

GRIDDED MOS “Talking Points”

Slide 4: What is MOS?

Model Output Statistics or MOS uses statistics to develop relationships between a weather element (or predictand) at an individual station to appropriate variables (or predictors) that can be obtained from NWP forecast models, prior surface observations, or geoclimatic information such as mean temperatures or precipitation. If a weather element is rarely observed, groups of climatologically similar stations will be pooled together to obtain these statistical relationships.

The statistical model used is multiple linear regression, with forward selection. This means that potential predictors are regressed against each weather element/predictand, and for each pass through the variables, the predictor that explains the most variance in the predictand is used and its effects removed. Then another pass is initiated and the next variable that accounts for the most variance is chosen and its effects removed. This process continues until the remaining error after accounting for the correlated variables is essentially random noise.

Slide 5: Why was MOS developed?

MOS was developed in the early 70s to “add value”, using currently popular terminology, to NWP models. It allows for an objective interpretation of the models. Some systematic bias is removed by traditional MOS; whatever bias existed in what is called the “developmental sample”. It also helps to account for forecast uncertainty at long lead times by tending to the developmental sample mean. MOS predicts what models cannot, or what models cannot predict well (for example, 2-m temperature where the model elevation isn’t representative of the single station of interest). It also allows for the production of site-specific forecasts from an NWP model that likely doesn’t have sufficient resolution to give such a forecast; as such, MOS can be considered a form of “downscaling” technique.

MOS also provides an obvious assist to WFO forecasters, by giving a built-in “first guess” for expected local conditions, and a model climatological memory for new staff that might be unfamiliar with the behavior of the weather at a given WFO.

Slide 6: An example of MOS versus GFS model performance (Tmax 1999)

This is for both CONUS and Alaska. Beyond the obvious better performance of the MOS forecast at all forecast projections, note direct model output is not constrained by the developmental sample’s “climatological” error, while traditional MOS is so constrained. MOS will always have at least the skill of forecasting “climatology”, or at least an estimate of it.

Slide 7: Developmental Considerations

When creating MOS equations with a developmental sample, there are a number of things that must be considered.

Observational databases must be selected carefully and **quality controlled!** A good example of an observational database with problems is the ASOS observations with respect to snowfall in winter. There have been instances where liquid equivalences have been understated by as much as 80-90% when there’ve been strong winds during a significant snowfall. Other issues come into play with mesonet data. How does this information get quality controlled? Are the instrument errors consistent throughout the mesonet? And how do we use remote sensing data?

The weather element being forecast for must be **PRECISELY** defined, including the time span of interest. For example, low temperatures are defined for purposes of MOS as being the minimum temperature from 7 pm to 8 am LOCAL time.

The predictors chosen to relate to the predictand must make physical sense. For PoPs, we'd use precipitable water, vertical velocity, model QPF, NOT 1000-500 thickness or tropopause height.

Finally, sometimes the predictor or predictand need to be transformed into binary data or some other form in order to develop traditional MOS relationships.

Slide 8: Developmental Considerations (continued)

How do we choose the terms in the regression equations?

What kind of data do we want to use? Do the relationships between the predictors and predictands (and the choice of predictands) change from season to season? What can we do if the data sample proves to be too small? We can increase the sample size by combining climatologically similar stations. When will we want to make "categorical" forecasts? QPF would be an example.

Slide 9: Example of Linear Regression: KCMH for January 1994

Here, we note that for KCMH (Columbus, OH), a single predictor (the 18-hr forecast of NGM 1000-850 hPa thickness in meters, on the x-axis), can explain 93.1% of the variance in that day's maximum temperature (on the y-axis). While this single predictor can therefore forecast much of the value of the predictand, we note that in the middle of the distribution, we still have about a 10°F residual error that we would like to account for. To do so, we remove the value from the linear relationship above to obtain a "residual value", and go to the next variable that might have a statistical relationship to today's maximum. The process is repeated until we get down to an acceptable tolerance for the residual error, which by that point should resemble random error.

Slide 10: "Real" regression equations

After this process, what does this "multivariate, linear regression equation" look like? It sets the predictand value, Y, to a series of predictors or X's multiplied by coefficients, the a's, whose values minimize the error in Y, and include an a-zero; the best forecast value for Y when all the X's are 0. The predictors, X, are chosen in such a way that they make some kind of physical sense.

Slide 11: Seasons of MOS

Predictor relationships with predictands vary by season. This table shows the seasons set up for different model-based MOS forecasts. Note that NGM has more seasons, since there is more data available to develop predictor-predictand statistical relationships. Also note: Eta MOS will go away in January 2007, to be replaced with WRF MOS once enough data is available to develop statistically reliable MOS equations.

Slide 12: Pooling of data for "rarely observed" variables

If an event occurs relatively rarely at a single station, for example, thunder at Washington DC (about 35 days a year), we need to pool data from different stations to be able to obtain statistically stable relationships. Pooling is done, of course, for regions with similar climate characteristics.

The weather elements that use pooled station data included, but are not necessarily limited to, the list shown.

Slide 13: An example of GFS MOS pooled regions for cool season QPF and PoPs

The example here is regions for GFS MOS cool season PoP and QPF. If there aren't enough climatologically similar stations for a particular predictand, the area combined will have to be expanded until stable statistics are obtained. An example of this here is region 3, over the Great Basin, which stretches from border to border north to south.

Slide 14: MOS Best Category Selection example

How do we convert probability forecast to a categorical forecast? The example is for 48-hr forecast of 12 hour QPF for Washington DC for 31 Oct 1993. Note categories are 0.01", 0.10", 0.25", 0.5" and 1.0". Bars are probabilities for each category; the white line is the exceedance threshold for the category. The lowest category for which a predetermined threshold is exceeded winds up being the forecast; in this case 0.10". Note the probability is only about 38%, but the 35% threshold is indeed exceeded. Also note that there still is a small probability of even an inch of precip., but because the 1.0" threshold is not met, this is not reported to MOS output.

Slide 15: MOS advantages and limitations

So, what are the MOS advantages and limitations?

1st off, MOS recognizes the skill of NWP models and uses it to best statistical advantage by removing some systematic NWP model bias, and objectively choosing the best predictors for the MOS forecast. For probabilistic forecasts, RELIABLE probabilities are produced, that is, if a forecast of 40% chance of measurable precipitation is made, over the long run measurable precipitation is observed 40% of the time. We also can downscale the NWP models to allows for specific site forecasts of weather elements.

What are the limitations then?

MOS is a synoptic scale method to downscale NWP.

Performance degrades for extreme events, and at longer lead times (regression toward the developmental sample climatology).

Developmental data samples present limitations as well: changing NWP models change predictor/predictand relationships, extreme events are not well-represented (and may not be represented at all), and differing flow regimes are combined to make the MOS equations. Undersampling of any particular regime will result in problems for MOS when that regime is present.

Slide 17: The National Digital Forecast Database (NDFD)

NDFD is a 3-d representation of the weather from the present to several days into the future. The current NDFD grid is 5-km resolution; plans are to go to 2.5-km resolution. The NDFD is built from WFO local digital forecast databases, and updated as needed, particularly in the short range (1-3 days).

Slide 18: Example of NDFD from the CONUS central plains

The example is a graphic of NDFD forecast maximum temperature over the central US plains for 22 August 2006, made at 8 pm 21 August 2006.

Question the students about what they notice about the graphic. Discussion should center around how many WFO's CWA boundaries show up as discontinuities in the forecast (e.g. Goodland KS v Denver/Boulder CO, Tulsa vs Oklahoma City OK).

Slide 19: Why Gridded MOS was developed

Why was Gridded MOS conceived?

To produce forecast guidance for NDFD at the native NDFD grid resolution, generate forecast guidance with enough detail for initialization of forecasts at WFOs, and generate forecast guidance comparable to single station MOS. We also get smoother guidance; there are no WFO discontinuities, though anywhere regional boundaries for Gridded MOS exist, there is a risk of discontinuities.

Slide 20: What methods could be used to develop Gridded MOS?

How is Gridded MOS developed? There are two basic methods we COULD choose from:

- We can develop regression equations that apply to single-station sites, like in traditional MOS, and then interpolate the results to the Gridded MOS grid. This presents the challenge of gridding from quasi-random points to a regular grid.
- Alternatively, we can develop regression equations that are applied at grid points, and directly make forecasts at each of these grid points. That presents obvious challenges in terms of what data might be available to develop statistical relationships. There are almost NO data sets that provide sufficient detail to accomplish this, but there are some exceptions which we'll discuss later. So the first method is the one used now.

Slide 22: Single station datasets

What datasets are currently in use in developing the single station values to interpolate to the 5-km Gridded MOS grid?

- METAR
- Buoys/C-MAN locations
- Mesonet sites
- RFC-supplied sites
- NOAA cooperative observer network
 - Used for 24 hour snow, not shorter time intervals.
 - Co-op max/min data are problematic due to different obs times by different observers.

Under development by Meteorological Development Lab (MDL):

- Precip obs from NPVU (<http://www.hpc.ncep.noaa.gov/npvu/>)

Slide 23: How does the combined data look in the CONUS?

We have 1500 traditional GFS MOS stations, 1500 RFC and Mesonet stations, and 5500 coop data sites, for a total of 8500+ stations! These are compared to the GFS model data as in traditional MOS to come up with single station MOS values. These quasi-randomly spaced stations are what now must be interpolated to the 5-km grid.

Slide 25: Data that can be directly used at the resolution of Gridded MOS

Before we go on to the interpolation method, let's talk about data that can be used directly in Gridded MOS without interpolation. Remote sensing data is typically at such resolution, including lightning, radar, and satellite data. Lightning data is currently used for single station MOS for the Gridded MOS product; the others are under development.

Slide 26: Examples of satellite and radar data (use animation buttons)

Points: Satellite based effective cloud amount has issues with distinguishing thin overcast with thicker scattered to broken cloud. Radar estimates are problematic in situations where radar/precipitation relationships are problematic; this can be resolved to at least some extent by blending, for example, gauge data with radar data.

Slide 28: Data that can be used to interpolate from single station to grid

Points: Can use slope, aspect (the direction the downward slope faces), land use type, and climatological data from PRISM to interpolate single stations to the Gridded MOS grid. The examples here are for NW Montana, as developed from USGS 30' topographic data, NASA 1-km land use data, and 4-km Oregon State University PRISM data, respectively.

Slide 29: The starting point for Gridded MOS

Any analysis must start somewhere. Typically it is called a "first guess", and can be an average value of the field for all stations in an area of the grid, some specified constant, data from a forecast model, interpolated to the 5-km Gridded MOS grid, and so on. This latter possibility has shown promise as an improvement for Gridded MOS interpolations.

Slide 30: Objective analysis

How do we then get to the final Gridded MOS product from this first guess? We use what's called the method of successive corrections, based on the Cressman analysis. The method uses successively tighter circles (with specified radii or influence) with relative weighting of values based on the distance of each gridpoint from those values.

Additionally, land gridpoints are treated differently from water gridpoints. The analysis parameters for land versus water are different, and land points cannot affect water points and vice versa.

An elevation or "lapse rate" adjustment is made based on the single station values at different elevations, obtained from the 30" USGS topographic data set discussed previously. Because the lapse rate adjustment calculation is done "on the fly", the lapse rate can vary with weather element, time of day, season of the year, and synoptic situation, **as long as the single station values pick up on these differences.**

Slide 31: Land/Water distinction

Because data over water is sparse, the radii of influence over water are 3 ½ times larger than over land. Again, land points do not affect water points in the analysis and vice versa. Small water bodies in a domain are not dealt with AS water unless there is a water data point close enough to influence a grid box.

Slide 32: Objective analysis example (use animation tool)

We start with an idealized mixed land and water domain to look at how the analysis is done over land. The squares over land represent analysis grid boxes with grid points in their centers. Points in brown over land and in white over water indicate the locations of single station MOS values.

The first circle used to determine what points will be used in the grid analysis is 30 grid points in diameter. The weight each station value gets in the determination of the grid point value is inversely proportional to its distance from the grid point. Here, the 76 value northwest of the Gridpoint will get much less weight than the 72 value right next to the Gridpoint. Note that the 62

value over water, though it would be within the radius of influence, would not be used in the calculations.

Once the result is determined, we go to the next circle, which has a 20 grid point radius, and we repeat the process. Then we go to 10, 5, and 3 grid point radii to complete the analysis. Since we have a single station within all the circles (value of 72), this single station will be heavily weighted in the grid point analysis value. This has consequences to the appearance of the Gridded MOS which we will discuss later.

Slide 33: Objective analysis example in reverse

The same process is used in the analysis for water grid points. The only differences are that land stations are not included, and the radii of influence are 105, 70, 35, 17.5, and 10.5 grid points.

Slide 34: Lapse rate calculated for each station

The “lapse rate” for purposes of the Gridded MOS analysis is defined as the average paired difference of all single station site forecast values divided by the corresponding elevation difference. The closest stations within a 30 grid point radius circle are used with a maximum of 100 stations. Typically, about 60 stations can be used, with as few as about 20 out west, and the maximum in areas of dense networks.

The lapse rate will typically be negative, but may be positive where the single stations capture inversions, especially along the west coast, but sometimes in other situations such as wintertime Great Basin inversions or cold air damming situations.

Slide 35: Lapse rate example (use animation tool)

Here is an idealized example of single stations around some grid box in the center of the graphic. The red digits are elevations and the blue numbers are single station temperature values at the circles indicated. Each are paired and differenced with respect to the single station values and their elevations. We'll calculate the pairings with the 72°F, 1675-m height station found in the left-center of the graphic. It is first paired with the 74°F station at 1475-m. The first pairing calculation is shown in the lower right blue box, and results in a lapse rate of -0.01F/meter. The rest of the pairings for this station are then shown in the animation, without a calculation. Each other station is checked for unique pairings of temperature and elevation and included in the calculation. If you do all the possible pairings for all stations in this graphic, you'll come up with about a -0.009°F/m lapse rate.

All the lapse rates are obtained over a circular domain within 30 grid points of the grid box being analyzed, or out to a distance of 150-km. What might happen if we mix two areas where the lapse rate characteristics are different in the Gridded MOS product? Keep this in mind when we discuss a GFE example later on.

Slide 36: Smoothing of Gridded MOS data

After the objective analysis is done, including the application of the lapse rate, two other procedures are done to the gridded data. The first is smoothing. Now, the Gridded MOS uses a 5-point (like an “x”) or 9-point (like an eight-pointed star) stencil where the terrain is relatively flat. In more rugged terrain, low or high points are not smoothed and along any contour where a series of 3 out of 8 cardinal directions are more-or-less along an elevation contour.

Slide 37: Smoothing example

Here's a 5x5 example of a piece of the Gridded MOS product after the analysis and lapse rates are applied. If we assume flat terrain, we can use the stencils. The 5 point stencil is centered on

the 5x5 grid, and is used to smooth the points around the center grid box with value of 77°F. The 9-point stencil is used the same way, with 4 additional grid boxes surrounding the central one. Finally, let's throw in some topography, indicated by the contours. The three grid boxes running NW to SE will be smoothed, since they are at about the same elevation.

Slide 38: Nudging in the Gridded MOS analysis

Besides smoothing, the closest station to a single station value is nudged toward or set to the single station value, depending on the difference between the grid point and single station values. This allows for a slightly closer fit to the single station data with less risk of bull's eyes, and allows users of the Gridded MOS to almost always recover the single station value from the grid.

Slide 39: The result of all this analysis, and the importance of terrain and land/water distinction (use animation tool)

In the end, what do we wind up with? This is a temperature forecast from the Gridded MOS for southern CA, northwest AZ, southern NV, and southwest UT at 00 UTC 15 July 2004. The contours you see are terrain. The shading is temperature, with values indicated in the color bar to the left of the graphic. The numbers are Gridded MOS single station temperatures. We can see many terrain related features, including the Sierra Nevada in the upper left, the Owens valley just to its east, Death Valley, east of that in white (temperature >95F), and so on. The land/water distinction can be seen in the lower left where the green shading over the Pacific, indicating temps below 70F, meets the yellow shading of the LA basin, indicating temps in the 70-75F range. This temperature forecast looks physically reasonable given what we know (or what is indicated) about the terrain in this region.

(hit animation >) If we remove the effects of the terrain and the land/water distinction, much of the meaningful detail is eliminated. Death Valley and the Owens Valley “disappear” in the temperature field, the temperature contrast across the coastline in the lower left of the domain is removed, and elsewhere we see that the fine details of the effect of terrain on the temperature is essentially lost, like in NV where the ridges and valleys indicated by the terrain contours are ignored.

Slide 40: NDGD (Gridded MOS) versus NDFD side-by-side comparison

We go back to the prior example from the NDFD maximum temperature forecast from 21 August 2006 valid on 22 August. When we compare the Gridded MOS on its NDGD on the left to the WFO-derived NDFD on the right, we note that the WFO boundaries disappear and that the forecast grid is a lot smoother across the full domain in the NDGD. However, there **is** an artifact of the analysis that becomes apparent. What do you notice about the NDGD that might be problematic?

Answer: circles or “measles” in the analysis, reflecting the analysis gives too much weight to stations that lie near or on grid points. Examples can be seen across just about all of the domain, for example, just west of Kansas City, MO, east and northeast of Liberal KS, southwest of North Platte NE, and so on. This problem is being worked on by the Meteorological Development Lab, and reducing the number of passes in the analysis (eliminating the small radii of influence) seems to reduce the “measles” apparent on the map on the left. Keep this in mind for the GFE example to be shown later.

Slide 41: NDGD versus NDFD PoPs

Similarly, the NDGD give smoother results for precipitation forecasts such as PoPs. This example is from October 2005 for the Pacific Northwest.

Slide 42: List of variables found in the NDGD

This screen-capture of the Gridded MOS webpage is shown to give the list of weather elements included in the Gridded MOS products. These can be seen on the left and include:

- Max and min temperature
- 6 and 12hr probability of precipitation
- 3 hrly values for temperature, dewpoint, RH, wind direction and wind speed; and
- 3, 6, and 12-hrly probability of thunderstorms

Slide 43: Example from GFE screen capture from Monterey, CA (use animation tool)

Now here is an example where we can apply what we've learned so far about the Gridded MOS product. The graphics are screen captures from the GFE editor at the Monterey, CA forecast office. In all graphics, color shading represents the temperature in degrees F (level as in the color bar at the top), the CWA boundary is marked in orange, the coastline is in black, and areas above 1,500' in elevation are crosshatched. Question for students: What situation are we seeing that is reflected in the temperature pattern in this graphic?

Answer is on the second panel of the animation: it's a marine layer with inversion, capped by warmer air aloft. Temperatures are 10-15°F warmer above 1500' than near sea level, anywhere near the ocean. High areas are "poking up" above the inversion. Note marine layer intrusion east of San Francisco into the Sacramento River delta, and east of Big Sur (south of MRY), all related to gaps in the topography. Inland low-lying areas in the Central Valley of CA are warmer than corresponding elevations along the coast.

Panel 3: How does Gridded MOS handle the marine layer and its overlying inversion? That will depend on how the analysis handles this type of situation. Recall that the "lapse rate" for the Gridded MOS interpolation uses the single stations from as far as 30 gridpoints away; in this case that would be 150 km. encompassing the Central Valley of CA, where the marine layer has minimal or little effect. Also, the circular analysis method may get us into difficulty, in that single stations from the Central Valley will get into the analysis for many of the grid points even near the Pacific Coast.

Panel 4: Dewpoints look more reasonable, however, from the Gridded MOS, even with the difficulties with the marine layer temperatures.

Page 44: With any Gridded MOS forecast product, one of the keys is the capturing of the effects of the synoptic situation through the single station MOS first, before the analysis is done. Without such a capturing, even the best interpolation will be meaningless. Here is an example from Central NY. Concentrate on Elmira and Binghamton (which is the next station to the east). In the first panel of forecast winds from 5 a.m. EDT 29 September 2006, note that the wind speeds indicate that the PBL according to Gridded MOS is not expected to significantly decouple. In such circumstances, we'd expect Elmira, at a lower elevation than Binghamton, to be warmer.

Panel 2: Minimum temperature forecast for period ending 8 a.m. EDT shows that indeed Elmira is a degree warmer than Binghamton.

Panel 3: The next night shows that the Gridded MOS has decoupling. Winds predicted to become nearly calm at Elmira, and 4 mph at Binghamton.

Panel 4: Radiation cooling is captured by the Gridded MOS as well, with Elmira being 35°F and Binghamton at 39°F.

Page 45: The Gridded MOS product was evaluated over the period Dec. 2005-September 2006, so that both cool and warm seasons were covered. Since it was expected that the western CONUS would be most challenging for the Gridded MOS, that was where the objective evaluation was done. NDFD, HPC, and MOS grids were compared for max/min temperature and dewpoint

only. Sites selected for the evaluation were not used in the development of the single station MOS equations, and were generally RAWS (Remote Automated Weather Station) sites.

Page 46 (use animation): We concentrated on the products available at approximately the same time: 17Z GMOS (from the 12z GFS), 15z HPC grids, 18z NDFD grids. Days 4 through 7 were compared. Max/min temperature AND dewpoints (panels 1 through 3 respectively) were generally better predicted by the Gridded MOS than HPC or the consolidated NDFD grids. The difference is most apparent with dewpoints, performing as much as 1°F better throughout the medium range forecast period.

Page 47: Reviews where further information on the Gridded MOS product can be obtained, with URLs provided.

Page 48: Issues with Gridded MOS similar to issues with “traditional” MOS. Predictors and especially predictand data sets need to be chosen carefully and quality control must be carefully applied. Certain synoptic situations are not dealt with well in GMOS. These are similar to situations where traditional MOS does not perform well either. The grids cannot be verified directly against a 5km analysis, since none is available yet. Once RTMA is available, that might be able to be used.

Page 49: Time line and schedule is as shown. AWIPS 7.1 has Gridded MOS available.

Page 50: RTMA will be important for verification, but in the meantime single stations can be verified in the same manner as traditional MOS. Further improvements to Gridded MOS are in the pipeline, including use of remote sensing data on the grid, refinement of the MOS equations, and use of interpolated GFS data as a first guess for the Gridded MOS analysis.

Page 51: Use the NWP forum for communications on Gridded MOS!

Page 52: How to contact the COMET NWP team.