Forecast Methodologies that Improved Probability of Precipitation Forecasts at WFO Billings

Don Moore
Brian Tesar
Joe Lester
William Rasch
WFO Billings, Montana

I. Introduction

Through the 2005-2006 winter, awareness of a dry forecast bias was raised across NWS Western Region (WR) forecast offices. A local study at WFO Billings (BYZ) validated WFO Billings also exhibited a dry bias for multiple years. A large part of the problem was zero probability of precipitation (PoP) being forecast too often in medium range forecasts. Another problem was PoP forecasts were often lower than the GFS MOS or MEX at the medium range when measurable precipitation was observed. So in many cases when the weather pattern was forecast to be more favorable for precipitation than climatology and the MEX provided a PoP near climatology, a zero forecast (or near zero) PoP was still provided because of lower forecaster confidence. This lower confidence was often a result of model run to run discontinuity and/or differing solutions among numerical models.

To help alleviate the regional dry PoP bias, a team was formed in WR that developed climatology PoP grids for use in the graphical forecast editor (GFE) by fall 2006 (Figure 1). Training was also provided to WR forecasters on the best utilization of PoP climatology in the forecast process. Locally at BYZ, a study was completed by fall 2006 that found that the GFS MOS often exhibited a dry bias when the GFS forecast quantitative precipitation. BYZ had an even greater drier bias in these situations.

Since a few cool or wet seasons have passed since awareness of these dry forecasts, the authors thought it would be enlightening to perform a comparison of PoP verification between the 2004-2006 cool seasons and 2006-2008 cool seasons. This paper summarizes the details of the study performed at BYZ as well as PoP verification results from the two aforementioned time periods. The verification results will focus only on the following:

1) The wetter months October to January; February to April was excluded since those months were not available in 2008 at the time this paper was written
2) 12 hour PoP through seven days
3) GFS MOS points Sheridan, WY (SHR) and Billings, MT (BIL), since All Weather Precipitation Accumulation Gauges (AWPAG) were installed at these sites in 2005, allowing for more accurate winter precipitation measurements.

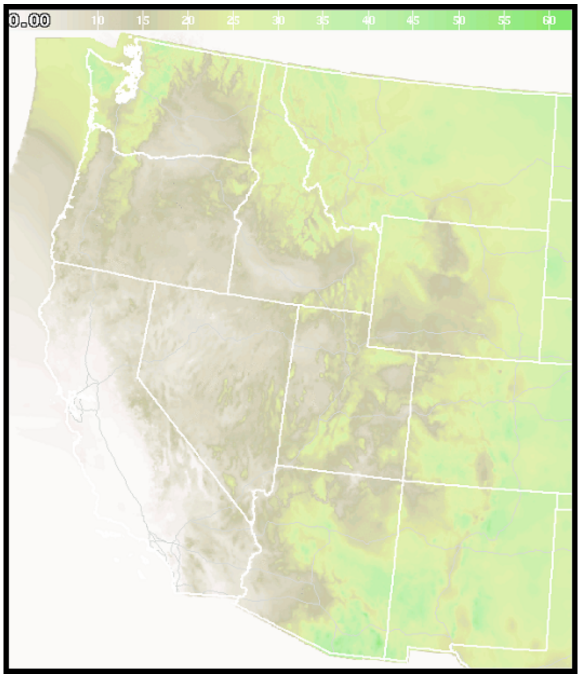
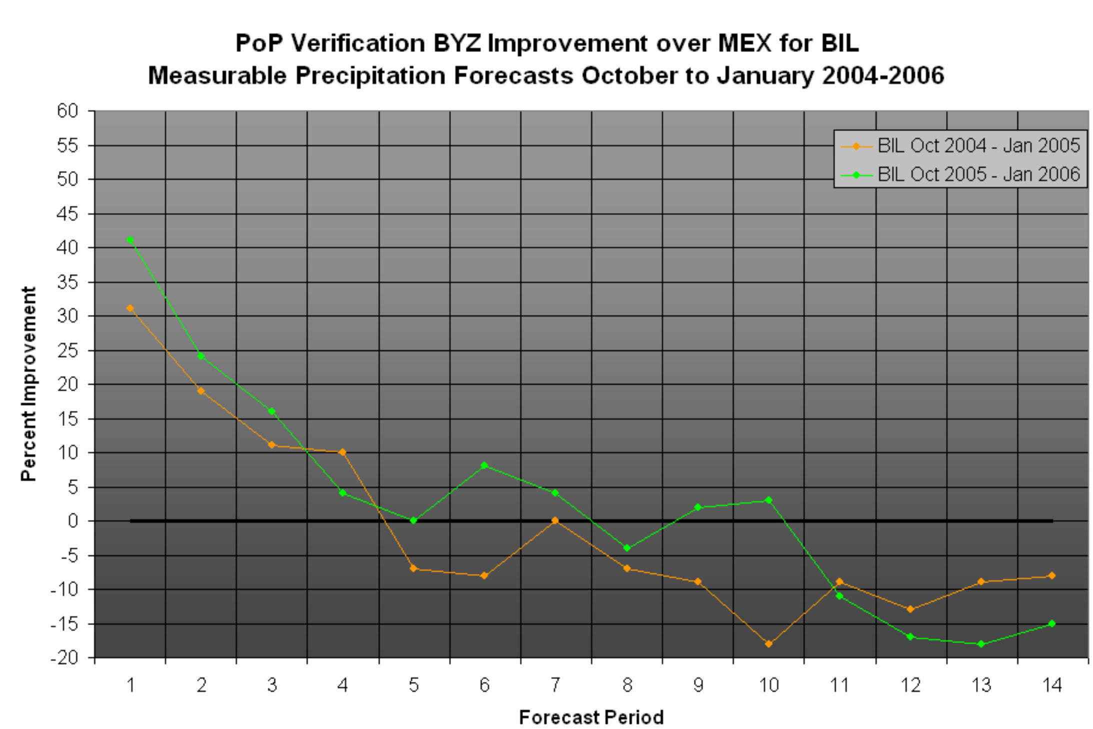


Figure 1. Climatology PoP for July developed for use in GFE (from WR Climatology PoP team).

II. 2004-2006 PoP verification and local techniques developed to improve PoP forecasts for later seasons

The local study completed at BYZ by fall 2006 revealed BYZ forecasters were often degrading MEX PoP forecasts per the Brier scores. Figure 2 shows this was frequently the case beyond forecast period seven and especially beyond forecast period ten for measurable precipitation events. The poorer Brier score for precipitation events in the latter part of the medium range were often from zero PoP forecasts while the MEX provided values closer to climatology. For example in October to January 2004 to 2005 and 2005-2006, BYZ forecast a zero to five PoP at forecast period ten 7.2% of the time when precipitation occurred in contrast to the MEX at 4.8% (Figure 3). Similar forecasts were provided for all forecasts beyond period ten. In addition, when BYZ forecast PoPs of 30% and 40% at both SHR and BIL precipitation occurred just 17% and 25% of the time respectively for period ten. So, not only did a dry forecast bias exists for precipitation events with low PoP forecasts, but higher PoP forecasts had a wet bias. This suggests that forecasters were unable to properly identify the wet days in medium range forecasts. The MEX on the other hand exhibited a near zero bias for PoP forecasts of 20% or less and a dry bias for 40% PoP forecasts (Figure 3).



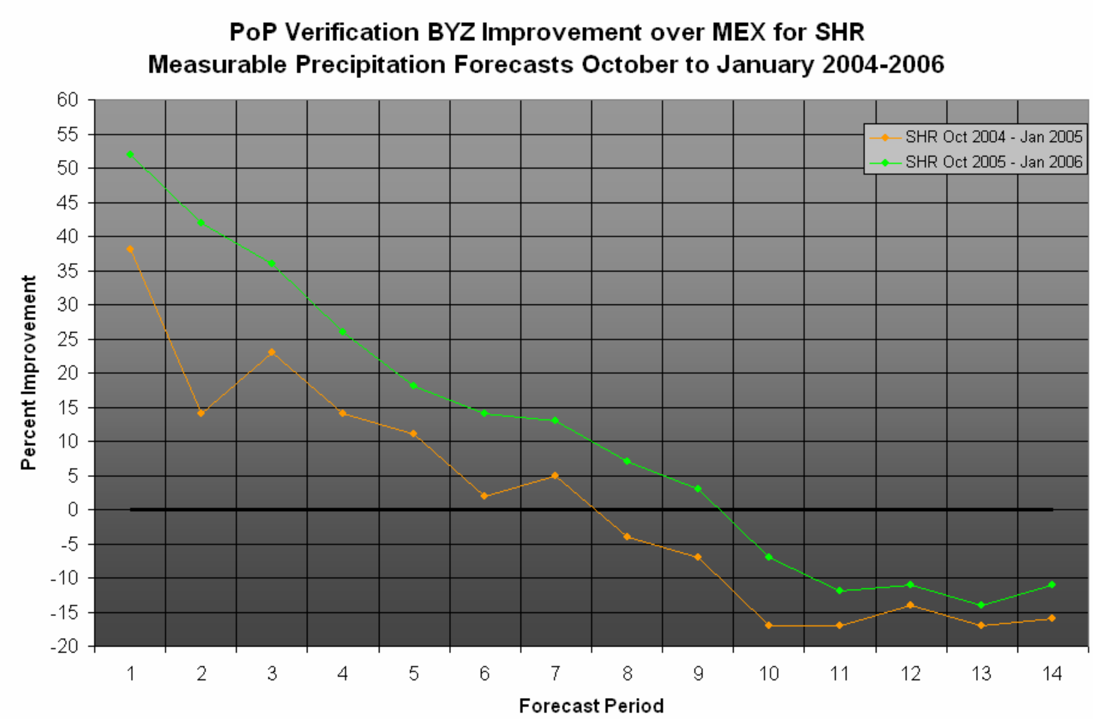


Figure 2. BYZ PoP forecast improvement over the MEX for precipitation events at BIL and SHR from Oct 2004 to Jan 2005 and Oct 2005 to Jan 2006. Positive values indicate a lower (better) Brier score than the MEX while negative values indicate a higher (worse) Brier score than the MEX.

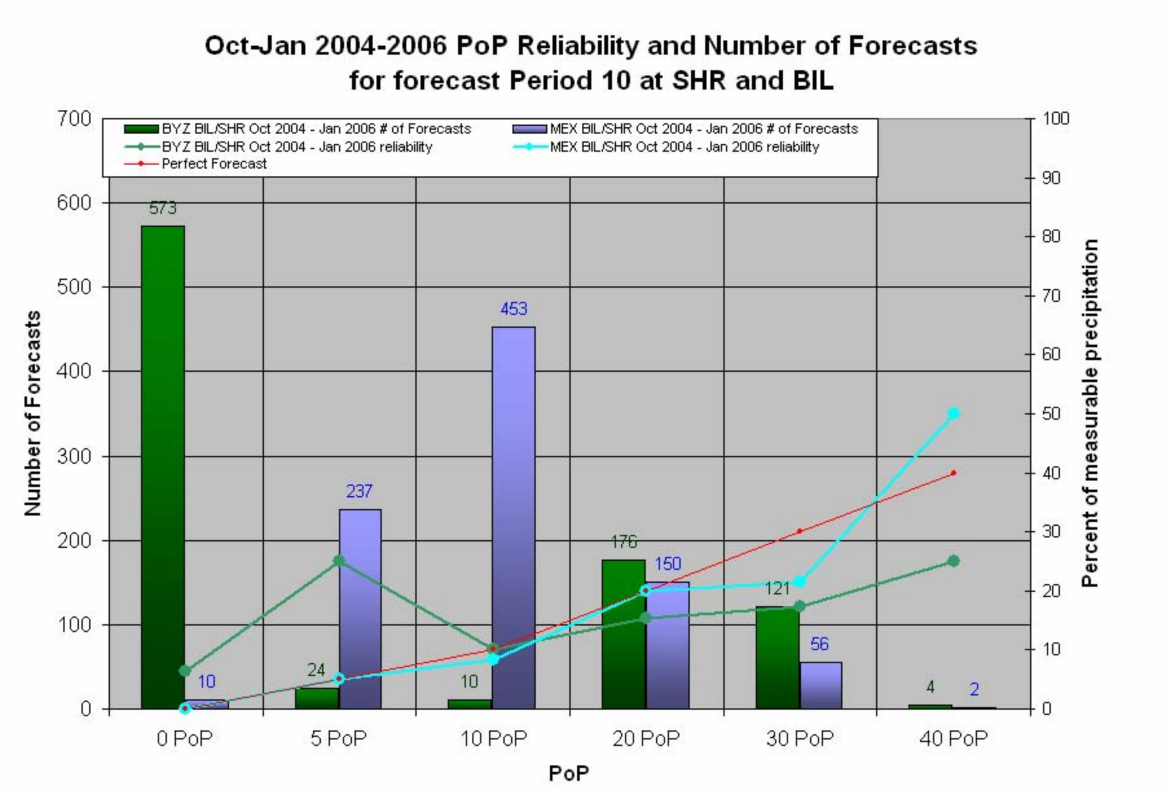


Figure 3. Green Bars indicate the number of forecasts provided by BYZ for the different PoP categories while blue bars are for the MEX with the scale shown on the left Y-axis. Lines represent the percentage of measurable precipitation for the different PoP categories where the green line represents BYZ forecasts, blue line MEX forecasts, and red line a perfect forecast. All data is for SHR and BIL combined. Points above (below) the red line indicate a dry (wet) bias.

To bring about improvements, the use of climatology PoPs were incorporated into the forecast process via climatology grids in GFE. A local study examining MEX PoP performance in relationship to climatology was conducted (study referred to hereafter as GFS PoP-Climatology Relationship Study) using data from 2003 to 2005. That study revealed the MEX 12 hour PoP had a significant dry bias when the MEX 12 hour PoP was at least 10% more than the climatological PoP. The signal was quite strong through the day five forecast, but relatively weak for six and seven day forecasts. BYZ often provided lower PoP forecasts than the MEX in those instances. Forecasters were therefore provided guidance in fall 2006 to: 1) consider starting with climatology PoPs in the medium range except in instances when the forecast weather pattern was very unfavorable for precipitation; 2) consider PoP forecasts 10% greater than the MEX when the MEX forecast a 12 hour PoP 10% or more over the climatology PoP; and 3) avoid PoP forecasts less than the MEX when the MEX forecast a 12 hour PoP 10% or more over climatology.

After the 2006 to 2007 cool season, an additional tool for GFE (hereafter referred to as PoP alert) was developed to alert forecasters when the MEX PoP was 10% or more above the climatological PoP. This was done to help raise forecaster awareness and allow for easier real time tracking of the MEX PoP in relationship to climatology. The PoP alert simply compares the MEX 12 hour PoP at all MEX points in BYZ's forecast area to the climatology PoP grid for the day. In instances of MEX PoP being 10% above the climatological PoP the following was provided to forecasters: 1) A red banner alarm on AWIPS alerted forecasters when and where the condition was met (figure 4); 2) a blank grid (MXCAlarm) appeared under the BYZ PoP grid to alert forecasters of the conditions (Figure 5); and 3) an additional PoP grid (ADJMXC) was created that increased the MEX PoP grid by 10% at the MEX points and nearby locations. However, it was left to the discretion of the forecaster, if the ADJMXC would be used.

Figure 4. Red banner alarm alerting BYZ forecasters of a MEX PoP 10% more than climatology for multiple sites and forecast periods in BYZ's forecast area.

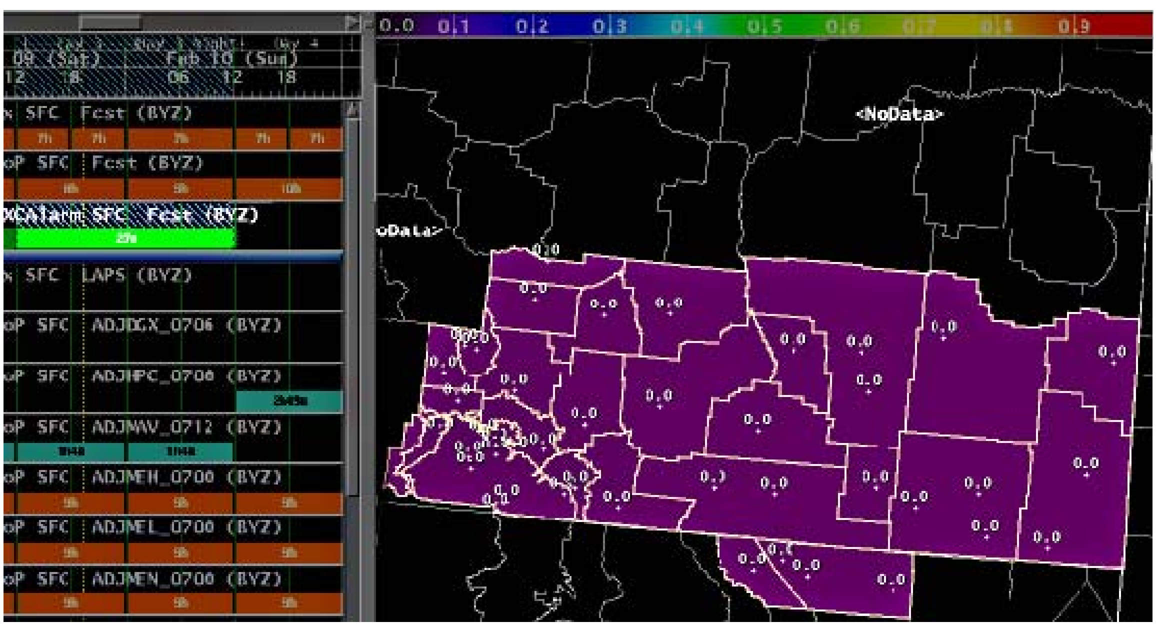
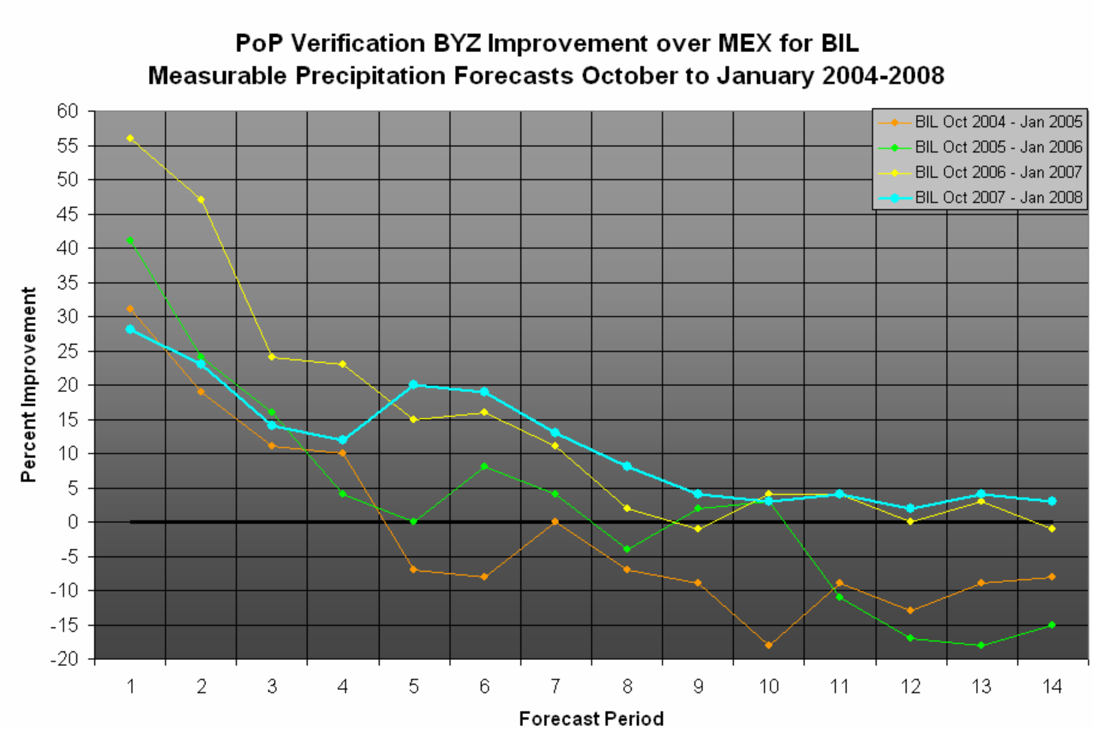


Figure 5. MXC alarm grid automatically created and displayed in GFE. The values in the grid are zero since the grid is just used to alert forecasters of a condition being met for those forecast hours.

III. 2006-2008 PoP Verification in comparison to 2004-2006

Verification scores of BYZ PoP forecasts for precipitation events from October to January 2006 to 2007 and 2007 to 2008 is significantly better than the previous two years (Figure 6). Part of this is likely from the use of climatological PoPs with fewer very low PoP forecasts for precipitation events, especially beyond forecast period 9 where the frequency of zero and five percent PoP forecasts by BYZ dropped significantly. Meanwhile the MEX frequency was similar in the two time periods. For example at forecast period 10, BYZ forecast zero and five PoPs 66% of the time from October to January 2004 to 2006 and just 38% from October to January 2006 to 2008. Meanwhile, the MEX frequency remained consistent around 27% of the time for both time periods (Figures 3 and 7). This resulted in a lower wet bias for zero and 5% PoP forecasts with both the MEX and BYZ having a 5% wet bias at forecast period ten (later forecast periods have similar results).

BYZ PoP biases in 2006 to 2008, also suggest better performance by BYZ with the higher PoP forecasts (Figure 7). For example at period ten, a neutral bias with 30% forecast PoPs is now shown and a dry bias with 40% forecast PoPs (still less than the MEX). Meanwhile, BYZ forecast four times more 40% PoPs at period ten than forecast in 2004-2006 while the MEX frequency only lowered by one occurrence (Figures 3 and 7).



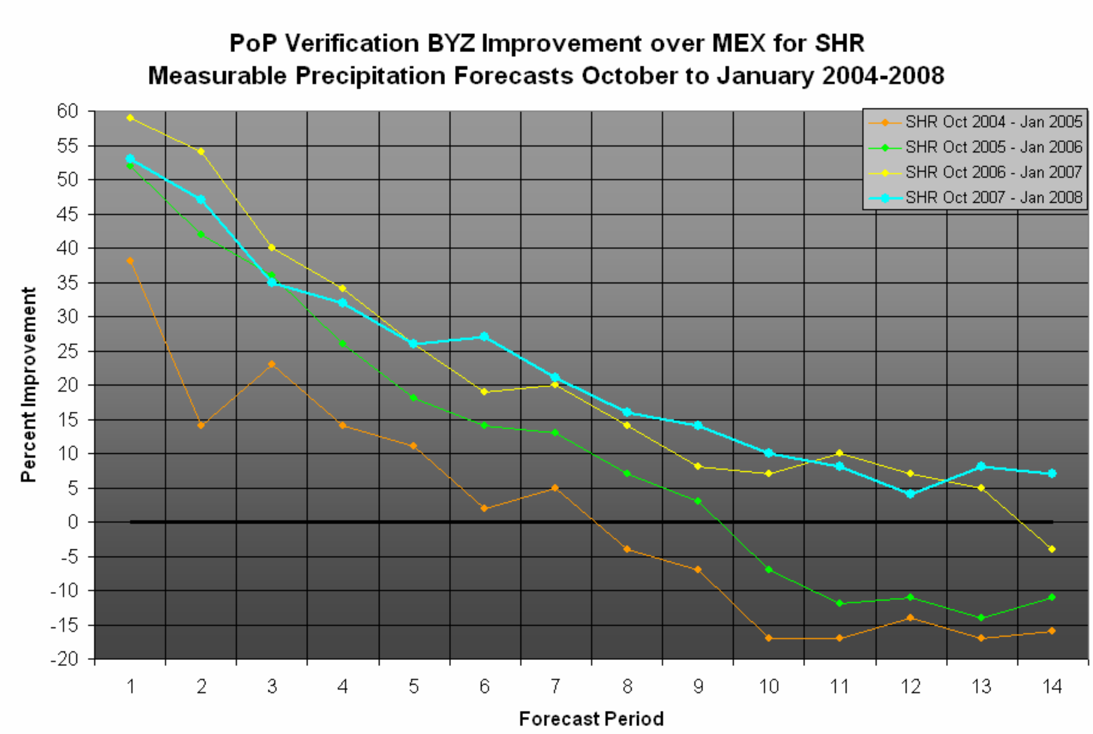


Figure 6. Same as figure 2 except also includes October to January 2006-2007 and 2007-2008.

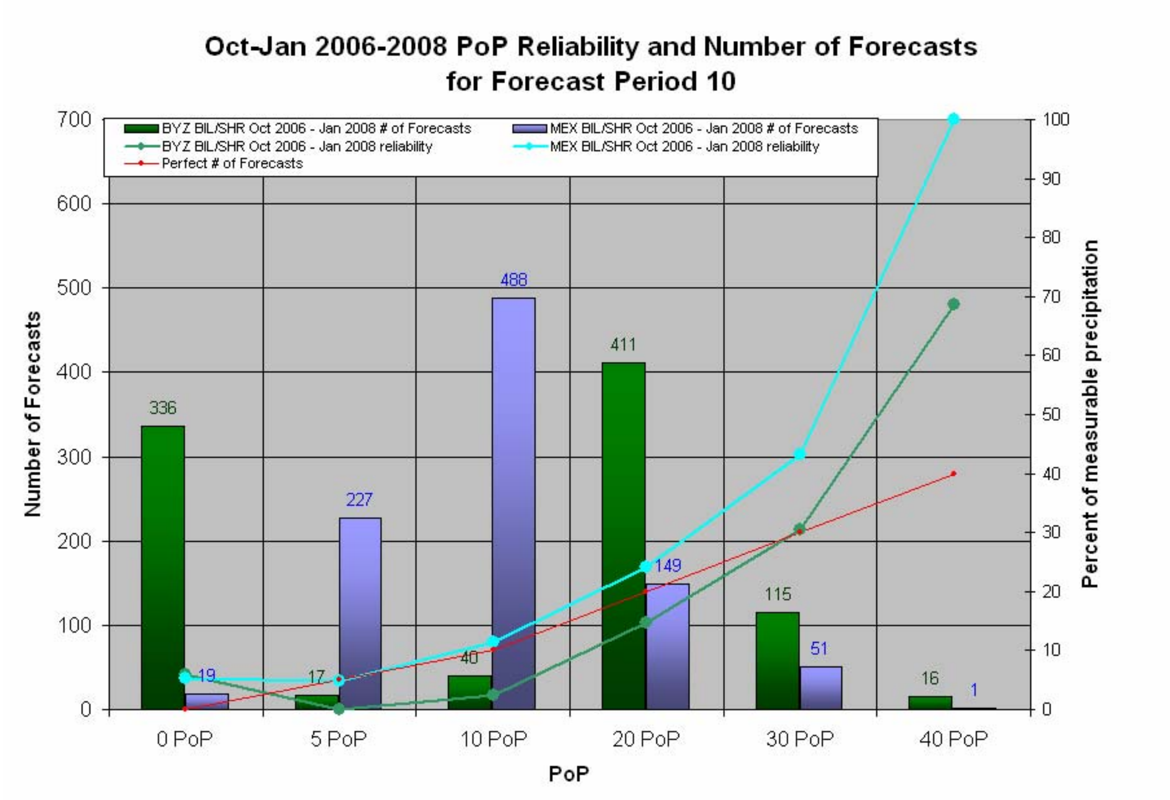


Figure 7. *Same as Figure 3 except for October to January 2006-2008.*

Separating the two years reveal BYZ more often provided higher PoPs from October 2006 to January 2007 compared to previous years when the MEX 12 hour PoPs were 10% or more than climatology. However, there were still many instances when BYZ forecasts were below the MEX. As a result, BYZ still had a drier bias than the MEX at forecast days four through six at both SHR and BIL when the MEX 12 hour PoP was at least 10% more than climatology (Figure 8). Also important to note is measurable precipitation was observed around 55 to 65% of the time when the MEX 12 hour PoP was 10% or more above climatology (peaks around 30%) for all forecast periods. Combined measurable and trace precipitation events, which in winter time may better represent public perception precipitation events since measurable snow can occur with just a trace of liquid precipitation, occurred nearly 70% of the time for all forecast periods (Figure 8).

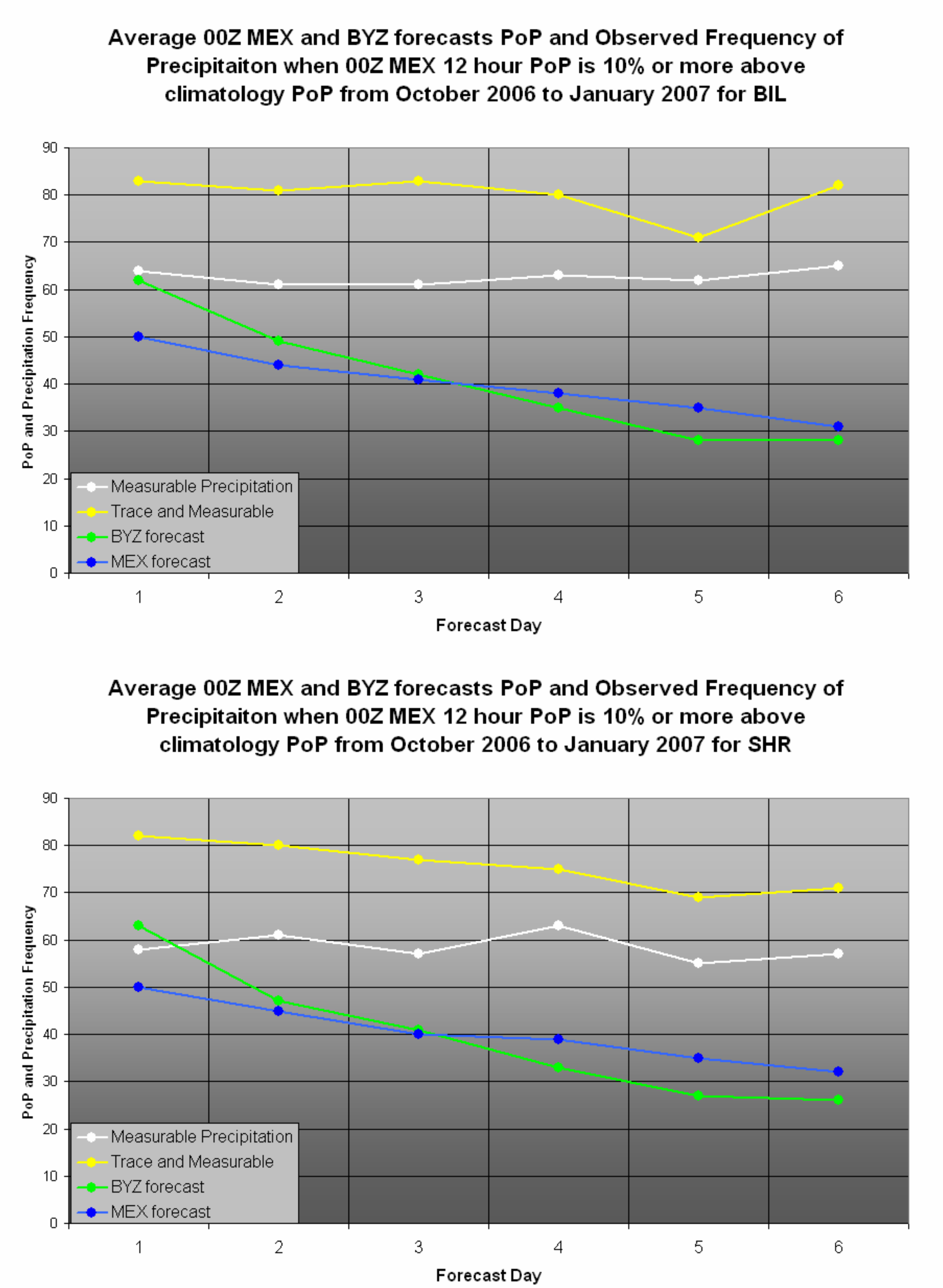
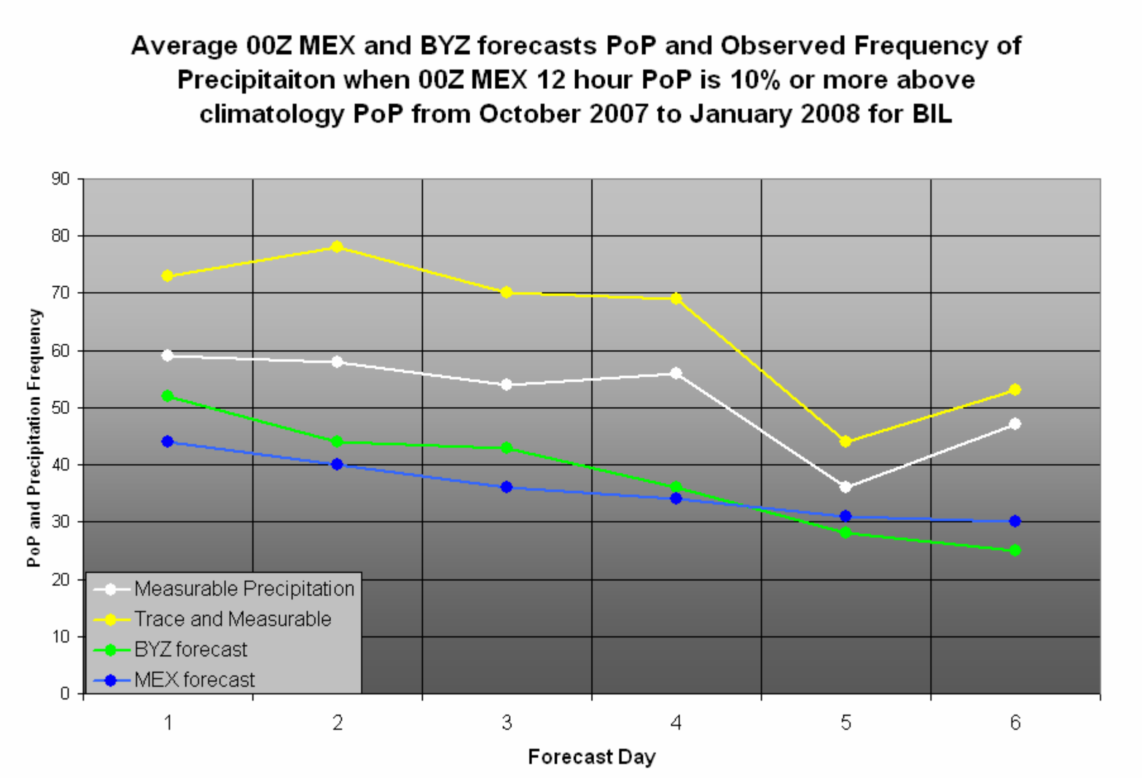


Figure 8. BIL and SHR average 00Z MEX PoP (blue line) and BYZ PoP forecast (green) and corresponding frequency of measurable precipitation events (white line) along with measurable plus trace precipitation events (yellow) from October 2006 to January 2007 when the MEX 12 hour PoP is 10% or more above climatology.

PoP Verification results for October 2007 to January 2008, after the incorporation of the PoP alert, show even greater improvement between forecast periods six and nine (Figure 6). The consistent improvement in two years from 2006 to 2008, compared to 2004 to 2006, further suggests the use of climatological PoPs and guidelines from the GFS PoP-Climatology Relationship Study likely brought about the improvements. This greater improvement is also seen in verification scores through forecast day four when the MEX 12 hour PoP was 10% or more than climatology (Figure 9). The continued dry bias of both the MEX and BYZ PoP forecasts is important to note, although not as significant as the previous year.



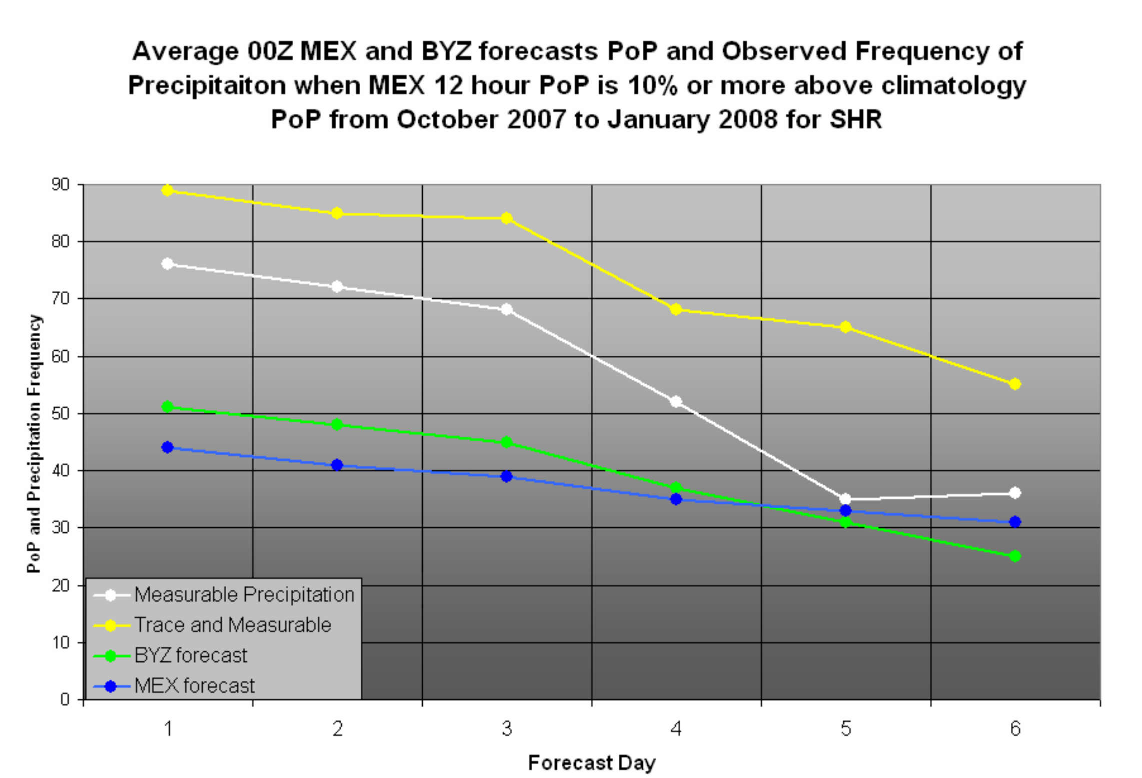


Figure 9. Same as Figure 8 except October 2007 to January 2008.

IV. Summary

Verification of PoP forecasts revealed the MEX was often outperforming BYZ forecasters in medium range forecasts in 2004 to 2006. This was due to dry forecasts being provided for precipitation events and not being able to properly identify precipitation events with higher PoP forecasts. Incorporating climatology PoPs into the forecast process did bring about improvements. Despite fewer zero and five PoP forecasts and a greater number of 20% PoP forecasts, biases of the lower PoP forecasts improved while the 20% PoP forecast biases remained unchanged. This shows that despite more frequent use of 20 PoPs, they were still being used an appropriate amount. As a result, the Brier score of BYZ forecasts compared to the MEX showed notable improvements at forecast days six and seven. Another substantial improvement came from recognizing the dry bias the MEX exhibits when the MEX PoP is 10% or more than climatology. Having a tool to alert forecaster of these instances made this improvement even greater (Figure 8 versus Figure 9). While the improvements beyond forecast day five were probably due to incorporation of climatology PoPs into the forecast process, a combination of the GFS PoP-Climatology Relationship Study and use of climatology PoPs likely brought about improvements for forecast days three to five.

Improvements in PoP forecasts at BYZ have been significant since fall 2006, but there is still room to improve. Many years of verification data suggest BYZ forecasters can be even more aggressive going with higher PoPs than the MEX at forecast days four, five, and six. While the initial study suggested BYZ consider forecasting PoPs 10% more than the MEX when MEX PoPs are 10% or more than climatology, climatological differences across BYZ's forecast area suggest some locations need even higher PoPs such as SHR compared to BIL. Livingston, MT (LVM) is another site where the MEX showed a significantly greater dry bias. Forecasters may need to consider forecasting PoPs 20% or perhaps even 30% higher than the MEX when MEX PoPs are 10% or more than climatology. Given the results of this study, the authors feel additional GFS PoP-Climatology Relationship Studies should be performed to better understand these spatial differences of MEX PoP performance.

Acknowledgements: The authors would like to thanks Gene Petrescu and Mike Staudenmaier for their work in developing climatology PoPs for use in GFE.