OBJECTIVE ANALYSIS OF HOURLY 2-M TEMPERATURE AND DEWPOINT OBSERVATIONS AT THE METEOROLOGICAL DEVELOPMENT LABORATORY

Jung-Sun Im and Bob Glahn

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

Abstract

As part of the Localized Aviation Model Output Statistics (MOS) Program (LAMP), the Meteorological Development Laboratory (MDL) is analyzing surface data reports on an hourly basis. The Bergthórsson-Cressman-Döös-Glahn (BCDG) objective analysis technique used for gridding MOS forecasts has been tailored to analyze surface observations. MDL is making the analyses to assess the spatial and temporal accuracy of gridded MOS and LAMP forecasts and to provide gridded LAMP nowcasts. The analyses are now available in the National Digital Guidance Database (NDGD); the grid is the same as that used in the National Digital Forecast Database (NDFD), a 2.5-km grid on a Lambert Conformal map projection covering the conterminous United States.

This paper describes the intensive effort needed to assure the metadata are correct for each station, to develop necessary quality control procedures, and to reduce spatial and temporal discontinuities in the analyses. One of the capabilities of the analysis package is to use an observation from the previous hour if the station did not report at the analysis hour. An adjustment is made to the previous hour's observation in order to account for typical diurnal changes from the previous hour to the analysis hour. In addition, the radius of influence computed on an individual station basis is incorporated to handle a heterogeneous distribution of the observations. This paper focuses on the analyses of temperature and dewpoint over the conterminous United States on the NDFD grid. The analysis product suite will be extended and modified to include most observed weather elements, including ceiling height, visibility, sky cover, and wind speed and direction.

Corresponding Author: Jung-Sun Im
Meteorological Development Laboratory, National Weather Service
1325 East-West Highway, Silver Spring, MD 20910
Email: Jung-Sun.Im@noaa.gov

1. Introduction

As part of the Localized Aviation Model Output Statistics (MOS; Glahn and Lowry 1972) Program (LAMP; Ghirardelli and Glahn 2010), the Meteorological Development Laboratory (MDL) is analyzing surface data reports on an hourly basis. The analysis scheme used by MDL for gridding MOS forecasts (Glahn et al. 2009) has been tailored to analyze surface observations. MDL is making the analyses to assess the accuracy of gridded MOS and LAMP forecasts. In addition to providing verification grids for gridded MOS and LAMP forecasts, our goal is to add gridded LAMP nowcasts to the gridded LAMP forecast suite. These fine-resolution analyses will eventually help forecasters create and verify the National Digital Forecast Database (NDFD; Glahn and Ruth 2003).

Real-time and retrospective analyses at both a fine spatial and temporal resolution are required to establish an Analysis of Record (AOR; Horel and Colman 2005), and to create the NDFD forecasts as well as to verify their accuracy. As a first step, a prototype Real-Time Mesoscale Analysis (RTMA; De Pondeca et al. 2007; Benjamin et al. 2007) is produced at the National Centers for Environmental Prediction (NCEP) in collaboration with the Earth System Research Laboratory (ESRL). It represents a fast-track, proof-of-concept of the AOR program and establishes a benchmark for future AOR efforts (De Pondeca et al. 2011). In addition to the RTMA, MDL analyses can be used to judge the quality of an AOR.

High quality surface weather observations and effective quality control processes are critical to generate fine-resolution objective analyses. The hourly surface observations for the analyses are obtained from NCEP in real time and are additionally quality controlled at MDL. While performing analyses of these observations, we found various issues such as inconsistent site information for stationary stations, stations reporting data at the same locations with different station names and types, multiple reports at the same time with different station types, stations repeatedly reporting the same values, and spatial and temporal discontinuities in the analyses. These issues were, of course, not unexpected in climate data and especially with real-time data in operational environments.

In this paper, we describe the intensive effort needed to: 1) assure the metadata are correct for each location; 2) develop and implement necessary quality control procedures; 3) assign a representative land, ocean, or inland water flag to each station; and 4) reduce spatial and temporal discontinuities in the analyses. This paper focuses on the analyses of 2-m (called in this paper "surface") temperature and dewpoint observations over the conterminous United States (CONUS) on the NDFD grid with grid length of 5 km.

2. BCDG Analysis Method

MDL has produced gridded MOS forecasts since 2006 (Glahn et al. 2009). The objective analysis scheme used to produce gridded MOS is based on the successive correction technique called Bergthórsson-Cressman-Döös (BCD; Glahn et al. 1985; Cressman 1959; Bergthórsson and Döös 1955). This successive correction technique consists of making multiple passes over the data, correcting each gridpoint on each pass with the data in its immediate vicinity. For gridded MOS, this BCD technique was extended by implementing the following features:

1) separate analysis processes for land, inland water, and ocean combined into one system to accommodate the different characteristics associated with land and water;
2) computation on-the-fly of change of a weather element with elevation, so that the vertical change varies with the location, time of day, day of the year, and synoptic situation;
3) a variable radius of influence R for each specific corrective pass to account for highly varying data densities;
4) error detection which employs a buddy check when a datum is in serious question; and
5) a contour-following smoother.

With these major extensions, all of which are either explained below or referenced, the BCD scheme was thereafter called Bergthórsson-Cressman-Döös-Glahn (BCDG; Glahn et al. 2009).

The BCDG analysis system has many options that can be used to tune the system based on data density relative to gridpoint density, variation in data density over the grid, choice of first-guess field, number of corrective passes, smoothness versus detail desired in the analysis, and error characteristics of the data. The smoothing algorithm adopted from previous work (Glahn et al. 2009) removes some grid-length noise introduced by the basic analysis procedure. It was modified, and termed "contour following," such that there would be no smoothing done across significant changes in elevation. (A full explanation is contained in the appendix to the cited reference.)

In analyzing surface observations especially, BCDG's error checking capability is an essential part of the analysis of the data. The BCDG software performs this error checking on each pass based on an acceptable difference (threshold) between the station observation value and the value interpolated from the analysis. Based on considerable testing and meteorological judgment, we have determined optimal threshold values for each pass.

The procedures of BCDG's error checking are summarized in Fig. 1. On each pass, the difference between a station's value (S) and the value interpolated from the first guess or previous pass analysis (I_S) is computed. If the difference is less than or equal to the threshold (Th) specified for that pass, S is accepted for that pass, but if it exceeds 1.5 times the threshold, S is not used for that pass; if it exceeds the threshold, but is less than or equal to 1.5 times the threshold, then the two neighbors closest to S (N_1 and N_2) are found and their observations are used to perform buddy checks to determine whether or not S will be discarded. The differences of N_1 from its interpolated value (I_{N1}) and N_2 from its interpolated value (I_{N2}) are computed. If either one of the two neighbors' differences is greater than 0.6 times the threshold, and the differences of both S and its neighbor are of the same sign, then S is accepted. If S is not yet accepted, one more check is performed. If either one of the two neighbors' differences is less than or equal to 0.6 times the threshold and the difference between S and the neighbor's value *adjusted for terrain* (A_{N1} or A_{N2} accordingly) is within 0.6 times the threshold, S is accepted. The intent of the buddy check is to determine whether neighboring stations also have values that differ in the same direction (plus or minus) by a substantial amount, and if so it is assumed the analysis for that previous pass is in error rather than the observation S. If none of these conditions is met, S is not used on the current pass. There can be a significant difference between the value S and its interpolated value I_S, especially on the first pass. The value "rejected" on one pass is still considered on the next pass.

More detailed information on the BCDG technique such as the gridpoint correction algorithm, determination of vertical change with elevation, and accommodation for land and water can be found in Glahn et al. (2009). Based on extensive experimentation performed at MDL, we adopted the BCDG options used in gridded MOS. These options incorporate a first-guess grid composed of the average value of the element being analyzed computed over all observations to be used in the analysis, four passes over the data to capture the desired detail in the analysis, limitation of the computed change with elevation when it is of the opposite sign than expected, and a contour-following smoother.

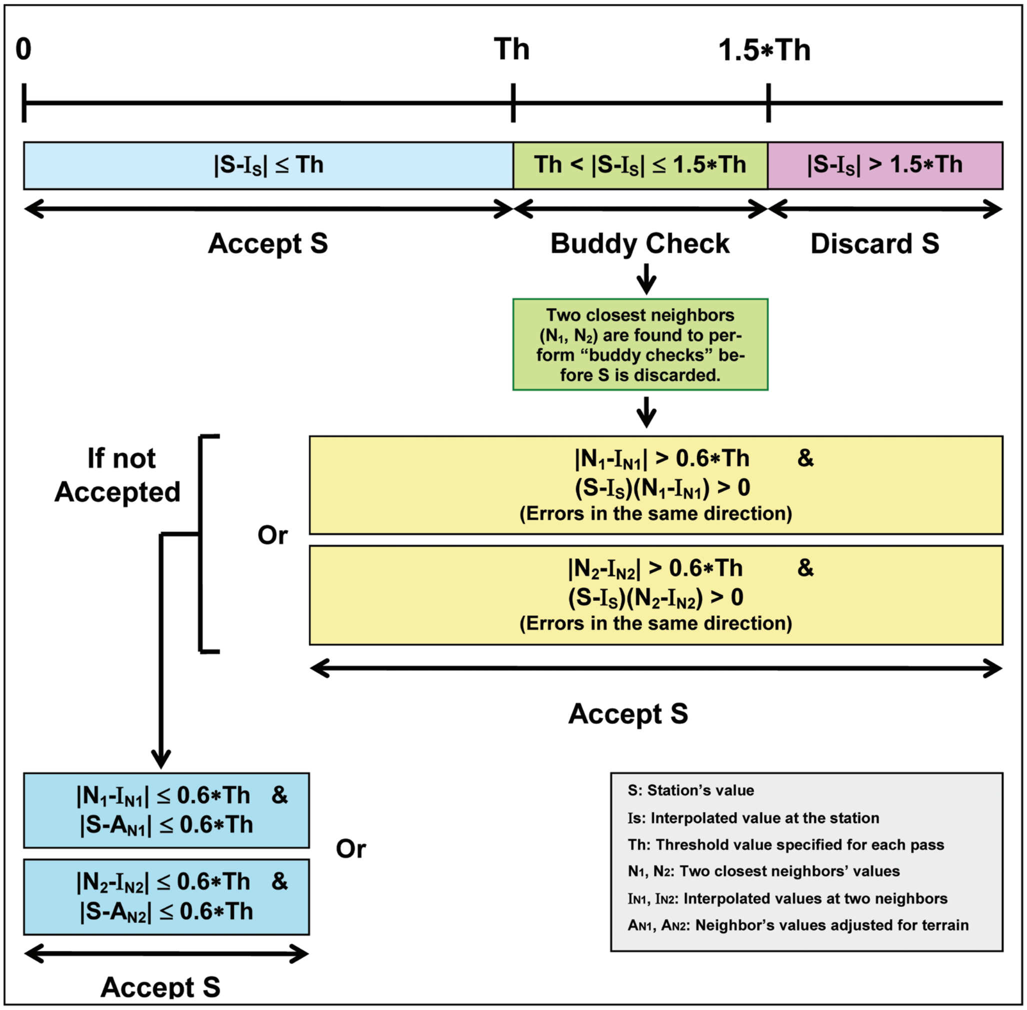

Fig. 1. BCDG's error checking procedures executed on each data pass.

3. Data Collection

Hourly surface observations are obtained from NCEP in real time and are additionally quality controlled at MDL. The first set of quality control checks at MDL ensures that all temperature and dewpoint observations are in an acceptable range for the station's geographical area, and each station's temperature is greater than or equal to the station's dewpoint. These checks, built into software, are necessarily subjective, and the acceptance interval is quite wide so as to not reject good data in unusual synoptic weather situations. Tables are used which vary with five sections of the United States and with four seasons defined as spring (March-May), summer (June-August), fall (September-November) and winter

(December-February). For instance, for winter and for the southeast section, defined to have northern and western boundaries of 35 °N and 100 °W, respectively, the acceptable temperature and dewpoint ranges are -26 to 102 °F and -31 to 85 °F, respectively (Glahn and Dallavalle 2000). These particular checks have been used by MDL for archival purposes for many years, and the resulting data are used in MOS and LAMP development and verification.

In preparing input observations to be used in the hourly analyses, we collect data observed between 15 minutes prior and subsequent to the analysis hour. If more than one observation is reported for a station, we select the report closest to 10 minutes prior to the analysis hour. The analysis system for temperature and dewpoint analyzes six types of observations, which are obtained from METAR (roughly translated as Aviation Routine Weather Report; OFCM 1995), mesonet, synoptic, moored buoy, Coastal-Automated Marine Network (C-MAN), and tide gauge stations.

METAR reports typically come from airports or permanent weather observation stations. Observations are taken by automated devices or trained personnel. Some stations have automated observations augmented by human observers. These sites have been well sited and are well maintained. Consequently, METAR reports are of high quality, and we have found that they are more reliable in terms of accuracy and reliability of receipt than all other observational data sets.

Mesonet observations are obtained from local, state, and federal agencies and private entities. These sites are quite dense compared to METAR sites. In fact, over 80% of the stations used in the BCDG analysis consist of mesonet type stations. However, because the mesonets are usually established for specific purposes, they furnish very dense reports in some areas, but not over large areas. They may report for some period of time, on the order of months, then are discontinued or the sites are moved. The siting and maintenance are many times not as good as at METAR sites.

Synoptic data are comprised of manual and automatic observations, and are available only every 3 or 6 hours. The quality is good, but they are not available every hour. In many cases, these data are redundant to the METAR data at the same location (this issue will be discussed in section 4.b).

Moored buoy, C-MAN, and tide gauge stations provide quality observations over the oceans, the Gulf of Mexico, and the Great Lakes. The CONUS analysis area covers land and near-coastal water areas; these are the observations that make acceptable analyses over water possible.

4. Quality Control

As indicated in the previous section, observations come from a variety of sources and vary in reliability of receipt and accuracy. Not all observations can be used in an analysis; it is necessary to quality control them. One critical aspect of the quality control is the development and maintenance of a station dictionary. This dictionary contains information such as station identifier, type, latitude, longitude, elevation, land/water flag, and quality flag. The total number of stations that can report weather elements of interest is on the order of 20,000 over the CONUS; however, on any given hour, only about half that number of stations report. Because site information changes from time to time (Allen 2001), upkeep of the station dictionary is required on a regular basis. This dictionary is designed and maintained specifically for the purpose addressed in this paper; other organizations have similar dictionaries that are designed for their purposes.

In quality controlling the data, we considered various issues such as inconsistently reported latitude/longitude/elevation for a stationary station (a station whose location is fixed, unlike a drifting buoy or ship), stations reporting data at the same locations with different station names and types, questionable land/water assignments on the coastlines, multiple reports at the same time with different station types, and stations which reported the same values for months. The following sub-sections describe the methods used to resolve these data issues.

a. Questionable metadata

The metadata associated with the observations collected in real-time can contain errors. Obviously, the location of an observation has to be known or it cannot be used. In order to not use an observation with bad location data, its associated metadata were compared with its past history of reporting. We investigated a 3-month data sample (August–October 2008) and found that 98.4% of the stationary stations (18,911 of 19,228 stations) had reported only at the same fixed locations. Of those that had different locations, 97.5% (309 of 317 stations) were mesonet stations.

As previously stated, MDL maintains a station dictionary containing all the stations with a reporting history over the previous year. A preprocessor was run on a recent year of data, and any station reporting as many as 10 times during the year was put into the dictionary with its metadata. When the analysis is performed, the metadata available with an observation in real time has to agree with the dictionary values within certain tolerance limits for that observation to be used. These limits were subjectively defined such that when exceeded and the observation was included, it made the analysis poor in that immediate vicinity. We judged acceptable limits for latitude, longitude, and elevation to be 0.01°, 0.01°, and 280 ft, respectively, for stations on land. With these limits, the maximum spatial error would be less than half the grid length being used for the analysis.[1] The elevation threshold of 280 ft was determined by making the assumption that 1°F is an allowable error range in a temperature analysis (1 °F corresponds to a change in elevation of 280 ft in the international standard atmosphere of 3.56×10^{-3} °F ft^{-1}). Because moored buoys provide valuable observations over the data-sparse ocean and lake regions, we relaxed the location threshold to 1°. When applying these rules, 61.5% of the questionable stations (195 of 317 stations) were included in the dictionary. The retained stations were further investigated to determine the true latitude, longitude, and elevation values. While searching for these values, the selection priority was given to: 1) matching with online sources of geographic information; 2) the most frequently reported values; and 3) the most recently reported values. Finally, with the site information available from the completed station dictionary, the real-time data were screened before starting the analysis. The screening procedure was executed in such a way that if the reporting location of a real-time observation deviated from its position in the station dictionary by a value greater than the threshold specified above for that station type, the observation was not used in the analysis for the analysis hour.

b. Redundant stations

Exploring the horizontal distributions of each type of station revealed that there were reports at exactly the same locations, but with different station names and types. As an example, stations KEYW (METAR type station) and 72201 (synoptic type station) were reporting observations at the same latitude, longitude, and elevation (24.55°N, 81.75°W, 3.3 ft). These observations were from the same reporting station, but with different station names, which would result in double weighting at that point in the analysis if both were used. During a 3-month period (August-October 2008), 786 synoptic stations reported, of which 355, or 45%, were clearly duplicates. We removed the synoptic stations because they report less frequently than METAR stations.[2]

Figure 2 shows the spatial distribution of the 18,024 observing stations for temperature and dewpoint in the station dictionary. Stations are heterogeneously distributed with highly variable density over the domain

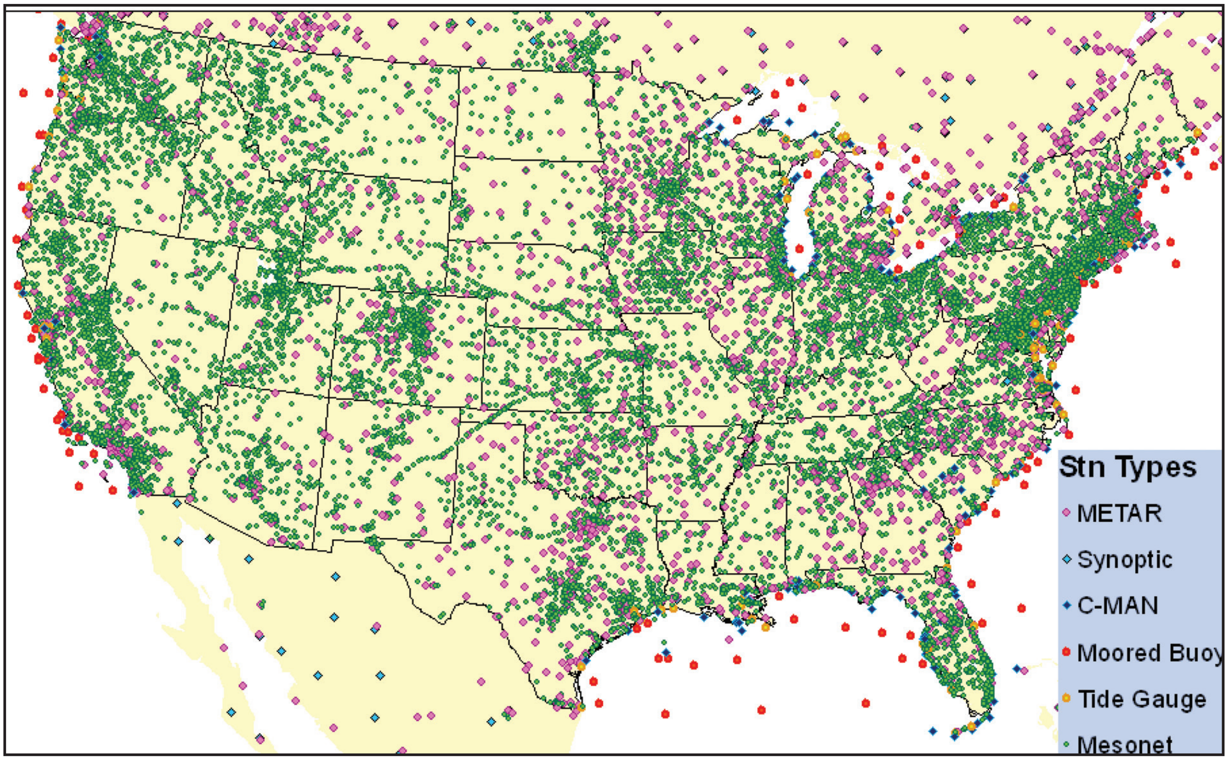

Fig. 2. Spatial distribution of surface observing stations for temperature and dewpoint.

---

[1] It will be less than one-half the grid length when the transition is made from 5-km to 2.5 km, as planned.

[2] More recent work indicates even more of the synoptic stations are redundant. For this reason, and because synoptic stations report only every 3 or 6 hours, all such stations will be removed in the future.

and are of the types: mesonet (82.5%), METAR (13.0%), synoptic (2.4%), C-MAN (0.8%), moored buoy (0.7%), and tide gauge (0.6%).

c. Stations reporting unchanging values

Each station in the station dictionary has its own quality flag for each element. To determine the quality flag, we used the "reject station lists" provided by the Global Systems Division (GSD) of ESRL and National Weather Service (NWS) Weather Forecast Offices (WFO) as part of the Advanced Weather Interactive Processing System (AWIPS; Seguin 2002) configuration. In addition to these master reject lists, we made a second reject list. This list included stations that continued to report unchanging observation values (e.g., zero values for temperature and missing for dewpoint simultaneously) for a long period of time (on the order of months).

d. Questionable land/water station assignments

The BCDG scheme restricts the influence of stations to gridpoints of the same type so that land station points influence only land gridpoints, ocean station points influence only ocean gridpoints, and inland water station points influence only inland water gridpoints. Geographic Information System (GIS) capabilities were used to access high-resolution coastal and lake shapefiles available in AWIPS. From this information, it was possible to flag each station as either land, ocean, or inland water. Then, for coastal sites, a detailed analysis was made to determine more exactly whether the observation would reflect temperature over water or over land. For instance, some stations are situated on spits of land or even on piers. Pictures of the sites were obtained where possible. In some cases, it was decided that the observation would not well represent either land or water, or might even be different in that regard depending on the synoptic conditions; in those cases, the station was omitted. This process is explained more fully in Sheets et al. (2005) and Sheets (2008). The integrity of the coastline in defining any contrast of temperature and dewpoint between water and land was better maintained by these omissions than keeping stations whose representation was ambiguous.

An example of the analysis of stations near the coastline is shown in Fig. 3. If we consider only the land/water gridpoints, black circled stations seem to be water stations; however, if we consider the coastline map overlaid on the land/water gridpoints and stations, these same stations seem to be land stations. This indicates the land/water assignments for these black circled stations would be questionable, so these stations were removed. It is not possible for a 5-km grid to accurately define a coastline. An important use of our analyses is at NWS WFOs where the AWIPS shapefiles are used, but the analysis should be as true to the actual coastline as possible, so some compromises had to be made.

e. Questionable station values

Inspection of real-time observation data revealed that some stations reported observations with different station types (e.g., both mesonet and C-MAN) at the same reporting time. Despite having different station types, observations reported from the same station at the same time should be identical. However, sometimes the differences between the observations were too large to be acceptable (see Table 1 for examples). Hence, another quality-check process was implemented. If the difference between the observation values from the same station at the same time was greater than 1 °F, all of the observations involved were removed. Observations are reported at varying degrees of precision. Sometimes a conversion from Celsius to Fahrenheit has been made. Such differences of reporting can result in differences of a few tenths of a degree Fahrenheit, but if the differences are more than one degree, then it is questionable which report is correct. Differences of up to 1 °F can be tolerated without serious consequences on 5-km grid. Most of the discrepancies were between mesonet and C-MAN stations.

5. BCDG Upgrades

a. Station-specific radii of influence

To handle highly variable data densities and to obtain the desired detail or smoothness over the analysis domain (Fig. 2), a specific radius of influence R was computed for each station. This was done in the following

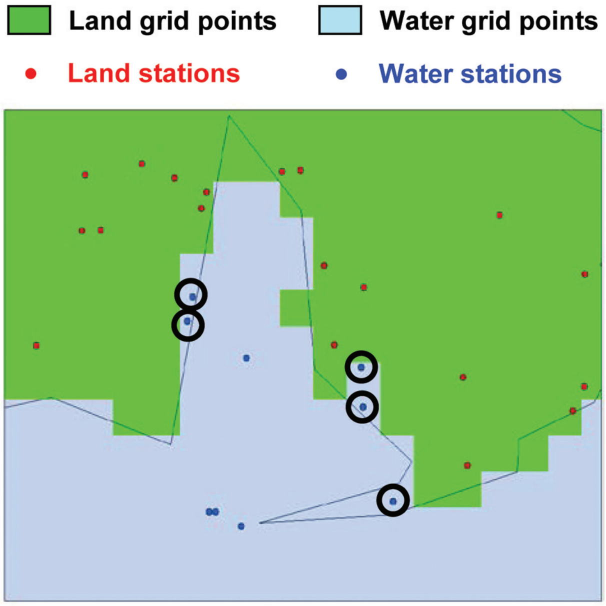

Fig. 3. An example of land/water designation and suspicious stations marked with black circles. The thin black line is the coastline from AWIPS shapefile.

Temperature Examples

Station ID	Station type	Day month year	HourMin (UTC)	T (°F)
ACXS1	Mesonet	22 Mar 2009	1245	48
ACXS1	C-MAN	22 Mar 2009	1245	45
SJOM4	Mesonet	22 May 2009	0350	76
SJOM4	C-MAN	22 May 2009	0350	62
NBLP1	Mesonet	15 May 2009	2248	32
NBLP1	Tide Gauge	15 May 2009	2248	73
KVDW	Mesonet	09 Jan 2009	2052	21
KVDW	METAR	09 Jan 2009	2052	32

Dewpoint Examples

Station ID	Station type	Day month year	HourMin (UTC)	Td (°F)
ELXC1	C-MAN	05 Dec 2008	2245	39
ELXC1	Mesonet	05 Dec 2008	2245	37
ACXS1	C-MAN	06 Mar 2009	1245	46
ACXS1	Mesonet	06 Mar 2009	1245	50
NAXR1	C-MAN	18 Apr 2009	0545	28
NAXR1	Mesonet	18 Apr 2009	0545	26

Table 1. Temperature and dewpoint examples for multiple reports at the same time with different station types for which the values are not identical.

manner. For every station, the first pass R (the largest R) was determined such that every gridpoint would have a correction made for it; the last pass R (the smallest R) must be such that the analysis shows the details that a skilled meteorologist would accept as real. The procedures to obtain an R satisfying the above requirements are as follows: for every gridpoint, up to 50 stations nearest it within a radius of 115 grid lengths are found along with the distances from the stations to the gridpoint. The largest station-to-gridpoint distance becomes the first pass R for that station. The subsequent values of R on the 2nd, 3rd, and 4th passes are determined by the products of the first pass R and 0.74, 0.54, and 0.41, respectively. Judgment is involved in any process that is dependent on highly variable data; this process guaranteed each gridpoint would be modified to fit the data, and the desired fine scale detail would be preserved.

The maximum number of stations (50) and the number of grid lengths (115) used in deriving R were determined by considerable experimentation performed for all available land stations. Different values were used, radii determined, and analyses made. These analyses were printed on large maps and flaws searched for and noted. The analyses are not very sensitive to minor variations in the values. They had to be large enough so that data sparse regions were treated effectively even when real-time data were missing; increasing them further was not necessary and that would lessen the detail in the analyses and increase computer time. This method works well for the land stations. For water stations, manual editing of the calculated values was necessary to accommodate very sparse observations and problems of frequently missing observations. To ensure that each water gridpoint has more than one water station within R on at least the first pass, the values computed were replaced subjectively by larger values. Also, a smaller R was assigned to some stations near ocean and lake shorelines and in Puget Sound and the Chesapeake Bay in order to better differentiate between open ocean and inlets and inland waterways. For instance, observations taken in Puget Sound should not affect the open ocean, and vice versa.

b. Quality control for inland water

As indicated in Section 2, a critical part of the objective analysis technique is to quality control data used in the analysis. As illustrated in Fig. 1, the BCDG scheme has an elaborate data checking mechanism which requires each datum to be within tolerance when compared to the existing pass of the analysis. If the tolerance is not met, before rejecting the datum, a buddy check is performed to see if at least one of the datum's two buddies agrees with it. The data throw-out (or acceptance) threshold criteria had initially been determined depending on analysis pass and month of the year, but not on the station land/water type (i.e., land, ocean, and inland water).

The data availability over the Great Lakes, in particular, is highly variable in space and time. In summer, buoy reports are available over the lakes; in winter they are not available, and consequently, the stations around the edges of the lakes are necessarily used. When observations are present from both the edge stations and buoys, differences between these observations may be such that the buoy data over deep water are rejected with the computed criteria. The buoy reports are much more representative of conditions over the lake than coastal stations, so in those situations the buoy reports should be used at the expense of the coastal stations. An example is provided in the left panel of Fig. 4. The data from the stations marked with red circles were rejected, which resulted in a poor analysis that only represented edge station characteristics.

In order to accommodate the larger variability

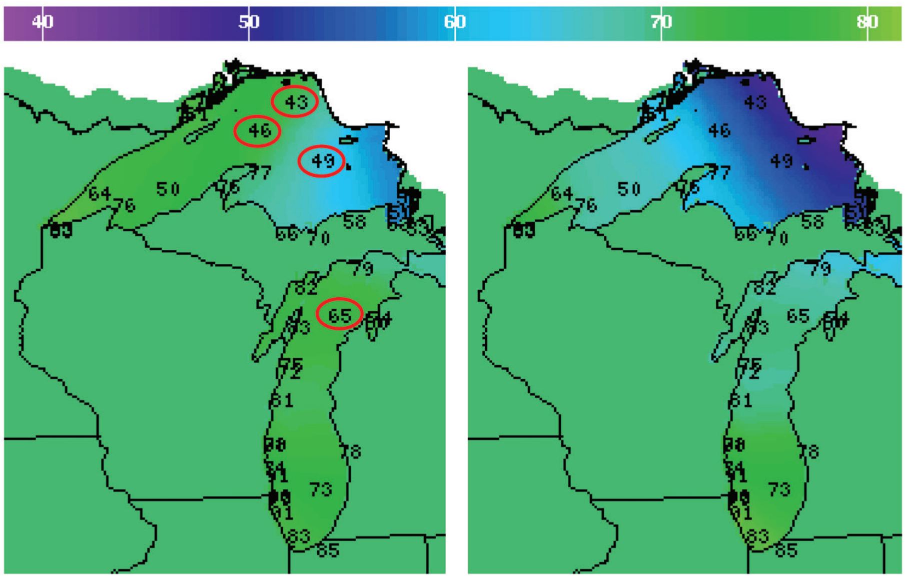

Fig. 4. Analyses of temperature (°F) with rejected data, marked with red circles (left) and using all the data (right), over inland waters of Lake Superior and Lake Michigan at 0000 UTC 25 June 2009.

in availability of observations over inland water, the threshold criteria needed to be increased, and it was found that a 50% increase was adequate. The altered criteria prevented undesired rejecting of data over deep water, and produced a more representative analysis (right panel of Fig. 4).

c. Augmentation with previous hour's data

As emphasized in Horel and Colman (2005), a real-time analysis should be available within roughly 30 minutes of the analysis time to satisfy the ongoing needs of the various communities. However, not all of the available surface observation data (observed within ±15 minutes of the analysis hour) are delivered by 30 minutes past the hour. To address the issue of observations missing at the analysis time, a new feature was added to the BCDG scheme: the capability to use an observation from the previous hour if the site did not report at the analysis time. An adjustment is made to the previous hour's observation in order to account for a possible temporal change from the previous to the analysis hour. The temporal change is computed by using five "closest"[3] stations which had both previous and analysis hour values. The average of the differences between the previous and the analysis hour values at these stations is added to the previous hour's observation to approximate the analysis hour's observation at the station whose real observation is missing. These adjusted observations are then used to augment the analysis hour observations which were available at the analysis time. Continuity from hour to hour is important for forecasting purposes, and harking back to the days when analyses were done "by hand," the previous hour's data were routinely consulted. The placement of fronts, for instance, can be facilitated when past data are viewed in conjunction with more current data. The process implemented in BCDG is designed to mimic that process.

Figure 5 shows examples of nighttime (0700 UTC 12 August 2009) temperature analyses for the western CONUS. Over the area shown, there were no strong frontal boundaries. The left panel shows the analysis which used only analysis hour data delivered by the analysis time (in which 10,537 reports were available for the whole CONUS domain). If we had waited for one more hour to collect more data, we would have produced the analysis shown in the middle panel (12,155 reports available by this time). This is more representative of the data reported at "observation time" (closer to the truth). The areas that indicate the most distinguishable differences between these analyses are marked with red circles. The right panel shows the analysis in which the augmentation method was used to handle observations that were missing at the analysis time. As can be seen from the right panel of Fig. 5, the analysis using both the adjusted previous hour and the analysis hour observations delivered by analysis time (total 12,464 reports) shows features more like the analysis shown in the middle panel than in the left panel. The augmenting capability implemented in the BCDG

---

[3] Candidate stations were identified by a preprocessor. First, a list of up to 20 stations within 35 grid lengths for land (125 for inland water, 175 for ocean) that had elevation differences of ≤ 45 m was prepared, in increasing order of distance from the base station (i.e., the station whose value is to be computed). Then, if < 20 stations were in the list, the list was augmented with stations that had elevation differences of > 45 m but ≤ 75 m. If the list still did not contain 20 stations, stations that had elevation differences of > 75 m were added to the list. This usually provided a list of 20 "nearby" stations for each station being analyzed. These were ordered in elevation bands according to distance from the base station. In an actual analysis, not all stations will be available, and only the closest 5 are used.

scheme improves the analysis by capturing more detailed features in the mountainous regions and depicting more representative temperatures over the Great Salt Lake.

6. Analysis Maps and Current Status

On the basis of the upgraded features and techniques described in the preceding sections, real-time hourly objective analyses of temperature and dewpoint are being produced for the CONUS on the NDFD grid. In addition, a post-processing step is necessary to ensure inter-element consistency. Specifically, the temperature must be greater than or equal to the dewpoint. Even though the temperature and dewpoint observations are consistent at each observation point, this does not guarantee consistency at each gridpoint. An inconsistency can be caused by either the temperature or the dewpoint being missing at a site (there are fewer dewpoint observations than temperature observations), the computed vertical change being generally different for temperature and dewpoint, or the analysis process not being perfect. BCDG checks each gridpoint, and in instances where the dewpoint exceeds the temperature, the dewpoint is set to the temperature. Among the possible methods to insure consistency, this was judged best because there are more observations of temperature than of dewpoint, which gives more credence to specific values of temperature at gridpoints.

Figures 6 and 7 display examples of the analyses[4] made for 0000 UTC 21 August 2009, and a corresponding surface weather map is provided in Fig. 8. The synoptic features included a low pressure centered in the Lake Superior/upper Lake Michigan area, with a weak cold front, mainly defined by surface winds and sea level pressure, extending southward through Illinois and the

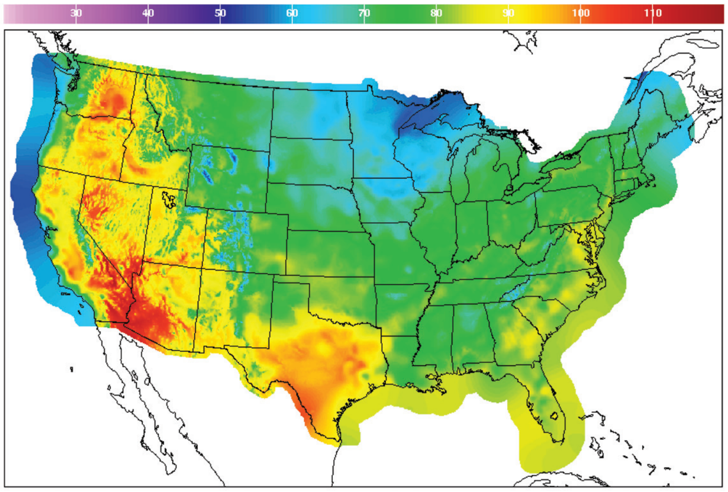

Fig. 6. Analysis of temperature (°F) valid at 0000 UTC 21 August 2009.

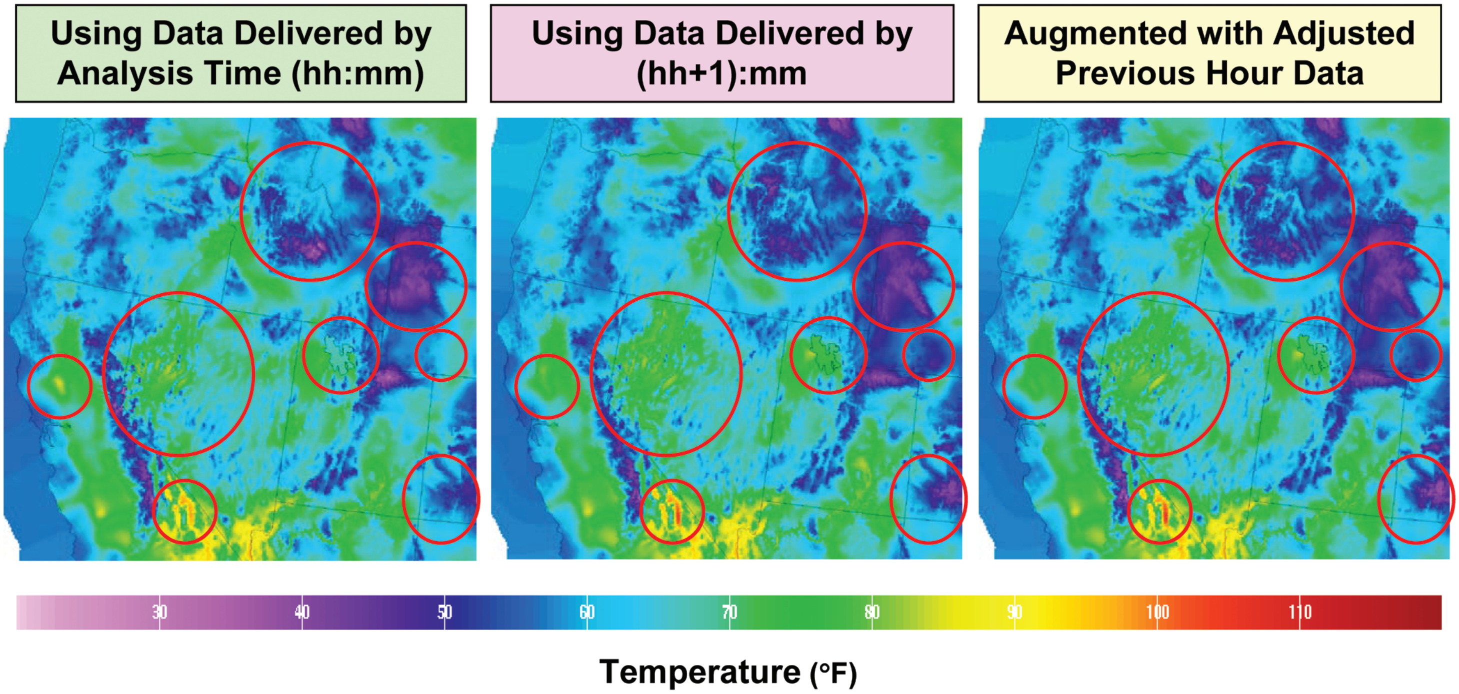

Fig. 5. Analyses of temperature for 0700 UTC 12 August 2009 in the western CONUS, with analysis hour data delivered by 0726 UTC (left), analysis hour data delivered by 0826 UTC (middle), and the adjusted previous hour data delivered by 0726 UTC as well as analysis hour data delivered by 0726 UTC (right).

---

[4] Actual analysis is performed in the NDFD rectangle shown in Fig. 2 and a final product is clipped to the CONUS area.

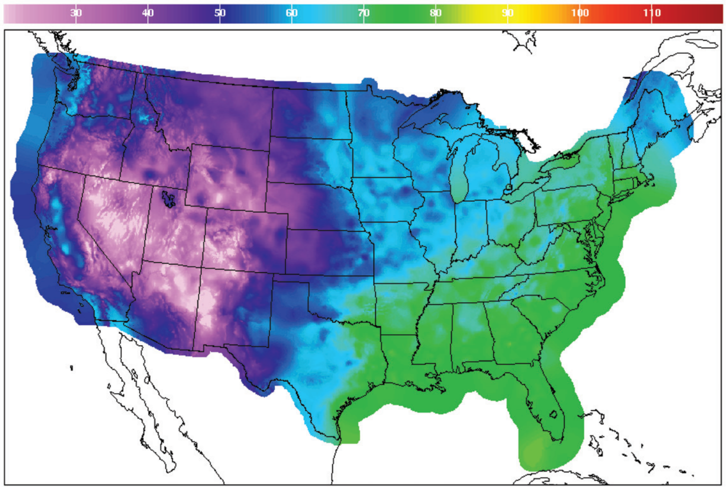

Fig. 7. Analysis of dewpoint temperature (°F) valid at 0000 UTC 21 August 2009.

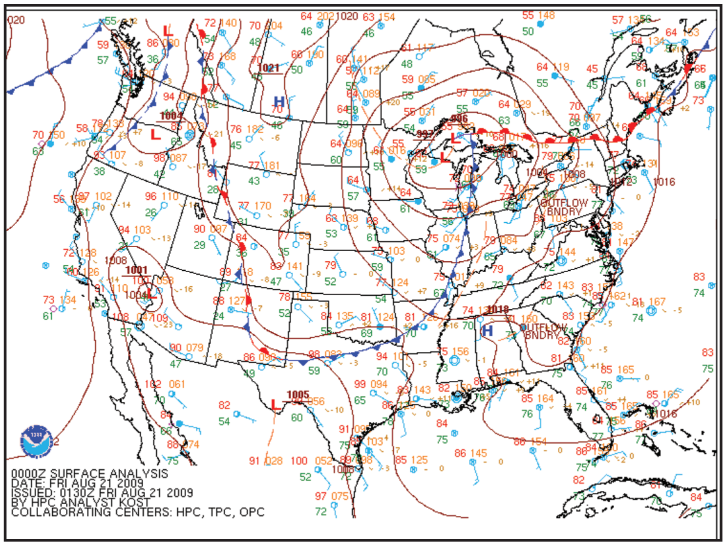

Fig. 8. Surface weather map valid at 0000 UTC 21 August 2009 (available online at http://www.hpc.ncep.noaa.gov/html/sfc_archive.shtml#CONUS).

boot heel of Missouri. This front is not evident in the temperature pattern, because there was only a very weak gradient across the front, but its influence, though ill-defined, shows up in the dewpoint analysis. These figures show that both the analyses of temperature and dewpoint are capturing well-defined terrain, major lakes, and coastal and ocean areas as well as local, mesoscale features. The predominant mesoscale features are due to terrain, but not all. Such features can be due to cloudiness over only a portion of a state, for instance, thereby lowering the temperature there during daytime. The opposite might be due to a sunny area in the midst of cloudiness. Patchy cloudiness at night can decrease the outgoing longwave radiation, and keep the temperature from lowering as it might without cloudiness. The analysis will not react strongly to one observation, but if several confirm a departure from the overall surroundings, then the feature conforms to the data values.

Real-time hourly objective analyses of temperature and dewpoint are now being produced and evaluated internally at MDL. In conjunction with the analyses, the errors involved in these analyses are being estimated by a method described in Glahn and Im (2011). The error estimation is intended to highlight areas where data that are suspect were used in the analysis, according to "predictors" defined to locate such problems. This is a statistical process in which analyses at five hour intervals over a year were made, and differences between each witheld datum and its value interpolated from the analysis grid were collected. Then a linear regression was formed to predict these differences in terms of such variables that could be calculated, such as terrain variability, data density, and data variability in space. These regressions, which are different for water and land and for temperature and dewpoint, were calculated from data at stations; then they were applied to each gridpoint to obtain a grid. At present, these analyses and analysis error estimates are available in the National Digital Guidance Database (NDGD).

7. Summary and Future Extensions

The BCDG analysis method developed to analyze point data in rough terrain and in regions with high data variability is being used by MDL to produce real-time analyses of hourly surface observations. A critical part of the analysis of the "surface observation" data is the quality checking procedure which ensures that incorrect data are not used in the analysis. This paper describes the extensive quality control procedures developed for pre-analysis (e.g., in making the station dictionary and preparing observation data), during-analysis (difference checks between station observation and analysis, and buddy checks), and post-analysis (inter-element consistency check) steps.

While making the station dictionary and preparing observation data, issues of questionable site information, stations reporting data at the same locations with different station names and types, stations repeatedly reporting the same values, suspicious land/water assignments near the coastlines, and multiple reports at the same time with different station types were identified and resolved. At the analysis step, the BCDG program performs efficient quality control procedures to decide whether to accept or throw out suspicious data.

In addition, in order to address spatial and temporal discontinuities of the analyses that are caused by observation data unevenly distributed over the analysis domain, data not delivered (transmitted) on time, and unpredictable data availability (missing data), new features were added to the analysis package previously reported in Glahn et al. (2009). One of the features is the capability of using an observation from the previous hour if the station did not report at the analysis hour. Adjustments are made to the previous hour's observations in order to account for typical diurnal changes from the previous to the analysis hour. These adjusted previous hour observations are then used to augment the analysis hour observations. To handle the heterogeneous distribution of the observations, a station-specific R for each individual station was implemented. A small R benefits the analysis in data dense regions where data are sufficient to define small scale features, and a larger R is necessary in data sparse regions and over deep waters.

This paper describes the analyses of temperature and dewpoint over the CONUS on a 5-km grid. The purpose of the BCDG analysis is to provide verification grids for gridded MOS and LAMP forecasts, and to add gridded LAMP nowcasts to the LAMP forecast suite. At present, the analyses are performed on the 2.5-km NDFD grid and transferred into the NDGD to be used by forecasters and for verifying the NDFD forecasts. These real-time hourly analyses on the 2.5-km grid can be found online at http://www.nws.noaa.gov/mdl/gfslamp/gfslamp.shtml. While only a few variables are currently being analyzed at MDL, the analysis product suite will be extended to include other weather elements. Analyses will also be made for Alaska and Hawaii.

Authors

Jung-Sun Im is currently a physical scientist and developer at NWS's Meteorological Development Laboratory (MDL) in Silver Spring, Maryland. Her prior affiliations include National Environmental Satellite, Data, and Information Service (NESDIS)'s Satellite Applications and Research (STAR) and NCEP's Hydrometeorological Prediction Center (HPC) in Camp Springs, Maryland. She earned a Ph.D. in atmospheric science from North Carolina State University in 2002. Her specialties are application of statistical methods to weather analysis/forecasting, uncertainty guidance, and satellite data analysis.

Bob Glahn was director of the Meteorological Development Laboratory of the National Weather Service until he retired in 2012; he currently is an NWS Scientist Emeritus at MDL. He holds a B.S. degree in Education from the Northeast Missouri State Teachers College, an M.S. degree in meteorology from the Massachusetts Institute of Technology, and a Ph.D. degree in meteorology with a minor in statistics from the Pennsylvania State University. His specialties are postprocessing numerical weather prediction data, mapping meteorological data into a digital database, and verification of weather forecasts.

Acknowledgments

The authors would like to thank Kari Sheets for her help with GIS data and usage, Scott Scallion for implementing the analysis products, and Judy Ghirardelli and David Ruth for their suggestions and comments.

References

Allen, R. L., 2001: Observational data and MOS: The challenges in creating high-quality guidance. Preprints, *18th Conf. on Weather Analysis and Forecasting/14th Conf. on Numerical Weather Prediction*, Ft. Lauderdale, FL, Amer. Meteor. Soc., 322-326. [Available online at http://www.nws.noaa.gov/mdl/synop/amspapers/18thwafobs.pdf.]

Benjamin, S. G., J. M. Brown, G. S. Manikin, and G. Mann, 2007: The RTMA background – hourly downscaling of RUC data to 5-km detail. Preprints, *23rd Conf. on Interactive Information Processing Systems,* San Antonio, TX, Amer. Meteor. Soc., P1.11. [Available online at http://ams.confex.com/ams/pdfpapers/119134.pdf.]

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

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

De Pondeca, M. S. F. V., and Coauthors, 2007: The development of the Real Time Mesoscale Analysis system at NCEP. Preprints, *23rd Conf. on Interactive Information Processing Systems,* San Antonio, TX, Amer. Meteor. Soc., P1.10. [Available online at http://ams. confex.com/ams/87ANNUAL/ techprogram/paper_118813.htm.]

_____, and Coauthors, 2011: The real-time mesoscale analysis at NOAA's National Centers for Environmental Prediction: Current status and development. *Wea. Forecasting*, 26, 593-612.

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

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, 11, 1203-1211.

——, T. L. Chambers, W. S. Richardson, and H. P. Perrotti, 1985: Objective map analysis for the Local AFOS MOS program. NOAA Technical Memorandum, NWS TDL 75, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, Washington, D.C., 34 pp. [Available from Meteorological Development Laboratory, Silver Spring, MD 20910.]

——, 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.

——, and J.-S. Im, 2011: Error estimation of objective analysis of surface observations. NOAA Technical Memorandum. NWS MDL 85, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, Washington, D.C., 22 pp. [Available from Meteorological Development Laboratory, Silver Spring, MD 20910.]

——, and J. P. Dallavalle, 2000: MOS-2000. *TDL Office Note* 00-1, National Weather Service, NOAA, U.S. Department of Commerce, Chapter 13, 48 pp. [Available from Meteorological Development Laboratory, Silver Spring, MD 20910.]

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

Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM), 1995: *Federal Meteorological Handbook* No. 1: Surface weather observations and reports. NOAA, U.S. Department of Commerce, 104 pp.

Seguin, W. R., 2002: AWIPS–an end to end look. Preprints, *18th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Orlando, FL, Amer. Meteor. Soc., J47.

Sheets, K. L., 2008: Employing geographic information systems for gridded model output statistics. Preprints, *24th Conf. on Interactive Information Processing Systems,* New Orleans, LA, Amer. Meteor. Soc., 5A.14. [Available online at http://ams.confex.com/ams/pdfpapers/ 131724.pdf.]

——, R. A. Trimarco, and K. K. Hughes, 2005: Using geographic information systems to develop gridded model output statistics. Preprints, *21st Conf. on Weather Analysis and Forecasting*, Washington, D. C., Amer. Meteor. Soc., 13B.1.