5.3

OBSERVATIONAL DATA AND MOS: THE CHALLENGES IN CREATING HIGH-QUALITY GUIDANCE

Rebecca L. Allen*
Meteorological Development Laboratory
Office of Science and Technology
National Weather Service, NOAA

1. INTRODUCTION

High-quality observations are essential for the creation of quality statistical forecast guidance. At the Meteorological Development Laboratory (MDL), we use the Model Output Statistics (MOS) technique to produce forecast guidance for a wide range of weather elements from 6 hours to 10 days in advance. The MOS approach (Glahn and Lowry 1972) statistically relates observed predictand data to predictors such as forecasts from dynamical models, surface observations, and geoclimatic information. While the content and stability of model archives are often discussed, the vital role of the observational data sample is usually not appreciated.

MDL has recently implemented new MOS guidance (Dallavalle and Erickson 2000, Erickson and Dallavalle 2000) based on the National Centers for Environmental Prediction's (NCEP) Global Spectral Model. In this new system, we use observational data as predictands, predictors, and for verification. In order to obtain the most representative sample possible, we have to deal with a range of issues including extreme values, quality control, and site changes. First, we strive to have a data sample large enough to include nearly the entire range of weather conditions one would expect to see at a given site. As with any statistical development, extreme values will skew the statistical relationship. The challenge is to remove values from the sample that are due to errors, and keep the true, valuable extremes, such as record breaking temperatures. Great care must be taken to ensure the overall quality of the data. Because the observational data are collected over time, changes occur in which stations report, what these reports measure, and how the reports are structured. All of this change must be managed so that the data can be used as one coherent archive, rather than many small sets of data with different characteristics.

Once observations are assembled, we further transform the data to more closely match our forecast needs. Sometimes, the weather element we need to forecast is not observed directly, so it must be estimated from other values. Other times, a single source of observations may not provide a complete description of the weather that occurred, so multiple datasets are combined to compute predictands.

In this paper, we discuss the challenges faced in creating a high-quality observational archive. The types of observational data used by MDL are outlined. We present techniques used to quality control the data and account for the ever-changing nature of observational data samples. Finally, we discuss predictand development, and how we use multiple datasets to create a comprehensive picture of the weather.

2. DATA COLLECTION

MDL currently collects six types of observational data for use in development, implementation, and verification of the MOS guidance. These include hourly surface observations, Supplementary Climate Data (SCD) reports, Satellite Cloud Product (SCP) bulletins, lightning detection reports, severe weather information, and summary of the day reports from the National Climatic Data Center (NCDC)'s cooperative network. TDL Office Note 00-01 (Glahn and Dallavalle 2000) describes the first five of these in detail.

Hourly surface observations from observing sites around the world are collected through the National Weather Service's (NWS) Telecommunications Gateway by NCEP. Since July 1996, these observations are sent in Aviation Routine Weather Report (METAR) format. NCEP decodes these METAR reports, and encodes them into Binary Universal Form for the Representation of meteorological data (BUFR) format. Once an hour, MDL decodes all the observations reported from 15 minutes before to 15 minutes after the hour from NCEP's BUFR file and creates an ASCII file in tabular format. This table contains temperature; dew point temperature; wind speed, direction, and gusts; pressure; altimeter; present weather; cloud amount and height; and visibility. At the appropriate hours, the table also contains 6- and 24-h maximum and minimum temperatures; 1-, 3-, 6-, and 24-h precipitation amounts; and minutes of sunshine. MDL only archives observations from sites in the United States, Canada, Mexico, the Carribean, Guam and the Marshall Islands, and Eastern Russia. Special observations (SPECIs) are not saved; when a correction is issued, the BUFR file only contains the final, corrected observation. The data are edited monthly for quality control purposes, and saved in our own binary format. We also maintain monthly files of the raw METAR reports to allow us to review the original reports, if necessary. Our current archive of edited hourly surface METAR observations is available from December 1996 to the present.

In addition to the METAR reports, MDL's hourly ASCII tables also contain Supplementary Climate Data (SCD) reports for depth of new snow during the past six hours, depth of snow on the ground, and 6-h liquid equivalent precipitation amount for select Weather Forecast Offices (WFOs) at 0000, 0600, 1200, and 1800 UTC. We have been archiving these data since December 1998.

* Corresponding author address: Rebecca L. Allen, W/OST22 Rm. 706 WWBG, 5200 Auth Road, Camp Springs, MD 20746-4304
email: rebecca.allen@noaa.gov

MDL archives the hourly Satellite Cloud Product (SCP) bulletins for use in developing sky cover guidance. Data are archived for the contiguous U. S., Hawaii, and Puerto Rico. Two satellites, currently GOES-8 and GOES-10, cover the eastern U. S. and Puerto Rico, and the western U. S. and Hawaii, respectively. The observations available for the contiguous U. S. are based on the sounder instrument, while those for Hawaii and Puerto Rico are based on the imager instrument. The SCP bulletins contain categorical cloud coverage for mid-level (631 to 400 hPa) and high-level (above 400 hPa) clouds, the effective cloud amount (ECA) in percent, and cloud top data. MDL currently archives the cloud coverage and ECA. Our SCP archive is available from September 1995 through the present.

In order to develop thunderstorm and severe weather guidance, MDL uses cloud-to-ground (cg) lightning data from the National Lightning Detection Network (NLDN) that we obtain from the Global Hydrology Resource Center at NASA's Global Hydrology and Climate Center in Huntsville, Alabama. These data contain information on the time of flash, total number of cg flashes, number of negative cg flashes, number of positive cg flashes, maximum signal strength, number of strokes, and the associated geographic location of the flash. Reports of severe weather, including the number of tornadoes, maximum tornado F-scale, number of hail reports, maximum size of hail, number of damaging wind reports, maximum wind speed, and the associated location and time, are provided by the WFOs, and obtained from the Office of Climate, Water, and Weather Services (OCWWS). We currently have lightning and severe weather data dating from April 1, 1994, to the present.

NCDC's cooperative observer network data are available for approximately 8000 sites in the United States. The daily reports contain maximum/minimum temperature, precipitation amount, snowfall, snow depth, wind movement, and soil temperature. We have already used the temperature information to develop maximum/minimum temperature guidance for the NWS's River Forecast Centers. Currently, we are in the process of developing an archive of these data for use in our development of a snowfall forecast system.

3. QUALITY CONTROL

Collecting the observational data is simply the first step of the archival process. No amount of data is useful if the quality is not good. Errors can be introduced into the data at many steps along the way. Besides the measurement error of the sensors, observing instruments can malfunction and report erroneous data. Reports entered manually can fall victim to typographical errors. The federal government has guidelines for reporting each weather element, but not all of the data we receive adheres to these guidelines. Decisions must be made about when to accept a report as valid, and when to throw it out.

MDL's quality control process starts during the creation of the hourly ASCII tables. Checks are performed to ensure that all temperatures and dew point temperatures are in an acceptable range for the station's geographical area, the temperature is greater than or equal to the dew point temperature, and the wind speeds are reasonable. The main quality control processing occurs when we edit the data after one month of reports is collected. Our code checks the hourly surface observations and the SCD reports for over 60 different errors. Some error checks are common sense. For example, was snow reported with a temperature much greater than 32 degrees Fahrenheit? Was blowing dust reported when the wind was calm? Was the maximum temperature greater than the minimum temperature for that same period; was the temperature greater than the dew point temperature? Elements like precipitation amount, pressure, cloud heights, and wind gusts are checked to see if they are within a predetermined range. All elements are checked to be sure that they are reported according to official guidelines (OFCM 1995). For example, cloud heights above 10,000 feet should only be reported to the nearest 1,000 feet. Several checks are made to ensure the temporal consistency of the elements. Temperatures, dew point temperatures, and pressures are checked for extraordinarily large changes over a 3-h period. The 6-h maximum (minimum) temperatures are checked to make sure that they are at least as large (small) as the six individual hourly temperatures reported for the same period. Currently, we do not check observations against those from nearby sites for possible inconsistencies.

The multiple decodings and encodings of the observations result in several chances for the data to be inadvertently modified. Many of the elements in the METAR reports are converted from their non-SI units of Fahrenheit, statute miles, knots, etc., to the appropriate SI units when they are put into BUFR. Then the values are converted back to their original units in our archive. In addition, the temperatures and dew point temperatures are reported in degrees Celsius in METAR, but are converted to Fahrenheit in our archive. As a result, the final archived values can be different than those originally reported. In order to minimize modification of the data, care is taken to use the same conversion factors to change from non-SI to SI units, and back again. The METAR present weather descriptions are converted from character representations to numerical descriptions when they are saved in the BUFR format. Unfortunately, some present weather groups, such as mixed precipitation events, are represented by the same number. For example, thundershowers of rain and snow are both given the same numerical representation. Consequently, when the present weather number in the BUFR message is decoded, we can not be exactly sure what the original report was. These situations can produce errors in the data that must be corrected by the quality control program.

Sometimes, corrections can be made to the data to account for vagaries of the reporting standards, or known deficiencies in the reporting equipment. In METAR, a distinction is made between mist and fog based on the horizontal visibility. We correct those instances where the present weather reported is mist (code BR), but the visibility is less than 5/8 statute miles, indicating that fog (code FG) is present, and vice versa. Also, we are aware that Automated Surface Observing System (ASOS) precipitation amount sensors often report false precipitation amounts when dew forms on the sensor overnight

(Fiebrich et al. 1997), or when previously reported snow melts into the sensor. Our quality control program tries to isolate these instances, and remove any precipitation reports that can be attributed to snowmelt or dew formation. Another subtlety of ASOS reporting can be seen in maximum and minimum temperature reports. Sometimes, the 6-h maximum/minimum temperature observation is recorded a few minutes before the hourly temperature observation, thus resulting in an inconsistency. For example, the temperature could continue to rise during those few minutes after the 6-h maximum was recorded, resulting in the 1200 UTC temperature report being greater than the 0600 to 1200 UTC maximum temperature report. As with the dew/snow melt cases, we look for these instances, and adjust the 6-h maximum/minimum temperature accordingly.

In the process of editing, any report deemed unreliable or questionable, but which cannot be corrected, is deleted. We err on the side of caution, preferring to eliminate suspect reports, rather than let bad observations remain in our archive. For a complete list of the quality control checks we perform, please see TDL Office Note 00-01 (Glahn and Dallavalle 2000).

In addition to these checks on the hourly surface observations and SCD data, the other observational datasets undergo some measures of quality control as well. The SCP data are screened for erroneous dates and reports that are out of order. The cooperative network data are examined for values outside of acceptable ranges. The lightning and severe weather reports have been screened for quality control purposes by the groups that collect the data, so no additional quality control is done by MDL. Finally, after the various data types have been quality controlled, we pack the data into our own binary format for use by our statistical processing software.

4. MANAGING CHANGE

One feature of the hourly surface observations has been the amount of change in the observational environment. Our complete hourly archive spans the period from January 1977 to the present. In that time, stations have closed, new stations have opened, the reporting standard has changed from the Surface Aviation Observations (SAO) code to METAR, equipment types have changed, and the advent of ASOS has meant that observations that were once predominantly taken by humans are now taken by machines. All of these factors must be taken into account when developing a statistical forecast system.

Keeping track of the reporting sites is imperative. In the past year alone, we have added 79 new stations to our hourly archive. Our users are constantly asking for more guidance for more sites, but until we have an established pool of predictand data for a site, we can not develop forecast equations. If we do not keep abreast of new stations coming on-line, we would not be able to increase the number of stations in the MOS system. Station closings are also an issue. Once we stop receiving observations from a site, the developmental sample becomes static, and the site must eventually be dropped from the MOS system. Occasionally, stations change their call letter identifiers. Adjustments must be made to our system to collect and use the most up-to-date observations. On rare occasions, a station will use call letters that were previously assigned to another station. In May 1999, the call letters for Bergstrom, Texas (KBSM) were changed to KAUS. Unfortunately, KAUS previously designated Austin City Airport in Texas. This change causes havoc with a system of observations that is catalogued by call letters. If one is not careful, all of the KAUS observations are combined into one developmental sample, when, in fact, the pool of observations is from two geographically different sites. To handle these station changes, we maintain a station dictionary - a file that contains information detailing the past and present location, station type, and call letters for all stations in our archive. This file provides a historical record of our archive, and allows us to construct the proper observational sample for sites that have moved or whose call letters have changed.

Changes in reporting standards can also affect the characteristics of a MOS predictand. When the NGM MOS was developed, the SAO format cloud cover observation made a distinction between opaque and non-opaque clouds. The MOS guidance was developed to forecast opaque clouds. After the conversion to METAR in 1996, the only available observation was total cloud cover, regardless of opacity. The precise weather element we were forecasting was no longer observed. In this case, we continued to produce the NGM MOS cloud guidance, with the caveat that the quality of the observational predictors had been degraded. Subsequently, when the AVN MOS cloud guidance was developed, the predictand was redefined to be total sky cover, regardless of opacity.

In addition to changes in standardization, the implementation of ASOS has also changed the observational landscape. Besides the quality control issues mentioned earlier, other changes have had to be accommodated. For instance, more observations are available. Sites that were previously manned only during the daytime now have automated equipment and report 24 hours a day. As a consequence, the MOS guidance for many of these sites has been expanded to cover all elements for the entire day. One downside of ASOS, however, has been the loss in the precision of observing precipitation type. According to the ASOS User's Guide (NWS 1998), light precipitation, such as drizzle, is hard for the sensor to interpret. In fact, the present weather category of UP was created to identify those cases where precipitation is falling, but the sensor can not accurately distinguish the form of precipitation. These indeterminate cases can not be used to develop precipitation type guidance, and this problem brings into question the accuracy of the ASOS reports for all light precipitation events. Efforts are ongoing to enhance the performance of various ASOS sensors, so the quality of the reports should improve with time.

5. PREDICTAND DEVELOPMENT

Once the irregularities in the data have been taken into account, we transform the data into appropriate predictands for the statistical development. One of the first issues that must be addressed is the time period the predictand covers. Some predictands are valid at a specific instant in time. In these cases, the predictand is based on individual hourly observations, that is, single snapshots of the weather. Other elements, like probability of precipitation (PoP), are valid over a period of time. In this case, the PoP predictand is based on the 6-h precipitation amount reports that cover the forecast period. Maximum and minimum temperature are more of a challenge. NWS forecasters must produce forecasts for daytime maximum and nighttime minimum temperatures; however, these values are not routinely observed. METAR reports contain maximum and minimum temperatures for the 6-h synoptic periods. NCDC cooperative sites report a 24-h maximum and minimum, according to the locally determined period. In order to produce the maximum/minimum temperature predictand, an algorithm was developed that combines the METAR 6-h maximum (minimum) with the hourly temperatures to estimate the maximum (minimum) temperature for 7am to 7pm (7pm to 8am) local standard time.

For the extended-range guidance, we forecast weather elements such as precipitation type and mean total sky cover for 12-h periods. Predictand algorithms for these elements combine the 13 hourly observations from the 12-h period to get a composite condition for the period. This illustrates one shortcoming of using only hourly observations. Each hourly observation is virtually a snapshot of the weather at the time the observation is taken. No information is saved from the SPECI observations to indicate what happened between the hourly reports. For example, if snow occurred for 45 minutes, but stopped by the time the observation was taken, the observation for that hour would report no weather. For this reason, when creating 12-h precipitation type predictands, we had a requirement that the event occurred at three or more of the observation times. We did not include small-scale, short-lived events in the developmental sample.

The sky cover predictand has required special attention since the advent of ASOS. The ASOS ceilometer cannot detect clouds above 12,000 ft. Therefore, on a day where 7/8ths of the sky is covered with clouds at 12,500 feet, an automated site would report clear. Given that most of the observation sites are automated, only using manual sites was not acceptable for producing sky cover and ceiling height guidance. Hence, an algorithm was developed to complement the METAR cloud observations with the corresponding SCP report for that station (Hughes 1996).

For thunderstorm and severe weather guidance, surface observations are not sufficient to detect the occurrence of these weather events. The most recent development of this guidance was done by using the NLDN lightning data and reports of severe weather from OCWWS (Hughes 1999). As opposed to the station-based hourly surface observations, these data are random in place and time. Therefore, the predictand algorithm must take these random reports and combine them into coherent fields of observed thunderstorms and severe weather. To do this, the predictand is defined on a grid, where any report that occurs in a 48-km box around the grid point is used to indicate the occurrence of the event at that grid point. In terms of time, any report from the 59 minutes preceding the hour counts for that hour. The end result is a report for each grid point for each hour indicating whether or not thunderstorms (or severe weather) occurred. The resulting forecasts are then also grid-based forecasts, not station-based like the other MOS guidance.

As a final illustration of how one must balance sample representativeness, quality control issues, instrumentation limitations, and predictand development, consider dew point temperature. According to the specifications, the lowest value an ASOS dew point sensor can report is -30° F, and most other sites won't report dew points lower than -35° F. Therefore, we set our quality control guidelines to delete any dew point below this -30° F limit when the dew point depression was greater than 15 degrees. Because the equations to predict temperature and dew point are developed simultaneously, all of the temperature cases where the dew point was missing (either due to a missing report, or our deletion of the data) were thrown out of the developmental sample. During testing of the MRF temperature equations, we discovered the equations did not perform well during an extreme cold outbreak in Alaska. Upon looking, we realized that most of the extremely cold cases in our developmental temperature sample were thrown out because the corresponding dew point temperature was missing. Our developmental sample did not represent the true variability of the temperatures in Alaska. To solve this problem, we reconfigured the dew point predictand dataset. In those cases where the air temperature was less than -30° F, and the dew point was missing, we set the dew point temperature equal to the air temperature. Although this will create some false dew point data, it allowed us to include these extreme cold events in our sample, and the resulting temperature equations performed much better. This is just one example where the data had to be manipulated properly to create the highest-quality guidance.

6. CONCLUSION

MDL frequently receives requests for MOS guidance for more elements and more stations. For these requests, the first question that arises is whether a sufficient sample of observational data exists for that element and that site. Users sometimes ask for MOS guidance for sites that have not begun to report, or have been closed for several years. To create high-quality guidance, one needs a sufficiently long sample of high-quality observations that can be used to build a developmental predictand sample. The data must be gathered, processed to ensure quality, and archived. Irregularities in the data must be accounted for, and a suitable predictand must be derived. Only then will we be successful in developing robust forecast equations to predict a variety of weather elements.

7. ACKNOWLEDGMENTS

The author would like to thank Paul Dallavalle and Mary Erickson of MDL for their assistance in the preparation of this paper. Additional thanks go to members of the Statistical Modeling Branch of MDL for providing details regarding our observational datasets.

8. REFERENCES

Dallavalle, J. P., and M. C. Erickson, 2000: AVN-based MOS guidance - The alphanumeric messages. NWS Technical Procedures Bulletin No. 463, NOAA, U.S. Dept. of Commerce, 9 pp.

Erickson, M. C., and J. P. Dallavalle, 2000: MRF-based MOS guidance - The alphanumeric messages. NWS Technical Procedures Bulletin No. 460, NOAA, U.S. Dept. of Commerce, 9 pp.

Fiebrich, C. A., V. J. Dagostaro, and J. P. Dallavalle, 1997: An algorithm to eliminate precipitation reports caused by surface condensation in AEV-ASOS data. TDL Office Note 97-2, National Weather Service, NOAA, U.S. Department of Commerce, 6 pp.

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.

_____, and J. P. Dallavalle, 2000: MOS-2000. TDL Office Note 00-1, National Weather Service, NOAA, U.S. Department of Commerce, 131 pp.

Hughes, K. K., 1996: Complementing ASOS cloud reports with the GOES satellite cloud product. Preprints 15th Conference on Weather Analysis and Forecasting, Norfolk, Amer. Meteor. Soc., 36-39.

_____, 1999: AVN-based statistical forecasts of thunderstorms and severe thunderstorms for the contiguous U.S. Preprints 17th Conference on Weather Analysis and Forecasting, Denver, Amer. Meteor. Soc., 201-205.

National Weather Service, 1998: Automated Surface Observing System (ASOS) User's Guide, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Office of the Federal Coordinator for Meteorological Services and Supporting Research (OFCM), 1995: Federal Meteorological Handbook No. 1: Surface Weather Observations and Reports, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.