# WFO La Crosse Research Series #8

# The NGM MOS Local Equations: What's Important to Know?

Kenneth Cook
September 1998

## 1. INTRODUCTION

National Weather Service forecasters have used Model Output Statistics (MOS) for several years now, but what understanding do we have of these? Since MOS is developed at the central level and used at the field level, not much information has been exchanged between the field and Headquarters. This document will attempt to briefly explain the methodology in developing the equations, provide some insight into the MOS equations themselves, and offer what is expected in the near future.

## 2. DEVELOPMENT

MOS is developed using a number of statistical regression processes. Data sets of predictand data (the observed data, temperature for example) and predictor data (the numerical model data) are built. Then using a number of trial and error procedures, a correlation[1] is found between a predictand and its associated predictors which best simulate it. This methodology should act to eliminate numerical model biases given the data set is large enough and the numerical model has not been altered over the time of the data sets.

## 3. THE EQUATIONS

Through this process, the best predictors for a given predictand are used. Of the equations explained below, max and min temperatures are the only predictands which have unique equations for each station. The remainder are developed using a regional approach (see Table 1 for a list of stations for each predictand). The reason for this is, by comparison, the data showed no significant increase in accuracy by developing regional equations versus station by station for these predictands. This could be due to the limited grid size (80 km) the data was developed from. It should be noted that this comparison has not been tested since the development of the LFM MOS equations. Additionally, this regional approach can and does contribute to missed forecasts of river valley fog and other small scale weather events.

In the future, developments may be more differentiated as computer power and model resolution have increased. It must be understood that MOS is developed using an approach

---

[1] The amount of correlation can and will be explained by a correlation coefficient where a correlation coefficient of 1 is perfect simulation of the predictand by its predictors and 0 is no similarity. In other words a correlation coefficient near 1 indicates that forecasts of a predictand, such as temperature, were found (through verification) to be reliable and accurate whereas a correlation coefficient near 0 would indicate the opposite is true.

which smooths data so as to minimize errors. (This same smoothing also eliminates extreme events from the data set which rarely occur (e. g. record high temperatures) so that forecasts of these rare events will likely not occur). The result from this has been (and will likely be in the future) that new MOS developments will not be using a grid size the same as or even similar to the one currently being employed by the numerical models unless the data proves otherwise.

## 4. PROBABILISTIC FORECASTS

NGM MOS can now be viewed using Ntrans under NGM_MOS. This "plan view" has advantages in that it shows probabilistic forecasts of several fields (predictands) where the MOS message only shows, in many cases, the best category. It should be known that many of these forecasts, such as PTYPE, CLOUDS, etc are additive. In other words, if you sum the probabilities of all the different forecasts of PTYPE, they will equal 100%.

Also, within these forecasts the same exact predictors are NOT used to predict clear, scattered, broken, and overcast skies. Instead, each individual predictand (clear, scattered, broken, and overcast) has its own set of equations. They are therefore individually correlated allowing for the best prediction possible of each individual field.

Table 2, provides an explanation of which are the most influential predictor for each predictand and defines how well the data sets are correlated. As previously explained, each predictand is a function of its associated predictors.

## 5. EXPLANATION OF THE PREDICTANDS

### A. Temperatures

For temperature predictions, the correlation coefficients are quite high for both LSE and RST. That is to say that temperature predictions should be reasonably accurate. Comparing the correlation coefficients, the fall and spring equations show the highest correlation with fall being the best. The winter season is third then ending with the summer.

For max temperature, the 00z is better correlated than the 12z and the opposite is true for the min temperature, though the differences between these are rather small. Station differences are negligible. The main fields of influence are the sun angle, low layer temperatures, dewpoints, winds, and thicknesses, and grid binary[2] precipitation (threshold of .05 in).

### B. Probability of Precipitation - 12 Hour

The 12 hour pop equations correlate marginally as the range of the correlation coefficients

---

[2] A binary predictor is a "yes or no" answer based on specific criteria (yes being equal to 1; no equal to 0). In MOS, for example, suppose a specific grid area at 850 mb is assigned a grid binary predictor of relative humidity with a threshold of 70 percent. The numerical model forecast for that grid area is 76 percent therefore, a 1 would be assigned to that grid space. This is how precipitation can be turned on and off in the MOS.

is from .540 to .670. There is better correlation with the earlier projections than the later. There is also a small but better correlation, in general, with the 00z cycle than the 12z cycle. The main fields of influence are the low layer relative humidity (several layers), mean (surface to 490 mb) relative humidity, moisture divergence, low and mid level relative vorticity, grid binary precipitation (01. and .05 thresholds).

C. Probability of Precipitation - 6 Hour

The 6 hour pop equations are not quite as well correlated as the 12 hour. The correlations are much better for earlier projections than later ones. The correlation coefficients average about .570 for the warm season and about .640 for the cool season. This better correlation of the cool season is noted throughout the equation data sets. The correlation falls as the projection time increases. In fact, by the 60 hour forecast, the correlation coefficients fall into the .300 to .400 range which is rather poor.

The main fields of influence are the low layer relative humidity (several layers), mean layer (surface to 490 mb) relative humidity, grid binary precipitation (several different thresholds), low and mid level relative vorticity and moisture divergence. It is interesting to note during the warm season, more the forecast contribution is from the moisture divergence field as opposed to the cool season where the greatest contribution is from the grid binary precipitation.

D. Clouds

The cloud equations are poorly correlated with the exceptions being predictions of 12 hours or less. For these equations, the correlation coefficients are better than .500, showing relative promise. The main fields of influence for these equations are the observed surface conditions at 03Z/15Z (03Z for the 00Z model run), low, mid and mean layer (surface to 490 mb) relative humidity. For the time period thereafter, the most significant fields of influence are the grid binary - mean relative humidity both 80% and 70% thresholds.

E. Ceiling Height

The ceiling height equations are also poorly correlated with correlation coefficients ranging from .300 to .450. However, on several occasions they fall below this range, especially in the outer projections. Therefore, the earlier projections or forecasts are better correlated than the later ones. The winter equations are correlated better than the summer. The main fields of influence are the observed ceiling height at 02Z/14Z, low layer relative humidity (several layers), and mean layer (surface to 490 mb) relative humidity.

F. Visibility

The visibility equations show the poorest correlation of all the equations discussed in this document. Most of the correlation coefficients are in the .250 to .400 range but a few are below and one above that range. The main fields of influence are 1000 mb grid binary relative humidity (90% threshold), grid binary mean relative humidity (90% threshold), surface to 490 mb mean relative humidity.  There is no significant difference between the 00z and 12z cycles however, the winter equations are better correlated than the summer.

**Table 1:**

| Predictand: | Stations Equations are Valid for: |
|---|---|
| Temperature | LSE and RST are done seperately |
| Probability of Precipitation - 12 hour | **COOL SEASON:** Y62, AUW, AXN, DLH, FAR, GFK, INL, LSE, MSP, RST, HON, OFK, SUX, FSD, EAU, RWF, BIS, MOT, Y26, ISN, SAW, MQT, CMX, RDR, HIB, 3SE, P75, CIU, IMT, P59, IWD, STC, BJI, OTG, BRD, D45, P39, P61, P11, JMS, DIK, MIB, P24, P67, ABR, ATY, Y22, REJ, 9V9.<br>**WARM SEASON:** PAH, SLN, IAB, ICT, CGI, COU, MCI, BLV, SZL, FRI, CNK, JLN, STL, SGF, TOP, MSN, MKE, PIA, NBU, AUW, GRB, LSE, MLI, RST, DSM, GRI, LNK, MCW, OFK, SUX, OFF, OTM, 3OI, CID, EAU, SPI, UIN, RSL, RFD, ORD, ALO, BRL, 3SE, DBQ, FOD, CIR, MWA, DEC, MDW, MTO, MDH, CMI, MMO, MHK, P28, FOE, FLV, CNU, EMP, T05, OTG, GVW, TBN, P02, SUS, MKC, STJ, VIH, IRK, P35, OMA, OVN, ATW |
| Probability of Precipitation - 6 hr | **COOL SEASON:** Same as cool season POP 12hour.<br>**WARM SEASON:** Same as warm season POP 12 hour. |
| Clouds | **COOL SEASON:** DEC, GVW, MHK, MCI, FOE, FLV, SZL, FRI, CNU, MKC, EMP, STL, TOP, MSN, PIA, AXN, P11, FAR, GFK, JMS, LSE, MSP, MLI, RST, STC, ABR, BRL, DSM, HON, IRK, LNK, MCW, OFK, OMA, SUX, FSD, ATY, OFF, OTM, BJI, 3OI, 3SE, CID, EAU, RWF, SPI, UIN, RFD, CMI, MMO, DBQ, ALO, OVN, RDR, OTG, FOD, BRD, P39, P35, P67.<br>**WARM SEASON:** DEC, GVW, MHK, MCI, FOE, FLV, SZL, FRI, CNU, MKC, EMP, STL, TOP, MSN, PIA, AXN, P11, FAR, GFK, JMS, LSE, MSP, MLI, RST, STC, ABR, BRL, DSM, HON, IRK, LNK, MCW, OFK, OMA, SUX, FSD, ATY, OFF, OTM, BJI, 3OI, 3SE, CID, EAU, RWF, SPI, UIN, RFD, CMI, MMO, DBQ, ALO, OVN, RDR, OTG, FOD, BRD, D45, P39, P35, P67. |
| Ceiling | **COOL SEASON:** DEC, GVW, MHK, TBN, COU, MCI, SUS, BLV, FOE, FLV, SZL, FRI, CNK, MKC, STJ, STL, TOP, VIH, MSN, PIA, AXN, LSE, MSP, MLI, RST, STC, BRL, DSM, IRK, LNK, MCW, OFK, OMA, SUX, FSD, OFF, OTM, 3OI, 3SE, CID, EAU, RWF, MTO, SPI, UIN, RFD, CMI, MMO, DBQ. ALO, OVN, OTG, FOD, P35.<br>**WARM SEASON:** P75, ATW, MTC, OSC, BTL, MDW, CLE, FDY, FNT, FWA, JXN, LAN, MSN, MKE, MKG, PLN, PIA, MBS, Y62, SBN, TVC, YNG, MBU, ERI, MFD, CAK, AUW, GRB, DLH, INL, LSE, MSP, MLI, RST, STC, CID, EAU, HTL, AZO, RFD, CIU, TOL, GUS, SAW, ORD, DTW, APN, MQT, GRR, CMX, LAF, MMO, IMT, MBL, P58, P59, DBQ, ALO, IWD, HIB, BRD, P39, P61, 3RN. |
| Visibility | **COOL SEASON:** P75, ATW, MTC, OSC, BTL, MDW, FDY, FNT, FWA, JXN, LAN, MSN, MKE, MKG, PLN, MBS, Y62, SBN, TVC, NBU, AUW, GRB, AXN, DLH, INL, LSE, MSP, RST, STC, BJI, DTW, APN, MQT, GRR, CMX, MMO, IMT, MBL, P58, P59, DBQ, IWD, HIB, BRD, D45, P39, P61, 3RN.<br>**WARM SEASON:** JHW, IAG, P75, ATW, GTB, RME, BUF, ROC, SYR, MTC, OSC, BTL, MDW, CLE, FDY, FNT, FWA, JXN, LAN, MSN, MKE, MKG, PLN, PIA, MBS, Y62, SBN, TVC, NBU, ERI, MFD, AUW, GRB, DLH, INL, LSE, MSP, MLI, RST, STC, BRL, MCW, CID, EAU, MSS, ART, UCA, HTL, AZO, RFD, CIU, TOL, GUS, SAW, ORD, DTW, APN, MQT, GRR, CMX, LAF, MMO, IMT, MBL, P58, P59, DBQ, ALO, IWD, HIB, BRD, P39, P61, 3RN. |

**Table 2:**

| Predictand: | Approx. Avg. Corr Coef: | Fields of Main Influence: | Best Corr: | Worst Corr: | Notes |
|---|---|---|---|---|---|
| Temperature | .945 | Sun Angle; Low layer Temps, Dewpts, Winds, Thicknesses; and Grid Binary precipitation (.05 cutoff). | Max; 00Z Cycle; Fall; Spring | Min; 12Z Cycle; Summer; Winter | All predictands well correlated. |
| Probability of Precipitation - 12 hour | .613 | Low Layer RH (svrl layers), Mean (sfc-490mb) RH, Grid Binary Precip, Moisture Divergence, Low and Mid Level Relative Vorticity. | Cool, 00Z Cycle | Warm, 12Z Cycle | Some decrease in correlation through projection time. |
| Probability of Precipitation - 6 hour | .492 | Low Layer RH (svrl layers), Mean RH, Grid Binary Precip, Moisture Divergence, Low and Mid Level Relative Vorticity. | Cool, <= 24 hr fcst | Warm, > 24 hr fcst. | Both cycles were equally correlated. |
| Clouds | .345 | Observed clouds at 03Z/15Z and Low, Mid, and Mean Layer RH. | 12Z Cycle, Warm, <= 12 hr fcst. | 00Z Cycle, Cool, > 12 hr fcst. | Correlation Coefficients until 12hr forecast better than .500. |
| Ceiling | .301 | Observed ceiling height at 02Z/14Z, Low Layer RH (several layers), Mean RH. | 00Z Cycle, Cool | 12z Cycle, Warm | Predictands are poorly correlated. |
| Visibility | .254 | 1000 mb Grid Binary RH (90% cutoff), Mean RH Grid Binary (90% cutoff), Mean RH. | Cool | Warm | All predictands poorly correlated. Little difference between 00Z and 12Z. |

## Questions

1) Which predictand (field) would you expect to produce the most accurate forecasts over time? The least accurate?

2) Suppose you were on a midnight shift and dense river valley fog developed. You went back and looked at the 00z MOS and there was no fog forecast. Why do you think this is?

3) Suppose you were looking at the numerical NGM model and noticed that we were saturated through 500 mb. Would you expect 12 hour NGM MOS pops to be high or low? What if only a layer or two were saturated? Can you think of a reason why there would be a column of moisture of 50-60% from surface to 500 mb indicated by the numerical NGM and no precipitation noted in MOS?

4) Go to Ntrans and open the NGM_MOS selection. Now go to the pcp selection from the latest run. What is the 24 forecast of Conditional probability of snow for the LSE area? Now load the 6hr snow amount >= trace. What is the 24hr forecast of probability of a trace of snow for the LSE area? Now load the 6hr qpf expected value. What is the 24hr forecast for the LSE area?

5) In the same NGM_MOS window load the 12hr conditional probability of a severe thunderstorm (CND SVR TSTM PROB) and next to it load the 12hr probability of a severe thunderstorm (UNCND SVR TSTM PROB). Why do you think the conditional probabilities are higher (If both windows show none, then imagine that they are higher)?

6) Select the latest arx_yymmdd_ngm_avn from the NGM_MOS selection window. Now load the best category of ceiling height. What is it for the 24hr forecast at LSE?

7) Load the probability of overcast clouds. What is the 24hr forecast probability of overcast clouds for LSE? Now imagine that it is 70%. Do you think the probability of scattered would be greater or less than this 70%?

8) You are on the evening shift and there has been an OVC130 deck across the entire northern portion of the country (including our forecast area). Since there are several ASOS sites across this same area, what do think the impact would be on the clouds and ceiling height forecast in the new 00z MOS forecast for this area?

9) Can you list 2 MOS Biases for the local area? (Reference the SDM pp 1.3.12)

10) Briefly explain why you would not expect MOS to forecast “big” or “extreme” events.

Thank you for your time. Feel free to comment below.