U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
OFFICE OF SCIENCE AND TECHNOLOGY INTEGRATION
METEOROLOGICAL DEVELOPMENT LABORATORY

MDL OFFICE NOTE 15-2

# OBJECTIVE ANALYSIS OF VISIBILITY AND CEILING HEIGHT OBSERVATIONS AND FORECASTS

Bob Glahn and Jung-Sun Im

October 2015

# OBJECTIVE ANALYSIS OF VISIBILITY AND CEILING HEIGHT OBSERVATIONS AND FORECASTS

Bob Glahn and Jung-Sun Im

## 1. INTRODUCTION

Observations of prevailing visibility and ceiling height are made at a variety of sites throughout the United States and other countries. These observations, and forecasts of them, are critical for aviation interests. Currently, most of these observations are made mechanically, such as at ASOS sites, but may also be made manually. Standards and requirements are specified in the Federal Meteorological Handbook No. 1 (FMH-1; OFCM 1995). Visibility and ceiling observations are most often made at major airports and are critical for aircraft takeoffs and landings. The observations are contained in METAR reports (OFCM 1995). However, there are many airports with significant air traffic where such observations are not made. Also, such observations are needed for helicopter operations at more or less random points, such as for search and rescue. This calls for analyses whereby estimates of the visibility and ceiling are made for each point on a high-resolution grid.

At locations where observations of visibility and ceiling are needed, forecasts are also needed, especially for just a few hours into the future. Such forecasts are made by a variety of numerical (i.e., NWP) and statistical techniques, as well as by forecasters. Manual (those made by forecasters) and statistical forecasts are usually made for sites where there are observations, the statistical forecasts requiring observations for development of the techniques specific to the sites or surrounding areas. The NWP techniques also require analyses (data assimilation) as initial conditions, but such analyses are made specifically for the purpose of initializing a sophisticated numerical model that concentrates heavily on the dynamics and physical processes not always well related to near-surface weather variables. Consequently, forecasts for a few hours of visibility and ceiling height are not well forecast by numerical models, which rely on algorithms based on other variables without adequate input from the all-important surface observations. Persistence (forecasting the same values as observed at initial time) is much better at an hour or two than numerical models (Rudack and Ghirardelli 2010; Glahn et al. 2014; Ghirardelli et al. 2015).

The Local Aviation MOS Program (LAMP; Ghirardelli and Glahn 2010) makes statistical forecasts every hour for hourly projections up to and including 25 h for most sites in the U.S. where visibility and ceiling observations exist. These forecasts have as input the current observation, output from simple advective models, and the synoptic scale MOS forecasts (Glahn and Lowry 1972; Dallavalle et al. 2004). Not only is the current observation used directly in the LAMP regression predictive equations, the LAMP models and MOS also directly consider the observation. The LAMP "station" forecasts of visibility and ceiling are distributed in text messages, and a web site (http://www.nws.noaa.gov/mdl/gfslamp/gfslamp.shtml) shows graphs and plots. Also, the forecasts are available in the National Digital Guidance Database (NDGD), the guidance companion to the National Digital Forecast Database (NDFD; Glahn and Ruth 2003). These gridded forecasts are made by analyzing the station forecasts with the BCDG method (Glahn et al. 2009; Glahn and Im 2011; Im and Glahn 2012; Glahn et al. 2012), which has been especially tuned with control parameters for discontinuous fields such as visibility and ceiling height. Analyses of visibility observations are also made to give the forecasters access to the existing conditions leading to the LAMP forecasts.

The purpose of this office note is to explain in some detail how the analyses of visibility and ceiling observations and forecasts are made. It is designed to not only explain in an overall sense how the analyses are made, but also to aid those using the BCDG software at MDL in the specifics. BCDG is comprised of a set of routines used to analyze several weather variables, and each application is handled by numerous parameters and specialized routines, some of which are specific to visibility and to ceiling; these specifics are explained in this office note. The lead subroutine for BCDG is **U155**, which performs housekeeping chores such as input and output and variable pre- and post-processing. The actual analysis is handled by **U405A** and its subroutines. The control file U405A.CN specifies the names of routines to be called for the specific variable being analyzed; each such subroutine can have eight parameters associated with it. These control variables, as well as many others, are explained here as they pertain to visibility and ceiling. Additional details are contained in the U155 and U405A write-ups in Glahn and Dallavalle (2000b); this office note is not a substitute for the write-ups, but rather particularizes some controls to visibility and ceiling height. The specific values herein are those we recommend be used.

In order to promote clarity, values and discussion that pertain only to visibility are in blue and that pertain only to ceiling height are in red. Subroutine names are usually in **bold**. The non-probabilistic observations and forecasts are, strictly speaking, neither continuous nor categorical, but are called categorical in this document to distinguish them from probabilistic.

## 2. THE BASIC ANALYSIS METHOD

The basic method is successive corrections, where each datum of the variable being analyzed is looked at in multiple passes, each time making a contribution to the correction at the gridpoints "nearby" by an amount indicated by the difference between the datum and the value interpolated from the then existing gridpoint representation of all the data. This method was put forth by Berghrossen and Doos (1955) and implemented at NMC (the forerunner to NCEP) for geopotential heights by Cressman (1959). A precept of this method, and actually of any method, is the assumption that given data at somewhat random points, values between those points can be reasonably estimated. This is a valid assumption for smooth fields like geopotential height and sea level pressure, where dynamic constraints are at play, and to a somewhat lesser extent for 2-m temperature or 10-m sustained wind. The assumption is in question for discontinuous fields like visibility; visibility can go abruptly from zero to 10 miles or more from one hour to the next and from one station to its nearest neighbor. However, to make an analysis, some method must be used to estimate a value at each gridpoint based on the available data in the vicinity; the BCDG method, tuned for the purpose, is probably as good as or better than any other. Many control parameters and techniques have been added since the method was first used by Cressman (1959).

In addition to the importance of *spatial* consistency of analyses of observations (obs) and forecasts valid at the same time, *temporal* continuity is important from the obs (0-h) to the 1-h forecast, and from the x-h forecast to the x+1-h forecast. The treatment of the data to maximize this continuity without impinging upon the accuracy of any individual analysis is part of BCDG. LAMP produces forecasts out to 25 hours, but forecasts are needed to 36 hours or more for making Terminal Aerodrome Forecasts (TAF) for some locations. These longer-range forecast grids can be made by analyzing MOS forecasts (Dallavalle et al. 2004), which are produced less frequently and at 3-h projection intervals, by interpolating in time from the appropriate MOS run to the hour needed. Merging these forecasts into an analysis around the 25-h projection and continuing with MOS is part of BCDG for non-probabilistic ceiling and visibility forecasts but not for probabilistic forecasts.

The values of many variables change with elevation, and most observations are made at stations at rather low elevations compared to the surrounding terrain. BCDG estimates the vertical change with elevation (called "lapse rate" here) from the observations being analyzed, as explained in Glahn et al. (2009), and applies that change in the correction algorithm.

## 3. TREATMENT OF THE DATA TO ANALYZE

### A. Observations

Visibility observations are predominantly used in the U.S. in miles. The "reportable" values are given in FMH-1 (OFCM 1995). Much detail is possible at low visibilities, and basically reports are at 1-mile increments from 3 miles upward to 15 miles, then at increments of 5 miles. With automation came the capping of most observations at 10 miles, although some may occur at higher values, mostly manual observations. For the analysis, all values > 10 are set to 10.

Similarly, ceiling observations are dealt with in FMH-1. Ceiling is defined to be the height above ground of the lowest cloud of $\geq$ 5/8 coverage. The reports are in hundreds of feet at increments of 100 ft below 5,000 ft, then at 500-ft increments to 10,000 ft, then at 1,000-ft increments. Other rules apply when the sky is obscured. Because of automated instrument capabilities, few observations occur above 12,000 ft. For the analysis, all values > 130 hundreds of ft and "unlimited" are set to 130.

### B. LAMP forecasts

LAMP visibility and ceiling forecasts are made in categories, and not as quasi-continuous values. This was necessary because of the violently non-normal distributions. Many attempts to statistically deal with highly non-normal variables as continuous, especially where the rare values are the most important, has met with limited success, if not outright failure; these variables include sky cover and precipitation amount, as well as visibility and ceiling. So, the REEP method (Wilks 2011) due to Miller (1958) is used to produce equations for each of six (seven) categories of visibility (ceiling) such that when applied yield the probability of the respective category occurring. These equations are produced by the program **U602**, which was written specifically to promote consistency from one projection to the next. It has been found that REEP works best when both the predictand and predictors are cumulative from above or below, rather than discrete. The six cumulative categories LAMP uses for visibility are the first six in the 2$^{nd}$ column in Table 1. When the probabilities of each of the cumulative categories is subtracted from the probabilities of the next higher one, the probabilities of the discrete categories in the 3rd column of Table 1 result. A 7$^{th}$ category can be formed by subtracting the probability of the 6$^{th}$ category from unity. In a similar manner, the ceiling categories are defined in the last two columns of Table 1. LAMP produces for distribution and archival the cumulative probabilities.

The LAMP regression equations were developed on a regional basis. That is, the data from stations within a region in which it was thought the predictand/predictor relationships were similar were grouped, and all such stations share the same equations. Forecasts are made each hour for hourly projections out to 25 h for about 2082 stations.

Table 1. Category definitions of visibility in miles and ceiling in hundreds of feet. The cumulative categories thresholds in the table are defined as they are used; these values were used to avoid possible roundoff errors. The discrete categories are as they are usually defined.

| Category Number | Visibility (mi.) Cumulative Categories | Visibility (mi.) Discrete Categories | Ceiling (hds ft) Cumulative Categories | Ceiling (hds ft) Discrete Categories |
|---|---|---|---|---|
| 1 | < 0.49 | < 0.5 | < 1.5 | < 2 |
| 2 | < .95 | ≥ 0.5 and < 1.0 | < 4.5 | 2 -4 |
| 3 | < 1.95 | ≥ 1.0 and < 2.0 | < 9.5 | 5 -9 |
| 4 | < 2.95 | ≥ 2.0 and < 3.0 | < 19.5 | 10 - 19 |
| 5 | ≤ 5.05 | ≥ 3.0 and ≤ 5.0 | < 30.5 | 20 - 30 |
| 6 | ≤ 6.05 | > 5.0 and ≤ 6.0 | < 65.5 | 31 - 65 |
| 7 | > 6.05 | > 6.0 | < 120.5 | 66 - 120 |
| 8 | | | ≥ 120.5 | > 120 |

While probability forecasts provide more information than non-probabilistic forecasts, the aviation community requires discrete values of visibility and ceiling, and generally in the same terms as the reportable values; the low values are important, while high ones less so. To make such forecasts, probability forecasts from the equations are made for each category for the developmental sample by **U700**, and a threshold is defined for each category such that the bias is in the range 0.98 to 1.20 and within that range the threat score (Palmer and Allen 1949; Wilks 2011), which is the same as the critical success index (Donaldson et al. 1975; Shaffer 1990), is maximized. These thresholds are calculated in program **U830**. Separate thresholds are derived for each of the regions used in equation development. Discrete category forecasts are then made by applying the thresholds to the cumulative probabilities in program **U710**.

Forecasting visibilities and ceilings in these operationally important discrete categories is necessary for airports and furnish guidance for producing the TAFs. However, such a set of values is not very robust when making an objective analysis. So, in order to have a more continuous range of values, each category is scaled within the limits of its definition by using the probability of that category occurring. In subroutine **SCLVIS** for visibility and **SCLCIG** for ceiling, the maximum and minimum probabilities for all stations with a specific categorical forecast are found. For the first six categories (seven for ceiling), the station with the minimum probability is given, for analysis purposes, a value near the high end of the category definition. Similarly, the station with the maximum probability is given a value near the low end of the category. In so doing, the assumption is made that a low probability indicates the threshold was just barely tripped, and the next higher category would have been made if the probability were a little lower. Stations with intermediate values are given scaled values within the category limits, based on the category probabilities compatible with the limiting values stated above. For the $7^{th}$ visibility category (8th for ceiling), which is cumulative from above, the opposite is true; a high probability for visibility indicates a forecast far above 6 miles and a low probability means a forecast of just above 6 miles.[1] Because the preponderance of observations above 6 miles is 10 miles, the scaling of this category is between 6 and 30 miles. After this scaling, the forecasts are truncated to 10 miles, resulting in a large number of forecasts of 10 miles, consistent with the distribution of observations. This essentially gives a continuous range of values to analyze. A similar process deals with ceiling; the scaling of the last category is between 6,500 and 130,000 ft, with the result many forecasts of 130,000 ft results. The control variables for SCLVIS and SCLCIG are given in Table 2. Analyses are made for the specific value forecasts of visibility and of ceiling, and for the cumulative probability forecasts of visibility and of ceiling.

The values in the following tables, unless otherwise stated, pertain to analysis of both observations and LAMP forecasts. If the values for the LAMP forecasts are different from those for observations, the values for LAMP are shown in parentheses. The values also pertain to both LAMP specific value analyses and probability analyses. If the values for probabilities differ from the specific value ones, they are shown in brackets and in green.

Table 2. The control values for subroutine **SCLVIS**, and as used in its subroutine **VISMBO** for visibility and **SCLCIG** and **CIGMBO** for ceiling. These routines are not used for analysis of observations or probabilities.

| Variable Name as Read | NCAT | NSCALE | CONST | IPREX1 | IPREX2 | PREX3 | PREX4 | PREX5 |
|---|---|---|---|---|---|---|---|---|
| Variable Name as used in SCLVIS/ SCLCIG | NCAT | NSCALE | CONST | IBSTRT | IBEND | CAP | N/A | N/A |
| Value for Visibility | 7 | 0 | 1 | 2 | 6 | 30 | | |
| Value for Ceiling | 8 | 0 | 1 | 2 | 6 | 130 | | |

The control variables read in U405A are generic, and do not necessarily have meaning as used in the subroutines. The six used for SCLVIS and SCLCIG are defined below:

NCAT— The number of categories = 7 for visibility and 8 for ceiling.
NSCALE—Before returning the data to U405A in XDATA( ), the processed values are further scaled by XDATA( ) = XDATA( )*CONST*10**NSCALE. With NSCALE = 0 and CONST = 1, the values are not changed.
CONST— See NSCALE.
IBSTRT— When the observed visibility or ceiling is in the same category as forecast, the forecast is given the observation value for projections < IBSTRT, in this case projection 1.

[1] The scaling would be more scientifically sound if the process were carried out for each of the regions over which the equations were developed and thresholds determined. However, it is believed this set of continuous values is better for doing the analysis than just assigning seven specific values arbitrarily chosen within each category.

IBEND— When the observed visibility is in the same category as forecast, and the projection is < IBEND and ≥ IBSTRT, then the forecast is given a value between the observation and the forecast, such that the forecast is weighted by W = (LAMP projection-IBSTRT)/(IBEND-IBSTRT) and the observation is weighted by 1-W. With the values used, at projection 3, the forecast will be weighted 1/4 and the ob weighted 3/4.

CAP– The value to use as the upper end of visibility category 7 (8 for ceiling) for probability scaling. The value of 30 will give many values > 10, which are then truncated to 10.01 (the 10.01 vice 10.0 is for BCDG; frequencies > 10 will occur). If more (less) values of 10 are desired, increase (decrease) the 30. The upper category of ceiling contains many "unlimited" reports (has no value), which are designated as "888." All reports in this category are capped at 130, a value greater than 120 in order to get a good demarcation in the analysis at 12,000 ft. If a value of 120 were used instead of 130 (the exact value of 130 is not critical), many gridpoints would end up in the lower category.

With these preprocessing steps, a reasonably continuous forecast is available for analysis at and between 0 and 10 for visibility (130 for ceiling) with a preponderance of 10s (130s).

## 4. DETAILS OF THE ANALYSIS

Most of the controlling parameters are contained in U405A.CN. The first such row of input contains values for 15 control variables. Table 3 gives those values, and their explanations follow. They are the same for analysis of observations and non-probabilistic forecasts and are only slightly different for probabilistic forecasts.

Table 3. The 15 control parameters in the first row of U405A.CN.

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Variable Name | NPASS | IFSTGS | IGUESS (1) | IGUESS (2) | IGUESS (3) | IGUESS (4) | IBACKN | IBACKL |
| Value for Visibility | 6 [5] | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Value for Ceiling | 6 [5] | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | | | | | | | | |
| Position | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| Variable Name | GUESS | TITLE | NSMTYP | I400ADG | LAPFG | LIMITX | IVRAD | |
| Value for Visibility | 10 [0] | VISIBILITY OBS (FCST) [PROB] | 8 | 1 | 0 | 2 | 1 | |
| Value for Ceiling | 120 [0] | CEILING OBS (FCST) [PROB] | 8 | 1 | 0 | 2 | 1 | |

NPASS– Six analysis passes (5 for probabilities) are made over the data, corrections being made each time.

IFSTGS– The first guess is neither gridprinted nor saved on output (= 0).

IGUESS(1)–The first priority for a first guess is a constant (=1).

IGUESS(2-4)–Because the first guess is a constant and is always available, other possibilities in case the first is unavailable are unnecessary (= 0).

IBACKN– Has no relevance for these analyses.

IBACKL– Has no relevance for these analyses.

GUESS– The constant used for the first guess = 10 for categorical visibility, 120 for categorical ceiling, and 0 for forecast probabilities.

TITLE– Visibility and ceiling are being analyzed. This is used for diagnostic print.. Titles include "OBS" for observations, "FCST" for categorical forecasts, and "PROB" for probabilities.

NSMTYP– **SPOTRM** is to be used for smoothing (=8), and the subroutine **SPOTRM** must be specified later in the U405A.CN.

I400ADG– Diagnostics are to be written to unit KFILDO, the standard FORT.12 output (=1). Use 0 to not write optional diagnostics.

LAPFG– The change of the element with height, the "lapse rate," will be calculated from the visibility and ceiling data rather than using upper air model forecasts (=0). Units are in terms of the element being analyzed per m.[2]

LIMITX– When analyzing several projections in one run, the input control data in U405A will be printed only twice (= 2) for that element.

IVRAD– Variable radii calculated in preprocessor **U178** will be used (= 1), unless **SWITCH** switches to constant radii (see below).

The second row of U405A.CN contains 26 values; they are defined in Table 4, and explanations are provided.

IQUALC– Column 1 is the column in the quality control information furnished in the dictionary used for visibility and ceiling (see Appendix I of the U405A write-up for details).

QUALWT(1-4)–All data will be weighed fully and equally (=1).

ISETP— 2 indicates that after the analysis (including smoothing), the gridpoint closest to a point being analyzed is set to the value of the point being analyzed. By doing this, the observations or forecasts can be retrieved from the grid except over the ocean; if a gridpoint is closest to two (or more) such points, only the closer (closest) can be retrieved. The ocean smoother **ORVWSM** follows, and smooths out these exact values.

ILS– For analysis of observations, ILS = 0 meaning water and land will be analyzed together. There are very few observations over water, too few to analyze over water separately. However, for LAMP, backup equations (equations without initial obser-

---

[2] A check is made in subroutine **LAPSE**, where the lapse rates are calculated, of the calculated values to make sure they are "reasonable," where reasonable is defined by judgment after testing. If the absolute value of a calculated value is deemed too large to be used, it is set to either the largest permissible value or zero, depending on element. The checks and decisions for ceiling and visibility are:
Limits for obs or LAMP categorical ceiling = - 0.0328 and +0.0328; if exceeded, set value to the value exceeded.
Limits for obs or LAMP categorical visibility = -0.001 and +0.001; if exceeded, set value to 0.
Limits for LAMP ceiling probability = -0.0001 and +0.0001; if exceeded, set value to the value exceeded.
Limits for LAMP visibility probability = -0.00005 and +0.00007; if exceeded, set value to 0.

vations as predictors) for coastal stations have been applied to nearby points over water. Therefore land and water can be analyzed separately (ILS = 1); this is being done primarily because the water forecasts are not as accurate as the land based ones, because the equations were not developed for water and do not have "persistence" as an input. However, this separation is only partial, because the land forecasts are allowed to influence water (WTLTW = 1) (but not vice versa, WTWTL = 0).

IBPKN— For visibilities, both positive and negative lapse rates are dealt with evenly; no preference is given to visibilities increasing or decreasing upward, as both are possible. (Low clouds would indicate visibilities decreasing upward from the surface, but with low stratus or fog, visibilities and ceilings may increase upward.) Ceilings can also increase or decrease upward, but it is believed decreasing is more prevalent. (A cloud deck would be lower relative to the ground as the elevation of the ground increased.) Therefore a positive increase is treated as unusual (+1). For probabilities, the opposite is most likely (-1).

Table 4. The 15 control parameters in the second row of U405A.CN.

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Variable Name | IQUALC | QUALWT (1) | QUALWT (2) | QUALWT (3) | QUALWT (4) | ISETP | ILS | IBPKN |
| Value for Visibility | 1 | 1 | 1 | 1 | 1 | 2 | 0 (1) [1] | 0 |
| Value for Ceiling | 1 | 1 | 1 | 1 | 1 | 2 | 0 (1) [1] | +1 [-1] |
| | | | | | | | | |
| Position | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Variable Name | LPNO | HGTTHA | HGTTHB | IWITH | ISEED | ITYPR | NBLEND | CSTSM |
| Value for Visibility | 99 | 20000 | 20000 | 0 | 1 | 1 | 0 | 0 (.25) [.25] |
| Value for Ceiling | 99 | 20000 | 20000 | 0 | 1 | 1 | 0 | 0 (.25) [.25] |
| | | | | | | | | |
| Position | 17 | 18 | 19-24 | 25 | 26 | | | |
| Variable Name | N4P | NCLIP | NSHLN (1-6) | WTWTL | WTLTW | | | |
| Value for Visibility | 36 | 1 | 111111 | 1 (0) [0] | 1 | | | |
| Value for Ceiling | 36 | 1 | 111111 | 1 (0)[0] | 1 | | | |

LPNO– The value 99 indicates all the points produced for a particular station by the preprocessor **U174** to form pairs with the base station will be used in computing the lapse rate. This can range up to 100. Internally, four or more pairs (set in subroutine **LAPSE**) are required before a computed lapse is used; otherwise a lapse of zero is assumed. Such zero values will, of course, water down any positive or negative values calculated for stations in the vicinity and reduce the "terrain effect."

HGTTHA, HGTTHB–These values of 20,000 m elevation are high so that the adjustment to gridpoints will not be curtained by the difference in elevation between the datum and the gridpoint.

IWITH– The number of data values to withhold from the analysis = 0. This has no current use in the analysis of ceiling and visibility.

ISEED, ITYPR–These have no relevance to operations as they relate to withheld data values.

NBLEND– 0 disables a first guess blending option.

CSTSM– The value 0.25 designates a very light smoother within one gridlength of a water/land boundary for LAMP. This recognizes the change from water to land may not be abrupt. When the land and water are analyzed together, no such smoothing is necessary, so 0 is used.

N4P– N4P is the number of gridpoints that can be used by the special "interpolation" routine **ITRPSX** for water or mixed lake/land points (= 36).

NCLIP– 1 indicates the output will be clipped to the NDFD grid.

NSHLN(1-6)–These indicate smoothing options in the terrain following smoother **SMOTHG** and should be 1.

WTWTL– When land and water are analyzed separately, as with LAMP forecasts, partial mix is still possible. WTWTL (weight water to land) = 0 indicates that LAMP water forecasts will not affect land. When land and water are analyzed together, the setting of WTWTL is not used, but set it = 1.

WTLTW– When land and water are analyzed separately, as with LAMP forecasts, partial mix is still possible. WTLTW (weight land to water) = 1 indicates that LAMP land forecasts will affect water. When land and water are analyzed together, the setting of WTLTW is not used, but set it = 1. Because obs over water are almost non-existent and LAMP forecasts are generally less accurate over water than land, land gridpoints should affect water gridpoints, but not vice versa.

The next two rows pertain to analyzing two or more cycles together; NORUNS = 0 indicate we are not doing that. The next 7 rows contain variable IDs, as shown in Table 5 (see Glahn and Dallavalle 2000a for a description if variable IDs).

Row 1 in Table 5 contains the ID of the desired analyzed field, and row 2 is the ID of the data being analyzed. For the non-probabilistic forecasts, row 6 contains the ID of the corresponding MOS forecasts used to augment the LAMP forecasts near and past the 25-h projection. Up to three cycles of MOS forecasts will be accessed to find the forecasts, starting with the most recent cycle. Row 7 contains the ID of the corresponding OBS used to augment the forecasts in the early projections. Merging MOS with LAMP beyond 25 hours is not done for the probability forecasts. Note that DD = 5 is not required (the DD comes from the variable being analyzed in U155.CN, but DD = 08 is required for MOS forecasts (they could be furnished by MOS from a different NCEP model).

Table 5. The CCCFFFBDD ID entries in U405A.CN.

| Row Number | Visibility OBS | Visibility Forecasts Non-Prob | Visibility Forecasts Prob | Ceiling OBS | Ceiling Fore-casts Non-Prob | Ceiling Fore-casts Prob |
|---|---|---|---|---|---|---|
| 1 | 728100085 | 228160000 | 228130200 | 728000085 | 228080000 | 228070200 |
| 2 | 708100000 | 208131000 | 208130200 | 708000000 | 208071000 | 208070200 |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | 208131008 | | | 208051008 | |
| 7 | | 708100000 | | | 708000000 | |

Several rows of control information follow in U405A.CN in groups of four rows each, each row, respectively, pertaining to one of the four first guess options; the definitions are contained to the far right in U405A.CN. The mesh length is designated as 3, meaning 2.5 km.[2] For the error limits, options are provided for each month. The error limits are set to zero, meaning no data will be tossed out. It is impossible to tell from either spatial or temporal continuity whether or not an observation or forecast is in error, so all are accepted. Type 3 corrections are made for each pass (see Glahn et al. 2009). The non-zero 4 in the 6th column of smoothing controls means that smoothing will be done on only the last (6th) pass (5th for probabilities). The constant radii of influence are set by pass at 84, 50, 25, 14, 9, and 5 gridlengths. Data are used outside the grid by only a small amount, except on pass 1 data are allowed 1 gridlength outside the grid for this constant first guess. (Note that for pass 1, interpolation will always give the constant; for other passes extrapolation far outside the grid is not advisable.) The type of interpolation is set to 1. Full elevation correction is used for the first three passes only; it has been found that at the later passes when only a few, even only 1, stations affect a gridpoint, the elevation correction from only a few stations is too erratic and can give too large an effect. The distance in gridlengths and the fraction of elevation an "unusual" lapse rate affects a gridpoint are not relevant for visibilities, because both positive and negative lapse rates are being treated the same (not unusual). The constant radii prescribed above (but not the variable radii) are multiplied by 3.5 for water points because of the low density of water data points.

Internal pre- and post-processors are indicated as subroutine names with up to 8 control parameters each. Because variable radii have been specified, the name of the file containing them is read in to use in the preprocessor RDVRHL. There are no other control variables for RDVRHL. The file will have been prepared by **U178A**.

The visibility (ceiling) obs are faired into the LAMP analysis by SCLVIS (**SCLCIG**) designated in the LAMP U405A.CN (see Table 2). The parameters for SCLVIS and SCLCIG are given in Table 2. If SCLVIS (SCLCIG) were not called here, the obs would not be blended into the

[2] When the AWIPS grid, that was later used for the NDFD, was designed with a basic gridlength of 80 km (Glahn 1988) with successive halving possible (e.g., 40, 20, 10), it was not envisioned the gridlength would ever be less than 5 km and fractional gridlengths be used. At 2.5 km, a proxy was necessary, and "3" is used. A "1" would be used for 1.25 km.

LAMP forecast, and the LAMP categories would not be given fractional values according to the LAMP forecast probabilities.

**SPOTRM** is a crucial subroutine, especially for discontinuous fields. It is really a post-processor, but it is listed in the preprocessor group. Its control parameters are shown in Table 6 and are explained below.

Table 6. The control values for subroutine **SPOTRM**.

| Variable Name as Read | NCAT | NSCALE | CONST | IPREX1 | IPREX2 | PREX3 | PREX4 | PREX5 |
|---|---|---|---|---|---|---|---|---|
| Variable Name used in SPOTRM | NPASSP | NSMNUM | NOPTN | DIFFA | LAKE/ OCEAN | DISTX | DPOWER | RAY |
| Value for Visibility | 6 [5] | 7 | 1 | 75 | 55 | 2 | 2 [1] | 1.10 [1.25] |
| Value for Ceiling | 6 [5] | 7 | 1 | 75 | 55 | 2 | 2 [1] | 1.10 [1.25] |

NPASSP– SPOTRM will operate on only the last pass, 6 for non-probabilistic and 5 for probabilistic.

NSMNUM–This indicates how many times, if any, the 9-point, terrain following smoother **SMOTHG** is used following SPOTRM. NSMNUM minus 4 indicates the number of passes, so in this case 7 will give three SMOTHG smoothings. This is to eliminate chatter left by SPOTRM. Note that SPOTRM will not smooth the four gridpoints surrounding a data point, but SMOTHG will. ISETP (see above) operates after this, so a datum is preserved at the closest gridpoint over land.

NOPTN– Missing values will be indicated in LTAG( ) (=1).

DIFFA– 75 is the maximum difference in terrain heights in meters between the point being smoothed and a point contributing a smoothing value. This keeps from smoothing across large terrain features.

LAKE/OCEAN–Two digits are read, the first pertaining to lakes and the second to ocean areas. The “5" means water and land are smoothed, land gridpoints contribute to water smoothing, but water gridpoints do not contribute to land smoothing.

DISTX– The value to multiply by R(1), the first pass constant radius, to define the radius over which to search to find the closest station to a gridpoint (= 2). For stations in Canada and over water, DISTX is doubled internally in SPOTRM.

DPOWER– The exponential weighting to use; a “2" indicates some smoothness with the major change about halfway between the two stations affecting a gridpoint, rather than a more gradual change. This is because there is likely no smooth transition from one discrete value of visibility or ceiling to another, but rather a “spatial persistence” factor. A smoother field is desisred for probabilities (DPOWER = 1).

RAY– Smoothing for each gridpoint is over a circle 1.10 (1.25 for probabilistic) times the distance to the closest station.

**SWITCH** can be used when IVRAD = 1. Its one parameter indicates on which pass and thereafter to revert from variable radii to constant radii. Here, the switch is made on the penultimate pass, (5 for non-probabilistic and 4 for probabilistic). The analysis seems not very sensitive to the switch-over point.

**SCLVIS** (SCLCIG) when called for analyzing LAMP will calculate the frequencies of the LAMP forecasts in categories. SCLVIS (SCLCIG) is not called when analyzing obs, so **VISFRQ** (CIGFRQ) does that for obs. VISFRQ and CIGFRQ have no control parameters.

**CVLMPM,** called only when analyzing LAMP categorical forecasts, merges the MOS forecasts with LAMP around the max LAMP projection of 25 h. Its two parameters are shown in Table 7 and then explained.

Table 7. The two control values for subroutine **CVLMPM**.

| Variable Name as Read | NCAT | NSCALE | CONST | IPREX1 | IPREX2 | PREX3 | PREX4 | PREX5 |
|---|---|---|---|---|---|---|---|---|
| Variable Name used in CVLMPM | N/A | N/A | N/A | IBSTRT | IBEND | N/A | N/A | N/A |
| Value for Visibility | | | | 22 | 26 | | | |
| Value for Ceiling | | | | 22 | 26 | | | |

The merging starts at IBSTRT+1 hours and ends at IBEND-1 hours. At IBSTRT hours, the grid is totally LAMP; at IBEND, the grid is totally MOS. At 23, 24, and 25 hours, LAMP has weights 0.75. 0.50, and 0.25, respectively, and MOS weights making the total weights unity. MOS analyses can be made as far out as desired at hourly, or other, projections.

After the terminator, analysis of both obs and LAMP use the lakes and ocean smoother **ORVWSM**. LAMP and MOS forecasts are sparse, and obs almost non-existent over water, so heavy smoothing is reasonable and necessary over the ocean. The controls and explanations for ORVWSM are in Table 8 and following.

Table 8. The control values for subroutine **ORVWSM**.

| Variable Name as Read | TLOA | SETLOA | THIA | SETHIA | CONSTA | NSCALA | EX1A | EX2A |
|---|---|---|---|---|---|---|---|---|
| Variable Name used in ORVWSM | SHOREA | SHOREB | NOL | SETNEG | CONSTA | CONSTB | IOCEXT | IOCINC |
| Value for Visibility | 5 | 30 | 1 | 0 [1] | 0 | 0 | 30 | 1 |
| Value for Ceiling | 5 | 30 | 1 | 0 [1] | 0 | 0 | 30 | 1 |

SHOREA– The smoothing starts 5 gridlengths from shore; no smoothing is done inside that 5-gridlength zone. A mask is used that contains, over water areas, distances from shore. The assumption is that the gridpoints close to shore are influenced by land station values and are more likely to be correct than gridpoints farther from shore.

SHOREB– Full smoothing is done 30 and more gridlengths from shore. From SHOREA to SHOREB, smoothing is gradual, weighted by distance from shore.

NOL– The "1" indicates ORVWSM smoothing will be over the ocean but not the lakes. This is probably too heavy a smoother to use over lakes. (Set to 0 to also smooth over the lakes as well as ocean.)

SETNEG– Negatives that may exist after the analysis is complete are not set to zero before smoothing categorical values (these are zeroed by POST, see Table 10). However, negative probabilities are set to zero before smoothing.

CONSTA, CONSTB–Have been disabled.

IOCEXT– ORVWSM is a ray smoother. For a gridpoint being smoothed, a ray is traced in the grid in each of 16 directions, and the average value is taken to get the smoothed value. The length of the ray is 30 gridlengths. The ray stops if land is reached.

IOCINC– Sampling along the ray is taken at 1 gridlength (IOCINC = 1) intervals.

Final postprocessing is done by **POSTPM** for categorical visibility and **POST88** for categorical ceiling grids. Their control parameters are explained in Table 9 and following; they are used for both analysis of obs and forecasts.

Table 9. The control values for subroutine **POSTPM** and **POST88**.

| Variable Name as Read | TLOA | SETLOA | THIA | SETHIA | CONSTA | NSCALA | EX1A | EX2A |
|---|---|---|---|---|---|---|---|---|
| Variable Name used in POSTPM/ POST88 | TLO | SETLO | THI | SETTHI | CONST | NSCAL | SET | PM |
| Value for Visibility | 0 | 0 | 9998.5 | 9999. | 1 | 0 | 10 | 0.9 |
| Value for Ceiling | 0 | 0 | 9998.5 | 9999. | 1 | 0 | 120 | 1. |

TLO– Any values less than TLO = 0 with be changed to SETLO = 0.

SETLO– See TLO.

THIA– Any value greater than 9998.5 will be changed to SETTHI = 9999. (missing).

SETTHI– See THIA.

CONSTA, NSCAL–Scaling is X = X*CONST*10**NSCAL. These scaling values (1 and 0) indicate no scaling.

SET– For visibility, any value that is within PM (= 0.9) of NSET (= 10) is set to NSET. This is to assure that values close to 10 are set to exactly 10. For ceiling, any value that is within PM (= 1.0) of NSET (= 120) is set to NSET. This is to assure that values close to 120 are set to exactly 120.

PM– See SET above.

Postprocessing is done by **POST** for forecast probability grids. Its control parameters are explained in Table 10 and associated text.

Table 10. The control values for subroutine **POST**.

| Variable Name as Read | TLOA | SETLOA | THIA | SETHIA | CONSTA | NSCALA | EX1A | EX2A |
|---|---|---|---|---|---|---|---|---|
| Variable Name used in POST | TLO | SETLO | THI | SETTHI | CONST | NSCAL | N/A | N/A |
| Value for Visibility | .01 | 0 | 1. | 1. | 1. | 2 (0) See below | | |
| Value for Ceiling | .01 | 0 | 1. | 1. | 1. | 2 (0) See below | | |

TLO– Any values less than TLO = .01 with be changed to SETLO = 0.
SETLO– See TLO.
THIA– Any value greater than1.0 will be changed to SETTHI = 1.0.
SETTHI– See THIA.
CONST, NSCAL–Scaling is X = X*CONST*10**NSCAL. These scaling values of 2 indicate the probabilities carried as fractions are changed to percent. This is necessary for producing grids that can be handled by ImageGen for plotting and viewing. For operations, the production of grids are in fractions, and the "2" should be "0."

CONVPR (CONCPR) assures consistency among the analyzed probability grids. Each grid is compared, point by point, with the grid of the next lower category. If a point value is less than the lower category value, it is set to the lower category value. This assures consistency among the probabilistic grids, although it implies a zero probability for the corresponding discrete grid. No control parameters are required. Not all categories of the grids would have to be analyzed for the subroutines to work correctly for the grids analyzed. If the category definitions were to be changed, then these two routines would have to be modified.

## 5. SUMMARY

Analyzing the non-probabilistic variables visibility and ceiling pose interesting challenges. They are highly discontinuous in space and time, and the observations and forecasts of them can change from their minimum to maximum value from one *data point* to its closest neighbor. Such a change is also reasonable from one *gridpoint* to another, even on a grid as fine as 2.5 km. The approach taken is that the value at a station is likely to persist spatially for some distance away, rather than the change to be linear from point to point. This makes for a somewhat spotty appearance, but emphasizes the "unusual" values. For instance, gridpoint values between ceilings of 0 and 120 (hds of ft) at neighboring stations will not exhibit a linear change, but rather both 0 and 120 will be emphasized.

Taken together, there are thousands of combinations of control values, and obviously not all can be tested on multiple cases. Even if done, the only reasonable criteria for judging goodness is meteorological eyeball experience. Error metrics, like MAE, are of little use, because (1) the variation of such metrics is extremely small from one (reasonable) control combination to another, and (2) they do little to represent or highlight problem areas. Maximum differences between station values and values interpolated from the analyzed grid are of some help. However, even a grid as fine as 2.5 km does not represent the terrain well in large regions of the U.S., and the grids cannot be expected to interpolate back to the station values exactly. Temporal consistency of forecasts is important, and animating the projections is a useful technique, but here, again, it is by visualizing the results that problems are spotted.

The problem of how to treat changes that might be expected due to elevation changes is acute for visibility and ceiling. Most observations, and LAMP forecasts, are at relatively low elevations, and there is no clear indication of what measurement would occur at higher elevations. It has been assumed, and discussed above, that in the treatment of change with elevation, no preference is given for increase or decrease for visibility. However, it has been assumed that ceiling height would be more likely to decrease with elevation than to increase.

A primary purpose of LAMP is to use the observations effectively to make short range forecasts. The skill of a 1- or 2-h forecast drops dramatically depending on whether the observation is used in the predictive equation or not. In operations, reports from stations may be missing, and in those cases, backup equations are used that do not contain the observation. This may create a spatial discontinuity, because forecasts for most stations will be made with observations, and one made without the observation may vary drastically from them. The question is, can a forecast at a gridpoint near the station without an observation contributing to the forecast be made better with or without that forecast. Because most visibility and ceiling observations are made at METAR stations, it is expected a low percentage will be missing on any one hour. However, the advisability of including forecasts made with backup equations at points were it is known there will not be an observation is in question, except over water; such forecasts over water are the only ones LAMP produces and are better than none. It is noted, that such a forecast without an observation is a calibrated interpretation of a combination of models, LAMP and GFS MOS; forecasts of visibility and ceiling directly from mesoscale models are similar in that respect, because they are interpretations of other variables directly forecast by the models, but are generally uncalibrated.

## REFERENCES


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.

Dallavalle, J. P., M. C. Erickson, and J. C. Maloney III, 2004: Model output statistics (MOS) guidance for short- range projections. *Preprints, 20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., **6.1**.

Donaldson, R., R. Dyer, and M. Krauss, 1975: An objective evaluator of techniques for predicting severe weather events. *Preprints, Ninth Conf. on Severe Local Storms*, Norman, OK, Amer. Meteor. Soc., 321-326.

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

_____, J. P. Charba, J.-S. Im, F. G. Samplatsky, B. Glahn, R. Yang, and A. J. Kochenash 2015: Improving Localized Aviation MOS Program (LAMP) guidance by utilizing emerging forecast and observation resources. Preprints, *Special Symposium on Model Post-processing and Downscaling*, Phoenix, AZ, Amer. Meteor. Soc., **3.2**, 15 pp.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1313-1329.

_____, 1988: Characteristics of map projections and implications for AWIPS-90. *TDL Office Note* **88-5**, National Weather Service, NOAA, U.S. Department of Commerce, 45 pp.

Glahn, B., and J. P. Dallavalle, 2000a: MOS-2000. *TDL Office Note* **00-1**. Techniques Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce, 16 chapters.

_____, and _____, 2000b: Computer programs for MOS-2000. *TDL Office Note* **00-2**. Techniques Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce.

_____, 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: Algorithms for effective objective analysis of surface weather variables. Preprints, *24$^{th}$ Conf. on Weather and Forecasting*, Seattle, WA, Amer. Meteor. Soc., **10A.4**, 11 pp.

_____, and _____, and G. Wagner, 2012: Objective analysis of MOS forecasts and observations in sparse data regions. Preprints, *21$^{st}$ Conf. on Probability and Statistics in the Atmospheric Sciences*, New Orleans, LA, Amer. Meteor. Soc., **8.4**, 11 pp.

_____, R. Yang, and J. Ghirardelli, 2014: Combining LAMP and HRRR visibility forecasts. *TDL Office Note* **14-2**. Techniques Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce, 20 pp.

Im, J.-S., and B. Glahn, 2012: Objective analysis of hourly 2-m temperature and dewpoint observations at the Meteorological Development Laboratory. *Natl. Wea. Dig.*, **36**, 103-114.

Miller, R. G., 1958: Regression estimation of event probabilities, U.S. Weather Bureau, Contract Cwb-10704, *Tech. Rep.* No. **1**, The Travelers Research Center, Inc., Hartford, CN.

OFCM, 1995: Surface weather observations and reports. *Federal Meteorological Handbook* **1**, Office of the Federal Coordinator for Meteorological services and Supporting Research, NOAA, U.S. Department of Commerce, 104 pp.

Palmer, W. C., and R. A. Allen, 1949: Note on the accuracy of forecasts concerning the rain problem. Weather Bureau Manuscript, Washington, D. C., 2 pp.

Rudack, D. E. and J. E. Ghirardelli, 2010: A comparative verification of Localized Aviation Model Output Statistics Program (LAMP) and Numerical Weather Prediction (NWP) model forecasts of ceiling and visibility. *Weather and Forecasting*, **24**, 1161-1178.

Shaffer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570-575.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 676 pp.