Local GFE "Poor Man's" Ensemble

October 1st 2004

Background

In response to the SOO/DOH whitepaper on IFPS several efforts are being made by different groups to provide better first-guess grids in the GFE. The SOO/DOH whitepaper outlined several areas of development and research needed to improve and somewhat streamline the grid editing process. In addition to downscaling model data, new methods and approaches to bias removal and statistical correction of raw model output were highlighted.

This informal write-up describes an approach being tested at WFO TWC since late 2002 to improve first-guess grids for maximum and minimum temperatures using both a simple bias correction and a time-lagged short-term ensemble to derive summary statistics and probabilistic information in the GFE. The method is limited to the data sources currently available to a WFO via AWIPS, but it is flexible enough to be extended to other sensible weather elements.

Poor Man's Ensemble

Several researchers have found that a multi-model and multi-analysis ensemble system is more skillful compared to a single model ensemble system with either multi-physics and/or perturbed initial conditions, especially in short range prediction. A cheap way to capitalize on these findings is to utilize the different deterministic models run by operational NWP centers and/or locally. This approach is referred to as a "Poor Man's" ensemble by researchers and modelers. To increase the effective ensemble size without adding additional models, a time-lagged (time phased) system can be constructed by using previous model forecasts all valid at the same projection time.

Please note that the term "ensemble" is used loosely here and should not be confused with a traditional NWP ensemble system. Rather, the method described here is an attempt to treat the different guidance sources as one entity and thereby present this collective information in summary products and probabilistic form. When formulating a decision, it makes sense to use as much diverse information as possible. The new era of gridded weather forecasting requires forecasters to rely heavily on model output from an increasing number of sources. The future addition of information from traditional NWP ensemble systems like NCEP's SREF and GFS will only add to the time consuming task of sifting through all sources before making a gridded forecast or decision on whether to issue a long fused watch, warning or advisory.

Adding to this volume of information is the increased cycle frequency at which new guidance products become available for use. While the most recent guidance might be more accurate, it can sometimes be wise not to ignore the previously issued guidance. This previous guidance whether converging or diverging from the latest guidance can still provide some useful information. This premise is part of the justification for incorporating the previous model cycles in the ensemble.

The current time-lagged ensemble being tested utilizes the capabilities of the GFE to store and manage the different models that make up the ensemble. GFE procedures calculate model forecast errors, and derive an ensemble weighted average based on each member's past performance along with the ensemble mean, range, spread, and high and low extreme grids. The models that currently make up the ensemble include bias corrected raw SmartInit grids as well as MatchMOS grids for both the ETA and GFS. For bias removal of the raw model output, a simple 7 day running mean error is applied to each model cycle and forecast lead time. By incorporating the 2 different but valid bias removal methods (mean error and MOS), the effective ensemble size can be increased even more. To determine each member's performance, the MatchOBS analysis grid is currently used as the "Record of Analysis". An additional step is taken to adjust the MatchOBS analysis of maximum and minimum temperatures with the daily COOP data from the Arizona Regional Temperature and Precipitation Tables (RTP) stored in AWIPS. Refer to figure 1 for the flow diagram of the GFE processes.

Operational Use of the Ensemble
The weighted average of the ensemble based on past performance of each member tries to nudge the forecast to the best possible consensus solution. This is based on the premise that a consensus or mean of the ensemble will be more skillful than any given member. Statistically, researchers have found this to be true in the big picture and the nature of the statistics favor the mean. However, it must be noted that the weighted consensus or the straight mean may not be more skillful than a single member for a given weather situation.

One way to judge the skill of an ensemble consensus or mean is to evaluate the spread or dispersion about the mean. Forecasts with a large spread and range between the two extremes are an indication of greater uncertainty in the model solutions. A smaller spread and range between the two extremes suggests more confidence in the forecasts. In the case of a lagged average ensemble, a smaller dispersion about the mean can also imply small run-to-run differences or consistency in the forecasts; hence, convergence on one solution.

The spread or dispersion about the mean can indicate problem areas that may reveal analysis errors that propagate through the modeling system. In the case of the method described here, a large spread can be an indication of errors in the "Record of Analysis" being used for bias removal. Large spreads may also reveal information about model bias, especially in a single model lagged average forecast. No matter what the cause is of the large spread or dispersion, there still is greater uncertainty in these areas.

Ensemble Grids
This GFE implementation is the first step toward introducing bias removed raw SmartInit grids and an ensemble forecast in the grid editing process. The following is a list of weather element names that will be in the GFE under the **"ModelBlend"** grid:

Ensemble Style Summary Grids:
MaxT: Is the performance weighted average.

MaxTmean: Is the average or arithmetic mean.
MaxTspread: Is the ensemble spread.
MaxTHi: Is the highest value in the ensemble.
MaxTLo: Is the lowest value in the ensemble.
MaxTRange: Is the range between the 2 extremes of MaxTHi and MaxTLo.

MinT: Is the performance weighted average.
MinTmean: Is the average or arithmetic mean.
MinTspread: Is the ensemble spread.
MinTHi: Is the highest value in the ensemble.
MinTLo: Is the lowest value in the ensemble.
MinTRange: Is the range between the 2 extremes of MinTHi and MinTLo.

Currently there are no tools for populating or editing using these summary grids. The best way to utilize these grids is to load them into the spatial editor and browse through them for guidance purposes. If you like what you see in the ensemble mean, high, low, or weighted blend, populate the FCST grids by using the "Copy Selected Grid From" option in the GFE. You can also use the ModelBlend smart tool to populate.

Probabilistic grids are also being generated for several select thresholds like the probability of 100F or greater. The probability for a threshold is calculated as the percentage of the forecast members that meet or exceeded the predetermined threshold value. There might be some seasonal utility in these probabilities. However, at this time it could be a bit overwhelming to incorporate these thresholds in the gridded forecast process other than using them as guidance to issue products like a Freeze Warning or Heat Warning. In addition to the probabilistic grids in the GFE, there is some point probabilistic guidance for low temperatures to support the Operation Deep Freeze program this winter. The tables are in html format and can be viewed using the internet browser on AWIPS. Figures 2 through 4 are screen shots of the probabilistic grids and threshold tables.

Threshold Probabilistic Grids:
MaxTgtrEq20- Is the probability of MaxT being equal to or greater than 20F.
MaxTgtrEq30 –""
MaxTgtrEq40 –""
MaxTgtrEq50 _ ""
MaxTgtrEq60 – ""
MaxTgtrEq70 – ""
MaxTgtrEq80 – ""
MaxTgtrEq90 – ""
MaxTgtrEq100 – ""
MaxTgtrEq110 – Is the probability of MaxT being equal to or greater than 110F.

MinTgtrEq10 – Is the probability of MinT being equal to or greater than 10F.
MinTgtrEq20 – ""
MinTgtrEq30 – ""

MinTgtrEq40 – ""
MinTgtrEq50 – ""
MinTgtrEq60 – ""
MinTgtrEq70 – ""
MinTgtrEq80 – ""
These probabilities have not been calibrated on the back end of the processing, but they're derived from the bias corrected grids and the gridded point MOS. In the next few years we should start seeing probabilities derived from the SREF and GFS ensembles in AWIPS. Until then, this approach should help in providing the forecaster with meaningful guidance for making watch, warning or advisory decisions like in the case of freezing temperatures or extreme heat.

Future Work

The RUC model and additional elements will be included in the ensemble this winter. The core elements will be Max/Min temp, hourly temp and dew point, POP6 and QPF6, wind, and cloud coverage. Several different packages will be used for the "Record of Analysis" in the calibration/bias removal of each element. For example, LAPS will be used for clouds and the NCEP stage IV precipitation grids for POP6 and QPF6.

Probabilistic World of Grid Editing

There is a push to provide our customers with probabilistic grids of sensible weather based on NWP ensembles. This will be a big problem for the current IFPS approach which is based on a deterministic forecast. Although more true to the science, probabilistic forecasts in the grid editing world will be a major challenge. How do you edit probabilistic intervals, thresholds or even quantiles? The best example of this would be QPF with probability forecasts in intervals/bins of .25 of an inch of precipitation (.01-.25, .25-.50, .50-1.0, 1.0-1.25, etc) that form a histogram. For spatial edits, this would be a lot of grids to view and adjust. One approach would be to edit the ensemble histogram derived from the entire grid domain or a local effect area and make relative adjustments to each grid point.

Another method would be to fit the ensemble distribution to a known theoretical Probability Density Function (PDF) (Gaussian or Gamma) and Cumulative Distribution Function (CDF). This approach cuts down on the number of grids necessary to represent the ensemble distribution since only the parameters and the type of theoretical distribution for a grid point need to be stored. This fitting is usually done as part of the statistical post-processing of ensemble output. Because of systematic biases in NWP models and the typical under dispersion of current ensemble systems, there will continue to be a need for some form of post-processing.

There are several post-processing techniques, all of which have pros and cons. Rather than trying to find the best approach, each method could be treated as a separate entity in the probabilistic forecast process even if derived from the same ensemble system. Several differently designed NWP ensemble systems exist today such as NCEP's SREF and GFS with overlapping time projections. More ensemble systems will likely evolve in the future. With this in mind, grid editing of ensemble based probabilistic forecasts could be done by the rather popular forecast process of subjectively blending or nudging several different deterministic models together. In the case of probabilistic forecasts, a forecaster will be presented with PDFs derived from the different NWP ensemble systems and/or different post-processing methods as guidance. The forecaster would then blend or nudge the PDFs to create the official probabilistic weather forecast. Diagnostic/verification tools similar in concept to the ones used for deterministic forecasts could aid forecasters in selecting the best ensemble system/post-processing method, and for determining the relative weights used in blending the multiple ensemble systems. This approach would fit nicely with the national centers which could provide the field forecaster with their best guess in the form of adjusted PDFs.

In light of multiple ensemble systems, what role will the current operational deterministic models play? Could they possibly form some type of "Poor Man's" ensemble that can be used as additional input to the probabilistic forecast methodology?

At any rate, hopefully the baby step approach presented here will provoke some discussion, and at the same time provide useful guidance information to the forecaster.

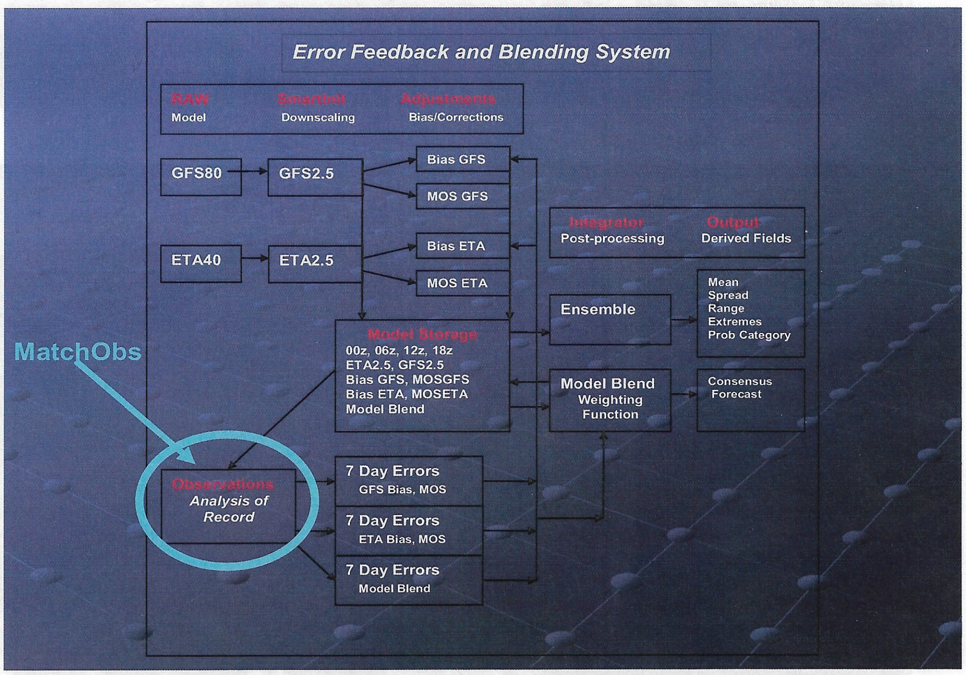


Figure 1.

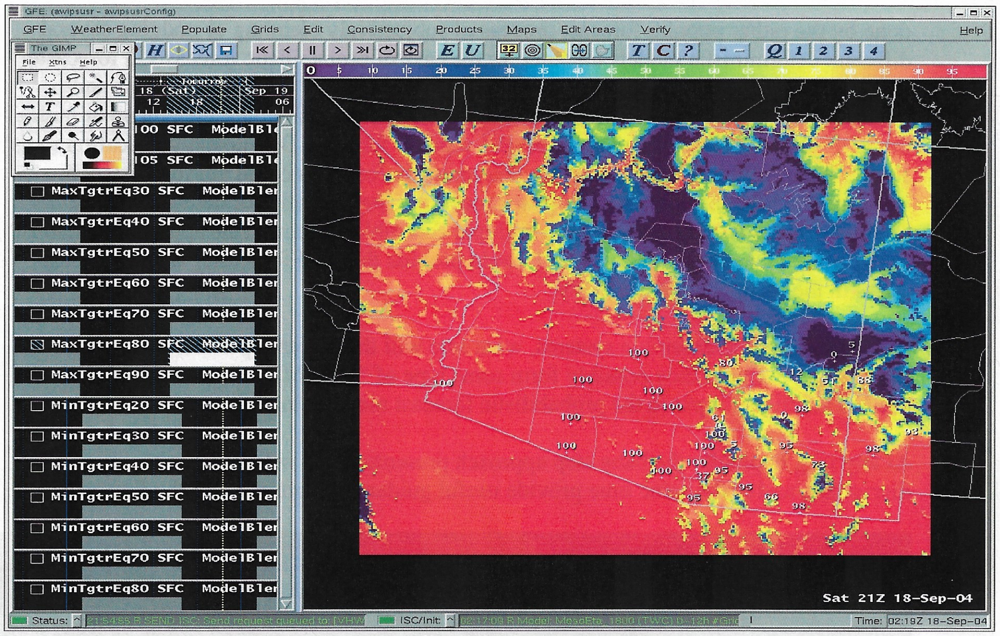


Figure 2.

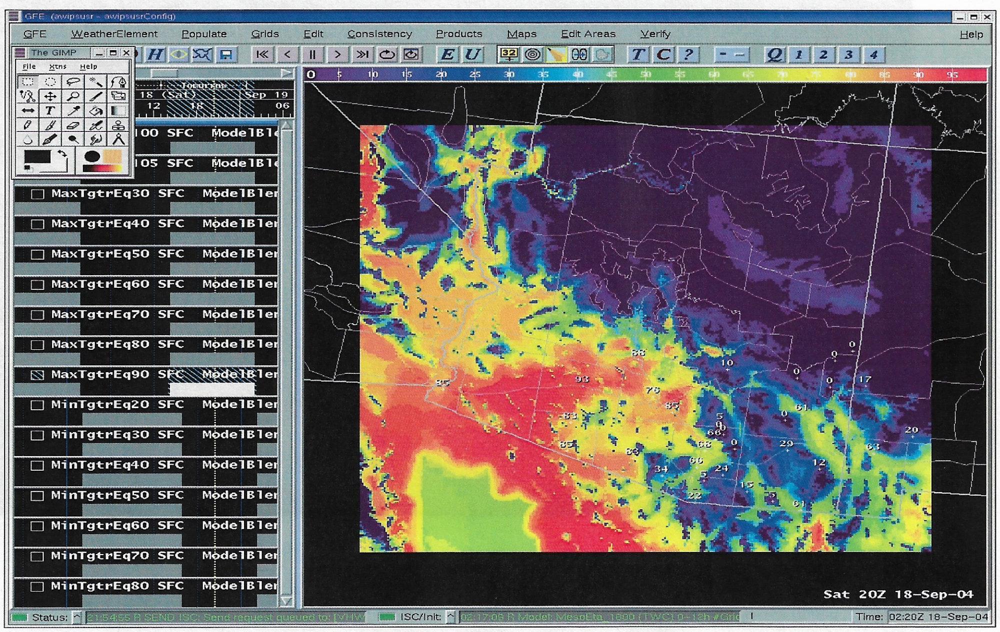
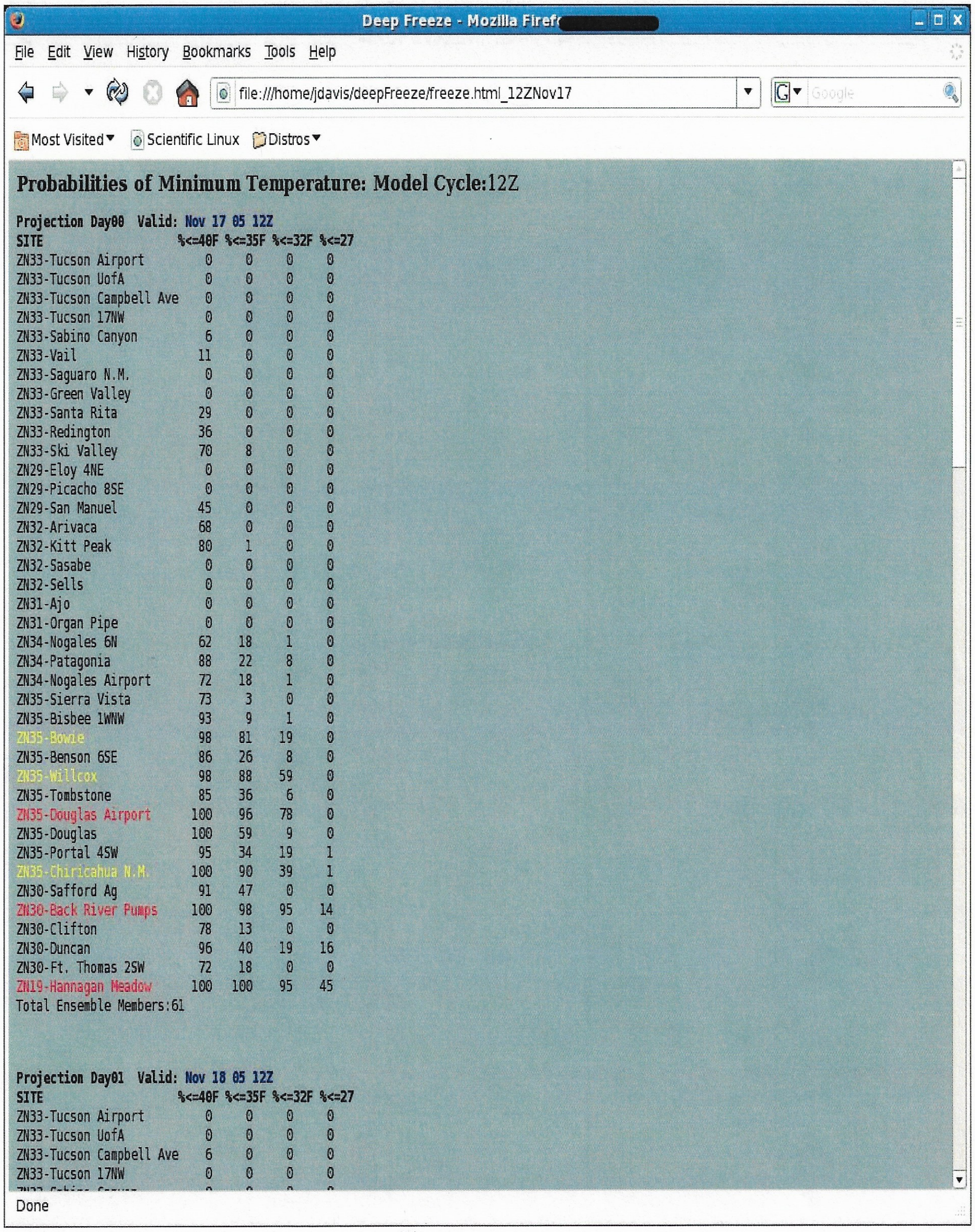


Figure 3.



Probabilities of Minimum Temperature: Model Cycle:12Z

Projection Day00 Valid: Nov 17 05 12Z

SITE	%<=40F	%<=35F	%<=32F	%<=27
ZN33-Tucson Airport	0	0	0	0
ZN33-Tucson UofA	0	0	0	0
ZN33-Tucson Campbell Ave	0	0	0	0
ZN33-Tucson 17NW	0	0	0	0
ZN33-Sabino Canyon	6	0	0	0
ZN33-Vail	11	0	0	0
ZN33-Saguaro N.M.	0	0	0	0
ZN33-Green Valley	0	0	0	0
ZN33-Santa Rita	29	0	0	0
ZN33-Redington	36	0	0	0
ZN33-Ski Valley	70	8	0	0
ZN29-Eloy 4NE	0	0	0	0
ZN29-Picacho 8SE	0	0	0	0
ZN29-San Manuel	45	0	0	0
ZN32-Arivaca	68	0	0	0
ZN32-Kitt Peak	80	1	0	0
ZN32-Sasabe	0	0	0	0
ZN32-Sells	0	0	0	0
ZN31-Ajo	0	0	0	0
ZN31-Organ Pipe	0	0	0	0
ZN34-Nogales 6N	62	18	1	0
ZN34-Patagonia	88	22	8	0
ZN34-Nogales Airport	72	18	1	0
ZN35-Sierra Vista	73	3	0	0
ZN35-Bisbee 1WNW	93	9	1	0
ZN35-Bowie	98	81	19	0
ZN35-Benson 6SE	86	26	8	0
ZN35-Willcox	98	88	59	0
ZN35-Tombstone	85	36	6	0
ZN35-Douglas Airport	100	96	78	0
ZN35-Douglas	100	59	9	0
ZN35-Portal 4SW	95	34	19	1
ZN35-Chiricahua N.M.	100	90	39	1
ZN30-Safford Ag	91	47	0	0
ZN30-Back River Pumps	100	98	95	14
ZN30-Clifton	78	13	0	0
ZN30-Duncan	96	40	19	16
ZN30-Ft. Thomas 2SW	72	18	0	0
ZN19-Hannagan Meadow	100	100	95	45

Total Ensemble Members:61

Projection Day01 Valid: Nov 18 05 12Z

SITE	%<=40F	%<=35F	%<=32F	%<=27
ZN33-Tucson Airport	0	0	0	0
ZN33-Tucson UofA	0	0	0	0
ZN33-Tucson Campbell Ave	6	0	0	0
ZN33-Tucson 17NW	0	0	0	0

Done

Figure 4.