Utilizing the New Email Stats Feature of BoiVerify 2.0 to Examine Min and Max Temperature Guidance

Mike Staudenmaier, Jr.
WFO Flagstaff, AZ

Introduction

With the advent of BoiVerify, NWS offices have been able to verify gridded forecasts, and also to automatically bias correct numerical guidance to potentially serve as better initialization in the digital forecast process. With the 2.0 upgrade, statistics can now be automatically generated and emailed to interested parties, which can serve as an easy and automatic way to determine which grids represent the best starting point for the upcoming digital forecast package. Offices in Western Region have spent the past 18 months examining output from BoiVerify and used it to determine a suggested methodology for utilizing bias corrected grids in the forecast process, primarily the MOSGuideBC grids. This short paper examines the statistics available in the email feature of BoiVerify to examine the variability of bias corrected grids under specific "regimes" to test the robustness of MOSGuideBC grids in an active winter season in the Flagstaff CWA, along with testing the new SREF/SREFBC guidance grids.

Methodology

Statistics from 6 December 2007 through 5 March 2008 were examined for an area covering approximately the southern half of the Flagstaff CWA (area with the most observations) for both min and max temperatures for the Day One period of the forecast (both 14-hr and 26-hr). Utilizing the email statistics feature of BoiVerify, it was easy to identify the two best performing model guidance grids available for each day and also for the past 30 days as verified against the MatchObsAll database. Bulk statistics for the entire period were examined and calculated. In addition, non-standard regimes were investigated including windy days (days/nights with avg. wind >= 12 mph at Flagstaff), strong inversion days (Winslow max temperature <= Flagstaff max temperature), and days with widespread precipitation (greater than 0.25" at Flagstaff and significant coverage on 24 hour precipitation chart). In addition to these regimes, the data were also examined for cases where the average anomaly rank was 1-8 (large deviation from climatology), and also for rankings 23-30 (very near climatology).

Data and Results

Bulk Statistics – 30 day running average

Examination of the entire period for 30 day average minimum temperature statistics showed that the MOSGuideBC was ranked either first or second best guidance 100% of the days examined. This was a surprising result as minimum temperatures can be very challenging to forecast over northern Arizona, yet the MOSGuideBC was continually in the top two best guidance products over the 30 day running average. GFS40BC was next at 38%. This table shows the results for all models that showed up in the top two best guidance products:

Min Temp Verification (30 day running average) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
MOSGuideBC	100%/100%	100%	**ADJMETBC**	26%/22%	24%
GFS40BC	54%/21%	38%	**SREFBC**	11%/7%	9%
ADJMAVBC	7%/49%	28%	**SREF**	3%/1%	2%

Interestingly, not all model guidance performed equally well for both the midnight shift (14-hr) and the day shift (26-hr) forecasts. Both the GFS40BC and the ADJMAVBC guidance grids were more skillful for one shift over the other.

Examination of the entire period for 30 day average maximum temperature statistics showed that scores were somewhat more variable with both MOSGuideBC and ADJMETBC performing similarly well, appearing in the top two best products approximately half of the days examined. This table shows the results for all models that showed up in the top two best guidance products:

Max Temp Verification (30 day running average) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
MOSGuideBC	44%/71%	58%	**SREFBC**	17%/33%	25%
ADJMETBC	65%/40%	53%	**ADJMET**	22%/18%	20%
NAM12BC	53%/22%	38%	**GFS40BC**	0%/12%	6%

Again, not all model guidance performed equally well for both the midnight shift (14-hr) and the day shift (26-hr) forecasts. The top three performing models showed significant variation between both shifts.

Windy Regime

Examination of the fourteen days which met the wind criteria defined above were also examined for the top two best performing models on those individual days for minimum temperature. Not surprisingly, almost all the guidance grids showed skill on one day or another, with the MOSGuideBC being ahead of the rest. Only the top six performing grids out of the eleven that verified best are shown in the table below. No guidance grids were best more than 30% of the time. Raw model grids showed some skill in this regime.

Min Temp Verification (Windy Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
MOSGuideBC	8%/50%	29%	**ADJMET**	21%/0%	10%
ADJMAVBC	8%/25%	17%	**GFS40**	17%/0%	8%
ADJMETBC	4%/17%	10%	**NAM12**	13%/0%	6%

Once again, not all model guidance performed similarly well for both the midnight shift (14-hr) and the day shift (26-hr) forecasts. In fact, the top three performing guidance grids at 26-hr were some of the worst grids at 14-hrs. It is unclear why this would be.

Examination of the same fourteen days for maximum temperature statistics continued to show big differences between the 14-hr and 26-hr performance, and also that no guidance grid was the best choice overall during windy regimes for max temperatures. This table shows the top six performing guidance grids of the eight total that showed up in the top two best guidance products:

Max Temp Verification (Windy Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
ADJMETBC	21%/21%	21%	**ADJMET**	21%/11%	16%
MOSGuideBC	7%/29%	18%	**SREFBC**	4%/21%	13%
NAM12BC	21%/11%	16%	**NAM12**	14%/11%	13%

Strong Inversion Regime

Examination of the eight days which met the strong inversion criteria defined above were also examined for the top two best performing models on those individual days for minimum temperature. As with the windy regime, almost all the guidance grids showed skill on one day or another, but none were overwhelmingly best. The two best performing guidance grids of the eight that verified best were the ADJMETBC and MOSGuideBC. Perhaps surprisingly, BC grids showed skill in this regime while raw model grids showed little skill.

Min Temp Verification (Strong Inversion Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
ADJMETBC	31%/25%	28%	**GFS40BC**	13%/13%	13%
MOSGuideBC	19%/13%	16%	**MOSGuide**	13%/13%	13%
ADJMAVBC	13%/13%	13%	**ADJMEXBC**	6%/13%	9%

Examination for maximum temperature is shown below. Once again, BC grids were better than raw model grids. Only the top six performing grids out of the eight that verified best are shown in the table below.

Max Temp Verification (Strong Inversion Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
ADJMETBC	19%/25%	22%	**NAM12BC**	6%/13%	9%
MOSGuideBC	19%/19%	19%	**ADJMAVBC**	6%/13%	9%
ADJMEXBC	13%/13%	13%	**GFS40BC**	6%/6%	6%

Precipitation Regime

Examination of the eleven days which met the widespread precipitation criteria defined above were also examined for the top two best performing models on those individual days for minimum temperature. Not surprisingly, almost all the guidance grids showed skill on one day or another, but none were overwhelmingly best. Of those, the raw model guidance tended to be slightly better. Only the top six performing grids of the twelve that verified best are shown in the table below. No guidance grids were best more than 20% of the time.

Min Temp Verification (Precipitation Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
GFS40	23%/14%	18%	**ADJMETBC**	14%/9%	11%
NAM12	18%/14%	16%	**NAM12BC**	5%/9%	7%
ADJMET	14%/14%	14%	**MOSGuideBC**	5%/9%	7%

Examination for maximum temperature is shown below. Once again, almost all guidance grids showed skill on one day or another, but none were overwhelmingly best. For maximum temperatures, BC grids were better than raw model grids. Only the top six performing grids out of the ten that verified best are shown in the table below. Again, no guidance grids were best more than 20% of the time.

Max Temp Verification (Precipitation Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
NAM12BC	20%/20%	20%	**ADJMETBC**	20%/15%	18%
MOSGuideBC	20%/20%	20%	**GFS40**	5%/10%	8%
ADJMET	25%/10%	18%	**NAM12**	5%/10%	8%

Large Deviation from Climatology Regime

Examination of the twenty days which had large average anomaly ranks (large deviations from climatology) were also examined for the top two best performing models on those individual days for minimum temperature. Again, almost all the guidance grids showed skill on one day or another, but none were overwhelmingly best. Both BC and raw model guidance showed similar skill, with the MOSGuideBC just slightly best. Only the top six performing grids of the fourteen that verified best are shown in the table below. No guidance grids were best more than 20% of the time.

Min Temp Verification (Non-Climatology Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
MOSGuideBC	15%/16%	15%	**ADJMETBC**	10%/8%	9%
GFS40	10%/8%	9%	**ADJMEXBC**	8%/8%	8%

ADJMET	10%/8%	9%	**SREF**	5%/11%	8%

Performance for the two best performing models for maximum temperature is shown below. Once again, almost all guidance grids showed skill on one day or another, but none were overwhelmingly best. For maximum temperatures, the raw NAM12 grids performed best, but only slightly. Only the top six performing grids out of the ten that verified best are shown in the table below. Again, no guidance grids were best more than 20% of the time.

Max Temp Verification (Non-Climatology Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
NAM12	15%/20%	18%	**GFS40BC**	10%/18%	14%
ADJMET	20%/13%	16%	**MOSGuideBC**	13%/10%	11%
NAM12BC	18%/13%	15%	**GFS40**	8%/13%	10%

Small Deviation from Climatology Regime

Examination of the twenty five days which had very small average anomaly ranks (near climatology) were also examined for the top two best performing models on those individual days for minimum temperature. In this regime, the BC grids seemed to have an advantage over the raw model grids, with the ADJMETBC the best performer. However the ADJMETBC showed significant differences in skill between the 14-hr and 26-hr forecasts. Only the top six performing grids of the twelve that verified best are shown in the table below.

Min Temp Verification (Near Climatology Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
ADJMETBC	10%/28%	19%	**GFS40BC**	16%/12%	14%
MOSGuideBC	20%/14%	17%	**MOSGuide**	8%/8%	8%
SREFBC	16%/14%	15%	**ADJMAVBC**	6%/6%	6%

Performance for the two best performing models for the twenty eight days of this regime for maximum temperature is shown below. Once again, in this regime the BC grids seemed to have an advantage over the raw model grids, with the NAM12BC the best performer. Only the top six performing grids out of the twelve that verified best are shown in the table below.

Max Temp Verification (Near Climatology Regime) Best Guidance for Day One

Guidance	14-hr/26-hr % in top two	Day One % in top two	Guidance	14-hr/26-hr % in top two	Day One % in top two
NAM12BC	18%/14%	16%	**ADJMET**	7%/14%	11%
ADJMEXBC	9%/14%	12%	**ADJMETBC**	9%/11%	10%
MOSGuideBC	13%/9%	11%	**GFS40BC**	7%/11%	9%

Conclusions

BoiVerify 2.0 has brought about the ability to more easily examine statistics related to various regimes. Examination of various regimes at WFO Flagstaff, albeit with just one winter season at this time, suggests the following for the Day One forecast:

- Overall MOSGuideBC is the best guidance to use for minimum temperature forecasts. MOSGuideBC and ADJMETBC had similar enough statistics that a blend of the two might be the best forecast for maximum temperature overall.
- During windy regimes, the data suggested MOSGuideBC for minimum temperatures and perhaps again a blend of ADJMETBC and MOSGuideBC for maximum temperatures. Raw model grids showed some limited skill in this regime, although BC grids were in the top three for both minimum and maximum temperature verification.
- Under strong inversion situations, BC grids again did well compared to raw model guidance grids or raw MOS adjusted grids. In this regime the ADJMETBC grids showed a slight edge over MOSGuideBC, but again perhaps a blend of the two might be better, as they showed skill under different days of the regime examined.
- Skill in precipitation events overall slipped to 20% or less for the best verifying grid showing the variability of guidance during this regime. For minimum temperature forecasts, the raw model guidance performed slightly better than adjusted MOS or the BC grids. For maximum temperatures, the NAM12BC and MOSGuideBC grids were equally skillful, with perhaps a blend being the best guidance.
- During events that are significantly colder or warmer than climatology, the skill of guidance slipped even more with the best model guidance being in the top two best category only 15-18% of the time. For minimum temperatures MOSGuideBC was slightly ahead of the other guidance, while for maximum temperatures the NAM12 was best. Raw model output showed up in the top six in this regime suggesting some skill.
- For events very near climatology, BC grids tended to perform the best. Surprisingly however, the skill of the best guidance was still low with the highest percentage of grids in the top two verifying only between 16-19% of the time. The ADJMETBC was best for minimum temperatures in this regime, while the NAM12BC was best for maximum temperatures, but only slightly.

While MOSGuideBC grids overall appear to offer the most consistent starting point for the forecaster to use in the forecast process for the Day One temperature forecast, the data presented here suggests some caveats. It has shown that in many regimes that are important to our customers and partners, the differences between guidance performance is so narrow as to make it challenging to determine a clear model to go with without further forecaster examination. In situations when it still isn't clear which guidance may be the best to choose from, blending several guidance packages may have the best benefit. Surprisingly, even in situations that were near climatology (which would be assumed to be less challenging), MOSGuideBC was not the clear winner, underscoring the challenge of guessing which model product may be best even in what appear to be benign situations. Additionally, variable guidance performance between the 14-hr and 26-hr forecasts under some regimes makes it more challenging to suggest a 'winner' for a particular regime. More investigation into these regimes and verification will hopefully shed more light on these issues to enable the forecaster to better choose the best guidance for the temperature forecast. With the new

email function, forecasters can much more easily see the variances of guidance to make better decisions based on regime and/or the best overall guidance over the past 30 days.