# Supporting Gridded Model Output Statistics Forecast Guidance System

Kari L. Sheets

ABSTRACT

The Meteorological Development Laboratory (MDL) of NOAA's National Weather Service is creating a National Digital Guidance Database at a fine resolution to complement the existing National Digital Forecast Database. To help accomplish this goal, MDL is producing a gridded Model Output Statistics (MOS) forecast guidance system. Currently, gridded MOS populates a 5-km grid with elements needed for weather forecast grids covering the contiguous United States.

This paper describes the use of Geographic Information System (GIS) to generate and quality control geophysical variables, create a station dictionary including land/water designations for the observing stations, and analyze or troubleshoot problem areas in gridded MOS weather elements. Stations used in traditional MOS development are unevenly distributed, leaving developers searching for additional data to aid in analyzing these values to the aforementioned grid. GIS is integral in supplying the data to move from station-based MOS to gridded MOS.

## 1. Introduction

The Meteorological Development Laboratory (MDL) of NOAA's National Weather Service (NWS) is developing a National Digital Guidance Database (NDGD) at fine resolutions to complement the existing National Digital Forecast Database (NDFD, Glahn and Ruth 2003). To help accomplish this goal, MDL is creating a gridded forecast guidance system. Current forecast guidance is produced for the United States and its territories at approximately 1800 hourly observing sites and over 5000 cooperative observing sites by using the Model Output Statistics (MOS) technique (Glahn and Lowry 1972). In the MOS approach, observed predictand data are statistically related to predictors such as forecasts from dynamical models, surface observations, and geoclimatic information. MOS guidance depends on a sufficiently long sample of high-quality observations to develop robust forecast equations for a variety of weather elements (Allen 2001).

Traditional observing stations used to develop MOS lack the spatial density desired to support a fine resolution grid, leaving developers searching for additional observational datasets as well as better predictor variables to capture the meteorological effects of elevation, land cover, and water. Efforts were made to gather, quality control, and archive data from additional meteorological observing systems, but these additional data sets did not bring the observed data resolution to the desired NDGD resolution of 5 km for the conterminous United States.

To supplement the meteorological data and adjust the MOS forecast guidance to terrain; we used a Geographic Information System (GIS) to generate additional geophysical variables at the proper NDGD grid resolution. For this purpose, grids of elevation, land cover, and a land/water mask were created. Additionally, GIS was employed to generate the map specifications for computational and output grids as well as to include land/water designations for the observing stations. This paper describes the use of a Geographic Information System (GIS) to generate and quality control geophysical variables, create a dictionary of station attributes including land/water designations, and analyze or troubleshoot areas questioned by local forecasters in gridded MOS products. Plans for the use of GIS to generate additional climatic and geophysical data sets for future gridded MOS development are also presented.

## 2. Data Preparation

Quality control is of utmost importance when developing a digital forecast guidance system. Successful analysis of MOS station-based forecast guidance to gridded forecast guidance requires that as many observing sites as possible be included in the MOS system and that the analysis scheme use high-resolution geoclimatic data to adjust for terrain, land use, or water influences (Dallavalle and Glahn, 2005). Therefore, the agreement of these data sets is critical for the quality of the final product.

The first data characteristic in need of quality control was the geodetic reference system. A geodetic datum is dependent on the assumed shape, ellipsoidal or spherical, and associated coordinate system; as well as a set of points and lines resulting from surveying (Bolstad 2002). Manipulating data from multiple sources also meant dealing with data with multiple datums and projections. The National Centers for Environmental Prediction's (NCEP) numerical weather products are based on a spherical datum unique to meteorological applications, which contributes greatly to the issue of disagreeing reference systems. NCEP model output is a primary data source for MOS guidance (Dallavalle et al, 2004), therefore the coordinate system of data acquired to improve MOS, is transformed to match that of the numerical weather products. This geodetic datum accepted for meteorological applications is described in NCEP Office Note 388 (National Weather Service 2002), as a spheroid with a radius of 6,371,200 meters. The geodetic datum is referred to as the NCEP Sphere (Sheets et al, 2005).

Projecting and resampling quickly became the basis for creating additional data sets to support gridded MOS development. Originally, Spatial Analyst's raster calculator's resampling tool and default projection techniques were used to simultaneously project and resample data. However, inconsistencies in these calculated data and other data already available in the appropriate projection and resolution raised concerns. Research showed a better outcome resulted by first using the project raster tool and then resampling the projected raster to the appropriate resolution. In the gridded guidance system, the values are valid at grid points, so nearest neighbor resampling was employed in order to preserve the characteristic of the co-located grid.

Point data is driven by a station database, which has been maintained as a text file with little ability to plot and check the attributes of the stations. GIS allows developers to quickly certify the location of MOS station data by plotting stations and evaluating locational accuracy. Time zones are the final piece of station specific information needed to produce both station and gridded MOS developments. New stations are assigned a time zone, while the current values are crosschecked using a spatial join with the world data provided with ArcGIS desktop installation pack.

After ensuring data alignment and location, consistency between gridded data and point vector data is the next area in need of attention. Weather observing station attributes include elevation. The analysis process from station to gridded MOS forecasts relies heavily on elevation values for many weather elements. Therefore, the station and grid elevations need to have similar values in order to minimize terrain-induced inconsistencies in weather element guidance fields. Extraction of raster data to point locations, followed by a comparison of the values, provides the quality assurance for height values. Height inconsistencies greater than 500 feet compose the group of stations in need of further point-by-point investigations (Figure 1). The extracted elevations from the digital elevation model are used for missing stations.

## 3. Station Database

Management of the point-based observations for the MOS system is critical to generating reliable guidance products. Previously the text versions of station information were maintained in a UNIX system, where each developer maintained element specific lists and often duplicated versions of the master database. Additionally, the maintenance of attribute data, as well as the ability to designate which weather elements the site observes, are techniques, which historically are not included in station data control. Employing a database makes these tasks possible, and simplifies and streamlines station data management.

As gridded MOS development began, the need for additional station attributes became obvious. A classic example is the requirement to designate stations sited over land versus water. Buoy designation is a simple task, but the designation of stations along the coasts and the Great Lakes is more complex. The solution is to add fields, based on a point extraction from raster data, to the station database to indicate the characteristics of the stations (Figure 2). The application continues to grow leading to the inclusion of additional descriptive fields to the database, such as indications of the station's observing network and the observed elements at the station.

The development of this station database coincided with the push for using geodatabases within ArcGIS. Originally, the coincidence seemed perfect. However, in MOS development, the numerical model used to initiate the guidance indicates the map projection of the development, which must be declared for a geodatabase. Therefore, the current ESRI geodatabase design requires a separate geodatabase for each MOS development, eliminating the benefit of the one-stop database editing capability, which makes a database the best method for managing station data.

Initially, MOS developers became acquainted with GIS as tool for creating geographic regions. Equations may be developed from a combined sample of data for stations in relatively homogeneous regions. This increases the sample size but may cause a slight loss in accuracy at the specific station location. Regional developments can compensate for this loss by including high-resolution geophysical data, such as terrain and climatic values, which are tuned to individual locations at desired resolutions. Digitizing these regions and performing spatial joins shortened the development time for MOS forecast guidance (Figure 3). Creating regions continues to be an integral GIS task for the MOS development team. Even considering the thousands of observing stations added to the MOS system in the last 5 years, the resolution of the network of quality weather observing sites is far from the resolution of gridded MOS, thus making regional developments a necessity for some elements.

## 4. Gridded MOS

Output grids appear to mimic the shape of the United States; however, they are actually rectangular grids with grid points set to missing giving the appearance of a shape. The data input into the code generating the grid from station MOS products needs to be available on a slightly larger area than the desired final output. Parameters for this computational grid, as well as the output grid, were generated using Spatial Analyst tools. High-resolution coastlines and marine zones were downloaded from the Advanced Weather Interactive Processing System (AWIPS) Map database, merged together, and buffered by 40 km to generate the computational grid mask (Figure 4). These same files were trimmed to include a 50 kilometer buffer to create the output grid (Figure 5).

Additional geophysical data sets generate additional information for analyzing the 10,560 MOS weather observation sites to the NDGD parameters. Terrain, land cover, land water cell designations, and Parameter-elevation Regressions on Independent Slopes Model (PRISM) precipitation and temperature climatologies (Daly et al. 1997) were processed to provide this additional information (Trimarco et al, 2005).

Finally, gridded MOS is a reality, but as with any new product, there are kinks to iron out and overall room for improvement. GIS has been integral in evaluating and troubleshooting the new products. These grid parameters are continually being modified to give a best fit to the available data for all forecasted guidance elements. Stations with bad element observations appear as irregular areas in the output grid. Comparing the raster data with the station data allows those stations to be tagged in the station database with a deficiency in observations. Comparing new versions of MOS developmental software is possible thanks to the ability to difference the resulting forecast guidance grids (Figure 6). As forecasters along the coasts began to use the products, they have commented on the land water masks, which has allowed for the evaluation of the generation of geophysical constants. As gridded MOS continues to expand and improve skill, GIS will be more and more imperative to visualizing the progression of the products and evaluating their usefulness.

## 5. Future Projects

As developers seek to improve and expand the gridded MOS system, the need for additional geophysical datasets will grow. Dew point and sky cover PRISM climate data sets, snow climatologies generated with Geostatistical Analyst, radar data, and satellite images have all been discussed as being of interest to developers. GIS will be critical in converting this data to a format that can be used in the MOS system. Work has already begun to provide additional station characteristic flags such as proximity to major highways to assist in the quality control of wind data. Furthermore, work is slated to produce gridded MOS at finer resolutions, requiring the generation of many of the data discussed at finer weather guidance.

## 6. Conclusion

Prior to use of GIS, MDL's ability to ingest, manipulate, and analyze high-resolution data was very limited, and station data management was tedious at best. GIS created data sets have played a critical role in the release and evaluation of gridded MOS products. GIS tools and resulting data will allow MOS developers to explore new analysis techniques and predictor data, which will translate to better gridded MOS forecast guidance.

## 7. Acknowledgements

The Environmental Systems Research Institute (ESRI) produces the ArcInfo GIS package with Spatial Analyst extension used for the tasks discussed in this paper as well as the accompanying lakes and states shapefiles. Coastline and marine zone data are offered as a free public service of the National Weather Service at the AWIPS map database http://www.weather.gov/geodata/. No recommendation is expressed or implied for the use of a specific commercial product.

## 8. References

Allen, R. L., 2001: Observational data and MOS: The challenges in creating high-quality guidance. Preprints, *18th Conference on Weather Analysis and Forecasting*, Ft. Lauderdale, FL, Amer. Meteor. Soc., 322-326.

Bolstad, P., 2002: *GIS Fundamentals: A first test on geographic information systems*. Eider Press, 412 pp.

Dallavalle, J. P., M. C. Erickson, and J. C. Maloney, 2004: Model Output Statistics (MOS) guidance for short-range projections. Preprints, *20th Conference on Weather Analysis and Forecasting,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, P6.1.

____, J. P., and B. Glahn, 2005: Toward a gridded MOS system. *Preprints, 21st Conference on Weather Analysis and Forecasting*, Washington, D.C., Amer. Meteor. Soc., 13B.2.

Daly, C., G. Taylor, and W. Gibson, 1997: The PRISM approach to mapping precipitation and temperature. Preprints, *10th Conference on Applied Climatology*, Reno, NV, Amer. Meteor. Soc., 10-12.

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

National Weather Service, 2002: ON 388 GRIB (Edition 1) The WMO Format for the Storage of Weather Product Information and the Exchange of Weather Product Messages in Gridded Binary Form as Used by NCEP Central Operations. [Available online at http://www.nco.ncep.noaa.gov/pmb/docs/on388/]

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

Trimarco, R. A., K. L. Sheets, and K. K. Hughes, 2005: Building a gridded climatological dataset for use in the statistical interpretation of numerical weather prediction models. Preprints, *15th Conf. on Applied Climatology*, Savannah, GA, Amer. Meteor. Soc., JP 1.6.

## 9. AUTHOR INFORMATION

Kari Sheets
Meteorologist
Meteorological Development Laboratory
National Weather Service
National Oceanic and Atmospheric Administration
1325 East West Highway

Room 11330 W/OST22
Silver Spring, MD 20910
301-713-0023 x125
301-713-9316 (fax)
kari.sheets@noaa.gov

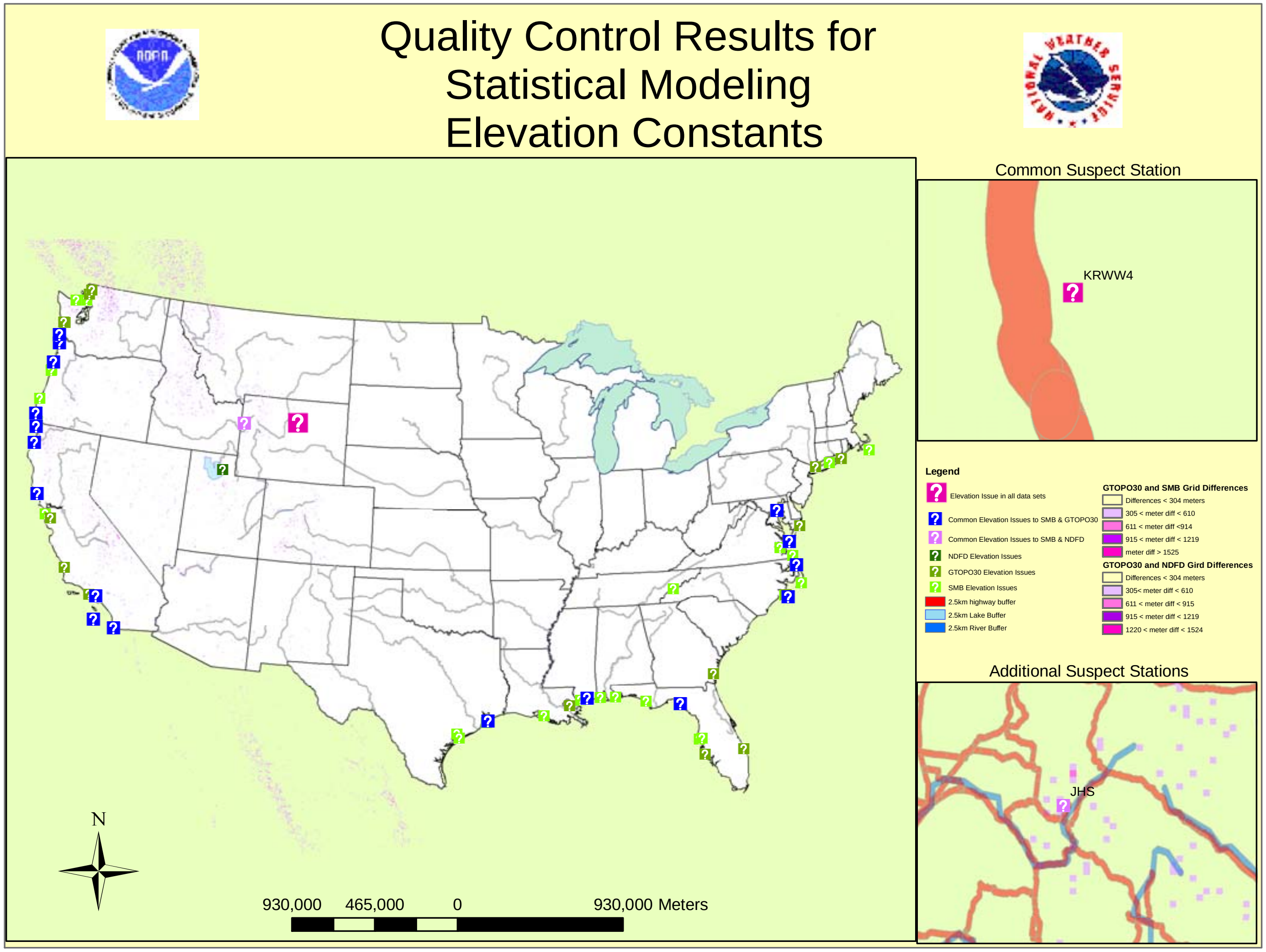


**Figure 1 –** Values of gridded elevation constant files were extracted to station locations and compared to the elevations reported by the stations. The question marks in the figure correspond to stations having greater than 300 feet differences in elevation when compared to two elevation sources.

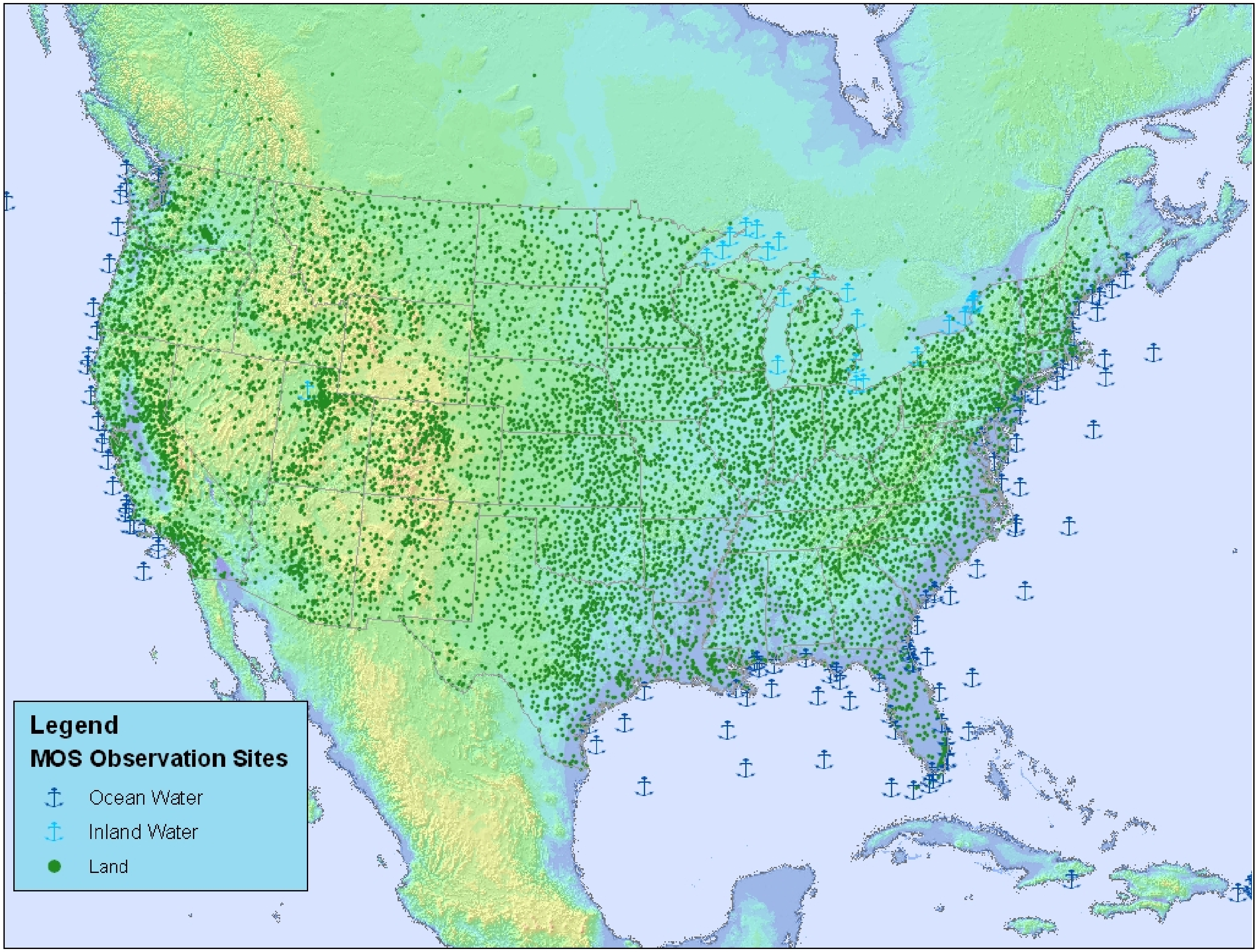


**Figure 2 –** MOS development stations designated as being located over ocean water, inland water (Great Lakes), or land.

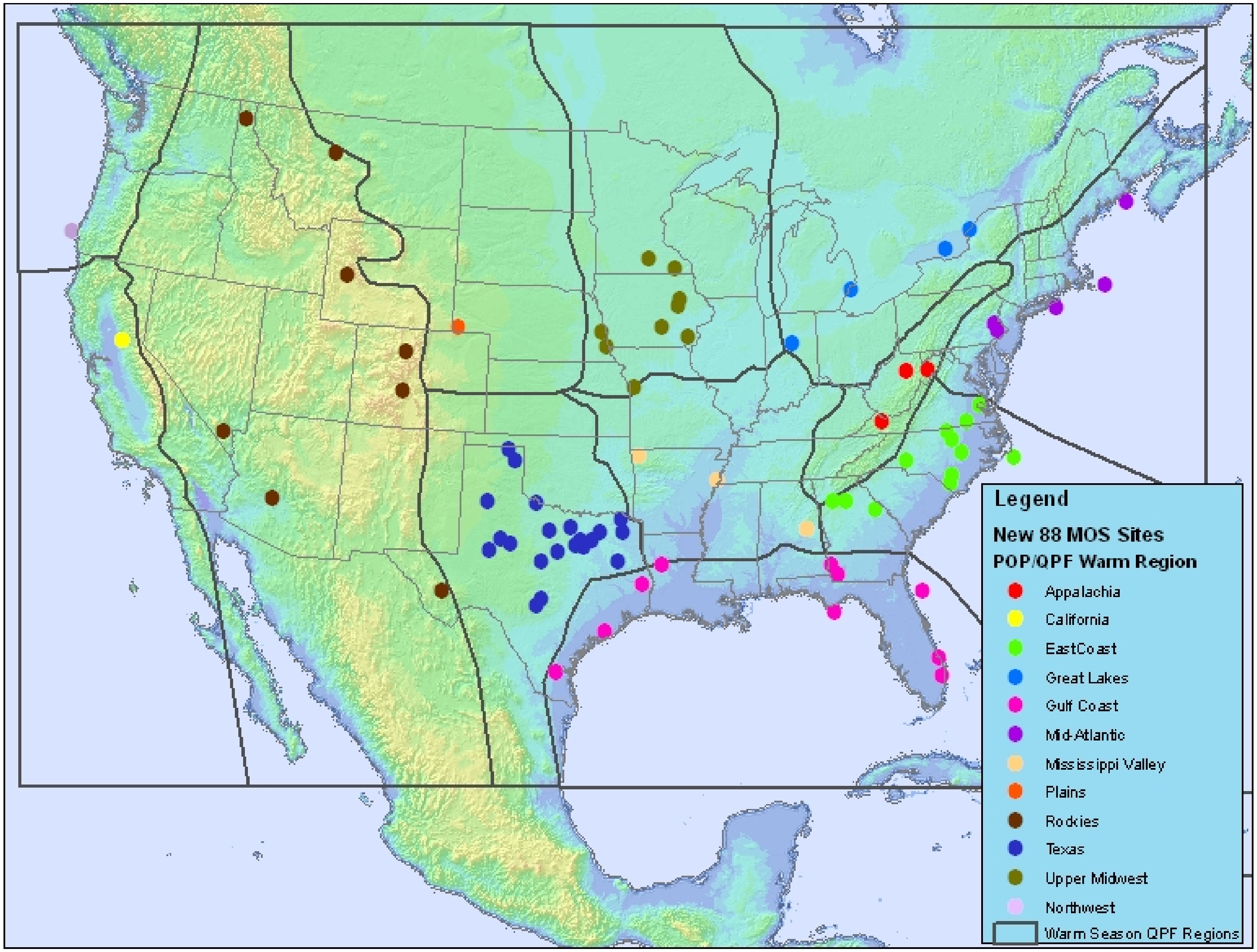


**Figure 3 –** 88 new development sites were added to the MOS system. GIS was used to assign these sites to existing probability of precipitation and quantitative precipitation forecasts regions.

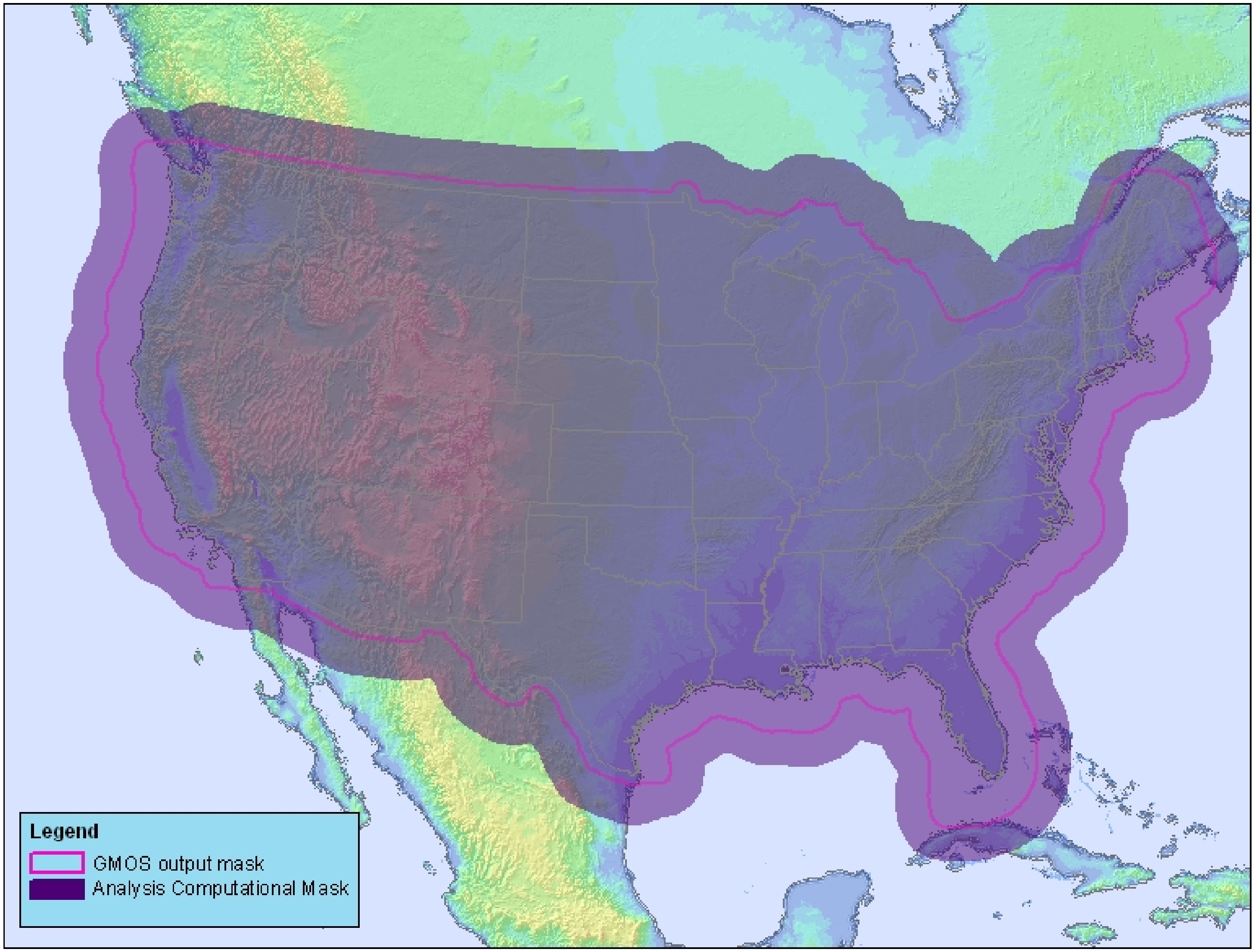


**Figure 4 –** The analysis area for gridded MOS needs to extend beyond area visible to users.  A 200 km buffer was created around the product extent to create the computational mask.

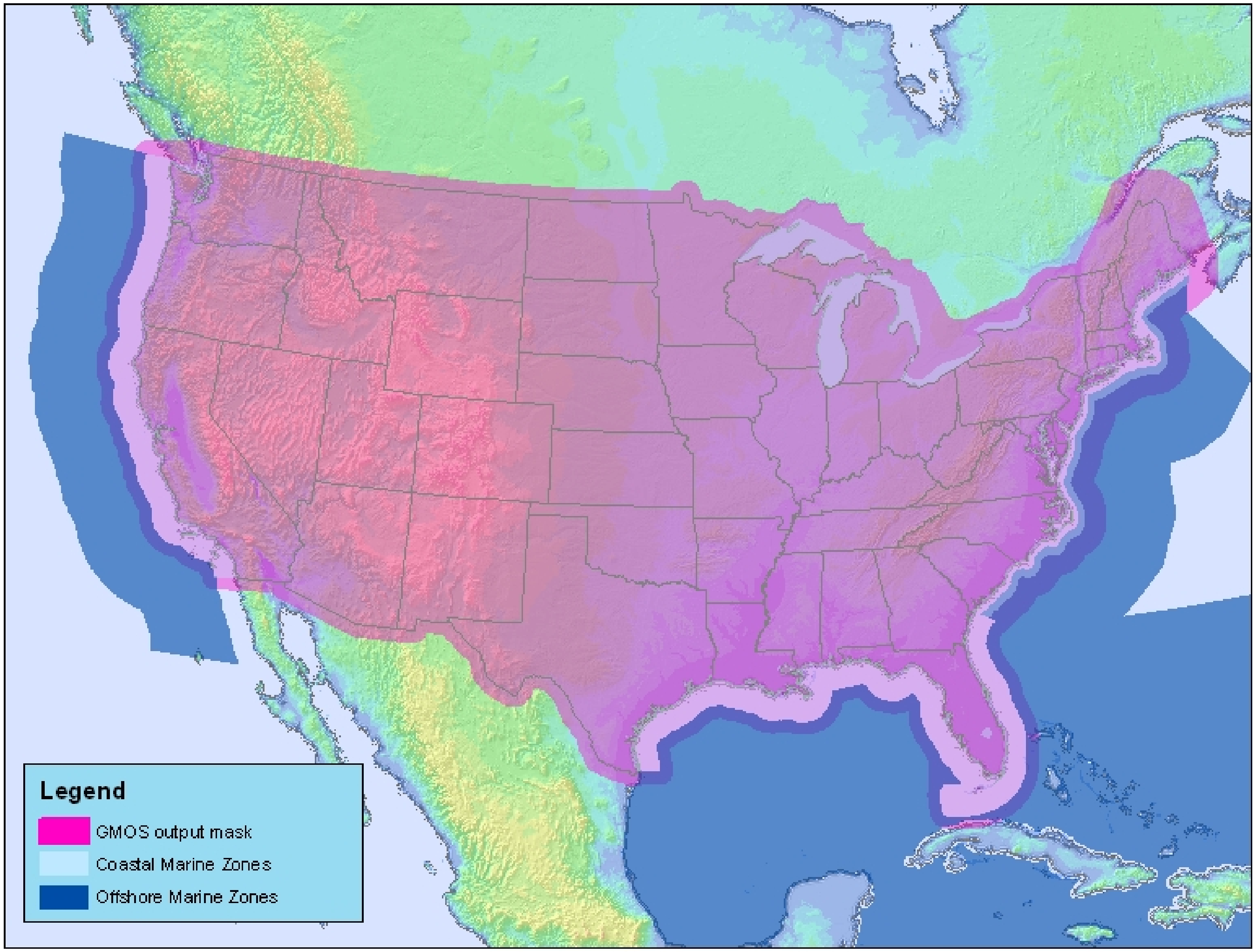


**Figure 5 –** Gridded MOS's operational output needs to include the areas for which Weather Forecasting Offices are responsible to forecast in addition to a small buffer for advecting weather systems. This output mask was created by merging National Weather Service marine zones, the land area of the conterminous US, the Great Lakes, and a 50 km buffer into Canada and Mexico.

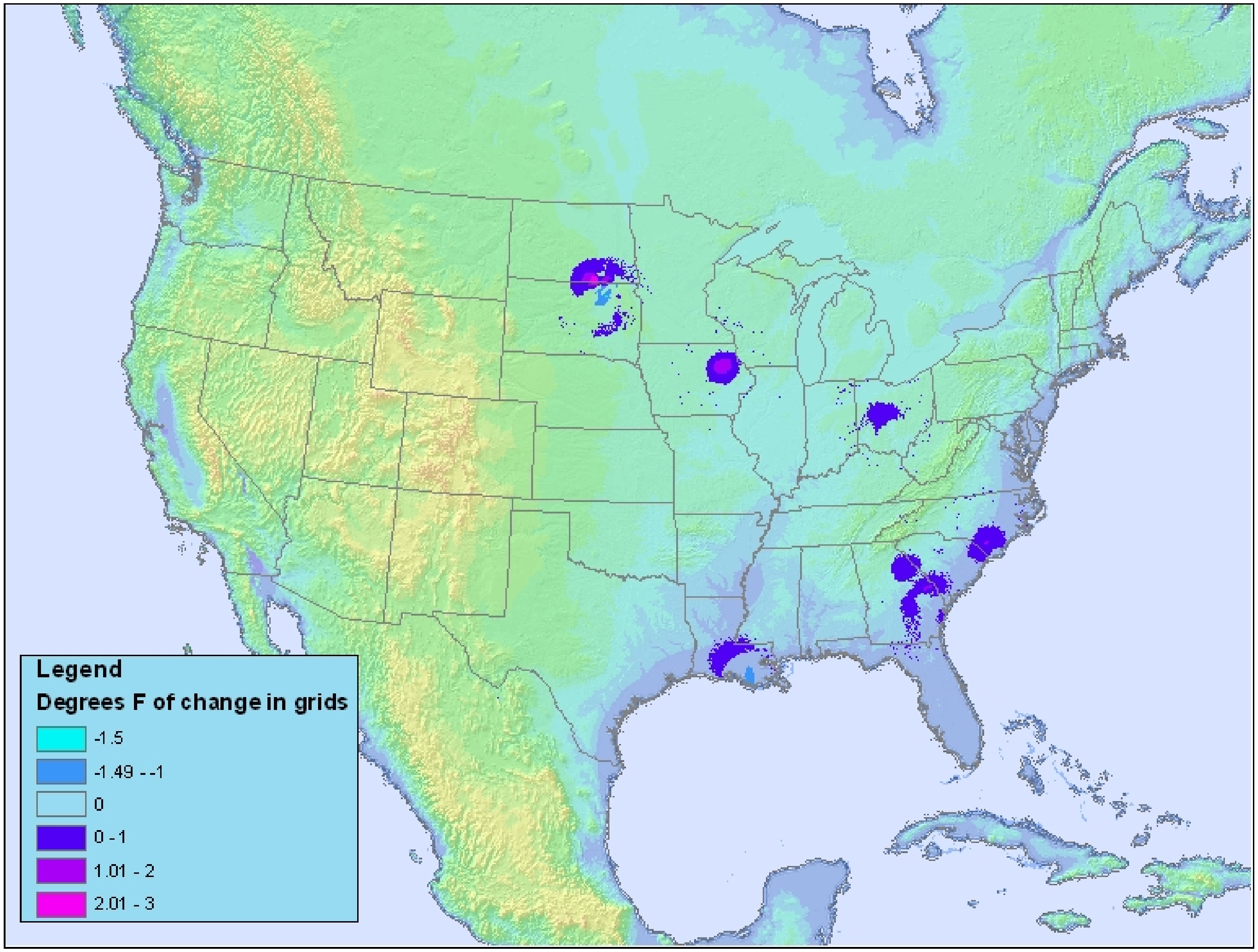


**Figure 6 –** A few stations were found to have bad dew point observations, so they were removed from the gridded MOS dew point temperature analysis. The figure shows the results when subtracting the grids including the erroneous sites from the grid excluding these sites.