al. (2009), there were a number of changes necessary and improvements for analysis of observations. These are described in Im et al. (2010, 2011) and Glahn and Im (2011).

While a considerable amount of effort has been placed on analysis methods, a more modest effort has been put on estimating the analysis error associated with a particular method. Characteristically, errors associated with analysis schemes have been estimated with idealized data (e.g., a combination of sinusoidal waves) and/or upper air data where the patterns are relatively smooth. Difficulties in making a good analysis and a good estimate of its error are evident by the exchanges between Smith and Leslie (1984) and Glahn (1987), Goodin et al. (1979, 1981) and Glahn (1981), and Fritsch (1971) and Glahn and McDonell (1971). The RTMA process (Pondeca et al. 2011) includes an estimate of the analysis error, the method being specific to the analysis process. Myrick and Horel (2008) have studied the sensitivity of surface analyses to a particular type of observation.

If one is going to estimate "analysis error," that error needs to be defined. One could be very interested in the location of fronts or other discontinuities, and not be overly concerned about "bland" areas. If such were the case, then a method would need to be defined that concentrated on that aspect. If one is concerned about making derivative calculations, then the method proposed by Achtemeier (1989) would be an option. Our use of the term "analysis error" is defined as a measure of the inability to recover the data values on which the analysis is based from the gridded analysis by linear interpolation anywhere within the extent of the grid. The measure used is absolute error (AE). Even though the definition is in terms of values at data points (that is where the error can be measured), it represents the difference between the true value (the value that would have been observed) and the analysis at *any* point. Note that this does not address the possible errors in the observations except to the extent they evidence themselves in the analysis.

For an analysis algorithm such as BCDG, one can think of different ways of making such an estimate of the error. The simplest one is to interpolate into the completed analysis grid and compute the error at each data point. The average of these errors would be an overall measure of error. This has two undesirable attributes. First, the interpolated value itself has potential error. If the gridpoint representation of the data were perfect, one would still not, in general, recover the value at the data point exactly by interpolation.[1] That is, the analysis process is not reversible. However, interpolation from a regularly-spaced grid to a random point is more exact than interpolating from randomly spaced points to a regular grid, especially when the data density relative to the grid spacing is uneven and/or sparse. This error of interpolation is unavoidable, although there are different methods of interpolation that could be used.[2] The second, and major, difficulty is that any good analysis process can fit the data points rather closely, but still be poor where the data are sparse; a calculation of error only at the data points may not well represent the error over the whole grid. In addition, this gives errors only at the data points, not at grid points.

To attempt to overcome the second difficulty, one could withhold a few data points when doing the analysis, then compute the AE only at those points. Then, the analysis would not be affected by the withheld points, and the AE would be a measure of the overall error at points on the grid between grid points where there were no data values. While the analysis is deprived of

---

[1] Suppose an analytic function were defined that could be evaluated at any point in the analysis area. Knowing the grid point values does not equate to knowing the values between grid points. Interpolation to a point from gridpoint values will give an estimate that will not, in general, be the same as the analytic function evaluated at that point.

[2] Biquadratic interpolation would work well in smoothly varying fields of data. Smooth fields could result from the data themselves not exhibiting much variablity, or the analysis process highly smoothing the data. For more variable data fields, bi-linear interpolation is probably best.

those withheld data, this is acceptable provided the number of withheld points is a very small fraction of the total points. Withholding data for error estimation was used as early as 1962 by Thomasell (1962). By replication with the same data, withholding different sets of points, one can estimate the mean absolute error (MAE) for a particular set of data. By performing analyses on many sets of data, with or without replication, one can estimate the MAE over that sample. But note that this is an overall error, and says nothing about the distribution of errors over the grid and its underlaying terrain.

Forecasters who ask about analysis errors are usually concerned about their specific area of interest, which may be rugged or not, be near water bodies or not, or be in sparse data regions or not. This paper describes a process for generating an error map for a particular analysis and gives results for surface temperature and dewpoint over the conterminous Unites States (CONUS). While the method is designed for BCDG, with some adjustment, it could be used equally well for other analysis methods.

2. THE ERROR ESTIMATION METHOD

As stated previously, our measure of error is the difference between a data value and the value obtained from the gridpoint analyses by linear interpolation. For a particular meteorological variable, we make analyses over many observation times, randomly withholding a very small percentage of the observations. A particular analysis over the CONUS and its immediate water surroundings will contain over 10,000 observations. We withhold 20 land reports and 1 water report, and compute the AE for each location and analysis. These values become a predictand data set for regression analysis.

For each of the withheld data points, we compute a set of predictors that might be related to the analysis errors. Then, two regression equations are computed by forward selection of predictors, one for land locations and one for water locations. The developed equations apply to observation points, but with certain assumptions can be applied to each appropriate (land or water) grid point for a specific case, thereby giving a grid of estimated errors for that observation time.

3. DESCRIPTION OF PREDICTORS

There are a number of factors that may cause error at specific locations. These are discussed below and the predictors are described. In all, 19 predictors were calculated.

A. Data Density

Obviously, if there are many data points relative to the spacing of the grid, the analysis will be better than if the data points are more sparse, other things being equal. We computed two predictors related to just data density, with respect to the grid spacing. They are:

1. The number of grid distances from the withheld station to the closest station within 110 grid lengths.
2. The number of grid distances from the withheld station to the 2nd closest station within 110 grid lengths.

The radii of influence were such that the first could always be computed. If the second could not be computed (there were not two stations within 110 gridlengths of the withheld station), that case was omitted. This rarely happened.

B. Data Variability

When data values are very nearly the same over some small region, say within a few grid lengths at most, then a gridpoint value should represent them very well, they should be highly recoverable, and the analysis error would be low. On the other hand, when there is high variability, one would not expect any particular value to be recoverable to a high degree of accuracy. Several potential predictors were computed as indicated below.

3. The data variability within a radius of 110 grid lengths of the station. Data variability is defined as the mean absolute difference between the data values and their mean. The withheld value itself is not included in the calculation.
4. Same as 3, except within 90 grid lengths.
5. Same as 3, except within 70 grid lengths.
6. Same as 3, except within 54 grid lengths.
7. Same as 3, except the vertical change with elevation (VCE) of the variable being analyzed between the withheld and the other stations is applied (the VCE is explained in Glahn et al. (2009), but because of its importance here, it is included as Appendix I).
8. Same as 7, except within 90 grid lengths.
9. Same as 7, except the distance between stations is weighted quadratically by the same weighting function used in the analysis (Glahn et al. 2009).
10. Same as 9, except within 90 grid lengths.
11. Same as 9, except within 70 grid lengths.
12. Same as 9, except within 54 grid lengths.

C. Roughness of Terrain

The values of most surface variables are influenced by the height of the terrain. For some variables, the affect of terrain can be anticipated. For instance, the observed temperature usually, but not always, decreases with altitude of the observing point within a local area. However, while the terrain may be a factor, its affect is not easy to anticipate for some variables. For example, cloud height above ground may well decrease with elevation in some restricted area, but cloud height of zero at an observing site may abruptly change to no clouds (clear or unlimited cloud height) above a certain elevation within the same local area. Several potential predictors were calculated, some of them also related to factors discussed above.

13. Roughness calculated on the grid centered on the grid point closest to the station within a radius of 8 grid lengths. Roughness is defined as the mean absolute difference between the terrain heights at the grid points and their mean.
14. Same as 13, except within 4 grid lengths.
15. Same as 13, except within 2 grid lengths.
16. Same as 13, except within 1 grid length.
17. Absolute difference in elevation between the withheld station and its closest neighbor. (Combines data density and roughness.)
18. Product of 17 and 1. (Combines data density and roughness.)
19. Absolute difference between the withheld station value and the value estimated from the closest station after applying VCE of the analyzed variable calculated at the closest station and with the elevation difference between the two. (Combines data variability, roughness, and data density.)

D. Land Use

Land use may also affect analysis errors. For instance, as one goes from a grassy location to a sandy or rocky one, the value of the variable may well change. We have not used predictors dealing with land use. It is likely any variation caused by land use is very localized, and is of a smaller scale than the analysis grid length. Water/land differences are dealt with by having separate relationships for water and land.

E. Land/Water Differences

In addition to there being no terrain roughness over water bodies, the different surface itself may cause a difference in the analysis error. As noted above, different relationships are developed for land and for water, so specific predictors are not necessary.

F. Synoptic Situation

While the atmospheric stability and wind flow characteristics can undoubtedly affect analysis error, most of these affects should be captured in the data variability and roughness of terrain.

4. COMPUTATION OF THE PREDICTAND

The predictand, the (absolute) error estimates at particular locations, is computed by making analyses over a large data set, randomly withholding a few stations from each analysis, and finding the absolute difference between the withheld station's value and the value interpolated from the analysis. (The method of random selection is explained in Appendix II.) This gives an AE at each withheld data point for each analysis. Specifically, we withheld 20 land stations per analysis and one water station, the latter from either ocean or Great Lakes buoys, or observing points judged to be more representative of a water location than land.[3] The total number of land

[3] Some coastal stations, while not over water, may be more representative of water than land. See Im et al. (2010) for a discussion this issue.

stations per analysis was on the order of 10,000, so the percentage withheld was about 0.2 %. The total number of water points was on the order of 300 for temperature and 200 for dewpoint, so the percentage of withheld points was about 0.3 and 0.5 %, respectively. Our sample consisted of every fifth hour for all days within a one-year period June 3, 2009, 2100 UTC to May 31, 2010, 2300 UTC.[4] The analyses were performed over the CONUS on a 2.5-km Lambert grid [essentially the grid used for the National Digital Forecast Database (Glahn and Ruth 2003)].

5. COMPUTATION OF THE PREDICTORS

The 19 potential predictors discussed above were computed for each withheld station, which incorporate data density, data variability, and terrain roughness. They are summarized in Table 1. Predictors dealing with elevation differences are not computed for over water. The grid lengths quoted are for land; for water they are double those quoted.

6. RESULTS

One year of data was processed, and a regression equation obtained for land and for water by screening the 19 predictors for land and 11 for water. The screening process consists of choosing predictors in order according to their additional reduction of variance (RV) of the predictand (Lubin and Summerfield 1951; Murphy and Katz 1985, Chapter 8). The development sample size for land was 30,100 for temperature and 30,060 for dewpoint; the sample size for water was 1,503 for temperature and 1,508 for dewpoint. These values should be reasonably independent and furnish stable equations, especially for a small number of predictors.

It became apparent that the best predictor by far is No. 19, which is the difference between the withheld station value and an estimate of it provided by its closest neighbor. This estimate includes the VCE procedure used over land in the analysis. No. 19 was selected first for all four equations (land/water, temperature/dewpoint) and provided the bulk of the total RV.

A. Land

The means, standard deviations, and correlations with the predictand are given in Table 2 for temperature and dewpoint over land. Of these, with a 0.001 cutoff for additional RV, three predictors were chosen in order 19, 14, and 5 for temperature and 19, 5, and 14 for dewpoint.[5]

---

[4] Using every 5th hour of the hourly sequence guarantees an even distribution of hours, and provides an adequate sample for stable regresson equations.

[5] Screening actually selected three more for temperature, but the coefficient was negative for the fourth one. The predictors were devised so that each one should logically contribute positively to the error estimate; a negative coefficient could easily give inconsistent results, even a negative absolute error. Negative coefficients can be caused by near multicolinearity among predictors, which occurs along with extremely low additional reductions of variance as additional variables are added to the equation.

For temperature, the total RV was 0.534–about half the total variance--and the standard error was 2.46 °F. For dewpoint, the total RV was less, 0.446, and the standard error was 2.53 °F.

Both coefficients and the mean and range of the variable itself have to be considered in assessing the influence of a predictor on the error. From Table 3, we see that if all three predictors had a value of zero, not likely but not impossible, the estimated temperature error would be only 0.29 °F; this is the lower limit for the temperature error estimate, and includes the interpolation error. If each predictor had its mean value, the error estimate would be 2.48 °F. If in addition, each predictor differed from its mean by one standard deviation in a positive direction, the error would be another 2.91 °F, for a total of 5.39 °F.

For dewpoint (see Table 4), the minimum error estimate is 0.20 °F, and the estimate if each predictor had its mean value is 2.77 °F. If in addition, if each predictor differed from its mean by one standard deviation in a positive direction, the error would be another 2.65 °F, for a total of 5.42 °F, very nearly the same as for temperature. Both temperature and dewpoint equations have a terrain roughness term (No. 14), a data variability term (No. 5), and No. 19, which embodies data density, roughness, and data variability. The only difference is that the order of selection for predictor Nos. 5 and 14 was reversed.

The correlations in Table 2 indicate the error is more highly (linearly) related to data density for dewpoint than for temperature (predictors 3-12). Also, it doesn't matter much over what area the density is calculated. Temperature and dewpoint are about equally related to roughness, mattering little over what area it is calculated. Predictor 19 being by far the best indicates the importance of the VCE in the analysis process. It may also indicate that if a better VCE process could be found, the error would decrease and therefore the analysis would be better, an obvious conclusion from logic alone.

B. Water

Table 5 similar to Table 2, and Tables 6 and 7 are similar to Tables 3 and 4, respectively, except for over water. Over water, there were only two predictors kept for temperature, (Nos. 19 and 3), and three for dewpoint, (Nos. 19, 2, and 5).

Tables 6 and 7 show that the minimum temperature and dewpoint error estimates are somewhat larger over water than over land, being 0.65 and 0.75 °F, respectively. The contribution to error from the constant and means of the temperature and dewpoint are 2.29 and 3.58 °F, respectively. If each predictor is different from the mean by one standard deviation in a positive direction, the error is 4.17 and 6.60 °F, for temperature and dewpoint, respectively. Data over water are much more sparse than over land, but the spatial variability is less.[6]

---

[6] As with temperature over land, more predictors were selected, but the next one selected had a negative coefficient.

7. IMPLEMENTATION

The equations were developed for stations–points where we had data. To implement, we could compute the estimated error for a particular time at each station where there is data. For instance, for the temperature/land equation, we could compute the absolute difference between the station's value and the value estimated by the closest station, taking into account the VCE (predictor No. 19), compute the roughness (predictor No. 14), and compute the data variability (predictor No. 5). These values can be used with the equation constant and coefficients to compute the error. However, this does not give values on a grid–what we really want. We could analyze these values with the BCDG analysis method, but that would give questionable error values in the same areas where we had a questionable analysis. This doesn't seem to be an acceptable solution.

Alternatively, we can, with some reasonable assumptions, apply the appropriate equations at grid points. We do it in the following manner. Predictor No. 19 is calculated by finding the absolute value of the difference of the analysis value at a grid point and the value for that grid point estimated by the closest station, taking into account the VCE at the station. The roughness (predictor No. 14) can be calculated at each grid point. Also, the data density at the grid point can be calculated in the same manner it was calculated at stations in the development.

Is implementation at grid points substantially different from implementation at stations? The data density calculation should not suffer. The number of stations within the specified radius will vary whether the calculation is at stations or at grid points. In development, whether the density was computed at stations or the closest grid point for a station on a 2.5-km grid would not vary much, the station itself not entering into the calculation. The roughness calculation is done at grid points in development, so there is no difference there. The major difference is for Predictor No. 19; in development the value at the station was known (observed), but in implementation the value is the analysis value.

The fact that predictor No. 19 is calculated at grid points in implementation and at stations in development may cause a low bias in the estimates for grid points. Because the density of grid points, in our application, is greater than the density of stations, the distance between a station and its closest neighbor will be, in general, greater than the distance between a grid point and its closest station. This may tend to underestimate the value of predictor No. 19, and the error at grid points, compared to errors calculated at stations.

Temperature and dewpoint analyses are shown in Figs. 1 and 3 and the corresponding error maps in Figs. 2 and 4. The temperature error maps show many errors in the eastern and central part of the U.S. are < 1.0 °F, as might be expected from Table 3, and only isolated spots where the errors are > 5.0 °F. In areas where the terrain is flat (terrain not shown), the larger errors are undoubtedly due to greater data variability. Examples of correspondence between the analysis "detail" and error can be found. For instance, Fig. 1 shows an area in southwestern Texas has analyzed temperatures of about 70 °F, whereas the surrounding areas are in the 80's and 90's. The lower values are legitimate, and are supported by four observations–67, 75, 71, and 71. However, the substantial data variability, combined with some terrain roughness, gives a rather high possible error, as indicated in Fig. 2. And indeed, the temperatures between a value of 71 °F and a neighboring value of 95 °F are uncertain.

A less noticeable example is a short, narrow band of potential error of approximately 6 °F in the midst of potential errors of about half that along the Mississippi River between Louisiana and lower Mississippi. This is caused by a temperature in Louisiana of 85 °F and one in Mississippi of 76 °F. These are the two closest stations to the narrow band along the river and about equally distant. Therefore, the values along the river are in more doubt than other surrounding values.

Notable in Fig. 4 is the estimated errors of near 6 °F in eastern Lake Erie. This is caused by the lack of adequate data there for this analysis time, even though one would not know that from the analysis shown in Fig. 3.

The largest temperature errors, for these analyses, are along the western seacoast and the nearby mountains. There is a sharp temperature contrast near the coast. Data are fairly dense, but the roughness is pronounced and the data variability is high. On the other hand, the largest dewpoint errors are associated with terrain in the West, and the error along the coast is not particularly high, indicating the less variable dewpoint there. The high dewpoint errors in high terrain, as compared to temperature, is likely due to temperature having a more consistent change with elevation than dewpoint and there being somewhat fewer observations of dewpoint than temperature for the analysis–11,218 for temperature and 9,807 for dewpoint.

Hourly analyses of temperature and dewpoint and the associated errors are in the National Digital Guidance Database (NDGD)[7] as part of the LAMP suite of forecasts (see Ghirardelli and Glahn 2010, 2011). Such analyses can be considered to be 0-h forecasts and compared to the LAMP 1-h forecasts for continuity. Areas of small analysis error should indicate where the consistency from analysis to forecast should be relatively good. However, because a station value not only affects the analysis but also contributes substantially to a 1-h forecast, both analysis and forecast may exhibit error and be indicated by the error map.

8. DISCUSSION AND CONCLUSIONS

A method to estimate the errors associated with the BCDG analysis of temperature and dewpoint has been developed and demonstrated. It should be recognized, any estimate of analysis error is just that–an estimate. The truth cannot be known (the values of the element being dealt with at each grid point) unless some data set is fabricated at both grid points (ground truth) and quasi-random points (data to analyze) with an analytic function. This fabrication route has been taken in analysis studies (e.g., Smith and Leslie 1984 and Goodin et al. 1979), but it is difficult to devise an analytic function that simulates the real world with elevation differences, data with unknown errors, and data densities that are variable and reasonable.

---

[7] The NDGD is the guidance counterpart of the NDFD, and can be accessed by the same methods as the NDFD.

This method, which we call BCDGE (BCDG Error), furnishes an estimate of error which is physically reasonable, is specific to the data set being analyzed, and is relatively easy to implement. To emphasize a previous point, the "error" used in the development included the interpolation error–the estimate of the station value from the regularly spaced grid. This itself, can be a considerable cause for error, especially in the West. It is also recognized that the development can be carried out only for the elevations where there are stations. For higher elevations, the estimated errors are essentially extrapolations from stations at lower elevations with similar terrain roughness and data density. This is also true of the analysis; the true values at high elevations are not known.

The error maps look reasonable in terms of pattern, and also in terms of absolute value, although there is no way to know how close the estimates really are. It is believed these error maps will help pinpoint where the problems are with the associated analyses. The actual values are probably not as important as the variability; high values can indicate possible data problems and unusual analysis error.

However, the detailed pattern, which is not obvious at the scale shown in Figs. 1-4, is more "choppy" than desired. This is due to the discrete nature of some of the predictors. For instance, predictors 1, 2, 17, 18, and 19, which include a contribution form the closest or second closest station, can switch abruptly from grid point to grid point because the closest station to the grid point switches abruptly. The result is a boundary about halfway between two stations. If different predictors could be derived that have a less discrete nature, the error estimates would not switch so abruptly. Because of this discrete nature, the analysis errors should be viewed as highlighting an <u>area</u> of possible error rather than focusing on individual gridpoint values.

This method for determining error is applicable to wind speed, possibly with some adaptation. However, the high variability of some variables, like ceiling height and visibility, make application of this method, or actually any method, questionable for these variables.

REFERENCES

Achtemeier, G. L., 1989: Modification of a successive corrections objective analysis for improved derivative calculations. *Mon. Wea. Rev.*, **117**, 78-86.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3,** 396-409.

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

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

Fritsch, J. M., 1971: Objective analysis of a two-dimensional data field by the cubic spline technique. *Mon. Wea. Rev.*, **99**, 379-386.

Ghirardelli, J. E., and B. Glahn, 2010: The Meteorological Development Laboratory's aviation weather prediction system. *Wea. Forecasting*, **25**, 1027-1051.

_____, and _____, 2011: Gridded Localized Aviation MOS Program (LAMP) guidance for aviation forecasting. Preprints 15 *Conference on Aviation, Range, and Aerospace Meteorology*. Los Angles, CA, Amer. Meteor. Soc., **4.4**.

Gilcrest, B., and G. P. Cressman, 1954: An experiment in objective analysis. *Tellus*, **6**, 309-318.

Glahn, H. R., 1981: Comments on "A comparison of interpolation methods for sparse data: Application to wind and concentration fields." *J. Appl. Meteor.*, **20**, 88-91.

_____, 1987: Comments on "Error determination of a successive correction type objective analysis scheme." *J. Atmos. Oceanic Technol.*, **4**, 348-350.

_____, and J. E. McDonell, 1971: Comments on "Objective analysis of a two-dimensional data field by the cubic spline technique." *Mon. Wea. Rev.*, **99**, 977-978.

_____, 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**, NOAA, U.S. Department of Commerce, 34 pp.

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

_____, K. Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, 2009: The gridding of MOS. *Wea. Forecasting*, **24**, 520-529.88-91.

_____, J-S Im, 2011: Algorithms for effective objective analysis of surface weather variables. Preprints, *24th Conference on Wether and Forecasting /20th Conference on Numerical Weather prediction*, Seattle, WA, **10A.4**.

Goodin, W. R., G. J. McRae, and J. H. Seinfeld, 1979: A comparison of interpolation methods for sparse data: Application to wind and concentration fields. *J. Appl. Meteor.*, **18**, 761-771.

_____, _____, and _____, 1981: Reply. *J. Appl. Meteor.*, **20**, 92-94.

Horell, J., B. Colman, 2005: Real-time and retrospective mesoscale objective analyses. *Bul. Amer. Meteor. Soc.*, **86**, 1477-1480.

Im, J.-S., B. Glahn, J. Ghirardelli, 2010: Real-time objective analysis of surface data at the Meteorological Development Laboratory. Preprints *20th Conference on Probability and Statistics*, Atlanta, GA, **219**.

_____, _____, and _____, 2011: Real-time hourly objective analysis of surface observations. *NOAA Tech. Memo.* NWS MDL **84**, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 19 pp.

Lubin, A., and A. Summerfield, 1951: A square root method of selecting a minimum set of variables in multiple regression: I. The method. *Physchometrika*, **16**, 271-284.

Manikin, G. S., and M. F. V. De Pondeca, 2011: Recent upgrades to and ongoing challenges for the real-time mesoscale analysis (RTMA). Preprints *15th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans and Land Surface (IOAS-AOCS),* Seattle, WA, **J7.3**.

Murphy, Allan H. and R. W. Katz, 1985: *Probability, Statistics, and decision making in the atmospheric sciences*. Westview Press, Boulder, CO,. 545 pp.

Myrick, D. T., and J. D. Horel, 2008: Sensitivity of surface analyses over the western United Satates to RAWS observations. *Wea. and Forecasting*, **23**, 145-158.

Panofsky, H. A., 1949: Objective weather-map analysis. *J. Meteor.*, **6**, 386-392.

Pondeca, M. and G. S. Manikin, 2009: Recent improvements to the real-time mesoscale analysis. *Preprints 23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction*, Omaha, NE, Amer. Meteor. Soc. , **15A2**.

_____, and Coauthors, 2011: The real-time mesoscale analysis at NOAA's national centers for environmental prediction: Current status and development. *Wea. and Forecasting*, **26**, 593-612.

Smith, D. R., and F. W. Leslie, 1984: Error determination of a successive correction type objective analysis scheme. *J. Atmos. Oceanic Technol.*, **1**, 120-130.

Thomasell, A., Jr., 1962: The areal-mean-error method of analysis verification. The Travelers Research Center, Inc., Contract FAA/BRD-363, *Tech. Pub.* No. **5**, Hartford, Connecticut, 18 pp.

Table 1. Definitions for nineteen potential predictors.

Predictor Category	Predictor No.	Definition	Within Grid Lengths
Data Density	1	Distance to the closest station	110
	2	Distance to the 2nd closest station	110
Data Variability	3	Data variability	110
	4		90
	5		70
	6		54
	7	Data variability with application of VCE	110
	8		90
	9	Data variability with application of VCE and use of the distance between stations weighted quadratically by the same weighting function used in the analysis	110
	10		90
	11		70
	12		54
Terrain Roughness	13	Roughness	8
	14		4
	15		2
	16		1
Data Density and Roughness	17	Absolute difference in elevation between the withheld station and its nearest neighbor	110
	18	Product of 17 and 1	110
Data Variability, Roughness, and Data Density	19	Absolute difference in value between the station value and its estimate from the closest station	110

Table 2. The variable means and standard deviations in °F, and correlations with the predictand for temperature and dewpoint over land.

Variable No.	Temperature			Dewpoint		
(See Table 1)	Mean	Std. Dev.	Correlation	Mean	Std. Dev.	Correlation
1	8.49	6.64	0.049	9.11	7.22	0.055
2	12.70	7.82	0.054	13.66	8.30	0.057
3	4.64	2.10	0.158	4.53	2.06	0.232
4	4.33	2.04	0.170	4.18	1.95	0.242
5	3.97	1.98	0.182	3.81	1.86	0.255
6	3.64	1.96	0.180	3.47	1.78	0.253
7	3.42	1.67	0.156	3.60	1.74	0.218
8	3.10	1.49	0.173	3.29	1.60	0.234
9	1.19	0.61	0.160	1.27	0.65	0.216
10	1.08	0.55	0.172	1.16	0.59	0.222
11	0.98	0.52	0.180	1.06	0.55	0.223
12	0.91	0.53	0.179	0.99	0.56	0.210
13	78.31	95.46	0.146	76.82	93.54	0.155
14	54.23	72.60	0.148	52.54	70.09	0.149
15	36.31	52.86	0.150	34.94	51.35	0.138
16	26.06	41.30	0.142	24.96	40.81	0.123
17	128.48	227.48	0.149	125.75	221.06	0.140
18	1310.01	2817.16	0.118	1483.99	3385.04	0.121
19	2.97	4.36	0.726	3.36	4.23	0.655
Predictand	2.48	3.60		2.77	3.40	

Table 3. The constant and coefficients, means, and standard deviations for the three predictor variables in the temperature equation over land, together with the predictor contributions to the total estimate. Total RV for the 3-predictor equation is 53.4%. Units are °F.

Variable No. (See Table 1)	Coefficient (Constant)	Mean	Contribution from Mean and Constant	Std. Dev (sd.)	Contribution from 1 sd.
Constant	0.2913		0.291		
19	0.5891	2.974	1.752	4.364	2.571
14	0.0026	54.234	0.141	72.595	0.189
5	0.0755	3.966	0.299	1.980	0.149
Total			2.483		2.909

Table 4. Same as Table 3 except for dewpoint equation. Total RV = 44.6%.

Variable No. (See Table 1)	Coefficient (Constant)	Mean	Contribution from Mean and Constant	Std. Dev (sd.)	Contribution from 1 sd.
Constant	0.2009		0.201		
19	0.5048	3.363	1.698	4.229	2.135
5	0.2033	3.814	0.775	1.859	0.378
14	0.0019	52.539	0.100	70.090	0.133
Total			2.774		2.646

Table 5. Similar to Table 2 for over water. Predictors involving the VCE, Nos. 7 and 8, and 13 through 18, do not exist over water because the elevation does not vary.

Variable No.	Temperature			Dewpoint		
(See Table 1)	Mean	Std. Dev.	Correlation	Mean	Std. Dev.	Correlation
1	27.00	16.99	0.042	32.38	27.78	0.050
2	38.61	21.44	0.003	54.43	39.86	0.065
3	3.73	1.77	0.218	3.73	2.15	0.143
4	3.44	1.71	0.217	3.54	2.14	0.184
5	3.14	1.71	0.203	3.30	2.39	0.216
6	2.76	1.77	0.191	2.88	2.22	0.170
9	1.73	0.92	0.177	1.76	1.14	0.148
10	1.55	0.90	0.178	1.59	1.08	0.160
11	1.35	0.87	0.176	1.45	1.26	0.184
12	1.19	0.89	0.172	1.30	1.27	0.187
19	2.98	4.96	0.518	4.00	4.59	0.701
Predictand	2.29	3.13		3.58	3.79	

Table 6. The constant and the coefficient, mean, and standard deviation for the two variables in the temperature equation over water, together with the predictor contributions to the total estimate. Units are °F. Total RV for the 2-predictor equation = 27.9%.

Variable No. (See Table 1)	Coefficient (Constant)	Mean	Contribution from Mean and Constant	Std. Dev (sd.)	Contribution from 1 sd.
Constant	0.6466		0.647		
19	0.3121	2.983	0.931	4.956	1.547
3	0.1898	3.729	0.708	1.773	0.337
Total			2.286		1.884

Table 7. Same as Table 6, except for the 3-predictor dewpoint equation. Total RV = 49.7%

Variable No. (See Table 1)	Coefficient (Constant)	Mean	Contribution from Mean and Constant	Std. Dev (sd.)	Contribution from 1 sd
Constant	0.7492		0.749		
19	0.5685	4.003	2.276	4.587	2.608
2	0.0059	54.429	0.321	39.862	0.235
5	0.0718	3.296	0.237	2.392	0.172
Total			3.583		3.015

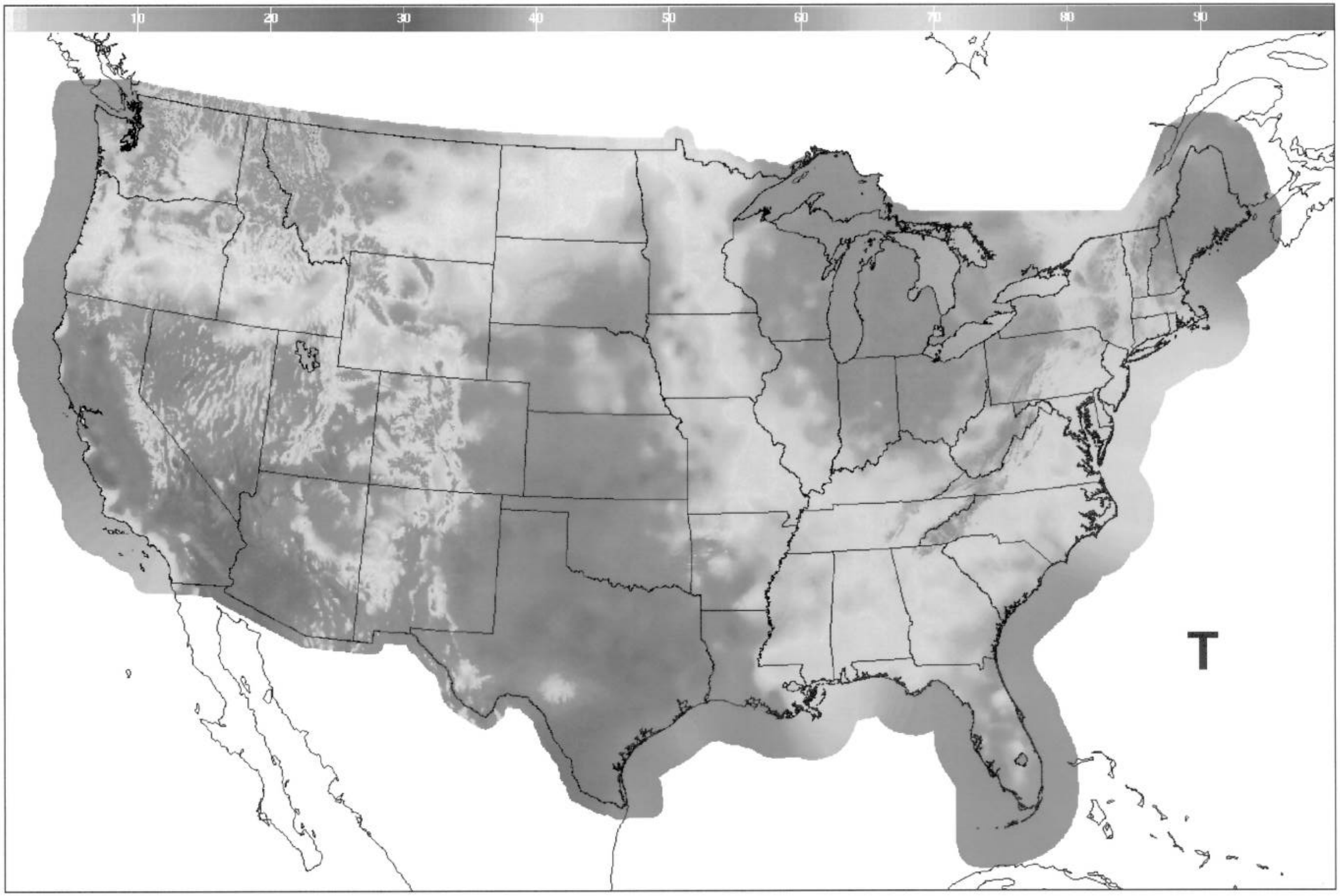

Figure 1. BCDG analysis of temperature (°F) produced for 0000 UTC 29 September 2011.

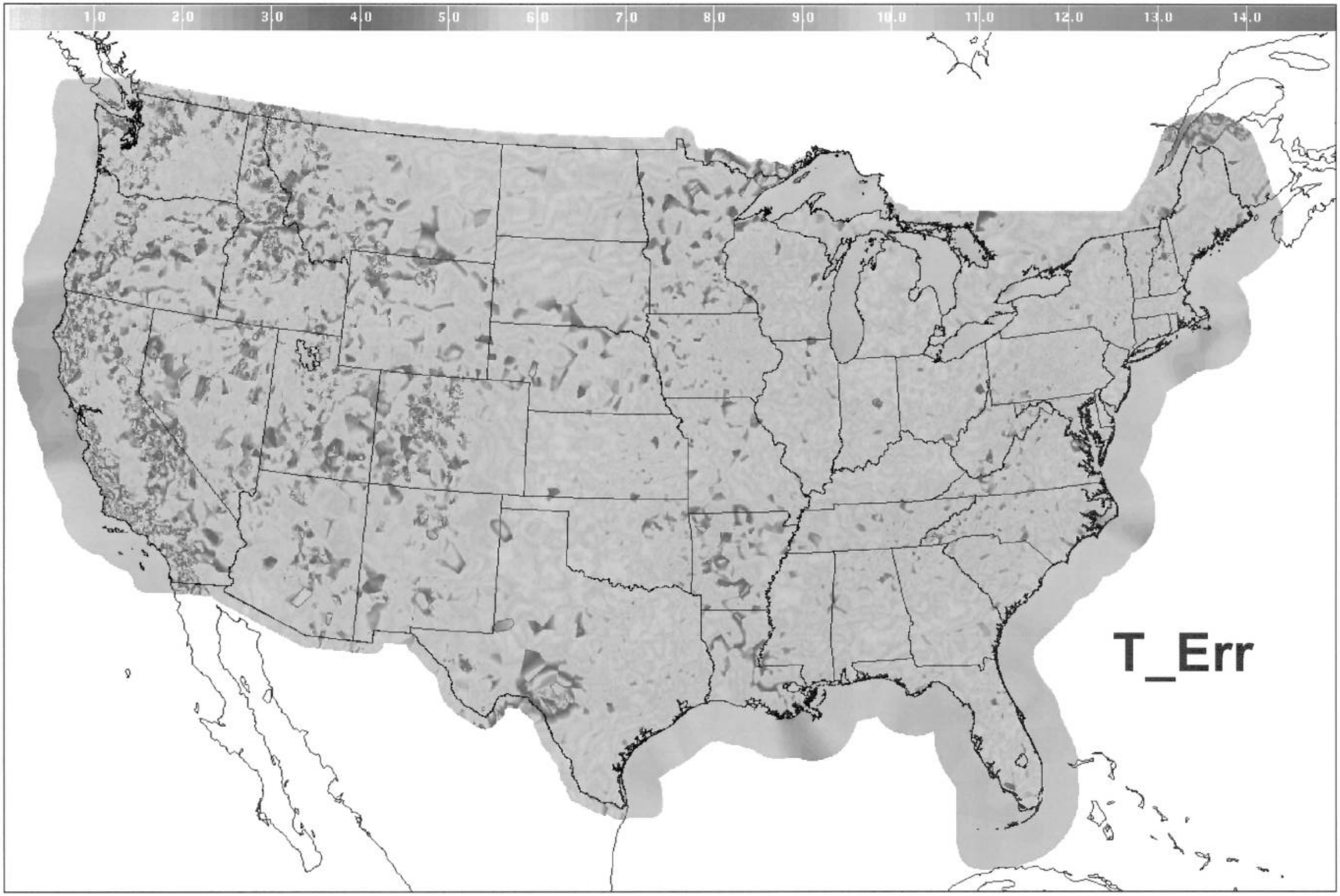

Figure 2. Error estimation (°F) of the BCDG analysis for 0000 UTC 29 September 2011.

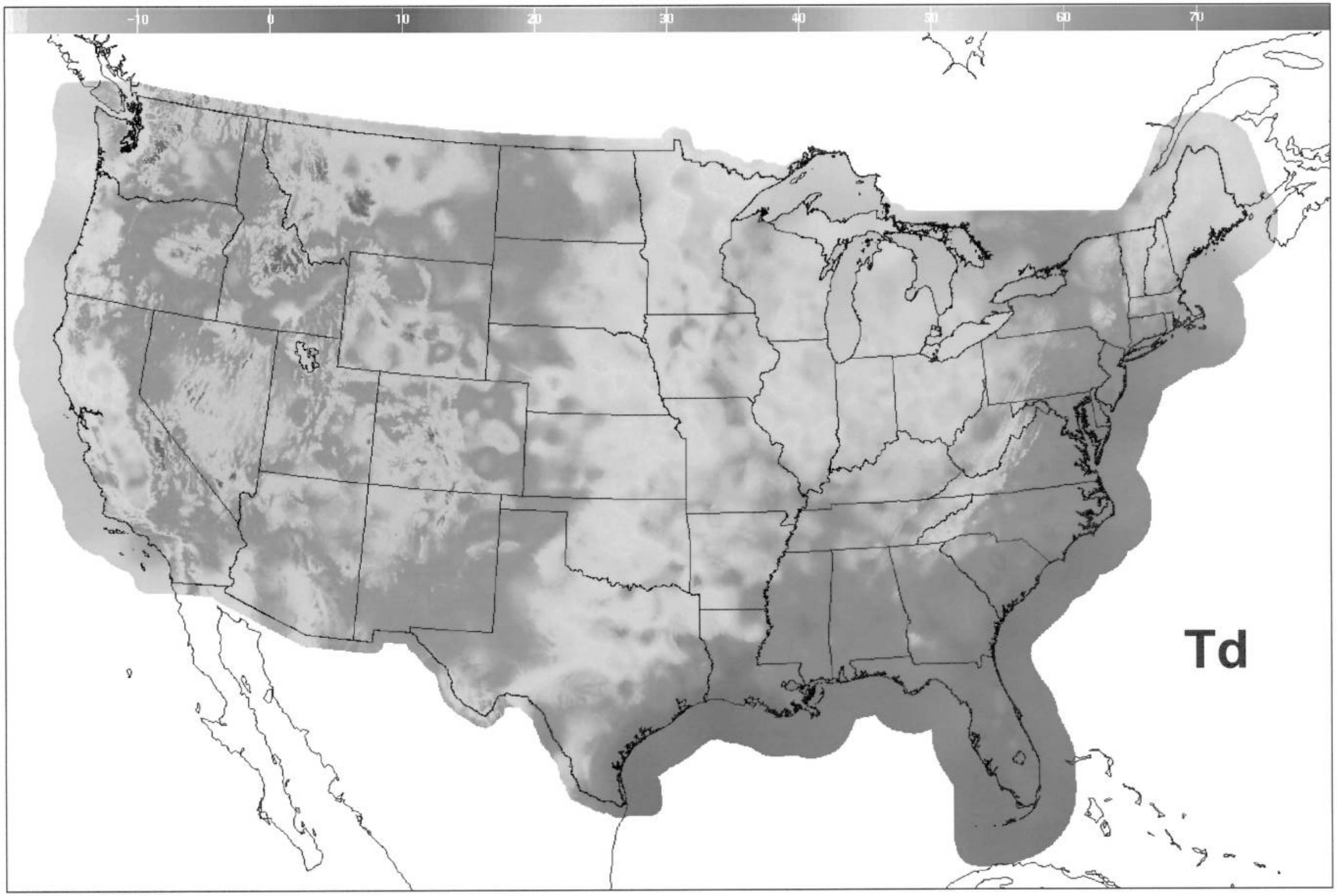

Figure 3. Same as Fig. 1 except for dewpoint.

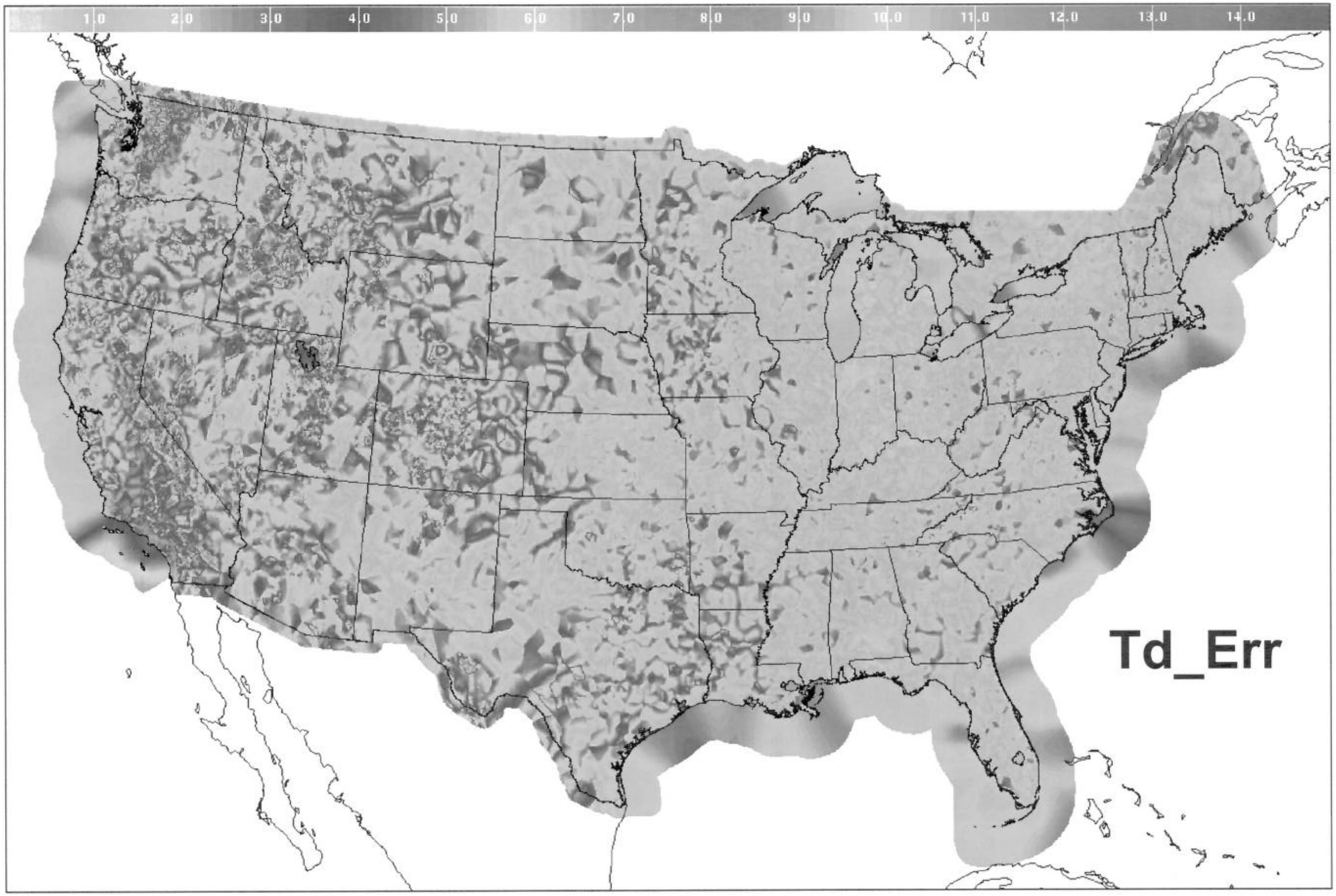

Figure 4. Same as Fig. 2 except for dewpoint.

APPENDIX I

The VCE Method

The Vertical Change with Elevation (VCE), sometimes called the "lapse rate," plays a major role in how a datum affects a gridpoint. 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 gridpoint near station B, one needs to consider that the correction should be based on the VCE, defined as

VCE = (TB-TA)/(EB-EA).

The VCE is computed for each station for each analysis. The specific value for a station 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,

$$\text{VCE(A)} = \sum(TB_i - \text{TA})/\sum(EB_i - \text{EA}),$$

over all designated close stations B_i. This process is quite robust, computing only one statistic from several pieces of data. It is important to find a list of stations that are close in horizontal distance but far apart in vertical distance. A preprocessor finds such lists of stations by making several passes over the vertical and horizontal station locations searching for the desired combinations.

In the analysis, the modifications to gridpoints from the stations within the radii of influence use not only the observation, but also the individual VCEs.

APPENDIX II

Method of Selection of Withheld Points

In order that the regression equations relating error to predictor variables represent the map equally and not be based on, for instance, data density, a latitude and longitude within the confines of the grid were each selected randomly. Their combination defined a point on the analysis grid. To be used, the point was required to be within the NDFD confines. Canadian stations were not used.

One disadvantage of this selection method is that, while regions of the country are represented approximately equally, a data value in sparse data areas might be selected to contribute to the computation much more than one in a dense data area. It is not possible to weight each geographic region <u>and</u> each datum equally. It was thought that using a datum oftener than others would not invalidate the results, unless the observation had unusual error characteristics.

Note that this selection process is in distinction to randomly selecting from the station list. If that were done, the areas of dense data would be weighted too heavily and create a bias toward areas of high density observations.

The seed for the pseudo random number generator can be the same in each checkout run to get constant results. However, it must be different in actual withholding runs, or the same stations will be withheld on each run. Rather, the seed can be based on the system clock to get different withheld stations.