SRH SSD 2006-5
11/2006

Technical Attachment

A Forecast Improvement Project at WFO Austin/San Antonio

G. Clay Anderson and Jon W. Zeitler
WFO Austin/San Antonio, Texas

1. Introduction

During the period June 1, 2005 to May 31, 2006, verification scores for maximum/ minimum temperatures (Max/MinT) and probability of precipitation (PoP) at WFO Austin/San Antonio showed significant improvement over previous years relative to the Global Forecast System Model Output Statistics (GFS MOS). In addition, forecasters at WFO Austin/San Antonio spent 30%-50% less time editing grids than during the previous year. These improvements were the direct result of changes implemented to local forecast operations. This paper will discuss the details of the project so similar results might be duplicated at other WFOs.

2. Background

The need to streamline the forecast process at WFO Austin/San Antonio became obvious once the Point Forecast Matrix (PFM) "Stats On Demand" verification data (Table 1) were made available in 2004, especially given the amount of time forecasters were spending performing edits in the Graphical Forecast Editor (GFE).

Month	MaxT	MinT	PoP
Jan 2004	-2.1	-7	-16.6
Feb	0.3	-5.6	-1.3
Mar	14.7	4	-9.1
Apr	-1.4	-5.6	-3.9
May	-2.9	-6.9	-6.4
Jun	-4	-4.1	5.7
Jul	9.4	0.4	0.9
Aug	-9.6	-8	-3.1
Sep	5.7	-17.6	0.7
Oct	14.3	1.6	-0.1
Nov	-4.5	-7.7	-8.5
Dec	4.4	3.8	-6.9
Jan 2005	-3.6	8.9	-10.9
Feb	-2.4	1	-2.3
Mar	-3.9	-2.8	-6.9
January 2004 – March 2005	**1**	**- 2.5**	**- 3.9**

Table 1: WFO Austin/San Antonio percent improvement relative to GFS MOS for all Point Forecast Matrix sites, all forecast periods, and all model cycles for January 2004-March 2005. Negative numbers indicate the MOS guidance was better than the official forecast.

Forecasters at WFO Austin/San Antonio spent an average of three to four hours editing grids every major model cycle (0000 and 1200 UTC), for a combined average of six to eight hours each day. This time investment resulted in little improvement, or more often detriment, to the forecast as compared to the GFS MOS (Table 1).

These verification statistics support the results of by Baars and Mass (2005) that "...it is getting increasingly difficult for human forecasters to improve upon MOS..." and "Humans cannot consistently beat MOS...and are only superior to MOS for short-term temperature forecasts..." Furthermore, Mass (2003) adds "...use of human beings to laboriously alter deterministic forecasts for a week into the future would be a serious mistake that would lessen forecasters' time for more productive work." Indeed, prior to the project, had the forecasters used the GFS MOS for the forecast and only deviated from the MOS guidance during 1st period "targets of opportunity" they would have made substantial improvements over the GFS MOS, while spending very little time editing grids.

There are many reasons why forecasters have not improved on the GFS MOS in recent years. Maglaras (2004) cites the introduction of new technology and changes to procedures in forecast operations, but notes the most important reason was a flawed methodology, whereby forecasters used a previous gridded forecast as the starting point for a new forecast cycle, only making modifications when "significant" changes were deemed necessary. Others (Entremont and Gagan, personal communication) have recognized flaws in this methodology by noting that forecast errors are carried from one cycle to the next, while MOS errors generally become smaller with successive cycles as a particular forecast period approaches 0-12 hours from the forecast issuance. In addition, the transition to digital grid forecasting and significant increase in number of verification points likely had a negative impact on forecast performance. Regardless of those reasons, forecasters should consistently beat model guidance, or risk losing involvement in the forecast process.

The staff at WFO Austin/San Antonio chose to seize this opportunity as a challenge to improve. The hypothesis was, forecasters *can* consistently show improvement over the GFS MOS, if only given a method for success. That method was developed over a period of two years and was inspired by the success documented by Maglaras (2004) and the suggestions of Mass (2003). The project offered a solution to the problems of under-performance relative to the GFS MOS, and excessive time spent manipulating grids in the GFE. The project was implemented several months before the NWS Southern Region Grid Preparation Policy was disseminated in August 2005.

3. The Forecast Improvement Project

The crux of the project methodology was to populate the forecast database with GFS MOS guidance from the 0000 and 1200 UTC model cycles and make adjustments only where the forecaster saw "targets of opportunity". This was accomplished using a locally created procedure within the GFE, appropriately named "Populate_GFS(X)_MOS", which would populate forecast periods 1-14 with locally created GFS MOS-adjusted model grids (i.e., ADJMAV, ADJMEX). Although the focus of this project was on improving Max/MinT and PoP verification, the grid populatation procedure also loaded ADJMAV/ADJMEX for temperature (T), dew point temperature (Td), wind, and sky elements (Figure 1).

In addition to running the Populate_GFS(X)_MOS procedure, forecasters were asked to fill out a project worksheet (Figure 2) on which they recorded grid preparation time and significant departures from the GFS MOS based on locally defined thresholds. Small deviations from guidance did not require documentation, only departures larger than one half the National Digital Forecast Database collaboration thresholds (NWS, 2004). Theia documentation requirement was in anticipation of defining the need for collaboration with neighboring WFOs. Forecasters could also provide comments on the worksheet that might be used by forecasters on subsequent shifts to identify model or MOS trends and biases. This worksheet was not intended to be burdensome, but instead promote decision-making accountability and forecast confidence.

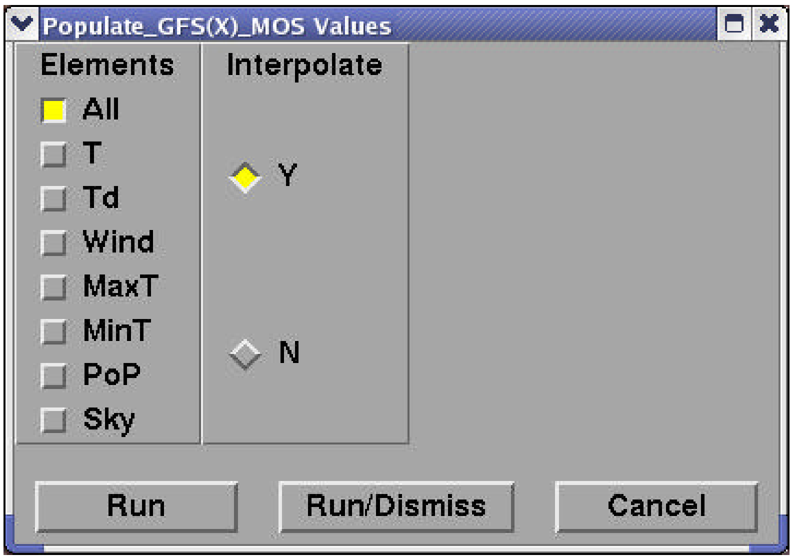

Figure 1: Locally-developed Populate_GFS(X)_MOS Graphical User Interface.

Each month during the project, verification statistics were obtained from the PFM Stats-on-Demand Web site and shared with forecasters via informal reports. These reports highlighted areas of forecast improvement and offered tips for additional gains based on a thorough analysis of the data. The GFS MOS and individual forecaster performance were tracked using an AWIPS local application (Cook, 2004). The output was made available to forecasters as a text product in AWIPS and was used to highlight recent GFS MOS biases in occasional e-mails to forecasters. All of this information, including project-related documentation, was made available to forecasters on a local Intranet page.

4. Results

At WFO Austin/San Antonio 11 paired PFM/MOS sites are used for Max/MinT verification, and eight for PoP. The project yielded positive results as forecasts substantially improved (Table 2). Figures 3 through 5 show monthly improvement over the GFS MOS for MaxT, MinT, and PoP, including smoothed lines to indicate trends obscured by month-to-month variability. (Data from April and May 2005 were not included in either period since that was a transition period when two of the ten forecasters prototyped the methodology.) Note WFO Austin/San Antonio improved over the GFS MOS in nearly all forecast periods. In addition to the improvement in verification scores, grid preparation time was reduced from three to four hours per major model cycle, to less than two hours; a time savings of 30-50%.

Project Results

	MaxT	MinT	PoP
Pre-Project *Jun 2004 – Mar 2005*	+1.0	-2.5	-3.9
Project *Jun 2005 – May 2006*	+6.1	+4.8	+1.1
CONUS WFOs *Jun 2005 – May 2006*	-0.9	-4.0	-4.3
Difference *(Project – Pre-Project)*	**+5.1**	**+7.3**	**+5.0**

Table 2: Forecast verification scores (percent improvement over GFS MOS guidance) for WFO Austin/San Antonio during the pre-project and project periods, the difference between the two, and the verification scores for all CONUS WFOs combined during the project period included as a benchmark.

5. Concluding Remarks

In an era of increasing atmospheric modeling skill and technological efficiencies it is difficult, but vitally important, to add value to the available model guidance. Doing that efficiently demonstrates forecaster value in the human-machine mix and allows more time for analysis/diagnosis and critical decision-making in the forecast process. Additional time savings can be used for enhanced support/services, outreach, professional development, and applied research studies.

The results of the project demonstrate that the prescribed methodology works. Successes following similar methodologies have also been demonstrated at WFO Albany, NY, (Maglaras, 2004) and WFO Jackson, MS, (Entremont and Gagan, personal communication). The methodology should be appealing to WFOs that are currently not adding value to the model guidance, or are spending an inordinate amount of time editing grids for little improvement over the guidance. Judging from the CONUS WFOs verification results in Table 2, the methodology might be of use to many WFOs.

Further improvements might be achieved through meaningful, individual forecaster verification, coupled with a policy to improve consistently poor forecast decision-making. Local studies should to be conducted to determine when GFS MOS performs poorly so forecasters can easily identify and capitalize on “targets of opportunity.”

6. Acknowledgements

We thank the forecasters at WFO Austin/San Antonio for following the prescribed methodology and adeptly adjusting to never-ending changes in forecast operations. We also thank Joe Arellano, Meteorologist-in-Charge at WFO Austin/San Antonio, for his encouragement and unfailing support. Bernard Meisner, Southern Region Headquarters Scientific Services Division, provided a thorough and helpful review. Kurt Vanspeybroeck, Science and Operations Officer at WFO Brownsville, played a key role in development of the project methodology.

7. 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.

Cook, K. R., 2004: AWIPSVerify. AWIPS Local Application Database, NOAA/NWS. [Available online at: http://140.90.90.253/~applications/LAD/index.php3]

Maglaras, G. J., 2004: Improving temperature verification results within the IFPS/GFE framework. Eastern Region Technical Attachment No. 2004-01, 10 pp. [Available online at http://www.erh.noaa.gov/er/hq/ssd/erps/ta/ta2004-01.pdf]

Mass, C. F., 2003: IFPS and the future of the National Weather Service. *Wea. Forecasting,* **18,** 75-79.

Young, D., 2004: *Digital Data Products/Services Specification – Appendix A*. National Weather Service Instruction, 10-506, NWS Directives System, 25 pp. [Available online at: http://www.weather.gov/directives/sym/pd01005006curr.pdf]

Project Worksheet

For use in verification documentation and as a guide for collaboration.

Element	Deviation From MOS Documentation/Collaboration Threshold
Max/Min T	> +/- 2ºF
Td	> +/- 2ºF
PoP	> +/- 10%
Wind	> +/- 30º for Direction, > +/- 5 kts in speed

Weather Element Deviation from GFS(X) MOS by Period

Period	Max T	Min T	PoP	Td	Wind
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					

Forecaster: *Shift/Date:* *Total Grid Preparation Time:*
Comments:

Figure 2: Project Worksheet for tracking grid prep time and deviations from GFS MOS.

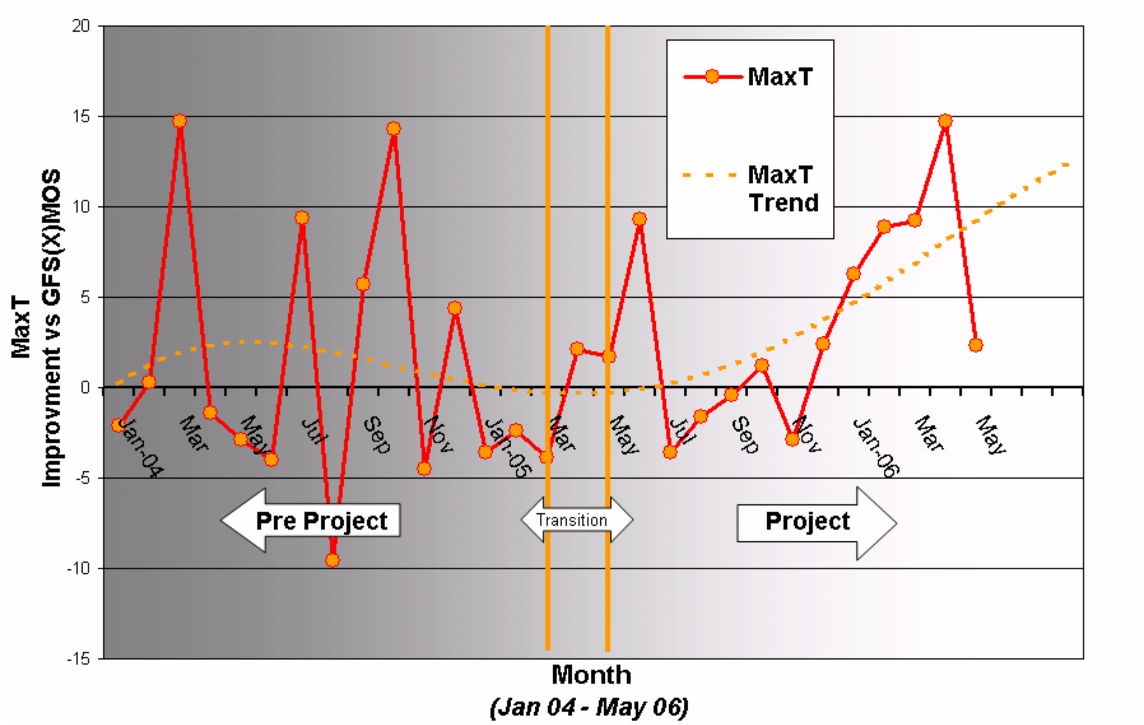

Figure 3: Maximum temperature verification relative to GFS MOS for 11 Point Forecast Matrix sites with all periods combined.

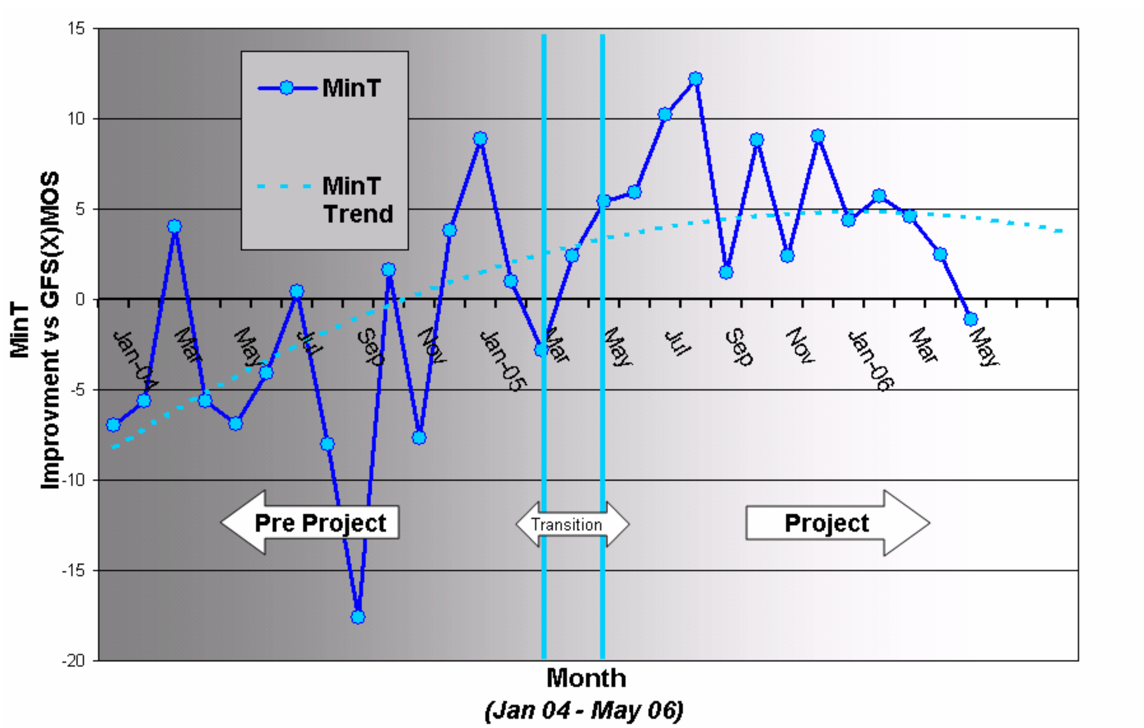

Figure 4: Same as Figure 3 except for minimum temperature.

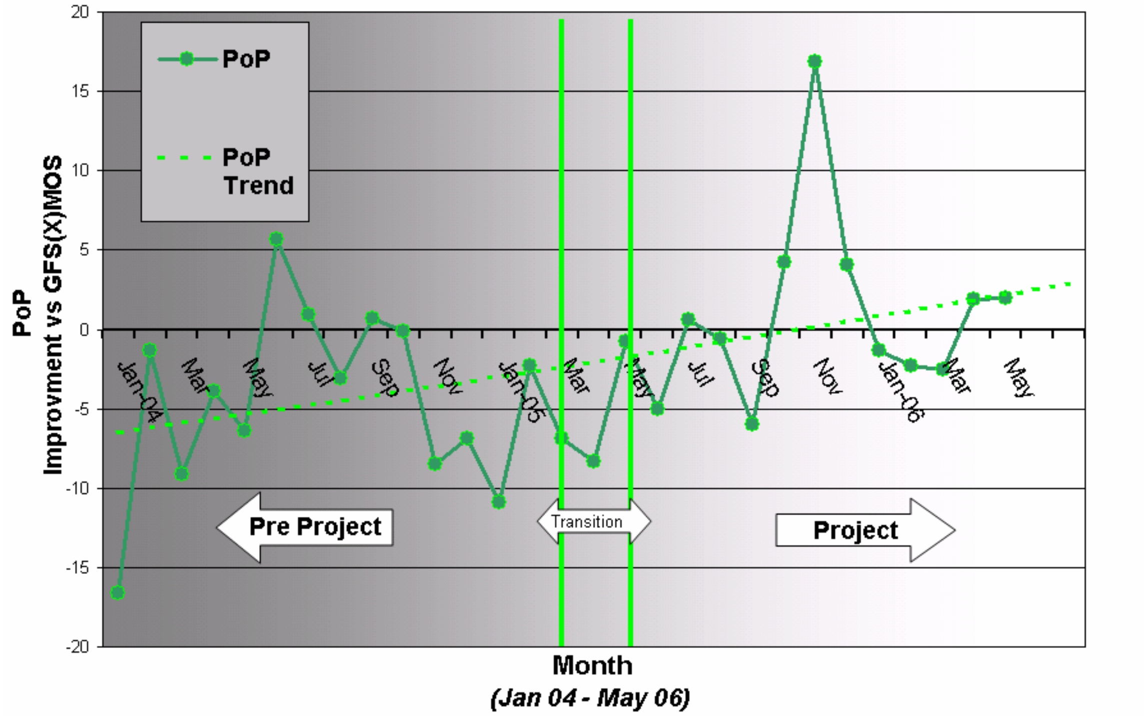

Figure 5: Probability of Precipitation verification relative to GFS MOS for eight Point Forecast Matrix sites with all periods combined.