Western Region Technical Attachment
No. 07-08
July 13, 2007

Post-Processing MOS Using a Moving Average Approach

By Donald Britton
NOAA NWS Great Falls, MT

1. Abstract

Numerical guidance, such as the Model Output Statistics (MOS), has proven useful in the preparation of forecasts, particularly in the extended period. Some studies (Baars & Mass, 2005) have shown that consensus forecasts, comprising a variety of numerical guidance, outperform a single guidance product. Currently, though, only the MOS derived from the Global Forecast System (GFS) model extends throughout the forecast period covered by WFO forecasters. Additionally, some forecasters criticize the GFS-based MOS (MEX) guidance for its tendency to fluctuate from one model run to the next, leading to decreased confidence on the part of forecasters. Algorithmic methods for reducing signal fluctuations might be applied to the MEX to improve forecaster confidence. Widely used in network, IT security, and financial analysis to dampen signal fluctuations, an exponentially weighted moving average (EWMA) reduces oscillations and applies more weight to recent measurements than to older measurements (Winters 1960, Ye 2003, Xi 2006). Initial results indicate that guidance processed using the EWMA outperforms the raw MEX guidance, and increases the number of forecasts that verify within 3°F and 5°F.

2. Introduction

Algorithmic methods for reducing signal fluctuations might be applied to the MEX to improve forecaster confidence. For example, a moving average can be used to dampen signal fluctuations. Such a moving average approach has both advantages and disadvantages: the averaging process will reduce the amplitude of forecast changes, which will result in a forecast that "lags" behind MEX trends. This lag will likely be reflected in a degradation of the Mean Absolute Error (MAE), but this may not be significant. Brooks and Doswell have described the inherent weaknesses of MAE, and suggest that a distributions-oriented approach tallying the number of forecasts that fall within some range centered upon the verification temperature may better illustrate the significance of forecast errors (Brooks and Doswell, 1996).

A simple moving average regards each point equally in calculating the average (Nau, 2006), and would not give greater weight to the most recent model run. This could lead to averages that may be too slow to pick up on modeled forecast trends. Better than a simple moving average would be to grant a greater weight to more recent forecasts to reduce lag.

Widely used in signal processing to reduce noise, and in the financial industry to smooth fluctuating market prices, the exponentially weighted moving average (EWMA) reduces lag by applying more weight to recent measurements than to older measurements. The EWMA might be a useful algorithm to solve the forecast problem at issue, since it can react more quickly to changes than a simple moving average.

3. Methodology and Discussion

To test the usefulness of applying a moving average to the MEX temperature guidance, thirty locations around the nation for which GFS MOS guidance is available (Figure 6), were selected to reflect as many climatological regimes as possible. Since atmospheric models are periodically adjusted, it was decided to limit the scope of this study to the calendar year 2005. This would reduce the effect of model changes on the guidance data, and would also capture the effects of seasonal parameters. Forecasts for each location were evaluated for the first 13 forecast periods available from the MEX guidance.

The GFS MOS products for 2005 were retrieved from National Weather Service's Meteorological Development Lab. At the time of the study, only GFS MOS data from the 00Z model cycle was available for the entire year, so only 00Z data was used in this study. A Perl program was written to separate the data by location, parse the data for each forecast period, and place data for each location into separate files in preparation for importing into a spreadsheet. The data were also adjusted so that all the forecasts for a given valid time appeared on a single line in the spreadsheet. A small amount of MEX data for the selected locations were missing (<1%). Data from other model cycles, previous or subsequent, for the same valid time, were substituted for any missing data, which had the effect of slightly enhancing guidance run-to-run consistency.

Observation data was collected from the XMACIS (Owen, 2005) online data repository (XMACIS, 2006) and a Perl script extracted the maximum and minimum temperature data, which was subsequently imported into the appropriate spreadsheets.

Criteria considered were the number of forecasts that exhibited flip-flopping (arbitrarily defined here as a 5°F change for a given forecast valid time followed by another 5°F change in the opposite direction), MAE, and the number of forecasts that verified within 5°F of the observed value. In an attempt to capture significant forecasts, the MAE was further broken down into cases where the 24 hour temperature change was 10°F or greater (similar to some existing National Weather Service forecast verification schemes).

A formula was developed from this information to derive the EWMA. Each forecast valid time was separated by 24 hours. The formula for each time step looks like this:

$$EWMA_t = (MEX_t - EWMA_{t+24})\, \alpha + EWMA_{t+24}$$

Where:

- t is the current time step
- t+24 is the previous time step
- α is the weighting factor defined as: $\alpha = 2 / (n + 1)$
- n is the number of averaged periods

Initially, the EWMA 192 and 180 hour periods must equal the raw MEX values for these periods since there are no previous values with which to average them. Therefore, the moving average process begins with the forecast at 168 hours. The EWMA forecast is calculated by subtracting the value of the previous EWMA forecast period from the value of the current MEX forecast period, multiplying by the alpha value, then adding the value of the previous EWMA forecast period.

As can be seen, the values of α and n are inversely proportional. Lower values of α will effectively average forecast values over a greater period of time. If the goal is to reduce so-called flip-flopping, averaging values over a greater number of forecasts would seem to be a more successful approach. To reduce flip-flopping *and* increase the number of forecasts within 5°F, though, the value for α must be carefully selected. It is likely that the optimal value for α will vary from location to location. For simplicity's sake and the purpose of the current study, a value that optimizes both goals over all 30 stations is most desirable.

The results included here are based upon $\alpha = 0.66$, which corresponds to a 2-period (i.e., $n = 2$) weighted moving average. Other values for n of 1.5, 2.5, and 3 were tested empirically, but yielded inferior results as determined by both MAE and the number of forecasts verifying within 5°F.

4. Conclusions

Several analyses were performed on the data. Since forecasters expressed concern about how run to run consistency affects the MOS guidance, the so-called "flip-flop" was analyzed. As stated, a flip-flop was defined as a change of 5°F or more, followed by another change of 5°F or more in the opposite direction. A quick examination of the verification data revealed that the MEX guidance may exhibit less run-to-run "flip-flopping" than forecasters believe. Of the 30 locations included in this study, guidance flip-flops of 5°F or greater only affect 10% or more of forecasts at three locations: Bismarck, ND, Kansas City, MO, and Rapid City, SD.

The results show that employing an EWMA scheme reduces fluctuations (flip-flops) while improving verification scores. The results averaged across all 30 locations indicate:

- a reduction in the percentage of forecasts that exhibit flip-flopping (Table 1 and Figure 1)
- MAE suffered most on maximum temperatures, but on average, the difference in MAE between the raw MEX and the EWMA was less than 0.05°F (Table 2, Figure 2)
- averaged MEX forecasts outperformed the raw MEX when the 24hr temperature change was 10°F or greater (Table 3, Figure 3)
- the number of forecasts that fell within 5°F of the observed temperature increased by 3.9% (Table 4, Figure 4) for all locations and forecast periods
- the number of forecasts that fell within 3°F of the observed temperature increased by nearly 6.3% (Table 5, Figure 5) for all locations and forecast periods

Using the information gathered in this study, the weighted moving average approach was applied in real time to forecasts at the Great Falls WFO, in July 2006. An exponentially-weighted moving average of the MEX guidance from the 00Z model cycle was created every evening. In addition to creating a text product available for viewing by forecasters, both the EWMA and MEX data were mapped to IFPS gridded forecast databases (ADJWMA and ADJMEX) using the MatchGuidanceAll SmartTool (Barker, 2005), along with a variety of other numerical guidance products.

To further examine the value of the EWMA approach, the BOIVeify program (Barker, 2006) was used. Using the gridded databases created by the MatchGuidanceAll SmartTool, BOIVerify verifies each grid point, rather than only the traditional MOS locations. BOIVerify thus provides a more complete picture of the accuracy of the digital temperature forecast in data sparse areas. A look at verification statistics for the Great Falls CWA for the months of July though September, 2006 also show the ADJWMA improving over the ADJMEX guidance. The number of EWMA maximum and minimum forecasts falling within both 5°F (Figure 7) and 3°F (Figure 8) was greater than the number for the unprocessed GFS MEX guidance, illustrating the potential of the moving average approach.

The most recent version of BOIVerify (v1.0) can perform bias-corrected calculations on gridded forecast guidance. A study is underway to determine the EWMA's performance during the summer of 2007 using a bias-corrected approach.

5. Acknowledgements

The author would like to thank David Bernhardt, Steve Brueske, Gina Loss, and Dan Reilly of the National Weather Service Forecast Office in Great Falls for their advice, support, and critical thinking.

For the tables below, orange indicates maximum temperatures and blue minimum temperatures.

Flip-Flop											
Fcst Period	72-24	84-36	96-48	108-60	120-72	132-84	144-96	156-108	168-120	180-156	192-144
Raw MEX	1.1%	0.6%	2.1%	1.1%	3.3%	1.9%	4.4%	3.0%	4.9%	3.4%	5.5%
EWMA MEX	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	0.2%	0.5%

Table 1 - Comparison of 5°F flip-flop between Raw MEX and EWMA MEX. The columns reflect the forecast hours over which a flip-flop occurs. For example, the first column shows the percentage of time the forecast flips from 72-48 hours, then flops from 48-24 hours.

Mean Absolute Error													
Fcst Hour	24	36	48	60	72	84	96	108	120	132	144	156	168
Raw MEX	2.67	3.03	3.12	3.22	3.53	3.44	4.00	3.76	4.55	4.07	5.04	4.47	5.60
EWMA MEX	2.70	3.01	3.14	3.20	3.54	3.43	4.04	3.74	4.57	4.08	5.08	4.47	5.63

Table 2 - Comparison of MAE between Raw MEX and EWMA MEX (highlighted cells indicate degradation of raw MEX guidance). Note that the on average, the EWMA performs better vs. the raw MEX on minimum temperatures than on maximum temperatures. In any case, the difference in MAE between the two guidance products is arguably negligible.

When 24 Temp Change >=10°F													
Fcst Hour	24	36	48	60	72	84	96	108	120	132	144	156	168
Raw MEX	4.10	7.12	4.59	7.06	5.15	7.15	5.88	7.43	6.93	7.29	7.60	7.34	8.25
EWMA MEX	3.76	6.90	4.28	6.83	4.97	6.92	5.83	7.05	6.74	6.94	7.45	7.12	8.06

Table 3 - Comparison of MAE when 24 hr temperature change is 10°F or greater

% Forecasts within 5F													
Fcst Hour	24	36	48	60	72	84	96	108	120	132	144	156	168
Raw MEX	82.7%	77.8%	76.4%	75.5%	71.6%	72.4%	66.3%	68.5%	59.8%	65.2%	55.8%	61.1%	51.9%
EWMA MEX	86.2%	82.4%	80.7%	79.9%	75.5%	76.6%	70.0%	72.8%	63.9%	69.1%	59.2%	64.4%	55.0%

Table 4 - Comparison of percentage of forecasts within 5°F of observed between Raw MEX and EWMA MEX

% Forecasts within 3F													
Fcst Hour	24	36	48	60	72	84	96	108	120	132	144	156	168
Raw MEX	59.6%	55.5%	52.5%	52.8%	48.3%	49.9%	43.7%	46.7%	39.3%	43.8%	36.2%	40.2%	32.8%
EWMA MEX	67.2%	63.5%	60.1%	60.3%	55.3%	57.0%	49.6%	53.2%	45.1%	49.6%	40.9%	45.1%	36.5%

Table 5 - Comparison of percentage of forecasts within 3°F of observed between Raw MEX and EWMA MEX

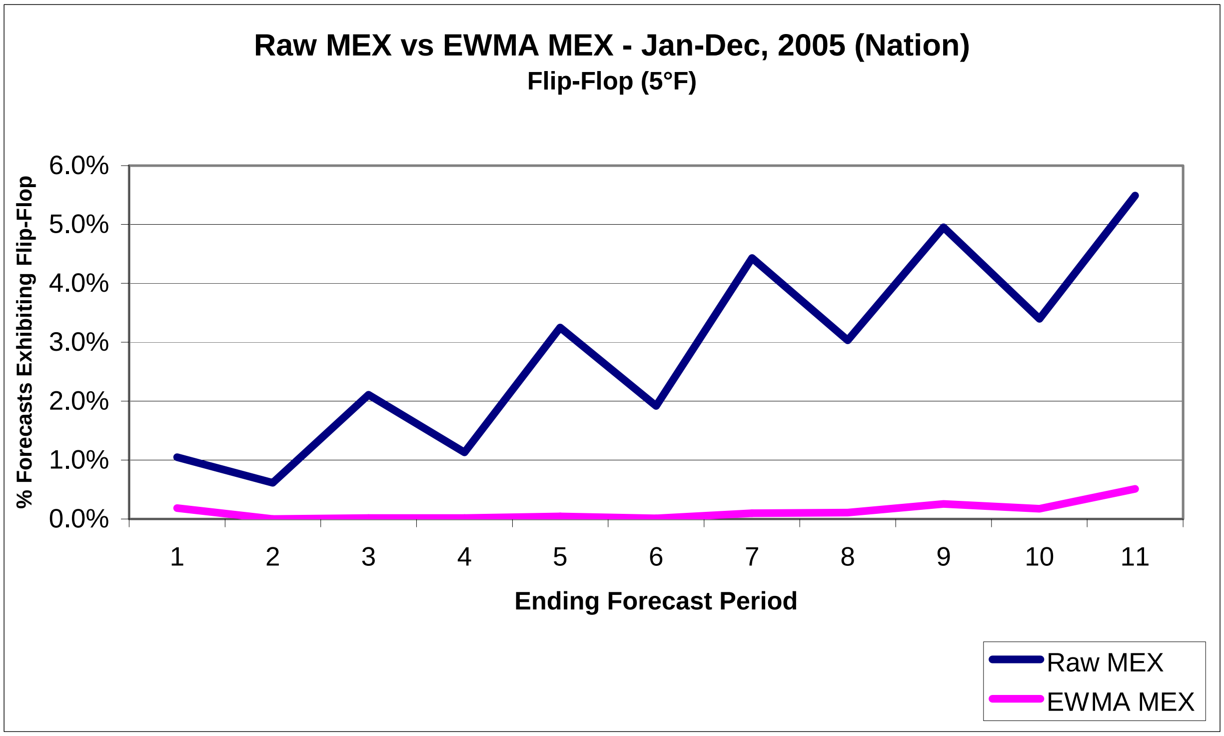


Figure 1 - Comparison of 5°F flip-flop between Raw MEX and EWMA MEX

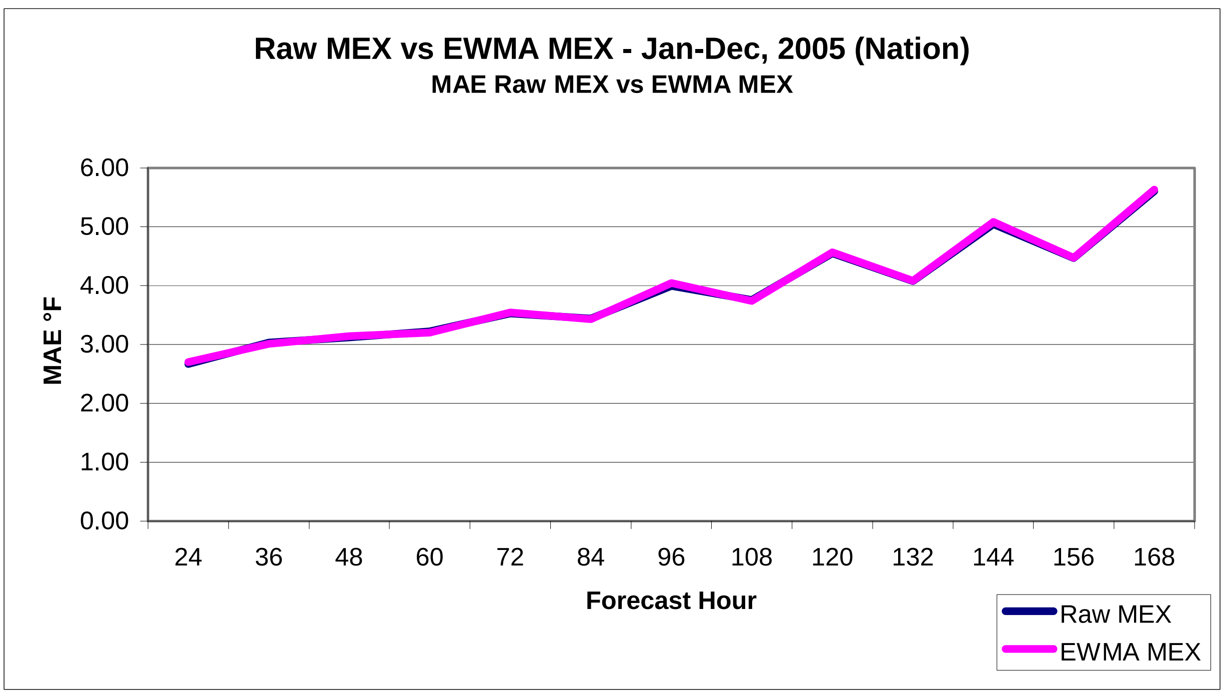


Figure 2 - Comparison of MAE between Raw MEX and EWMA MEX

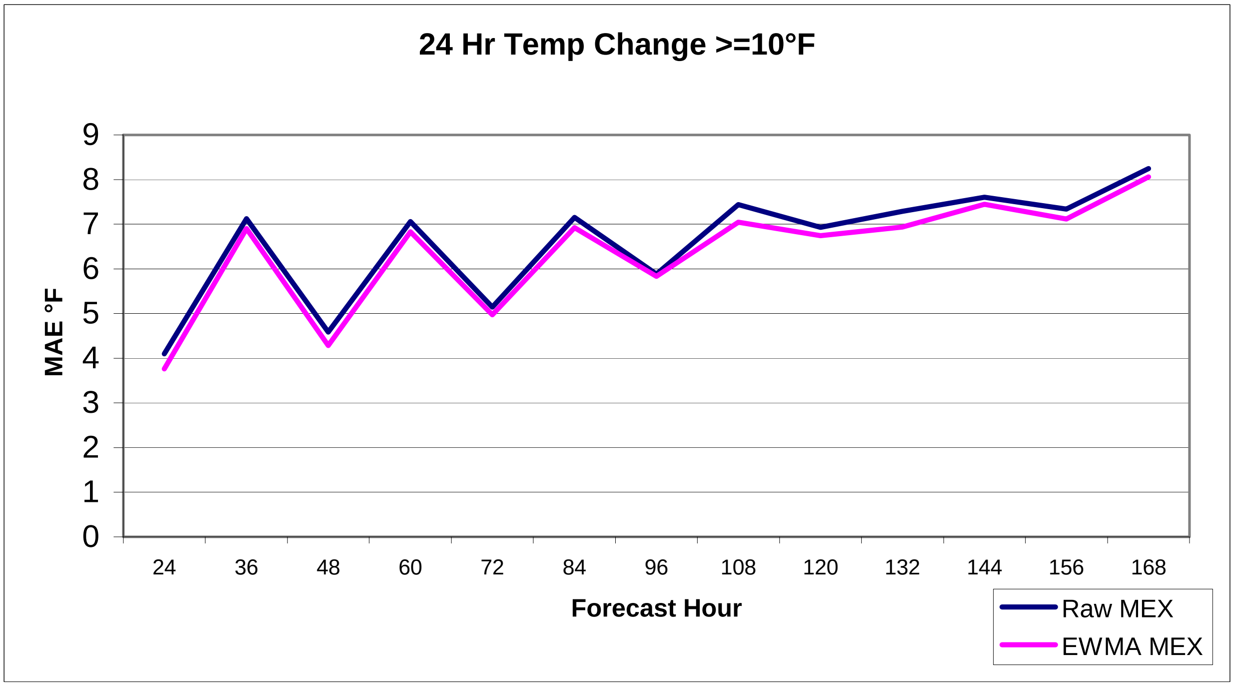


Figure 3 - Comparison of MAE when 24 hr temperature change is 10°F or greater

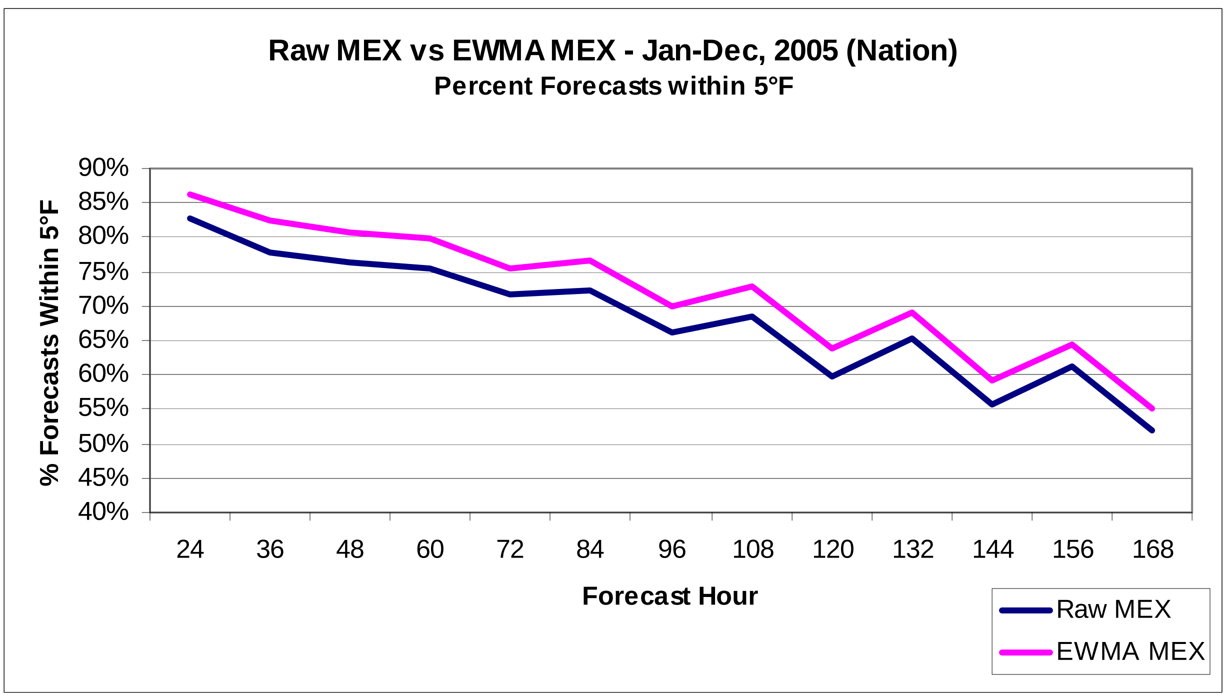


Figure 4 - Comparison of percentage of forecasts within 5°F of observed between Raw MEX and EWMA MEX

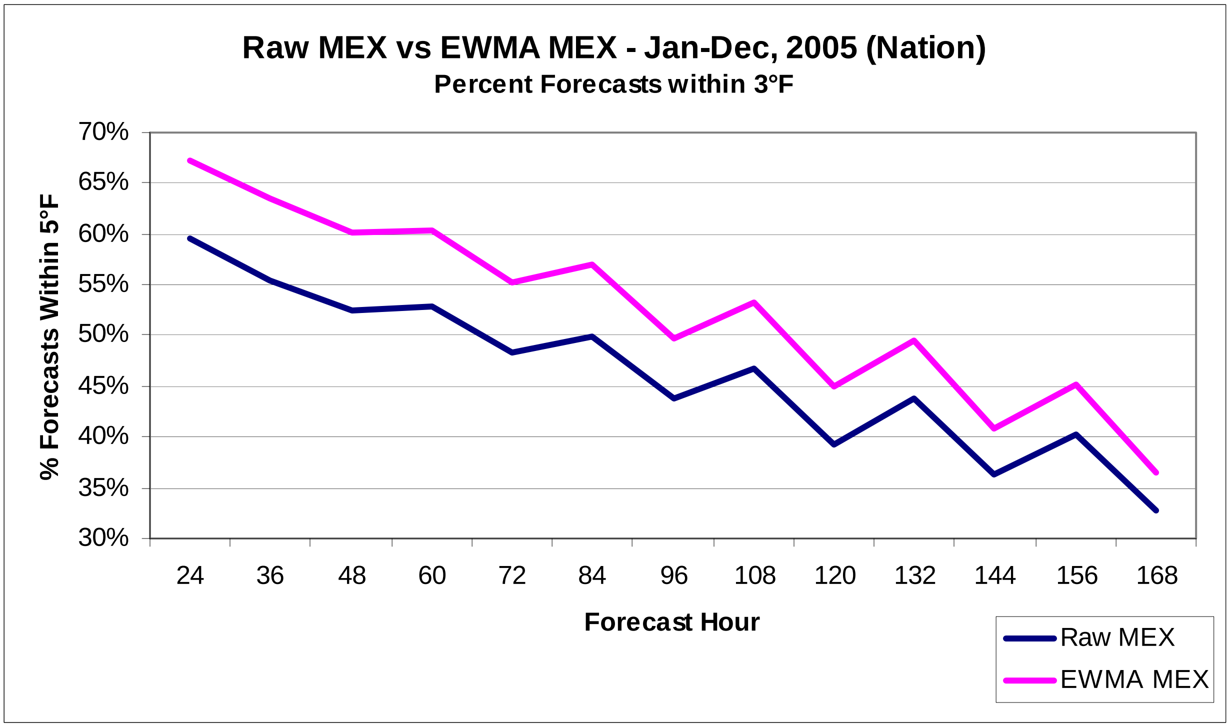
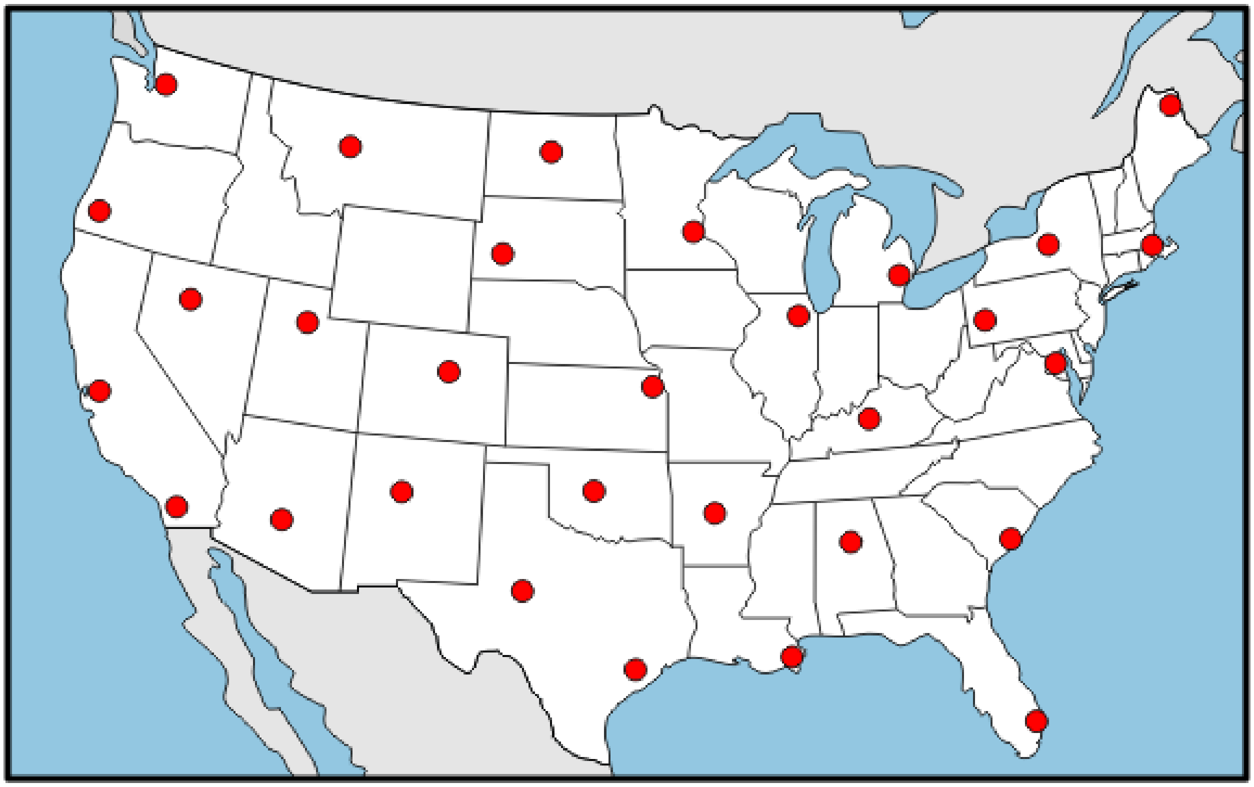


Figure 5 - Comparison of percentage of forecasts within 3°F of observed between Raw MEX and EWMA MEX



Albuquerque, NM	Great Falls, MT	Phoenix, AZ
Baltimore, MD	Houston, TX	Pittsburgh, PA
Birmingham, AL	Kansas City, MO	Rapid City, SD
Bismarck, ND	Little Rock, AR	Rochester, NY
Boston, MA	Louisville, KY	Salt Lake City, UT
Caribou, ME	Medford, OR	San Angelo, TX
Charleston, SC	Miami, FL	San Diego, CA
Chicago, IL	Minneapolis, MN	San Francisco, CA
Denver, CO	New Orleans, LA	Seattle, WA
Detroit, MI	Oklahoma City, OK	Winnemucca, NV

Figure 6 – MOS sites used for this study, 30 locations distributed fairly evenly across U.S.

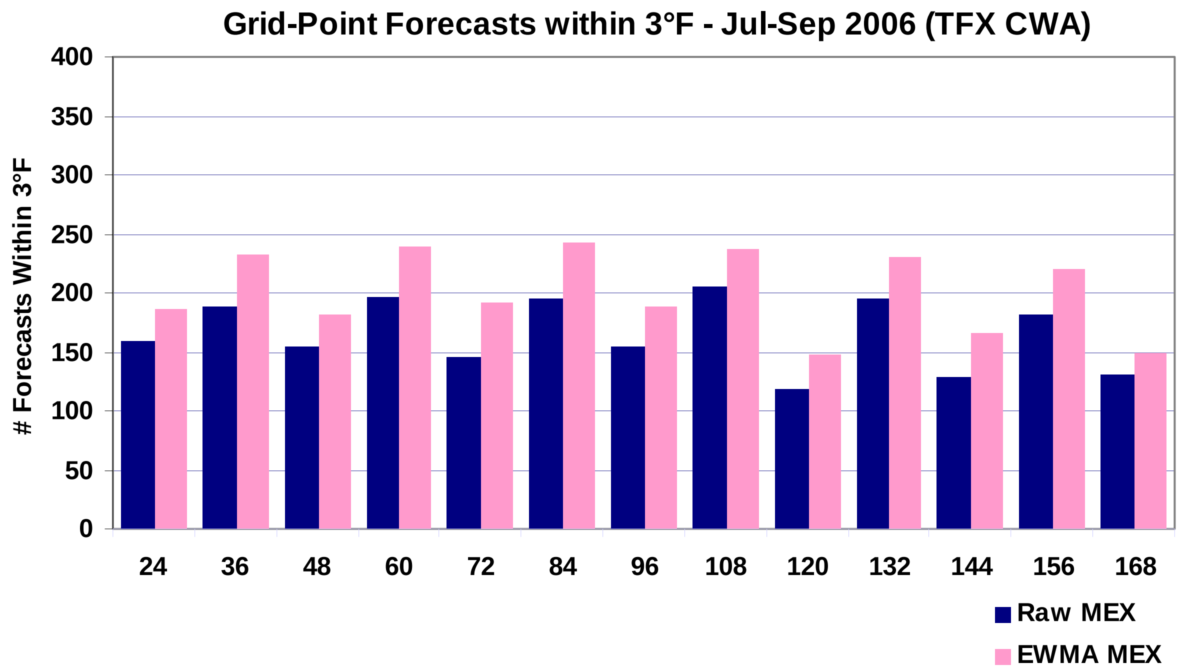


Figure 7 – Number of grid-point forecasts within 5°F for Great Falls CWA, July through September, 2006 Verification values obtained from BOIVerify (Total number of grid-points in Great Falls CWA = 6016)

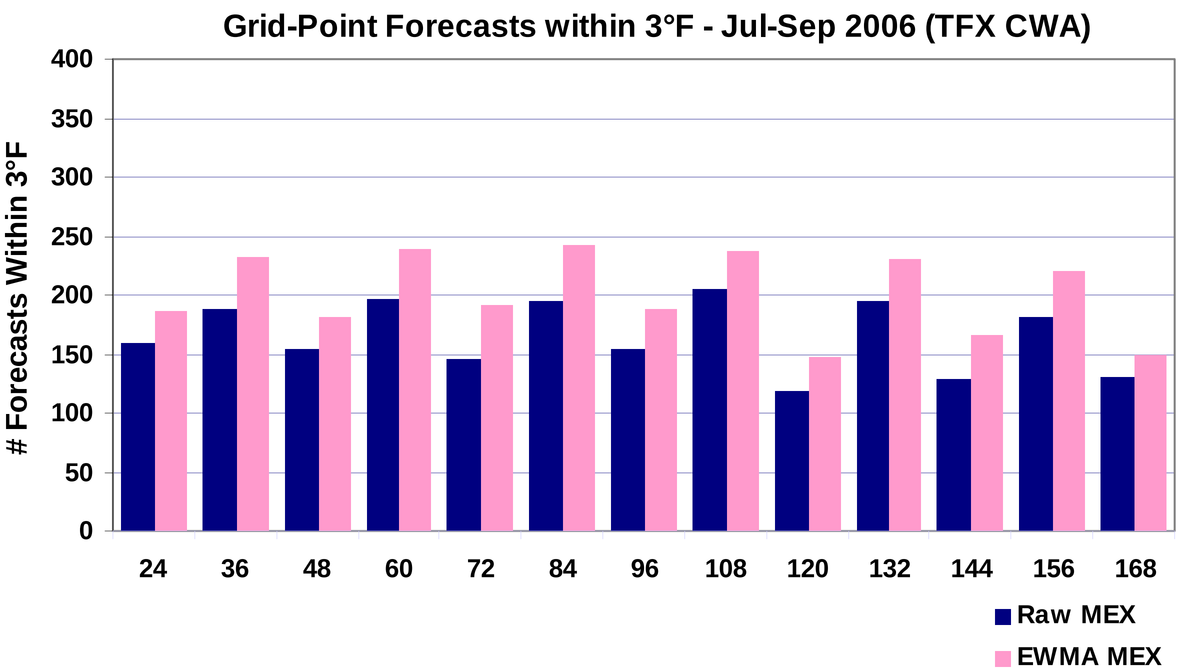


Figure 8 - Number of gridded forecasts within 3°F for Great Falls CWA, July through September, 2006 Verification values obtained from BOIVerify. (Total number of grid-points in Great Falls CWA = 6016)

References:

Baars, J. A. and C. F. Mass, 2005: Performance of National Weather Service forecasts compared to operational, consensus, and weighted Model Output Statistics. Wea. Forecasting, 20, 1034-1047.

Barker, T., 2006. BOIVerify [computer software]. Boise: National Weather Service.

Barker, T. 2005. MatchGuidanceAll [computer software]. Boise: National Weather Service.

Brooks, H. E., and C. A. Doswell, 1996: A comparison of measures-oriented and distributions-oriented approaches to forecast verification. Wea. Forecasting, 11, 288–303.

Meteorological Development Laboratory. GFSMEX MOS Archives. Retrieved January 28, 2006, from http://www.mdl.nws.noaa.gov/%7Emos/archives/mex.html.

Nau, R. 2006. Averaging and Exponential Smoothing Models. Retrieved January 29, 2007, from Duke University web site: http://www.duke.edu/~rnau/411avg.htm.

Owen, T.W, 2005. Cyberinfrastructure advances at NOAA's National Climatic Data Center and regional climate centers: Quality feedback, assurance, and dissemination. Proceedings of the 21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology. Retrieved February 6, 2007, from http://ams.confex.com/ams/pdfpapers/84170.pdf.

Winters, P.R., 1960. Forecasting Sales by Exponentially Weighted Moving Averages. Management Science, Vol. 6, No. 3. (Apr., 1960), pp. 324-342. Retrieved June 4, 2007 from http://links.jstor.org/sici?sici=0025-1909(196004)6%3A3%3C324%3AFSBEWM%3E2.0.CO%3B2-T

Xi, B. W., Michailidis, G. and Nair, V. (2006) "Estimating Network Loss Rates Using Active Tomography," JASA, T&M, 1430-1449. Retrieved June 4, 2007 from http://www.stat.lsa.umich.edu/~gmichail/technometrics_data_nets.pdf

XMACIS, 2006: Applied Climate Information System. Retrieved January 28, 2006, from http://www.rcc-acis.org/.

Ye, N. Computer Intrusion Detection Through EWMA for Autocorrelated and Uncorrelated Data. IEEE Transactions on Reliability, Vol. 52, No. 1, March 2003. Retrieved June 4, 2007 from http://enpub.fulton.asu.edu/ye/Published_Journal_Papers/Ye_41.pdf