VERIFICATION OF FORECAST GRIDS IN THE CENTRAL REGION OF THE NATIONAL WEATHER SERVICE

Paul Wolyn *
NOAA/NWS Pueblo, CO

1. INTRODUCTION

The Central Region (CR) of the National Weather Service (NWS) has an initiative to create a common set of forecast grids for days 4 through 7. These grids will be of good quality and consistent among the various NWS offices. It is hoped these grids will be "good enough" in most instances, and only minor adjustments will be needed. By reducing the amount of time spent creating and collaborating long term grids, the forecaster can spend more time on other activities, such as decision support services, During busy weather, these untouched grids could be used as the long term forecast, freeing another forecaster to perform other duties, such as those associated with severe convective weather.

The NWS forecasts are generated using the Graphic Forecast Editor (GFE). The GFE has various weather elements on a 2.5km by 2.5 km grid for the entire forecast domain and surrounding areas. Grids from a wide range of forecast inputs are available in GFE. These grids can be blended in the forecast process, and the forecaster has the ability to edit the grids before the forecast is issued. Nearly all of the sources for the forecast input grids, such as numerical models, are not available on a 2.5km grid, and these inputs have to be downscaled to the 2.5km GFE grid.

This study will examine verification of maximum temperature (MaxT), minimum temperature (MinT) and dew point (Td) for the entire CR using verification data for each Weather Forecast Office (WFO). Temperature was not verified since the hourly temperatures from the various forecast inputs are typically not used. Instead, the hourly temperatures in the forecast grids are derived from the MaxT and MinT based on various schemes such as using the implied diurnal cycle in a forecast input or based on the observed diurnal cycle.

* *Corresponding author address:* Paul Wolyn, NOAA/NWS, 3 Eaton Way, Pueblo, CO 81001
email: paul.wolyn@noaa.gov

2. DATA SET

This study will examine two years of verification data from 1 March 2012 to 28 February 2014 for the entire CR of the NWS. The data are collected centrally at Central Region Headquarters and are in a database accessible by the WFO's. The verification data for Boulder, CO and Omaha, NE are excluded because they are AWIPS2 beta test sites. At these two beta sites, the GFE procedures, which created the various grids and performed the verification, were unreliable for significant portions of the study time period. The verification data base includes Mean Absolute Error (MAE) for the entire GFE grid of each WFO. In this dataset there is no way to calculate or infer the MAE for a subarea (or individual point) of each WFO's GFE grid.

There are four groups of forecast inputs in GFE: *deterministic models*, *guidance*, *consensus*, and *forecasts*. The first group, *deterministic models,* is the "raw" model output, which is mapped to the GFE domain. Since the grid size of the model is often much larger than the grid size in GFE, downscaling of the deterministic models is done, which includes accounting for the much finer topography in GFE. The deterministic models are listed in table 1.

Table 1. Deterministic models grids used in this study.

Name	Description
NAM12	12 km NCEP NAM Model
NamDG5	downscaled NCEP NAM to 5km grid
GFS40	GFS available on 40 km grid
GEMreg	regional GEM model from Environment Canada
GEMnh	Hemispheric GEM model from Environment Canada.
ECMWFPR	ECMWF model on a 0.5 °grid from the previous model run.

The ECMWF is typically obtained around 7 to 8 hours after model initialization. Often, this is too late for the latest ECMWF to be evaluated for the "routine" forecast package which is done twice daily around 4 AM local time (0800-1100 UTC in CR) and 3PM local time (1900-2200 UTC in CR). In addition, the ECMWF arrives too late for the latest version to be used in the various consensus forecast inputs, which will be discussed later. As a result the previous version of the ECMWF is used in the comparisons. The suffix “PR” (for PRevious) is appended to model name because the previous version of the model is used forecast process for the twice daily routine packages. For example, the routine forecast afternoon package will mostly use forecast inputs generated from the 1200 UTC models. Since the ECMWF arrives too late, the “previous” version of the ECMWF, initialized at 0000 UTC, is used in generating the forecast and in the various consensus forecast inputs.

The second group of forecast inputs is *guidance*, which consists of various statistical or human derived guidance products. Table 2 lists the guidance forecast inputs used in this study.

Table 2. Guidance grids used in this study.

Name	Description
ADJMET	NAM12 background field adjusted using MET guidance.
ADJMEN	NAM12 background field adjusted using MEN guidance
ADJMAV	GFS40 background field adjusted using MAV guidance
ADJMEX	GFS40 background field adjusted using MEX guidance
MOSGuide	NCEP digital guidance. 5km grid mapped to 2.5km grid
HPCGuide	Official HPC guidance

Many of the guidance products available to the NWS (for example MET and MAV) are statistical guidance at various points in the forecast area. GFE grids from these point statistical guidance are created by using the statistical guidance values used at various points in and near the forecast area to adjust a background. The guidance inputs will have the prefix “ADJ” (for adjusted) before the guidance name. For example, the ADJMET grid is the NAM12 grid adjusted using MET guidance at various points in and near the WFO forecast area.

The third group of forecast inputs is *consensus,* which is a blend of various forecast inputs. Table 3 gives a list of all the consensus grids used in this study.

Table 3. Consensus grids used in this study.

Name	Description
CONSRaw	Blend of all available deterministic models
CONSMOS	Blend of all available guidance
CONSAll	Blend of all available deterministic models and guidance.
AllBlend	Average of CONSAll and previous forecast.
RawBlend	Average of CONSRaw and previous forecast
WModel	Blend of the three forecast inputs with verified the best over the last 7 days. 50% best, 33% second best, 17% third best.

The blends of guidance or deterministic models (CONSRaw, CONSAll, and CONSMOS) use all the available guidance or deterministic models (except HPCGuide). In CONSRaw and CONSAll, ECMWFPR is given a double weight, and the remaining inputs are given a single weight. Other forecast inputs, in addition to the ones given in Tables 1 and 2, are used in deriving the consensus grids. Tables 1 and 2 only list the deterministic model and guidance inputs which will be analyzed in this study, and the criteria for choosing the forecast inputs in the analysis is given in section 3.

The fourth group of forecast inputs is *forecasts*, and a description of these inputs is given in table 4.

Table 4. Forecast group grids used in this study.

Name	Description
Official	WFO forecast
Previous	The previous forecast made 12 hours earlier with the routine forecast package.

The “Previous” forecast is one half of the AllBlend and RawBlend consensus forercast inputs. (For AllBlend, the other half is CONSAll, and for RawBlend the other half is CONSRaw.) The “official” forecast is not technically a “forecast input” as it is the output of the forecast process. To simplify the analysis of the data, the official forecast is included in this group.

Nearly all of the forecast inputs also have a “bias corrected” (BC) counterpart. These bias corrected forecast inputs have “BC” in their name. The bias correction adjusts the value at each grid point based on the previous 30 days of forecast inputs and observations. The calculation is done independently for each forecast hour and initial time. The core of the bias correction is a linear regression between forecast and observed values at a given point. If the forecast value is outside of the range of values used in the linear regression calculation, the bias correction will trend towards the mean difference between the forecast and observed values for the last 30 days. The bias correction is discussed further at:

http://www.werh.noaa.gov/SSD/verification/BOIVerify.2.0.pdf

Forecast inputs are verified using an "observed" grid in GFE. The "observed" grids are created using a blend of various initial fields in GFE. This initial grid is adjusted based on observed data. The observed data includes the data available in the MADIS dataset (http://madis.noaa.gov) as well as the ASOS and AWOS observations. These GFE "observed" grids are routinely examined every 2 to 4 hours, when the first 12 hours of the forecast is routinely updated, and bad data points are regularly identified and removed.

The CR extended forecast routine uses the AllBlend forecast input as the first guess grids for days 4-7. Ideally, these grids should not be altered much, if at all, when generating for forecast. For WFO's with significant mountainous terrain, BCAllBlend is used instead of AllBlend. The WFO's in Colorado and Wyoming (Pueblo, CO, Grand Junction, CO, Boulder, CO, Riverton, WY and Cheyenne, WY) are considered to have significant mountainous terrain. (Analysis presented later will show the increased benefits of bias correction in complex terrain.)

3. MAE ANALYSIS FOR ENTIRE CR

The dataset used in this study consists of the average MAE of the GFE grid, covering entire forecast area, at each WFO. Figure 1 shows a map of the CR and the forecast area for each WFO. The forecast areas have varying size. For simplicity, the size of the WFO's the forecast area is not considered when using data from various WFO's. Each WFO has the same weight in the analysis despite the varying forecast areas.

Only forecast inputs which are commonly available are used for the data analysis. The availability of a forecast input for a given initial time (0000 UTC or 1200 UTC) and forecast hour is determined for each WFO. The results for each WFO are combined with each WFO having equal weight. For each parameter (MaxT, MinT, Td), a list is created identifying which forecast inputs are available at least 75% of the time. This scheme does not require that a forecast input is available at every WFO for a given initial time and forecast hour on a particular date.

This list of commonly available forecast inputs for each initial time and forecast hour at a WFO is used to analyze the two years of data. The data is included in the analysis only if all of the forecast inputs in the list are available at the WFO on that date for a particular forecast hour, initial time, and parameter. Strictly requiring all the forecast inputs to be present at a WFO prevents including days with larger or smaller MAE, due to difficulty of the forecast, in only some forecast inputs, which could skew the results.

The number of data points for the analysis generally range from 11,000 to 16,000. Typically, fewer data points are available for the shorter forecast hours because more forecast inputs are being analyzed and there is a great chance for one or more of the forecast inputs to be missing. Given that 36 WFO's are used in the data analysis (since Boulder and Omaha were excluded), the percentage of possible data points included range from approximately 40% to 60%.

Figures 2-4 show the comparison of MAE's for MaxT, MinT and Td. The MAE for CONSAll is subtracted from the MAE for ease of display. If the unadjusted MAE is used, the significant increase in MAE with increasing forecast hour would dominate the graph. In these plots, the y=0 line is the verification for CONSAll. Lines below y=0 (negative MAE differences) indicate the forecast input has a lower MAE than CONSAll.

The colors of the lines identify the forecast input group. Green is the deterministic model group, yellow is guidance group, red is consensus group, and blue is forecast group. The data for the 0000 UTC and 1200 UTC initial times are combined in the plots. Since the MaxT and MinT grids are 24 hours apart, these plots have forecast hours which alternate between 0000 UTC and 1200 UTC initial times.

The verification for the various forecast inputs stratify by groups well. The best verify group overall is the consensus group with nearly all the members having an MAE difference from CONSAll of 0.5°F (0.28°C) or less. (The y=0 black line is the verification for CONSAll). The worst verifying group is the deterministic models and the guidance group is in the middle.

Figures 5-7 show the same plots, but only includes forecast inputs with MAE differences from CONSAll of 0.5°F (0.28°C) or less. For MaxT (Figure 5), BCCONSAll has the lowest MAE until the 36 forecast hour. Afterwards, BCAllBlend and BCRawBlend verify slightly better. Beyond 96 hours, CONSAll verifies the best as seen by all the points, except for RawBlend at 144 hours, being above the y=0 line. (Since the plotted values are the CONSAll MAE subtracted from the MAE, the y=0 line is the verification for CONSAll.)

For MinT (Figure 6), BCCONSAll is the best verifying forecast input for many of the forecast hours until forecast hour 96. Afterwards, CONSAll is the best verifying model (as seen by all the points being above the y=0 line). For dewpoint (Td), (Figure 7) CONSAll has the lowest MAE for nearly all the time periods, with a few forecast inputs being slightly better during the first 36 hours.

CONSAll is an unsophisticated blend of all available forecast inputs, which are not bias corrected, from the deterministic model and guidance groups (except for HPCGuide and WModel). (There are more forecast inputs than what is listed in Tables 1 and 2.) The ECMWFPR is weighted double. Any missing forecast input is simply ignored when creating CONSAll and there is no adjustments done to account for missing forecast inputs. BCCONSAll is similar to CONSAll except is uses bias corrected grids.

The unsophisticated blends forecast inputs generally verifies the best (lowest MAE) for many cases over a large area. Using only the forecast inputs which verified the best during the last 7 days (WModel) or blends of specific groups (CONSMOS and CONSRaw) do not verify as well.

Figure 8 shows that bias correction decreases the mean MAE for MaxT and MinT up to about the 96 forecast hour, and beyond this time CONSAll verifies the best. For Td, bias correction does not improve the forecast even in the shorter forecast hours.

4. TERRAIN EFFECTS AND BIAS CORRECTION

The forecast offices in Colorado and Wyoming have significant complex terrain with elevations of some peaks exceeding 4km MSL. To examine the importance of bias correction in complex terrain versus flatter terrain, MAE analysis was done for the WFO's only in Colorado and Wyoming: Pueblo, CO, Grand Junction, CO, Riverton, WY, and Cheyenne, WY. (WFO Boulder,CO was not included because it was an AWIPS2 beta site). The MAE analysis is also done for the remaining sites in CR (excluding Omaha, NE).

Figure 9 shows plots of the difference on MAE between BCCONSAll and CONSAll for the Colorado and Wyoming WFO's (Mountain WFO's). Negative values indicate that BCCONSAll has a lower average MAE than CONSAll. BCCONSAll is superior to CONSAll for MinT throughout the entire forecast period. For MaxT, CONSAll becomes superior to BCONSAll only from the 144 forecast hour and later. For Td, CONSAll mostly remains superior to BCCSONAll except for a couple of periods early in the forecast cycle.

Figure 10 is the same as figure 9 except it is for the remaining CR WFO (no mountains). BCCONSAll is still superior to CONSAll in the shorter forecast hours, but the magnitude is smaller when compared to Figure 8 when the "mountainous" WFO's were included. In addition, the forecast hour when CONSAll verifies better than BCCONSAll is 12 hours earlier compared to when the mountainous WFO's are included.

Examination of data from WFO Pueblo provides insights to the potential reasons why BCCONSAll performs better in complex terrain. Figure 11 shows the topography in GFE grid for WFO Pueblo and surrounding areas. The eastern third of the forecast area is comprised of the high plains with modest changes in topography. The western portion of the forecast area has very complex terrain with rapid changes in topography over short distances. The broad valley over the south central portion of the grid is the San Luis Valley, which is surrounded by high mountains. The central portion of the forecast area is the interstate 25 corridor. The main features are two pronounced west to east ridges. The northern ridge is the Palmer Divide and the southern ridge is the Raton Ridge.

Figure 12 shows the BCCONSAll MAE for the 36 hour MinT forecast from 0000 UTC data averaged over 250 grids during 2014. The MAE generally ranges from 3 to 4°F (1.7 to 2.2°C) over the entire domain with some subtle local maximums and minimums. Figure 13 is the same as figure 12 but for CONSAll. The most noticeable changes are the areas of large MAE in and near the complex terrain. One possible reason for bias correction improving the grids over complex terrains is that bias correcting the grids appears to help better map the data to complex terrain. Most of the forecast inputs into GFE are based on models with more coarse resolution. Inputs from these models are downscaled to the finer resolution GFE topography based on schemes which account for terrain differences. Apparently, the bias correction further refines this downscaling, accounting for the differences between the observation grid and input grid during the past 30 days.

Qualitatively, the results are similar for MaxT; however, the locations of the large MAE for MinT do not necessarily correspond to MaxT.

For Td (not shown) there does not appear to be much improvement with adjusting dew points in higher terrain. In general, bias correction, using the last 30 days of data, does not provide much improvement to the Td forecast, whether over complex terrain or flatter terrain.

5. ANALYSIS FOR FORERCAST INPUT RANKING

The verification of the various forecast inputs, using the same dataset as described in section 3, are ranked. A rank of 1 indicates the forecast input verifies the best (lowest MAE), and the highest rank indicates the input verifies the worst (highest MAE). Figures 14, 15, and 16 show the ranking of the forecast inputs for MaxT from the 1200 UTC cycle for forecast hours of 24, 72, and 120, respectively. The color coding for forecast inputs groups is the same as in earlier figures.

At 24 hours (Figure 14), the consensus forecast group generally have the lower ranks (lower MAE). For higher ranks (higher MAE), the number of occurrences for the consensus group is low, while the number increases for the guidance and deterministic groups.

As the forecast hour increases (72 hour in Figure 15 and 120 hour in figure 16), the lines for the deterministic model group members become more of a “U” shape. The deterministic models have the largest percentage of forecast inputs which verify the best or second best. For the middle rankings, the deterministic models group has the smallest number of occurrences. For the higher ranks (highest MAE), the number of occurrences increase rapidly.

The consensus group develops more of an “inverted U” shape with increasing forecast hour. In the longer forecast hours, the consensus models are less likely to be in the one of the best models, but they are also much less likely to be among the worst models. The guidance group also has more of a “U” shape with increasing forecast hour, but it is less pronounced than the deterministic model group.

This analysis suggests that the deterministic models have greater variability, especially in the longer forecast hours, than the guidance and consensus groups. Sometimes the greater variability in the deterministic models allow them to vary correctly resulting in a lower rank (lower MAE). However, this greater variability also allows for the deterministic model to vary incorrectly resulting in higher rank (higher MAE). The times the deterministic models vary correctly, as compared to the consensus group, does not counteract the times the deterministic group varies incorrectly. As a result, the deterministic models have a higher mean MAE compared to the other groups.

6. LIMITATIONS OF SINGLE DETERMINISTIC FORECAST

The analysis of mean MAE suggest that the “best” forecast for MaxT, MinT and Td, when averaged over many cases, will be either CONSAll or BCCONSAll, depending on the parameter (MaxT, MinT or Td) and forecast hour. While minimizing the MAE is a worthwhile goal for a single deterministic grid, this forecast can be of limited use to many users especially when there is a wide range of possible weather. The analysis of rank in the previous section shows that sometimes the deterministic models will perform better than the consensus model, and these could be instances where significant events or changes occur.

There appears to be a conflict between using the consensus forecast inputs, which minimizes MAE but can mask potential significant changes and extreme events, and using the deterministic models, which have higher MAE over many cases but can more likely highlight potential extreme events and significant changes. In the case of a potential significant weather event in the longer periods with a fairly low probability of occurrence, should forecaster use a consensus forecast or lean towards the more extreme event? An example will illustrate the difficulty.

On day 5, an arctic front may pass through a region. For locations ahead of the front, a high temperature of 50°F (10°C) and a low of 30°F (-1°C) are good forecasts. Behind the front, a good forecast for maximum temperature is 20°F (-7°C) and a minimum temperature of -10°F (-23°C). The models and ensembles have an equal probability for a city to be either ahead or behind the front. For this highly idealized situation, the forecast for this city which minimizes the MAE (and very likely would be CONSAll) would a high of 35°F (2°C) and a low of 10°F (-12°C). While this forecast would tend to minimize the MAE, given the wide range of possible solutions, it does not clearly convey the potential for a significant, sub-zero cold weather event.

Other scenarios could be thought of where a single deterministic grid, which over many cases would have the lowest MAE, would not provide information on the potential for significant, lower probability events. In these cases, other mechanisms could be used to convey the uncertainty in the forecast and the potential for significant events.

Finally, this study does not deter the use of local expertise in the grid preparation process. Significant local effects often cannot be accounted for in the GFE grid initialization process. One local example from last winter involved a late November snowfall in the San Luis Valley, which resulted in the development of a persistent cold pool for much of December and January. Most of the deterministic models and guidance were much too warm with daytime temperatures in early December, and the forecasters had to manually reduce the maximum temperature for a more accurate forecast. Eventually, bias correction (which is calculated from 30 days of data) was able to identify the bias and provided a more realistic temperature forecast. WFO Grand Junction, CO had similar problems last winter as well (Michael Meyers, personal communication).

7. COMPARISON TO ECMWF AND ECMWF GUIDANCE

As discussed in section 2, six hourly ECMWF grids are available on a 0.5° by 0.5° grid. These grids are typically unavailable in time for the routine forecast packages. In the comparison of the forecast inputs, the previous version of the ECMWF (labeled ECMWFPR) was used.

Additional ECMWF based guidance is also available in at the WFO's. The Meteorological Development Laboratory has developed MOS products based on higher resolution ECMWF data (Rudack et.al, 2014). The text products are only generated using 0000 UTC ECMWF data and they are available at the WFO's between around 0700 and 0900 UTC. One ECMWF based text product is the "ECS" guidance, which provides statistics up to the 72 hour forecast. The format of these products is similar to the MET and MAV guidance bulletins. Another ECMWF based text guidance is the "ECE", which has a similar format for the "MEX" text guidance providing statistical forecast up to 7 days in the future.

Similar to the "ADJ" products discussed in section 2, GFE grids utilizing these ECMWF text guidance bulletins are created using the ECMWF on a 0.5° by 0.5° grid as the background field. The "ADJECS" forecast input uses the "ECS" text products to adjust the background grid, and the "ADJECE" forecast input uses the "ECE" text products to adjust the background grid. Bias corrected versions of these grids are also created.

MDL guidance from the ECMWF (ECS and ECE) is only generated from the 0000 UTC run of the ECMWF. They typically are not available in time for incorporation into the generation of the routine forecast package on the overnight shift, which mostly uses information from the model runs initialized at 0000 UTC. Similar to the convention used for the ECWMF, the "PR" suffix in the forecast input name indicates that the data are from the previous run of the model.

The EC statistical guidance was not available in March of 2012, and an analysis of the ECMWF based forecast inputs was only performed on one year's worth of data from 1 March 2013 to 28 February 2014. The ADJECE, ADJECS, and their bias corrected grids are not routinely generated at all WFO's. A similar scheme is used as in section 2 to identify the forecast inputs commonly available during this period. Only forecast times at which all these forecast inputs are available at a WFO are used to verify the grids. For MaxT and MinT, the percentage of possible grids used was about 31% to 45%.

Figure 17 shows the verification for MaxT listing only the data points which verified well (0.5°F (0.28°C) or less difference from CONSAll). The forecast inputs derived from the ECS and ECE are in black. This plot also includes ADJECS, ADJECE, ECMWF, and their bias corrected counterparts, showing how well these forecast inputs from the ECMWF would verify if they are timely. That is, if the 0000 UTC ECMWF run based forecast inputs are available in time for the routine midnight shift package, which mostly uses data generated from 0000 UTC model runs, and 1200 UTC grids are available for the afternoon routine forecast package, which mostly uses data generated from the 1200 UTC model runs.

The forecast inputs ADJECEBC and ADJECSBC verify quite well. BCCONSAll verifies better than these forecast inputs up to 36 hours. At forecast hours 60 and 84, ADJECEBC verifies better than BCCONSAll, and at forecast hour 60, ADJECSBC verifies better than BCCONSAll. (These forecast inputs are only available from the 0000 UTC model runs.) The ADJECS and ADJECE guidance also performs well, and it is the best unaltered statistical guidance forecast input.

The ECMWFBC is the best deterministic model in this plot, and the MAE is comparable to the consensus forecast inputs. The maximum temperature of the ECMWF field is derived by taking the maximum temperature of 6 hourly temperature fields. The bias correction appears to be able to adequately determine the maximum temperature from a limited number of temperature fields.

Since the forecast inputs generated from the ECMWF are not timely, the previous versions (PR) of these ECMWF derived forecast inputs are used in the forecast process. Since ECMWF derived statistical guidance (ADJECS and ADJECE) is only generated from the 0000 UTC run of the ECMWF, the 0000 UTC ECMWF guidance is used in creating the "afternoon routine" forecast along with forecast inputs mostly derived from 1200 UTC initialized data. The ADJECEBCPR and ADJECSBCPR verify well, being only slightly worse than CONSAll through the 72 hour forecast cycle. This "12 hour old" guidance is still better than any other member of the guidance group available at that time. Similarly for the bias corrected raw ECMWF fields , the "12 hour old" bias correct ECMWF fields (ECMWFBCPR) are the best deterministic guidance available except for the GEMregBC (bias corrected Environmental Canada regional model), which is only available out to 48 hours.

Figure 18 shows the same plot but for MinT and Figure 19 shows the same plot for Td. For MinT the results are generally similar to MaxT. For Td, ADJECEBC and ADJECSBC are still the best verifying among the deterministic and guidance forecast inputs. However, CONSAll still is superior to all forecast inputs for nearly all forecast hours.

8. CONCLUSIONS

Two years of CR GFE verification data were analyzed to determine which forecast inputs have the lowest average MAE. For MaxT and MinT, BCCCONSAll generally has the lowest MAE up to the 84 forecast hour. For later periods, CONSAll has the lowest MAE. In complex terrain, bias correction appears to be more important. BCCONSAll was superior to CONSAll for nearly all forecast hours when averaging over the offices in Colorado and Wyoming. For Td, CONSAll is mostly the best forecast input for all forecast hours even in complex terrain.

For longer forecast hours, individual runs of the deterministic models tend to either verify well or verify poorly. The deterministic models tend to verify poorly much more frequently than they verify well.. As a result, the average MAE for the deterministic models is higher than other groups. The consensus forecast inputs tend to rank more in the middle for the longer forecast hours. While they infrequently are one of the best forecast inputs for a specific forecast, they are not among the worst verifying either. As a result, the average MAE for this group is the lowest.

The implications of using a single deterministic grid to convey a forecast was discussed. While a forecast which will have the lowest average MAE over many cases may be desirable, it will often not convey the potential for lower probability, higher impact events. Other means should be used to convey the possibility for lower probability events, which could have significant impacts.

Finally, forecast inputs derived from ECMWF based MOS products created by MDL was verified using 1 year of data. These forecast inputs verify well, and were comparable to the consensus group, epsecially if they were available in a timely manner.

9. REFERENCES

Rudack, D.E. , D. Ruth, K. Gilbert, and T. Curtis, 2014: A first look at the Meteorological Development Laboratory's experimental ECMWF MOS system. 26th Conference on Weather Analysis and Forecasting / 22nd Conference on Numerical Weather Prediction. Atlanta, GA, Amer. Meteor. Soc., J4.1. [Available online].

Figure 1. Map of central region showing forecast areas for each WFO.

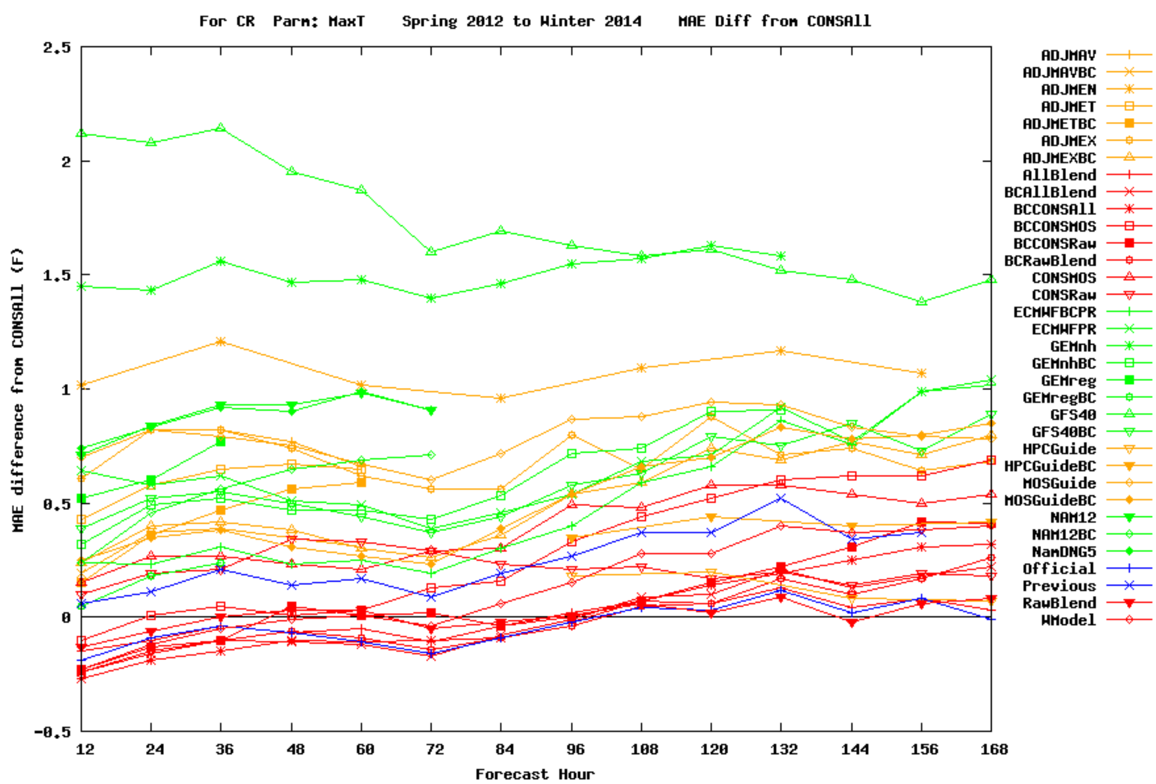

Figure 2. MAE for forecast inputs for MaxT for the entire CR from 1March 2012 to 28 february 2014. MAE is subtracted from the MAE for CONSAll for ease of comaprsion.

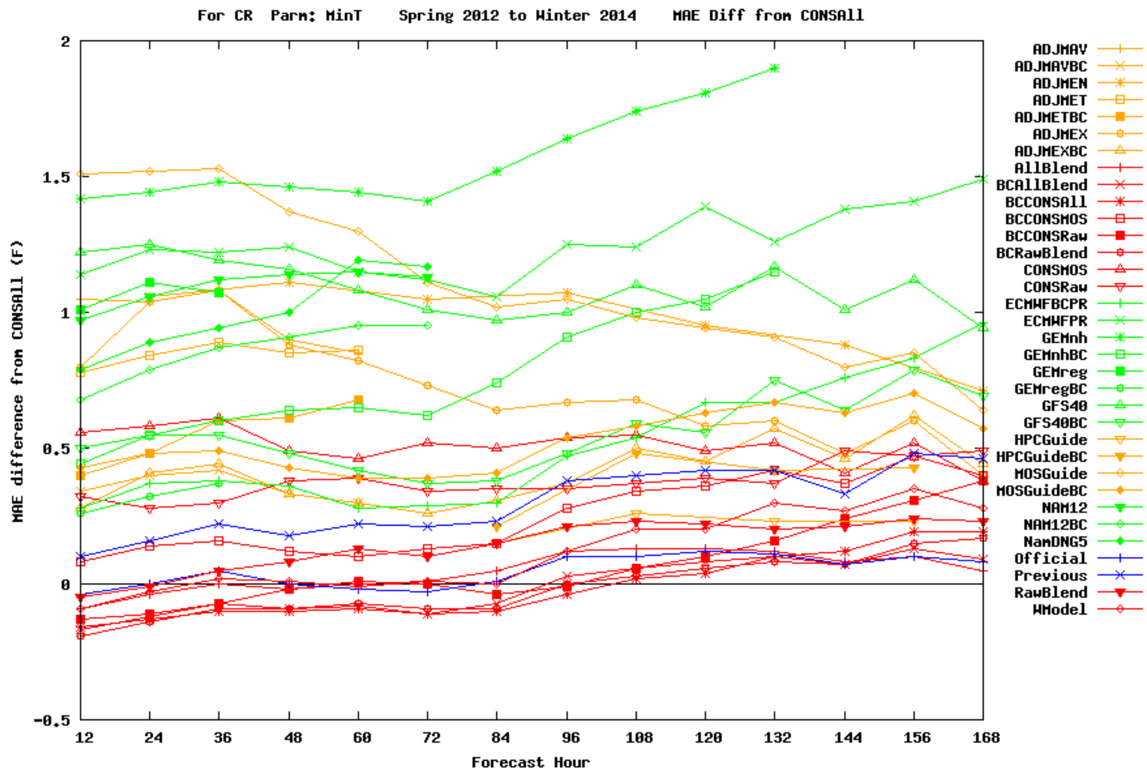

Figure 3. Same as Figure 2 but for MinT.

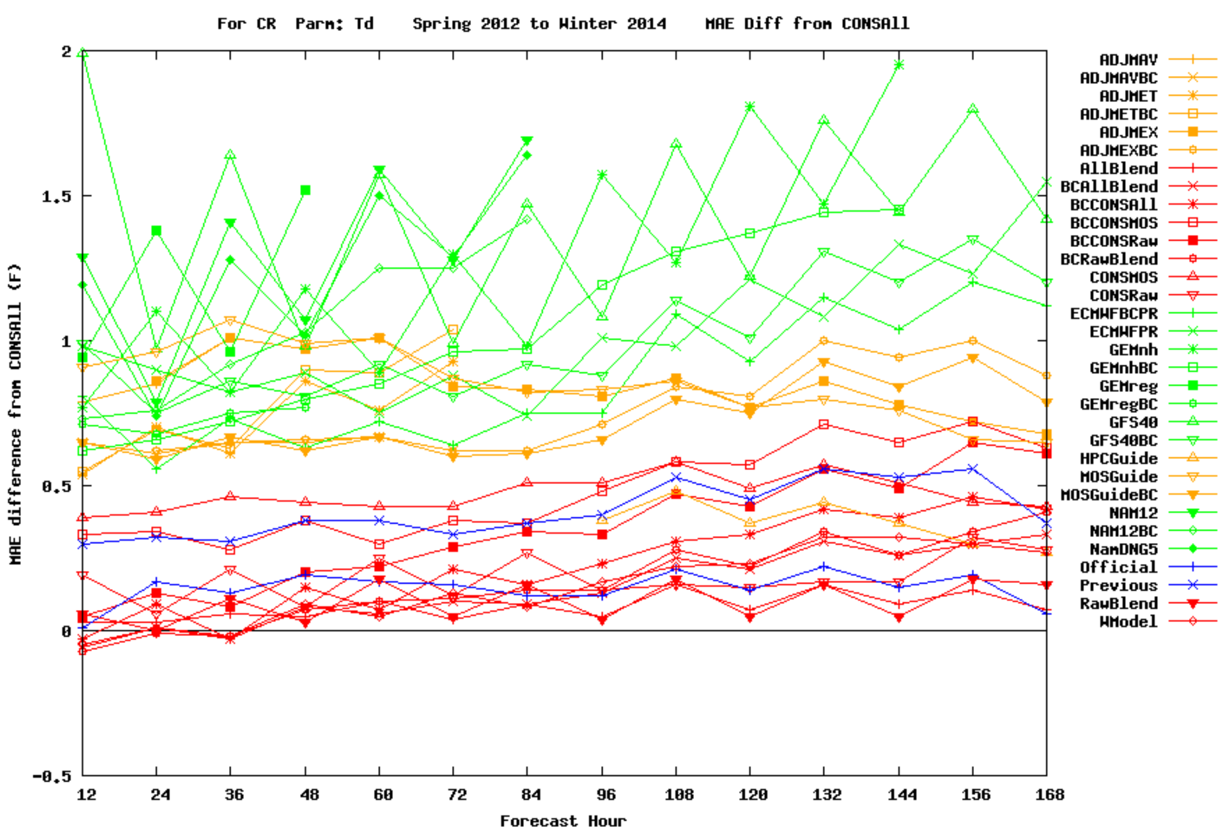

Figure 4. Same as Figure 2 but for Td.

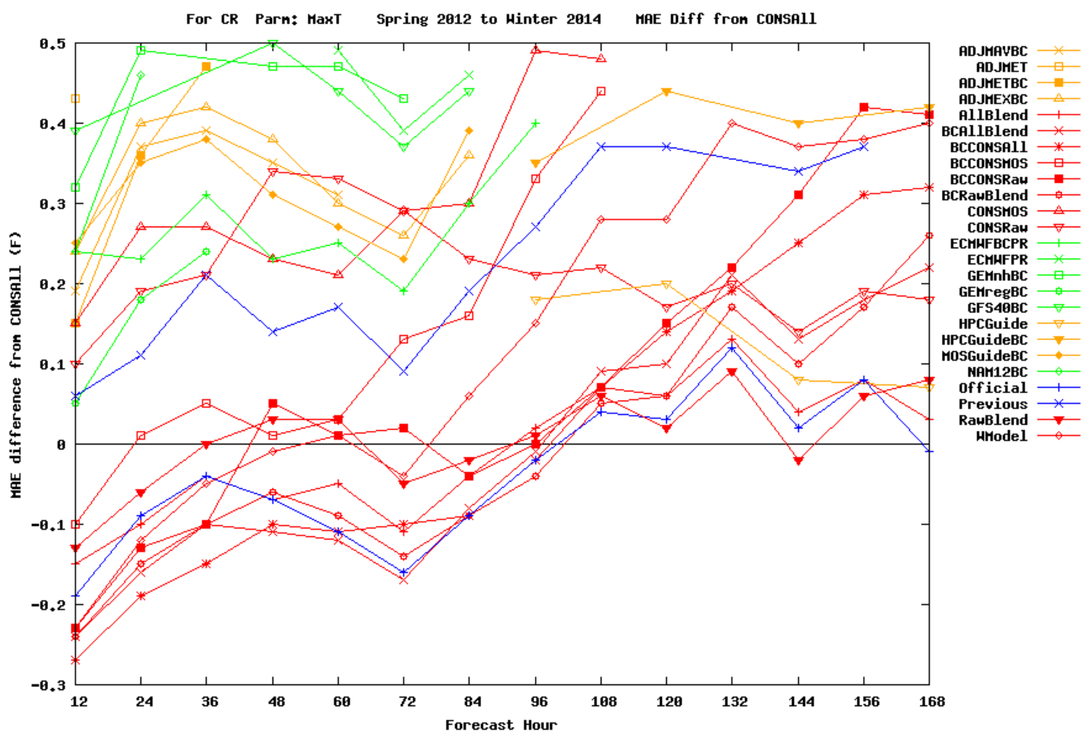

Figure 5. Same as Figure 2 but only displaying MAE differences of 0.5F or less.

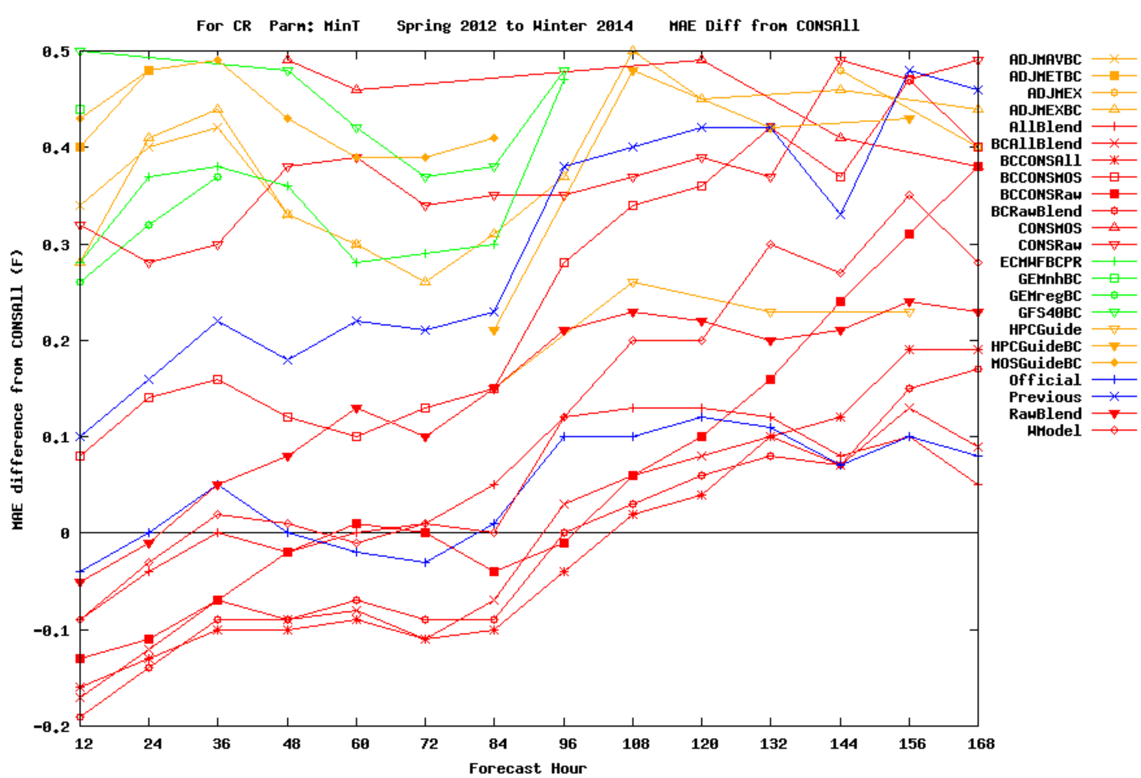

Figure 6. Same as Figure 5 but for MinT.

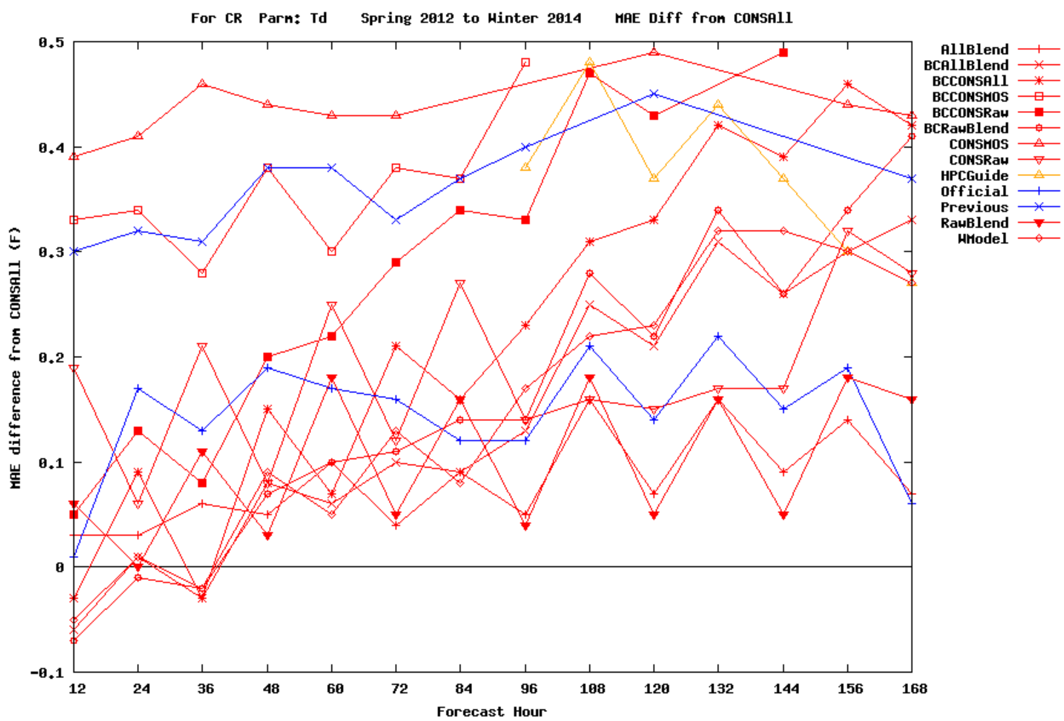

Figure 7. Same as Figure 5 but for Td.

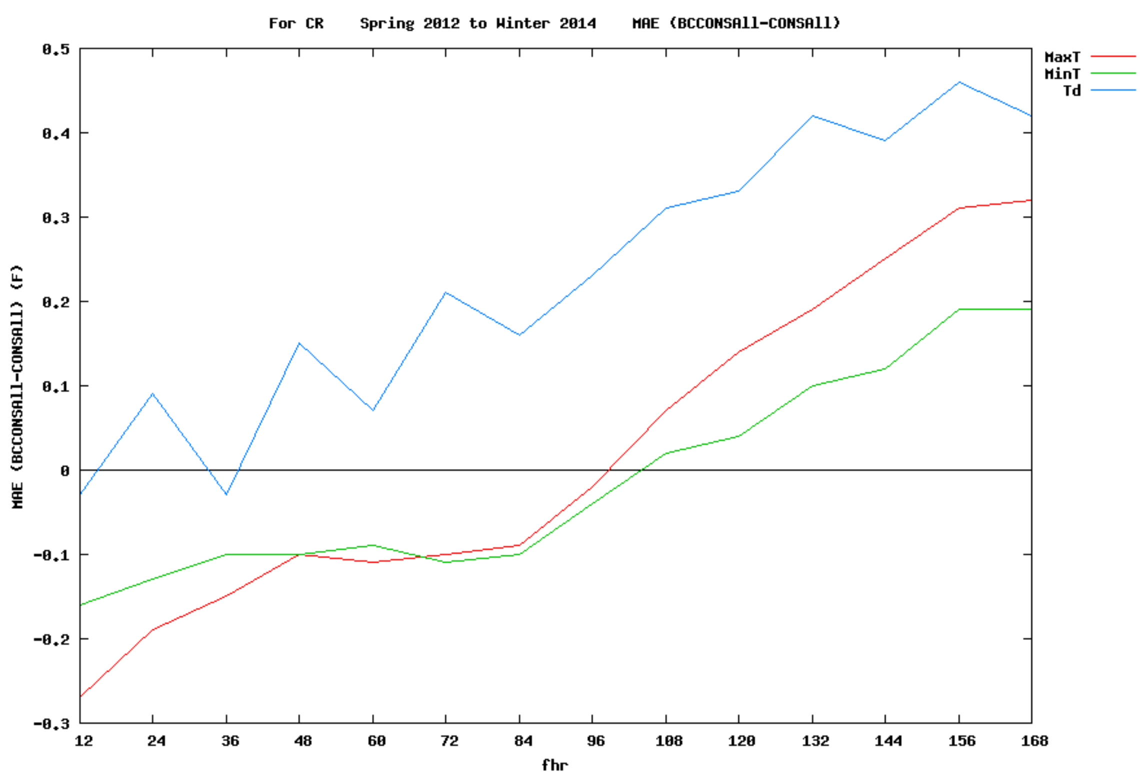

Figure 8. Difference between MAE for CONSAll and BCCONSAll for the entire central Region for MaxT, MinT and Td.

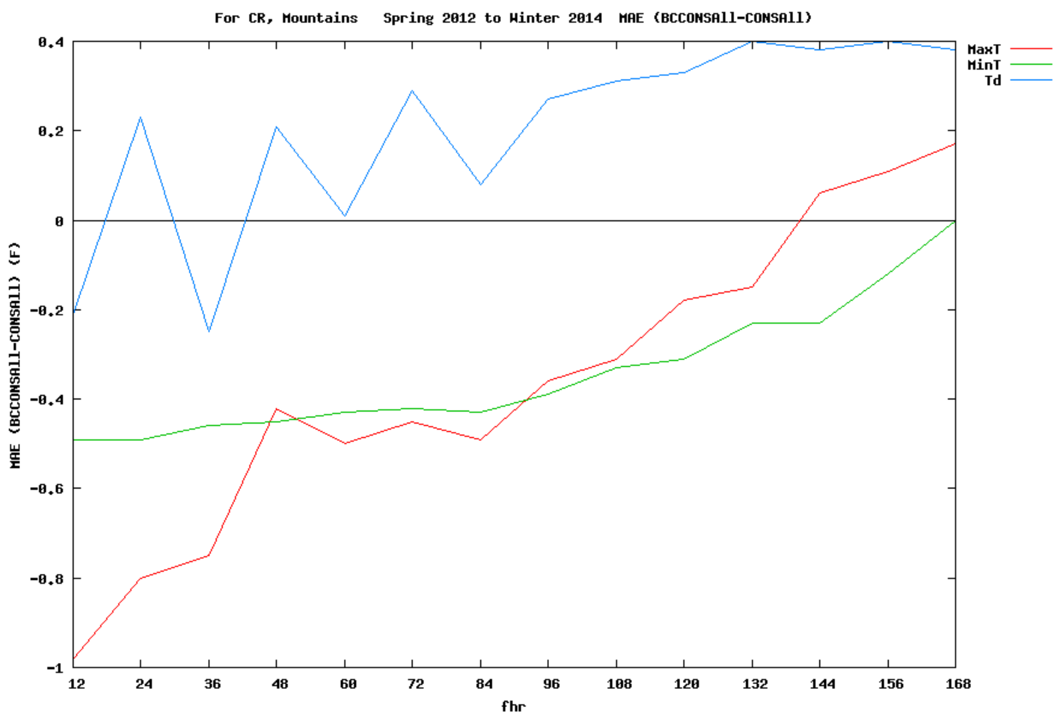

Figure 9. Same as Figure 8 but for the "Mountain" WFO's.

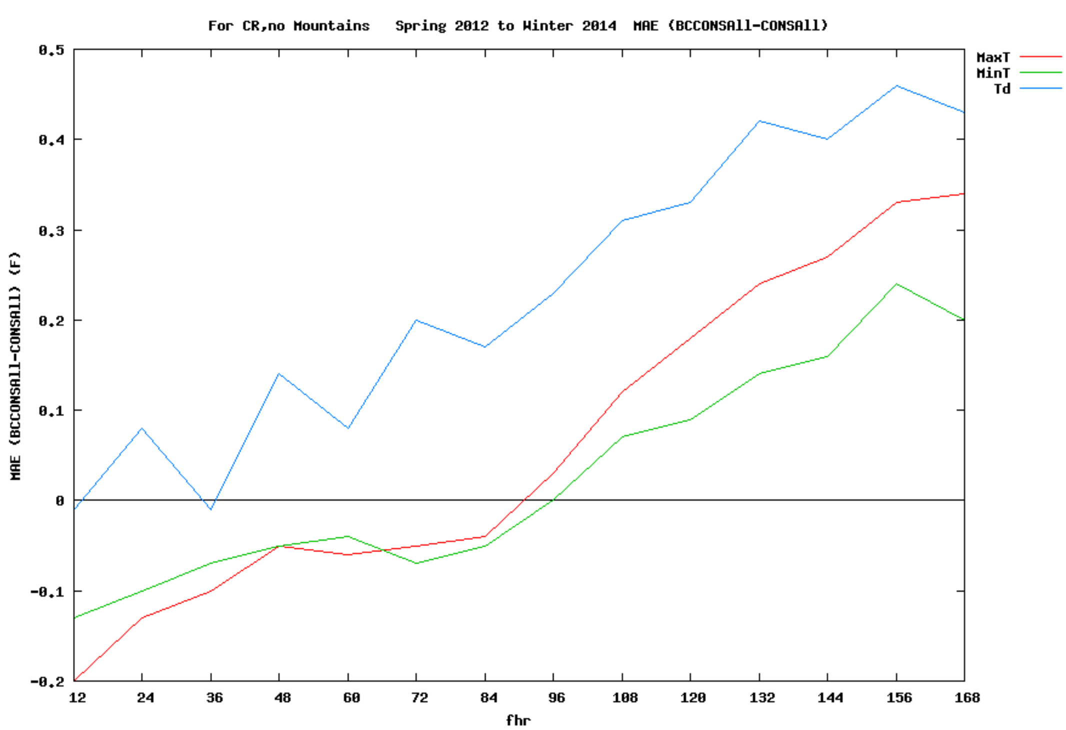

Figure 10. Same as Figure 8 but for the "no mountains" WFO's.

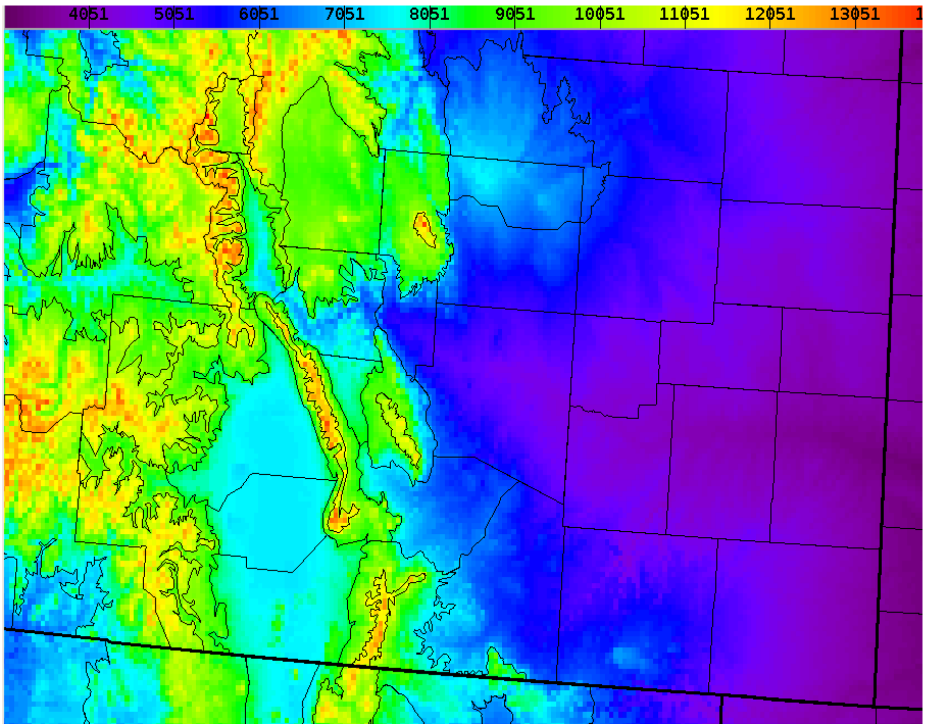

Figure 11. Topography in GFE grid for WFO Pueblo. Heights are in feet.

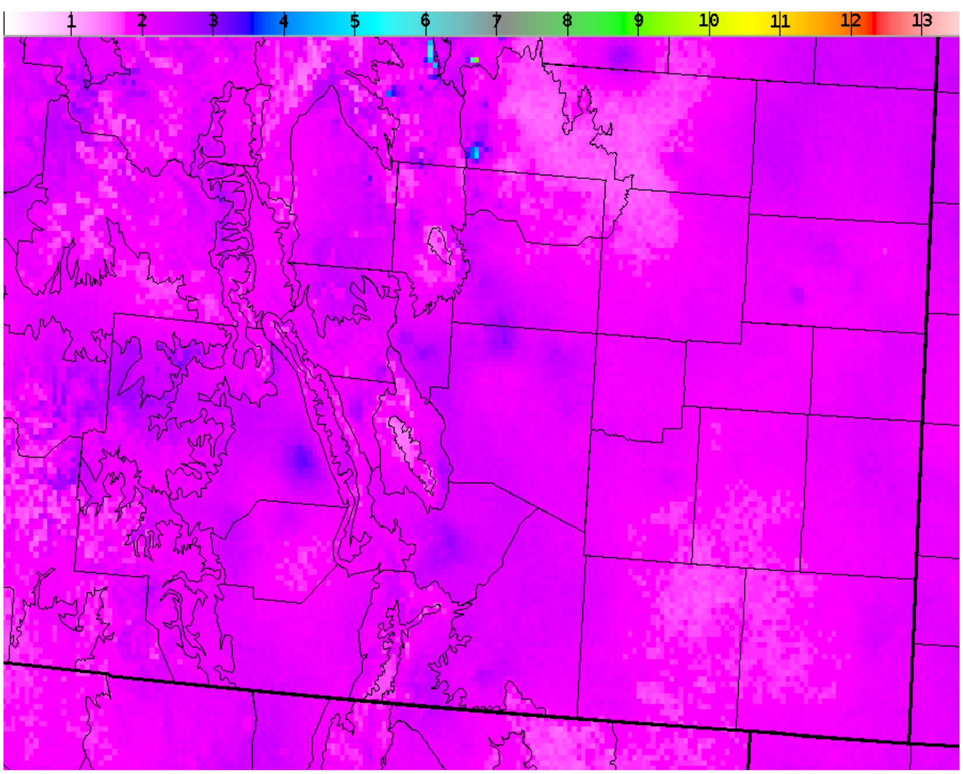

Figure 12. Mean MAE grid for BCCONSAll MinT at 36 forecast hour for WFO Pueblo. Values are in F.

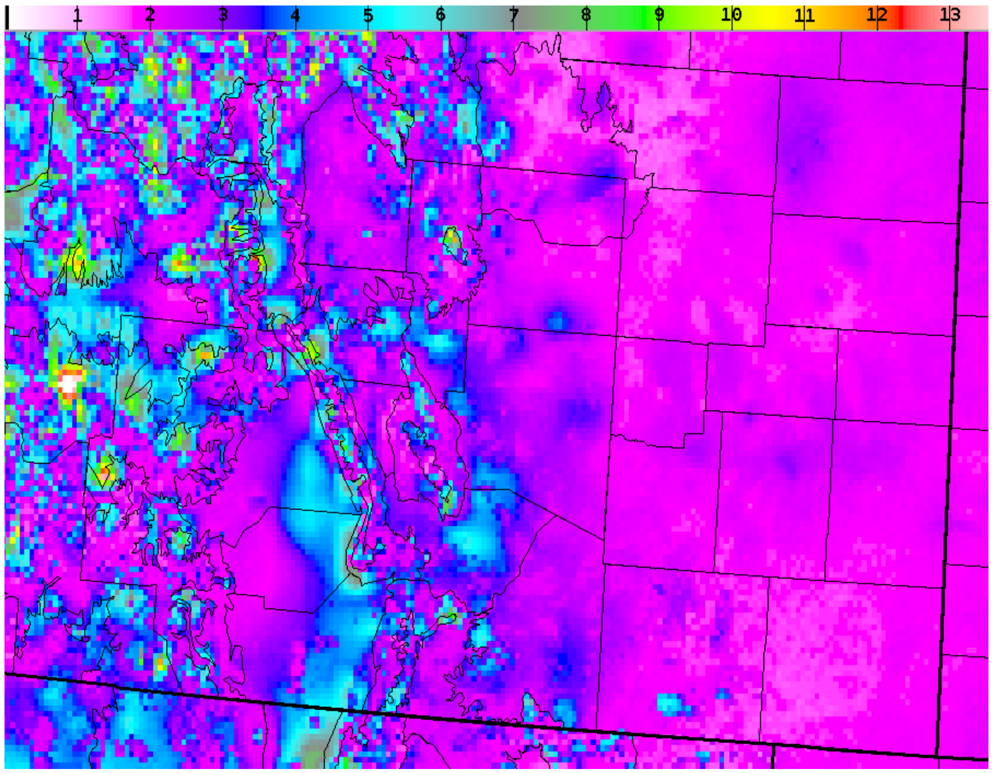

Figure 13. Same as Figure 12 but for CONSAll. Values are in F.

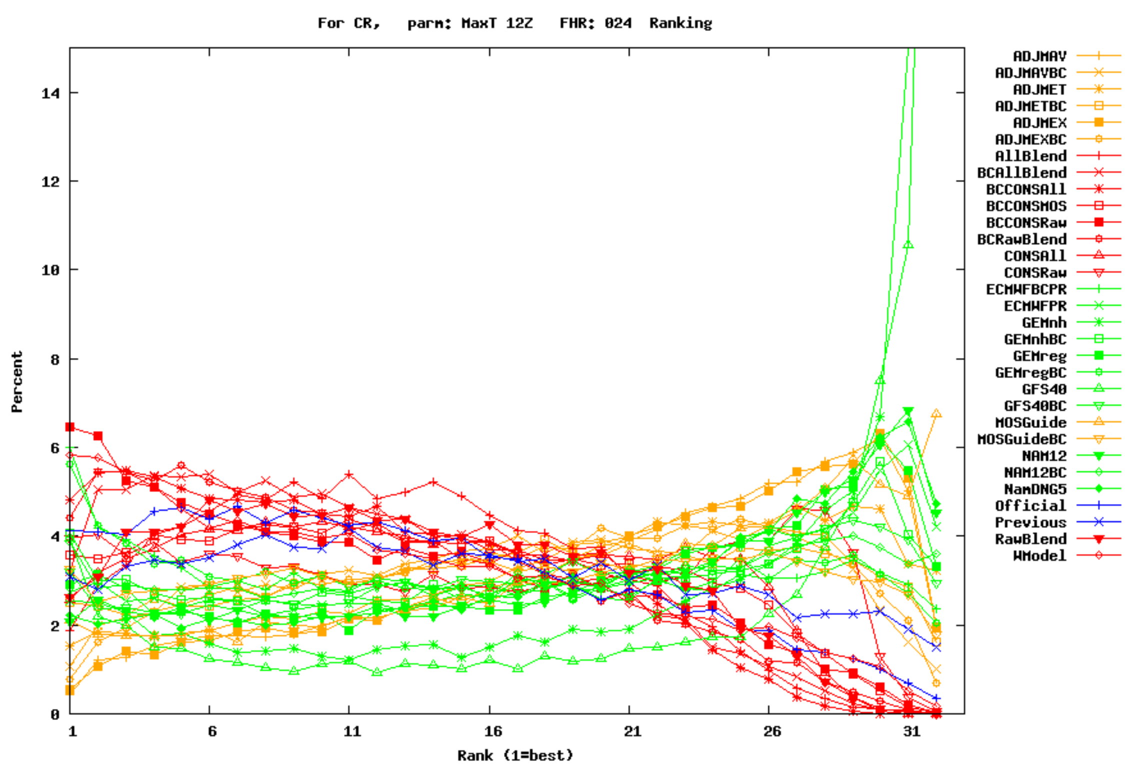

Figure 14. Rank for forecast inputs for MaxT from 1200 UTC initialized data at forecast hour 24.

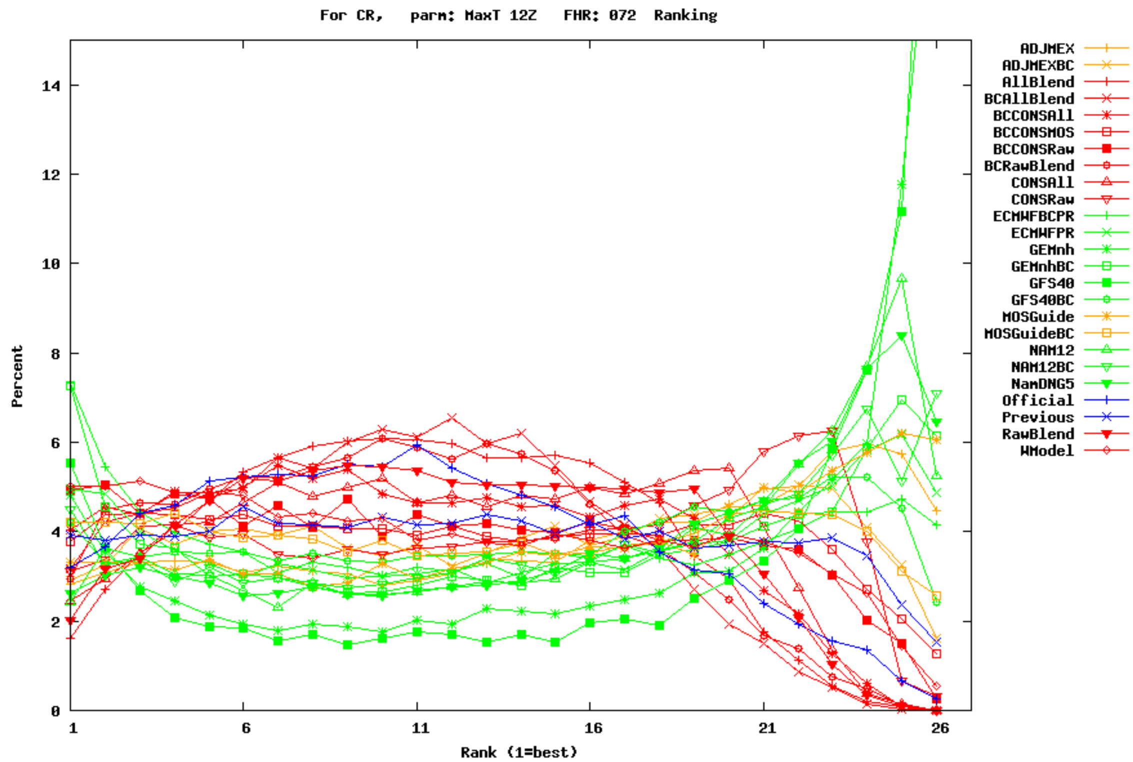

Figure 15. Same as Figure 14 but for forecast hour 72.

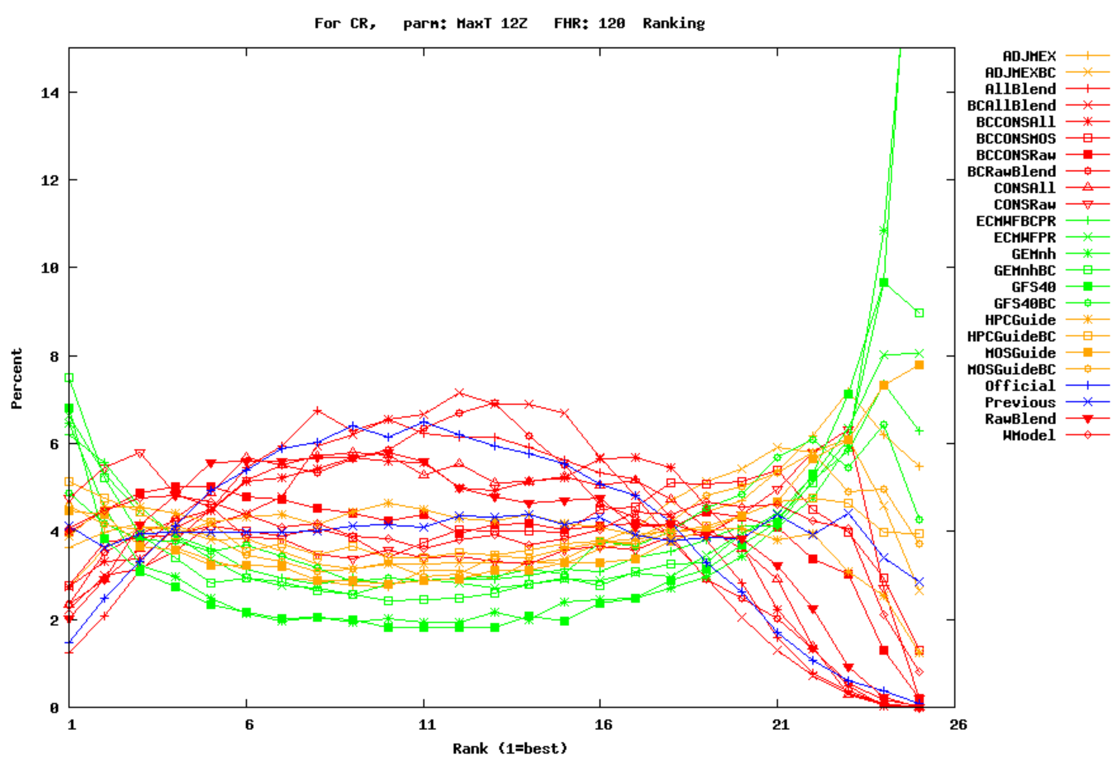

Figure 16. Same as Figure 14 but for forecast hour 120.

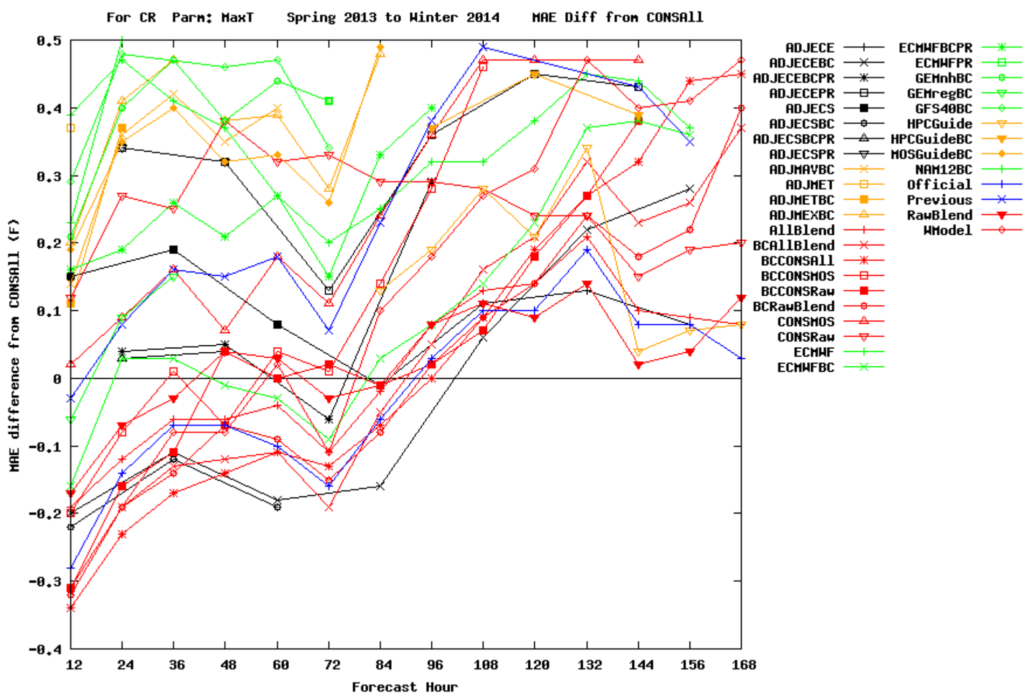

Figure 17. Rank of forecast inputs for MaxT. Rank of 1 is the best.

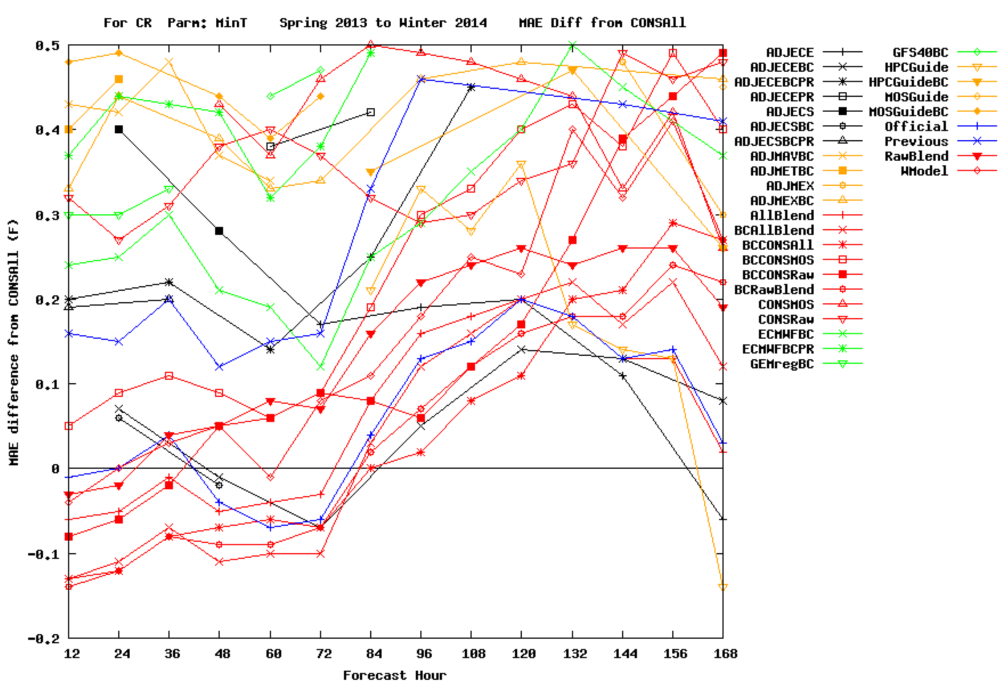

Figure 18. Same as Figure 18 but for MinT.

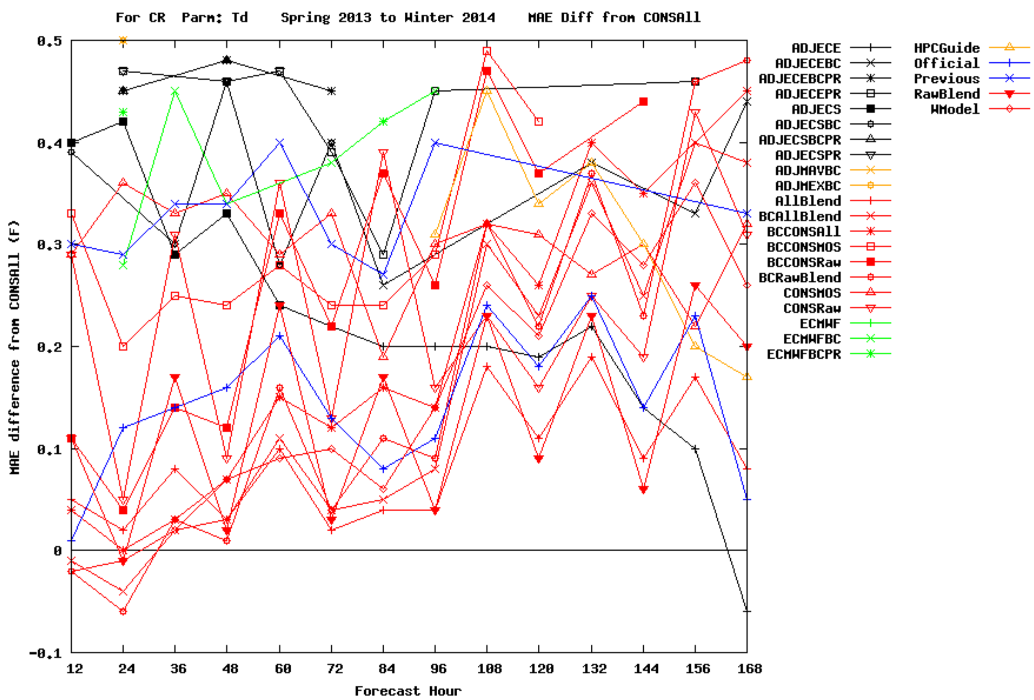

Figure 19. Same as Figure 17 but for Td.