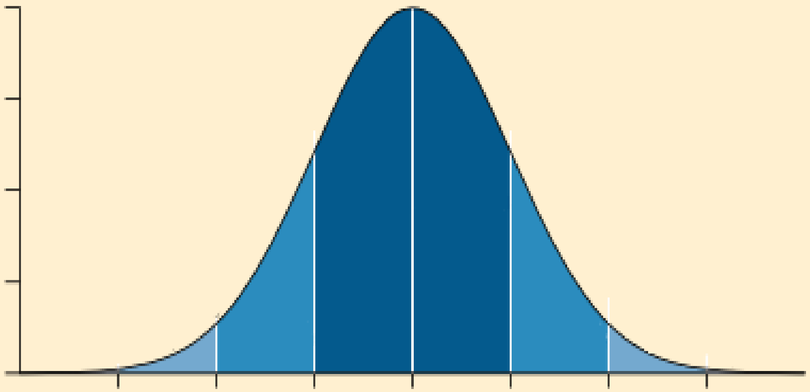
Probability Distribution Forecasts of a Continuous Variable

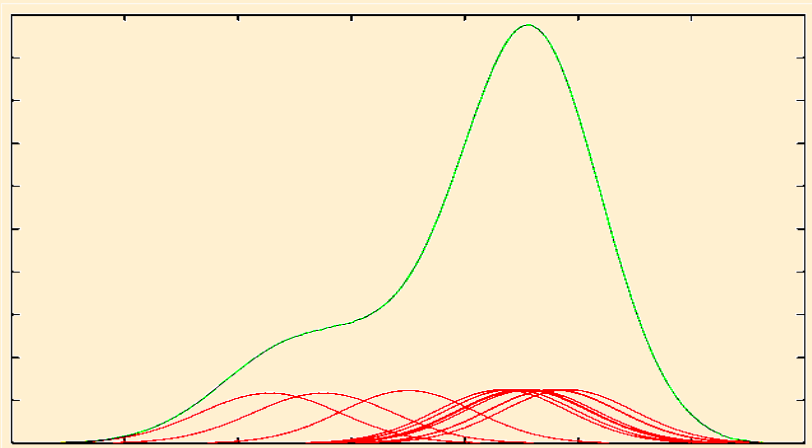
Meteorological Development Lab

October 2007

Overview

- Outputs
- Tools and concepts
- Data sets used
- Methods
- Results
- Case Study
- Conclusions
- Future Work





Uncertainty in Weather Forecasts

It is being increasingly recognized that the uncertainty in weather forecasts should be quantified and furnished to users along with the single value forecasts usually provided.

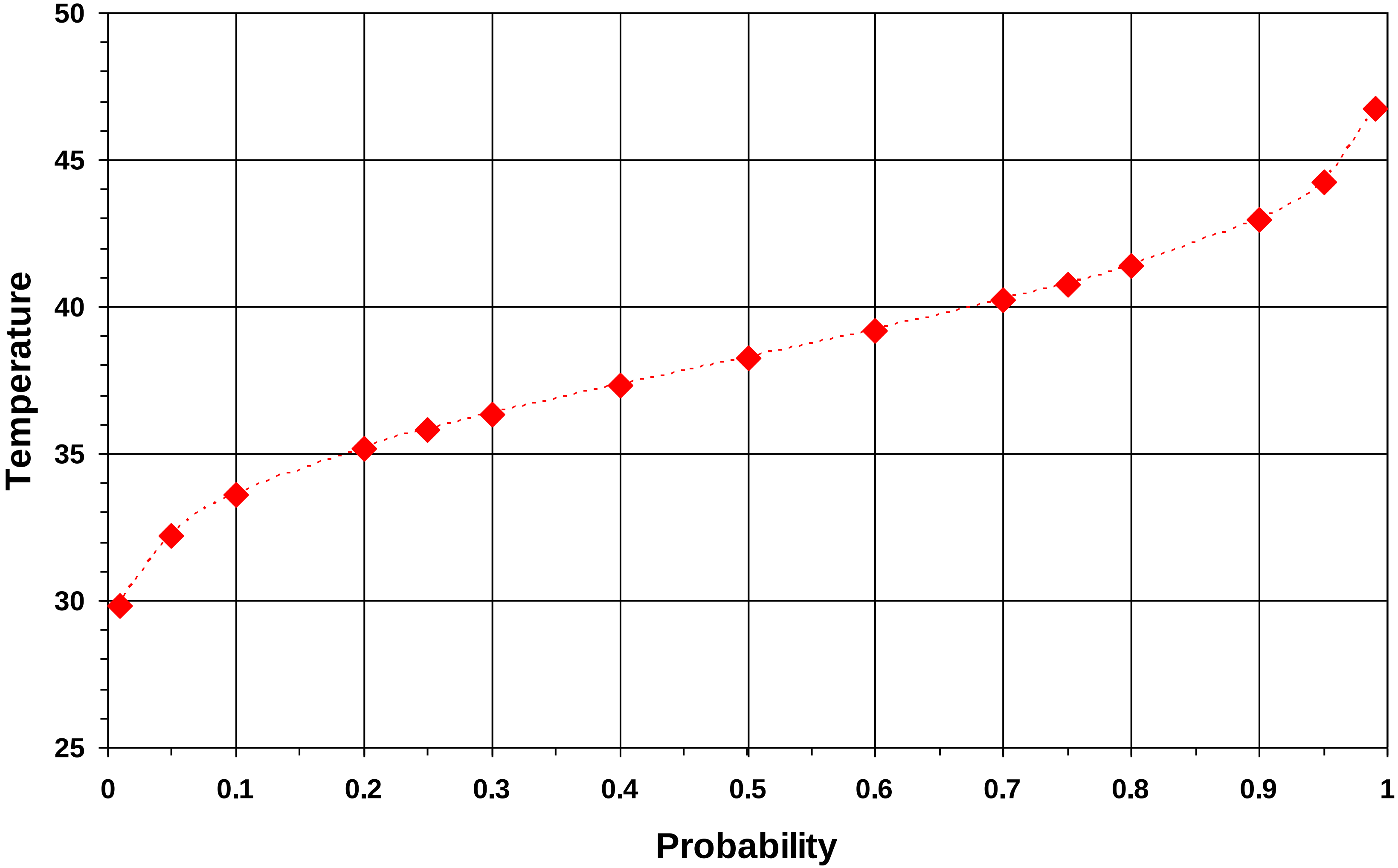
MDL's goal is to provide probabilistic guidance for all surface weather variables in gridded form in the National Digital Guidance Database (NDGD).

Outputs

How do we provide probabilistic forecasts to our customers and partners?

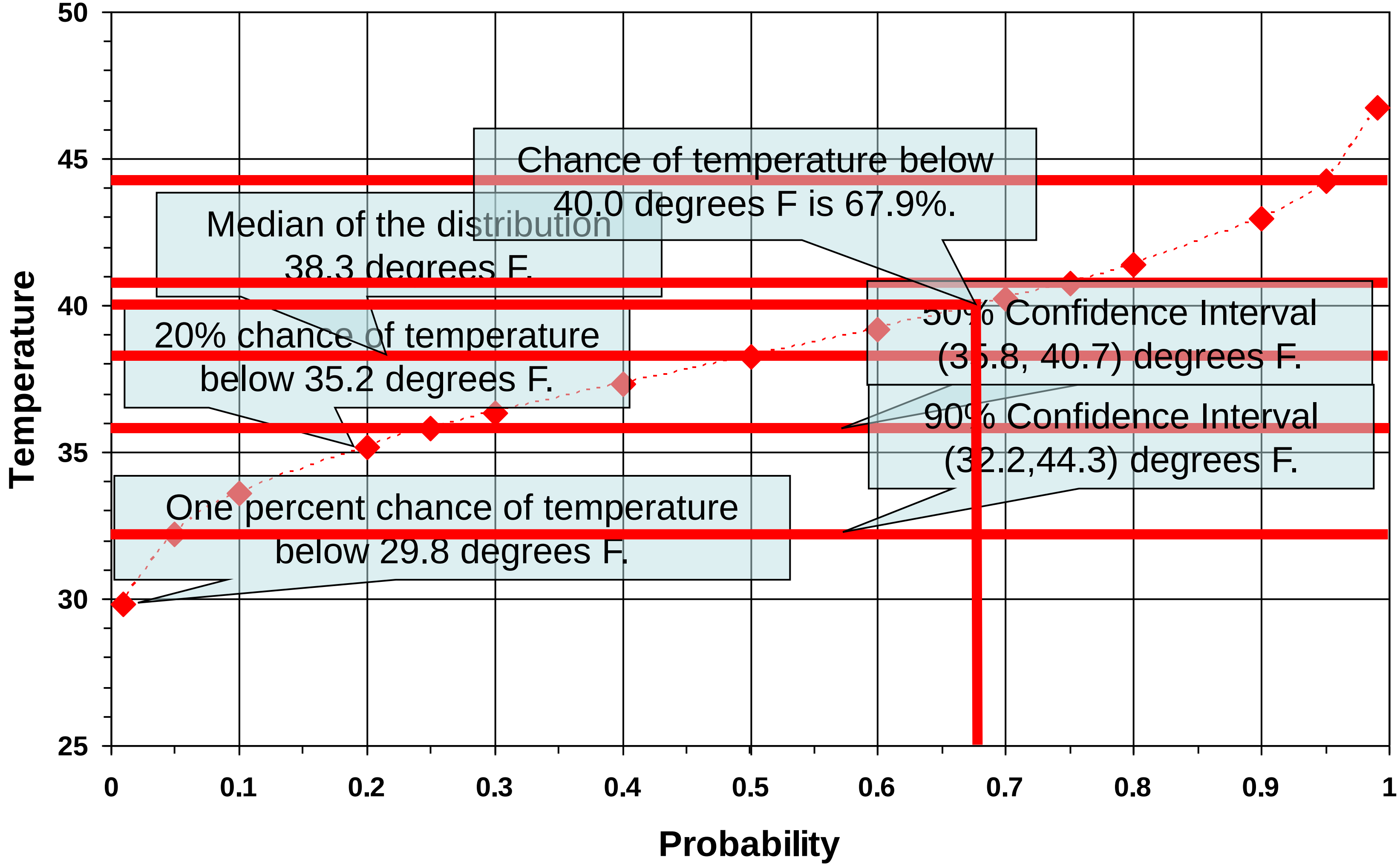
- Fit a parametric distribution (e. g., Normal).
  - Economical, but restrictive
- Enumerate Probability Density Function (PDF) or Cumulative Distribution Function (CDF) by computing probabilities for chosen values of the weather element.
  - Values must “work” everywhere
- Enumerate Quantile Function (QF) by giving values of the weather element for chosen exceedence probabilities.

Sample Forecast as Quantile Function



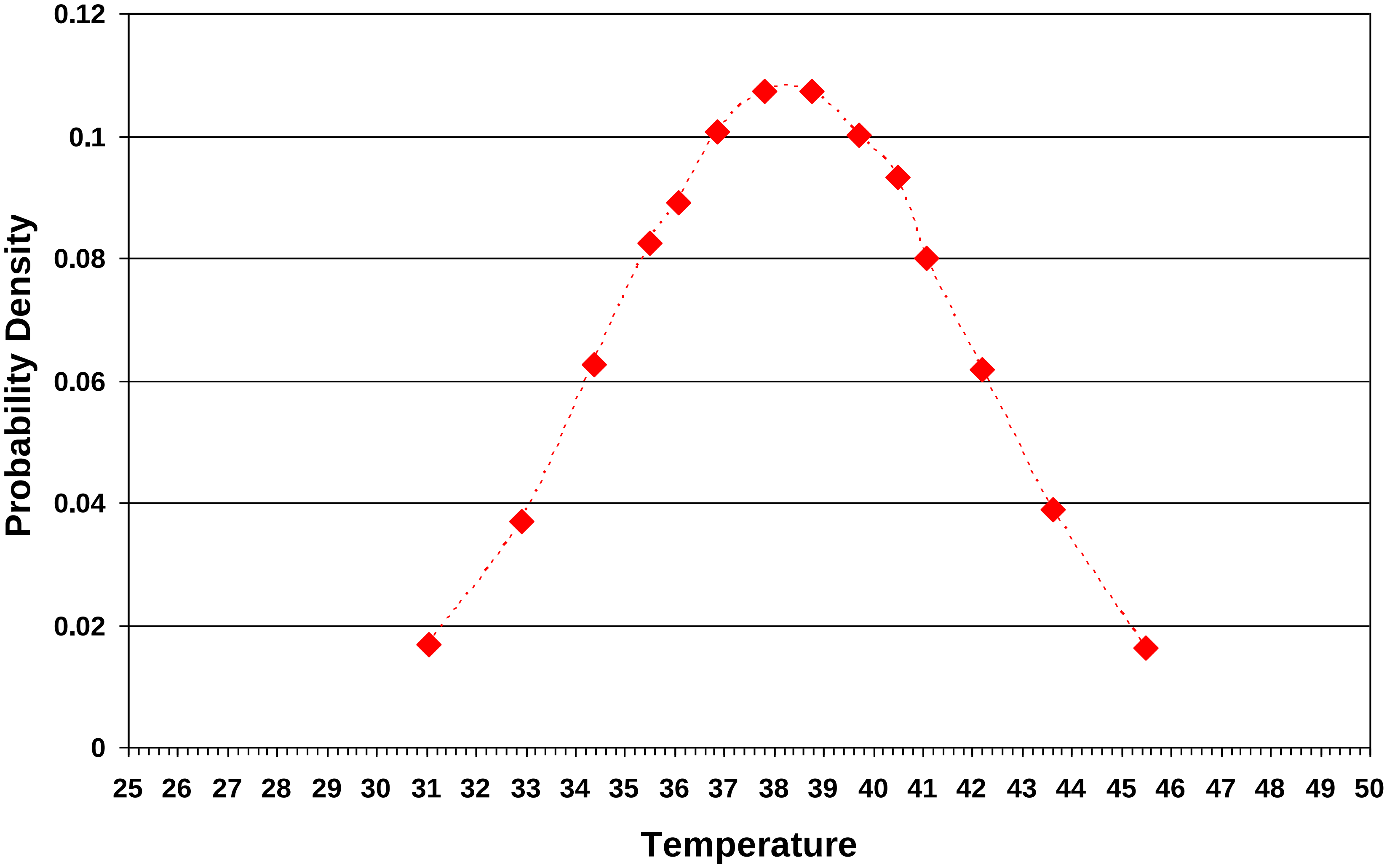
72-h T Fcst KBWI 12/14/2004

Sample Forecast as Quantile Function



72-h T Fcst KBWI 12/14/2004

Sample Forecast as Probability Density Function



Tools and Concepts

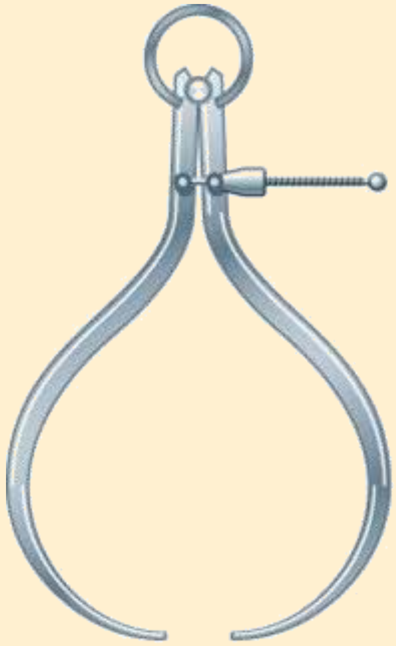
We have combined the following tools in a variety of ways to take advantage of linear regression and ensemble modeling of the atmosphere.

- Error estimation in linear regression
- Kernel Density Fitting (Estimation; KDE)

A brief overview of these tools follows.

Error Estimation in Linear Regression

- The linear regression theory used to produce MOS guidance forecasts includes error estimation.
- The *Confidence Interval* quantifies uncertainty in the position of the regression line.
- The *Prediction Interval* quantifies uncertainty in predictions made using the regression line.



The prediction interval can be used to estimate uncertainty each time a MOS equation is used to make a forecast.

Estimated Variance of a Single New Independent Value

- Estimated variance

$$s^2\left(\hat{Y}_{h(new)}\right) = MSE\left|1 + \frac{1}{n} + \frac{\left(X_h - \overline{X}\right)^2}{\sum\left(X_i - \overline{X}\right)^2}\right|$$

- Where

$$MSE = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}$$

Computing the Prediction Interval

The prediction bounds for a new prediction is

$$\hat{Y}_{h(new)} \pm t(1-\alpha/2;n-2)s\left(\hat{Y}_{h(new)}\right)$$

where

t(1-α/2;n-2) is the t distribution n-2 degrees of freedom at the 1-α (two-tailed) level of significance, and

s($\hat{Y}_{h(new)}$) can be approximated by

$$\sqrt{s^2(1-r^2)}$$

where

s^2 is variance of the predictand

r^2 is the reduction of variance

Multiple Regression (3-predictor case)

Predictand Vector

$$ {}_n\mathbf{Y}_1 = \begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{vmatrix} $$

3-predictor Matrix

$$ {}_n\mathbf{X}_4 = \begin{vmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ 1 & x_{41} & x_{42} & x_{43} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{vmatrix} $$

Coefficient Vector

$$ {}_4\mathbf{A}_1 = \begin{vmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{vmatrix} $$

Multiple Regression, Continued

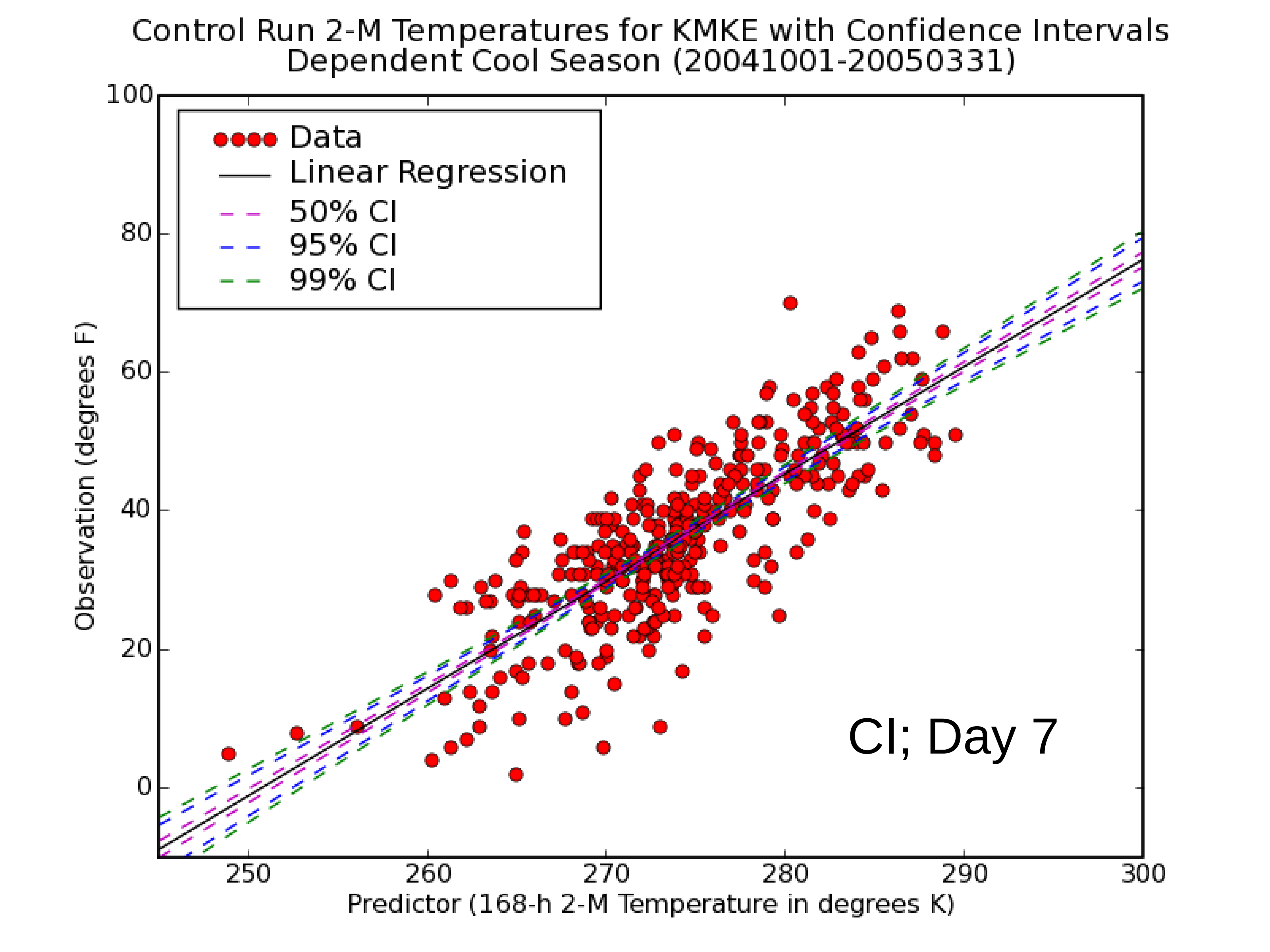
Error bounds can be put around the new value of Y with

$$\hat{Y}_{h(new)} \pm \left| s^2\left(1-R^2\right)\left\{1+ {}_1\mathbf{x}_4\left({}_4\mathbf{X}'_n\mathbf{X}_4\right)^{-1}{}_4\mathbf{x}'_1\right\}\right|^{1/2}$$

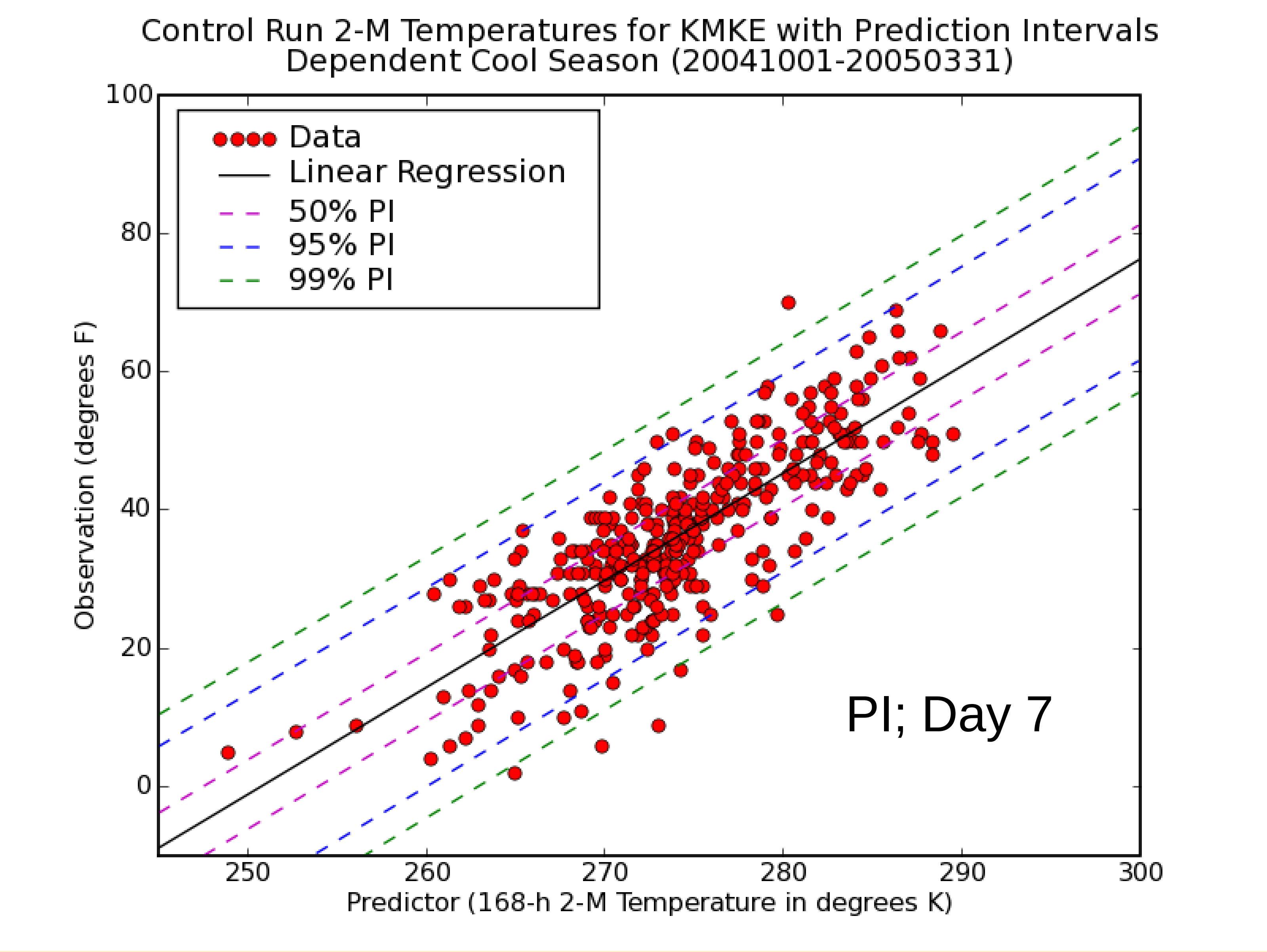
where

- s^2 is the variance of the predictands,
- R^2 is the reduction of variance,
- **X'** is the matrix transpose of **X**, and
- $()^{-1}$ indicates the matrix inverse.

Example: Confidence Intervals for Milwaukee, Wisconsin



Example: Prediction Intervals for Milwaukee, Wisconsin

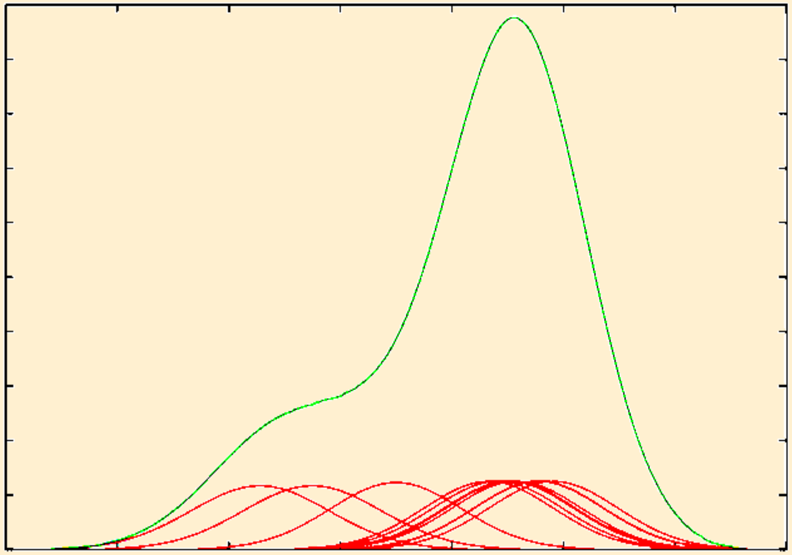


Advantages of MOS Techniques for Assessing Uncertainty

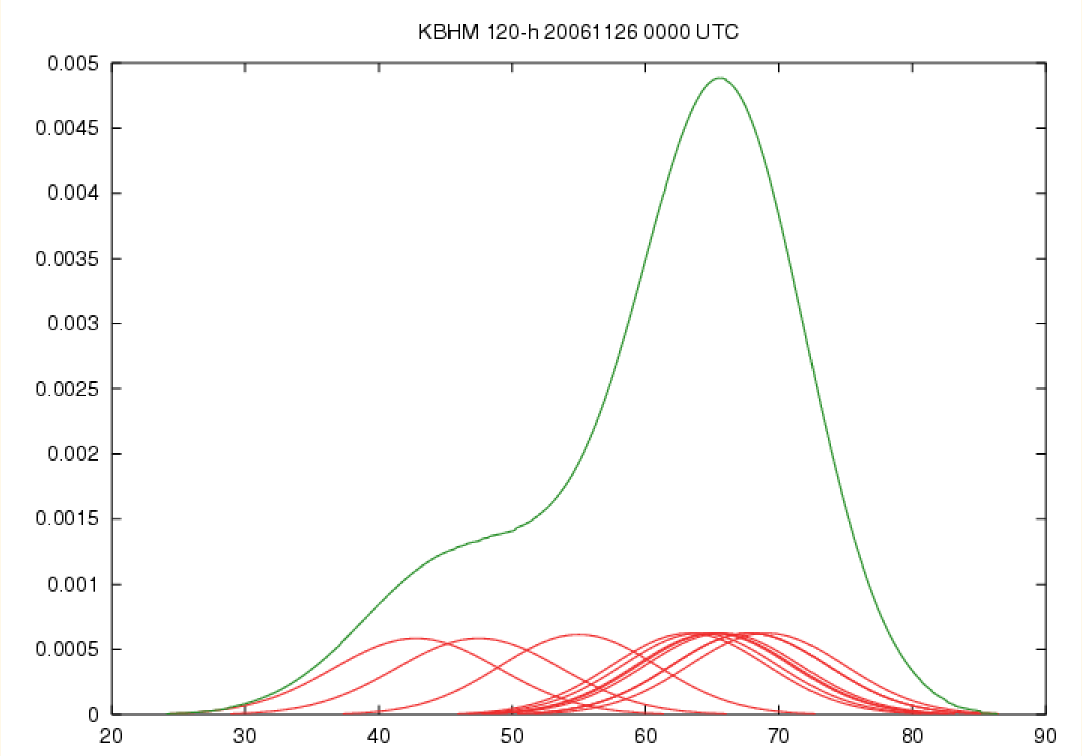
- Single valued forecasts and probability distributions come from a single consistent source.
- Longer development sample can better model climatological variability.
- Least squares technique is effective at producing reliable distributions.

Kernel Density Fitting

- Used to estimate the Probability Density Function (PDF) of a random variable, given a sample of its population.
- A kernel function is centered at each data point.
- The kernels are then summed to generate a PDF.
- Various kernel functions can be used. Smooth, unimodal functions with a peak at zero are most common.

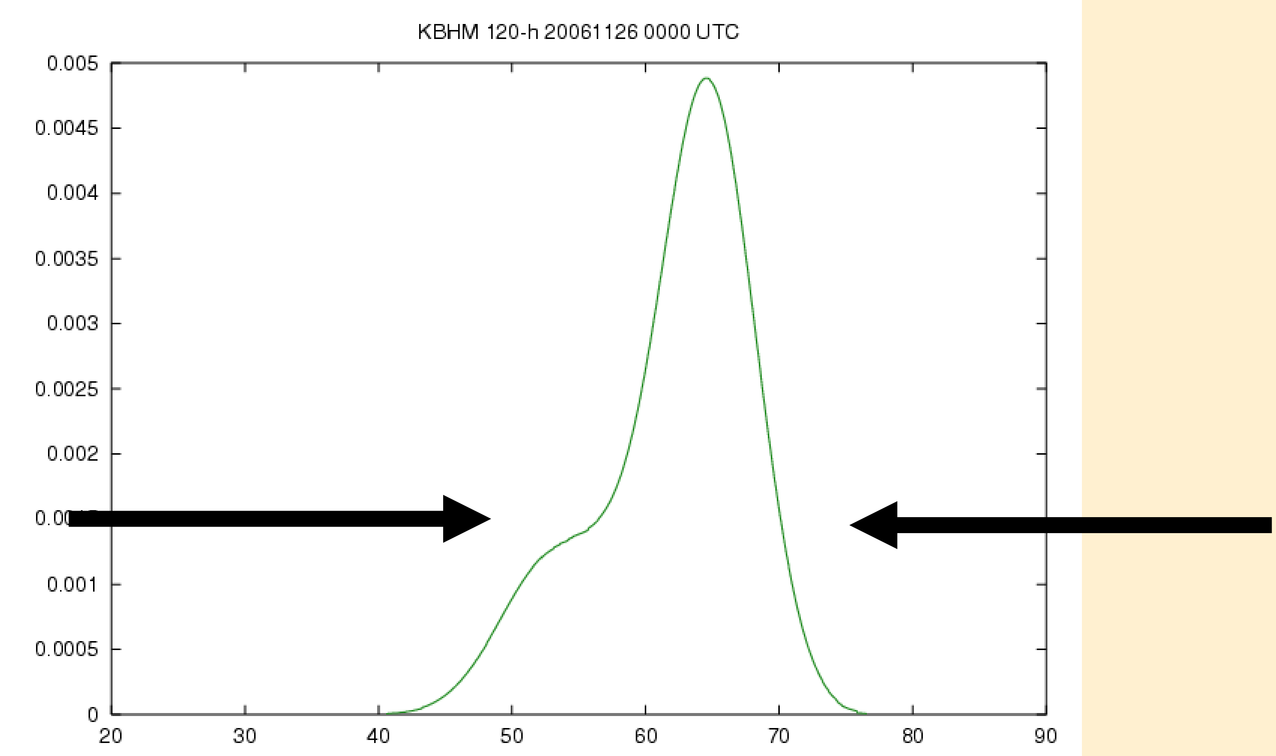


Kernel Density Fitting



A common problem is choosing the shape and width of the kernel functions. We've used the Normal Distribution and Prediction Interval, respectively.

Spread Adjustment



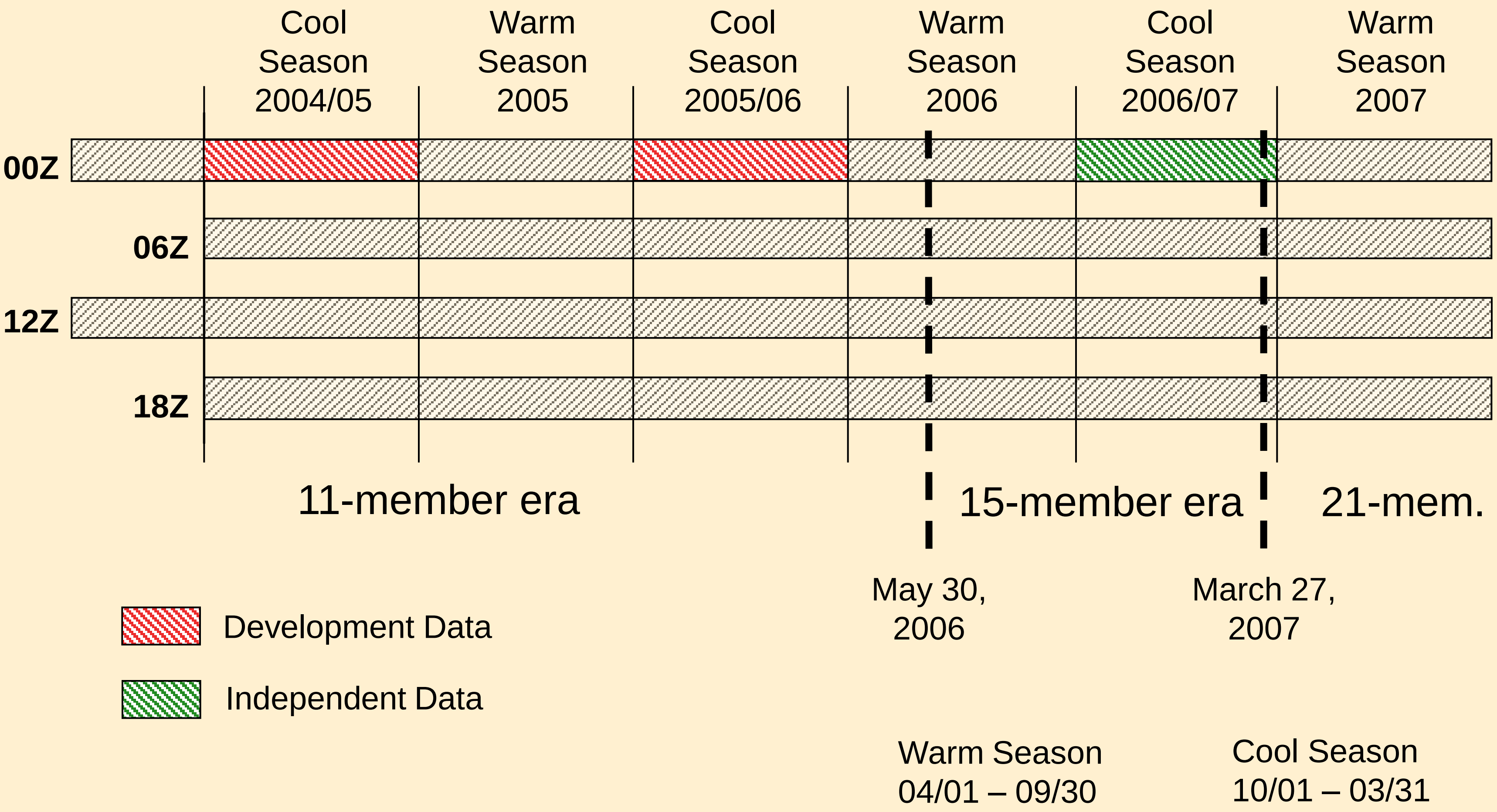
Combination of prediction interval and spread in the ensembles can yield too much spread.

Spread Adjustment attempts to correct over dispersion.

Weather Elements

- Temperature and dew point, developed simultaneously
  - 3-h time projections for 7 days
  - Model data at 6-h time projections
  - 1650 stations, generally the same as GFS MOS
- Maximum and minimum temperature
  - 15 days
  - Same stations

Global Ensemble Forecasting System Data Available for Ensemble MOS Development



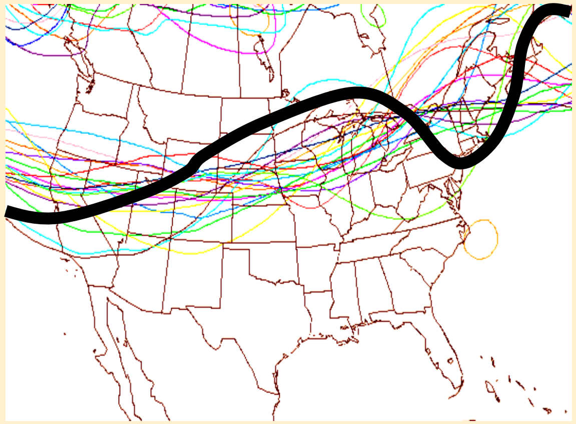
Methods

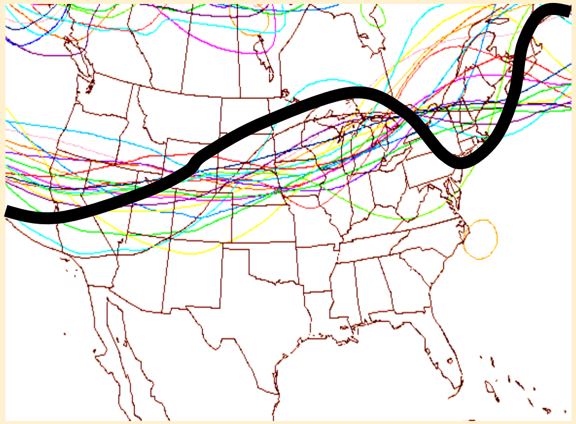
We explored a number of methods. Three are presented here.

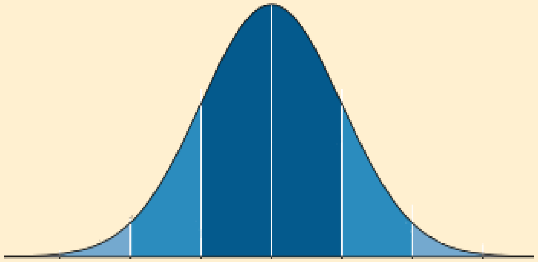
Label	**Equation Development**	**Equation Evaluation**	**Post Processing**
Ctl-Ctl-N	Control member only	Control member only	Use a Normal Distribution
Mn-Mn-N	Mean of all ensemble members	Mean of all ensemble members	Use a Normal Distribution
Mn-Ens-KDE	Mean of all ensemble members	Each member individually	Apply KDE, and adjust spread

Ctl-Ctl-N

Equation Development	**Equation Evaluation**	**Post Processing**
Control member only	Control member only	Use a Normal Distribution





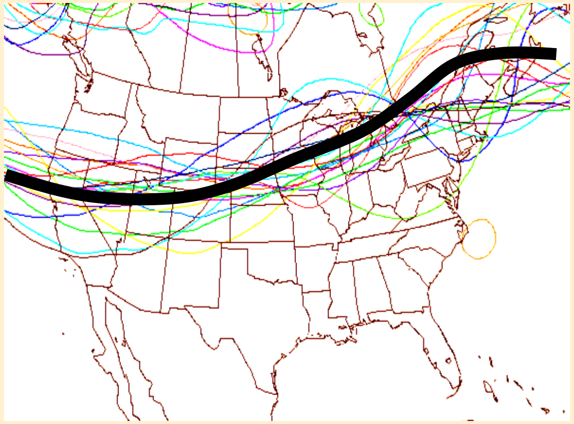


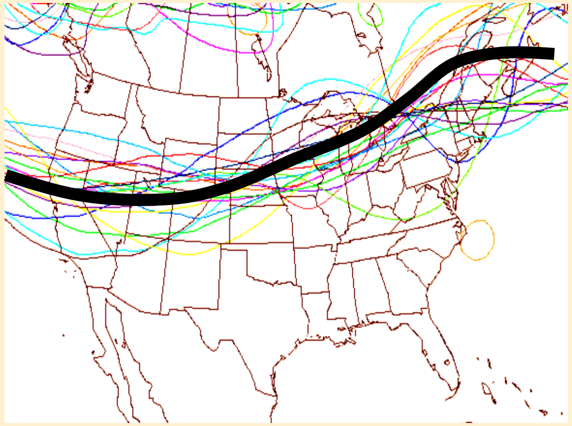
Mn-Mn-N

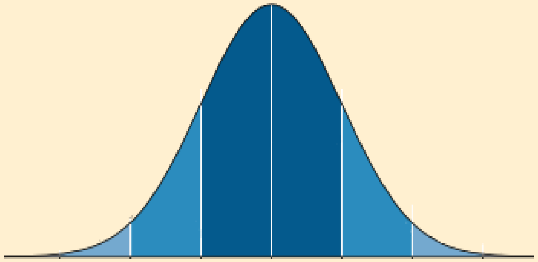
Equation Development
Mean of all ensemble members

Equation Evaluation
Mean of all ensemble members

Post Processing
Use a Normal Distribution

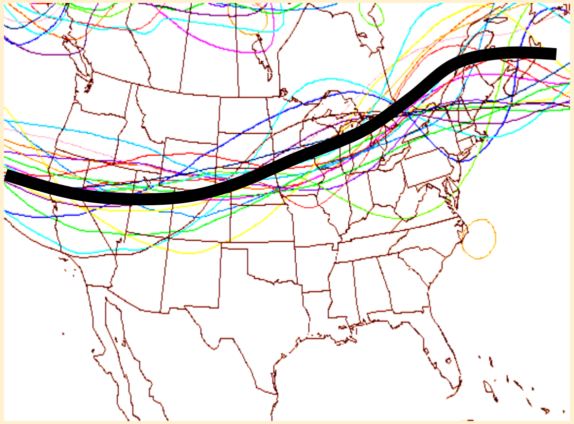


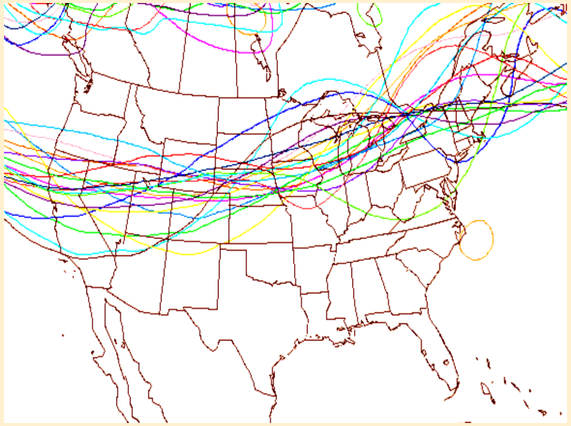


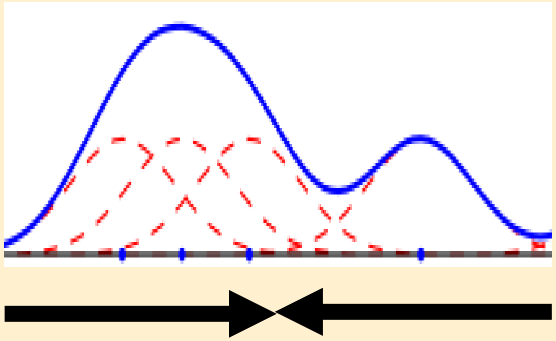


Mn-Ens-KDE

Equation Development	Equation Evaluation	Post Processing
Mean of all ensemble members	Each member individually	Apply KDE, and adjust spread





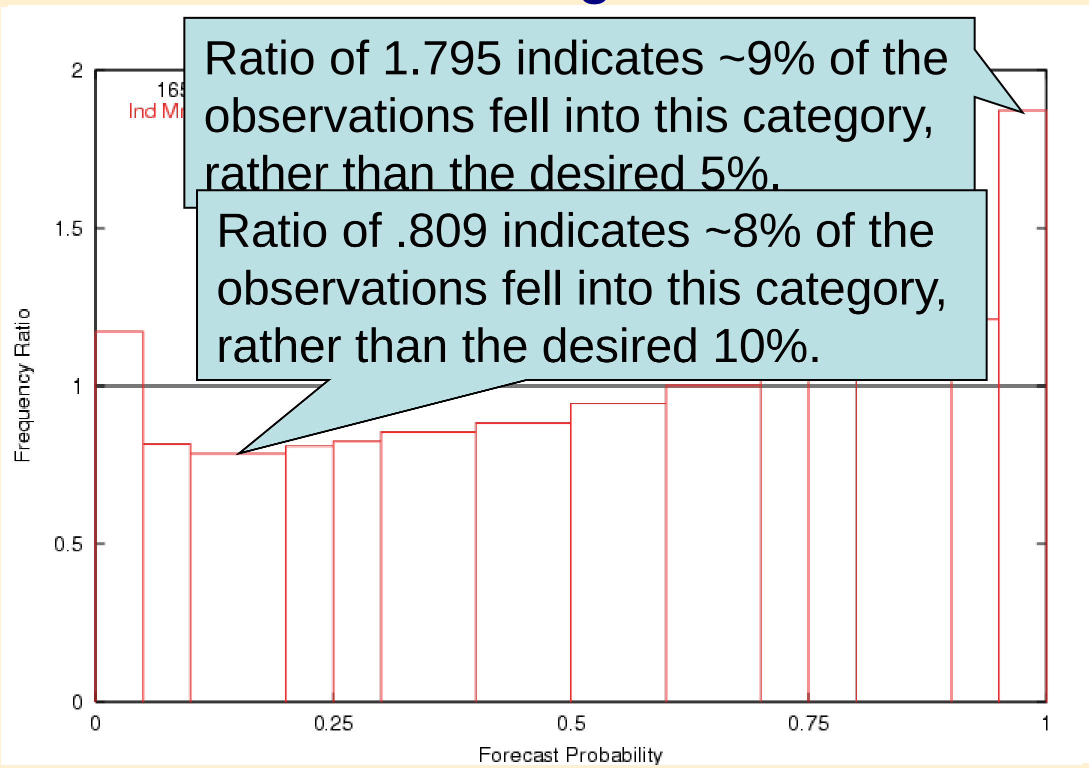


Results

- Will present results for cool season temperature forecasts developed with two seasons of development data and verified against one season of independent data.
- Results center on reliability and accuracy.
- The 0000 UTC cycle of the Global Ensemble Forecast System is the base model.
- Results for dew point are available and very similar to temperature.
- Results for maximum/minimum temperature are in process, and they are similar so far.

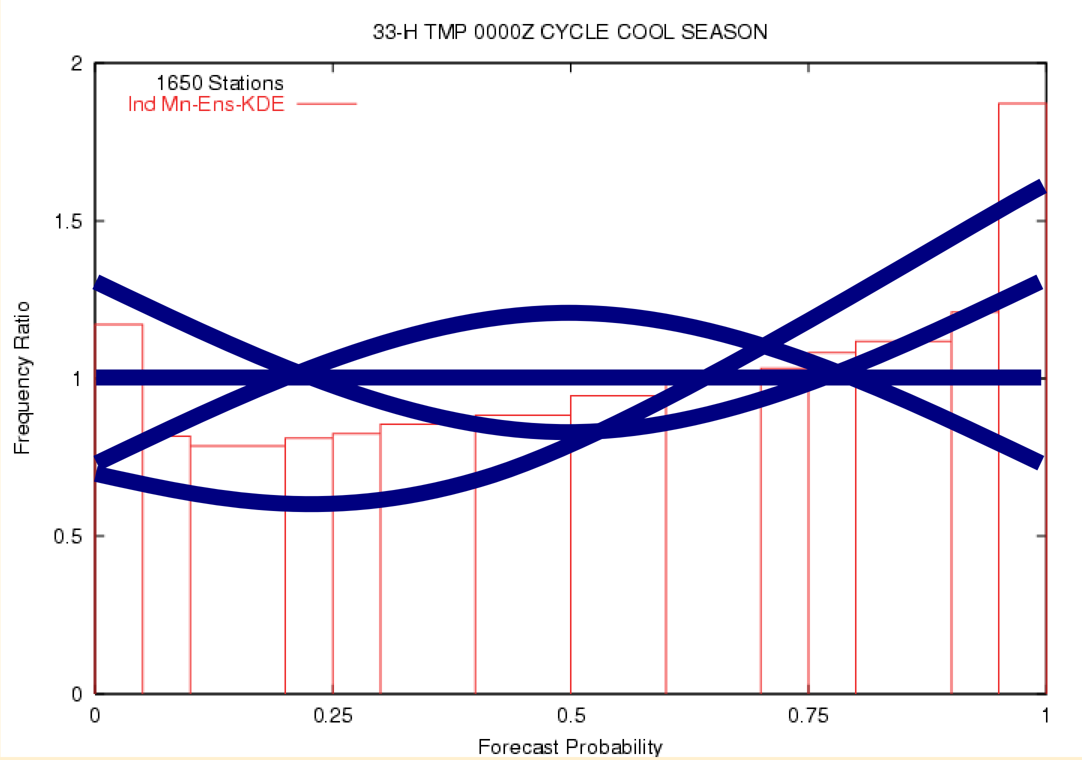
Probability Integral Transform (PIT) Histogram

- Graphically assesses reliability for a set of probabilistic forecasts. Visually similar to Ranked Histogram.
- Method
  - For each forecast-observation pair, probability associated with observed event is computed.
  - Frequency of occurrence for each probability is recorded in histogram as a ratio.
  - Histogram boundaries set to QF probability values.



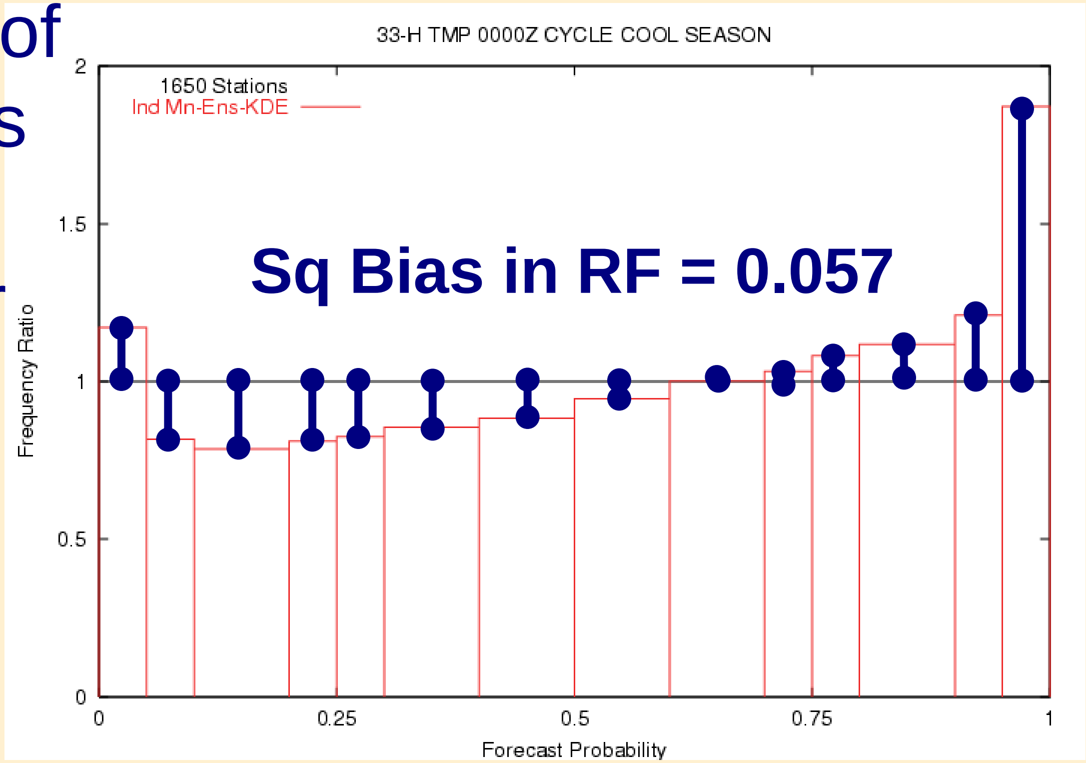
Probability Integral Transform (PIT) Histogram, Continued

- Assessment
  - Flat histogram at unity indicates reliable, unbiased forecasts.
  - *U*-shaped histogram indicates under-dispersion in the forecasts.
  - *O*-shaped histogram indicates over-dispersion.
  - Higher values in higher percentages indicate a bias toward lower forecast values.

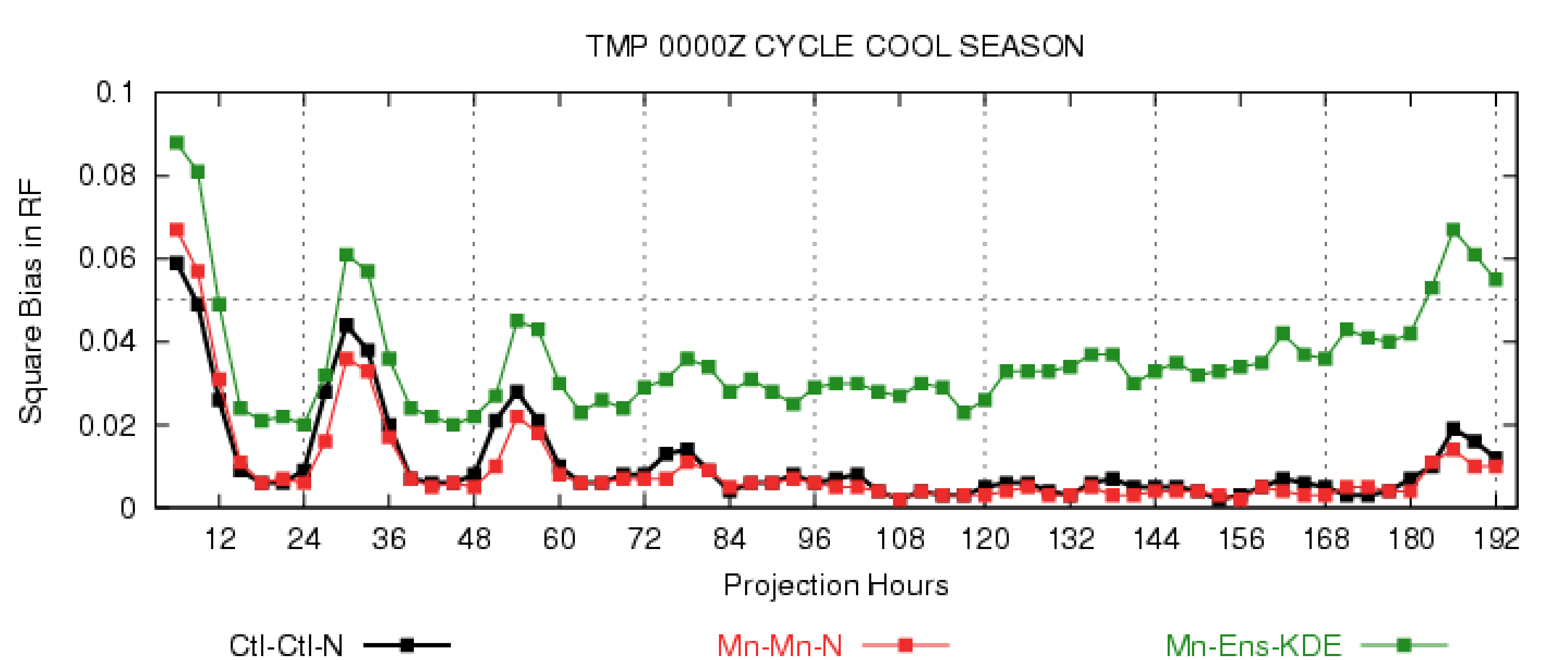


Squared Bias in Relative Frequency

- Weighted average of squared differences between actual height and unity for all histogram bars.
- Zero is ideal.
- Summarizes histogram with one value.

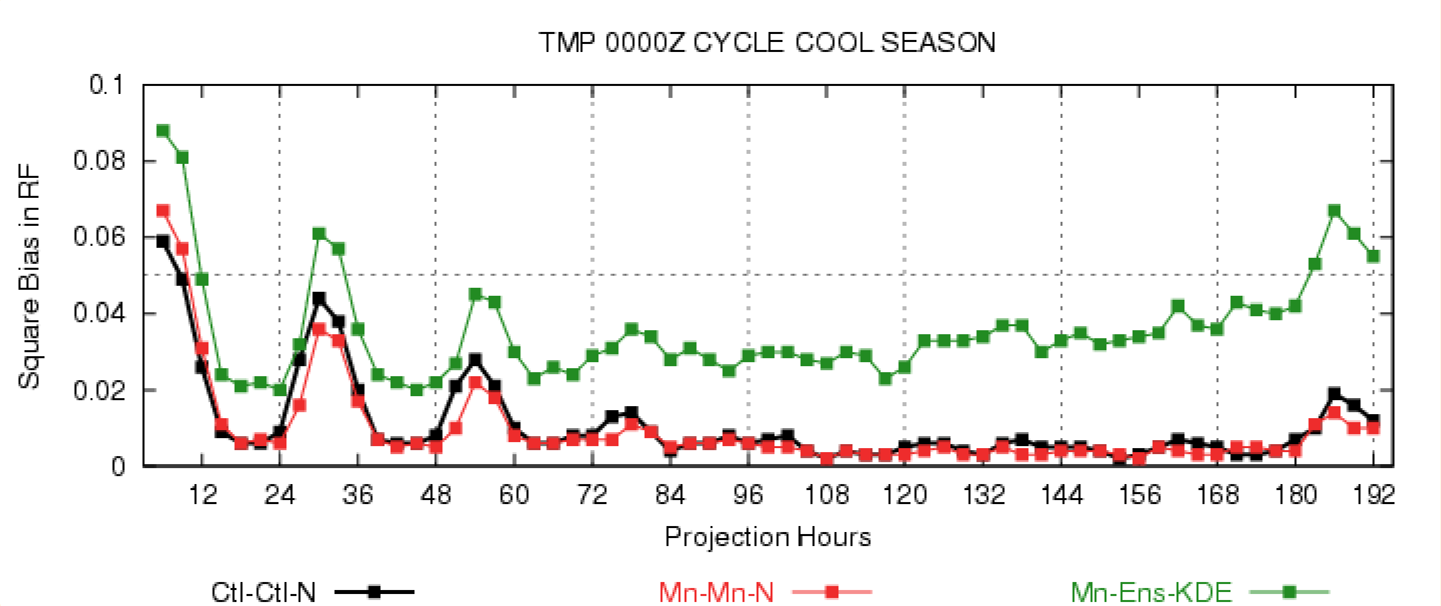


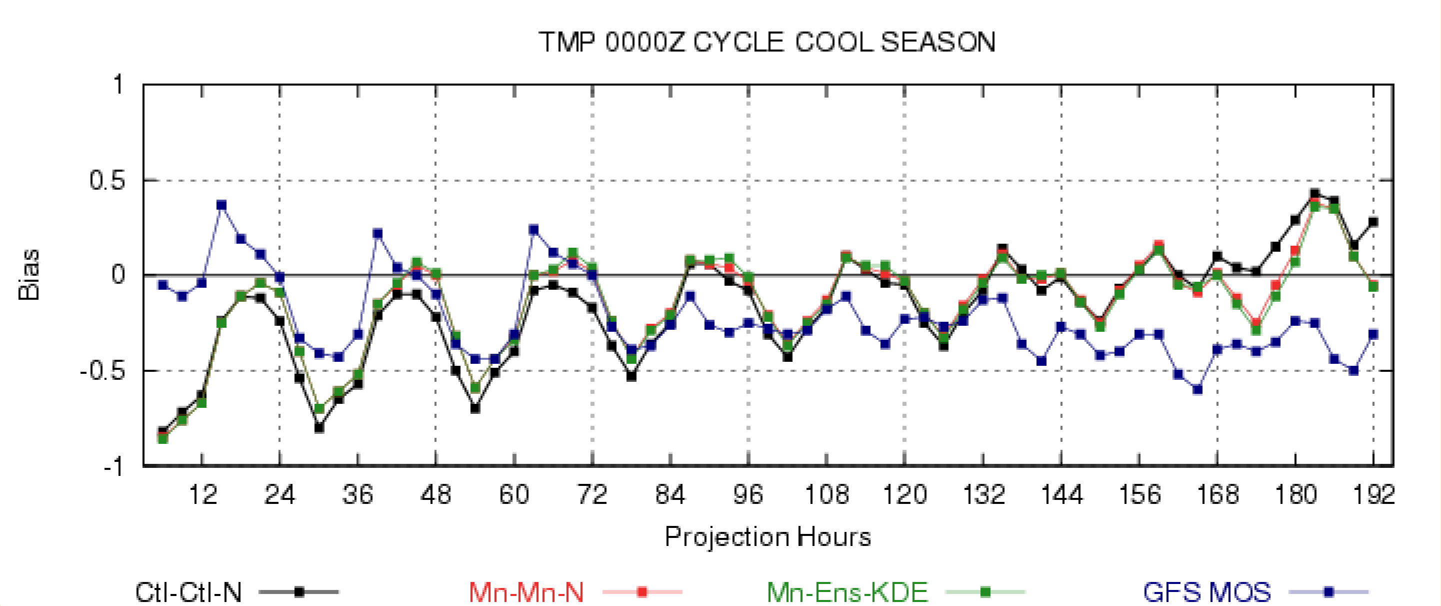
Squared Bias in Relative Frequency



- Diurnal cycle evident in early projections.
- Use of ensemble mean as a predictor improves reliability at most time projections.
- KDE technique seems to degrade reliability.
- Model resolution change evident in latest projections.

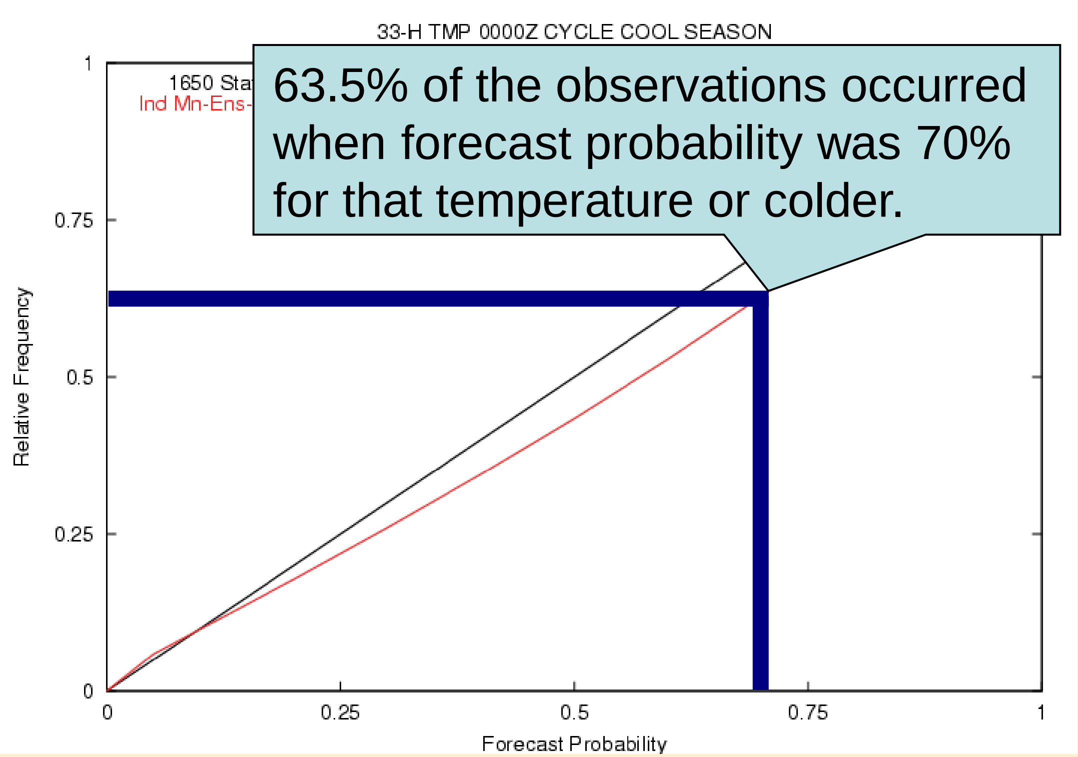
Bias Comparison



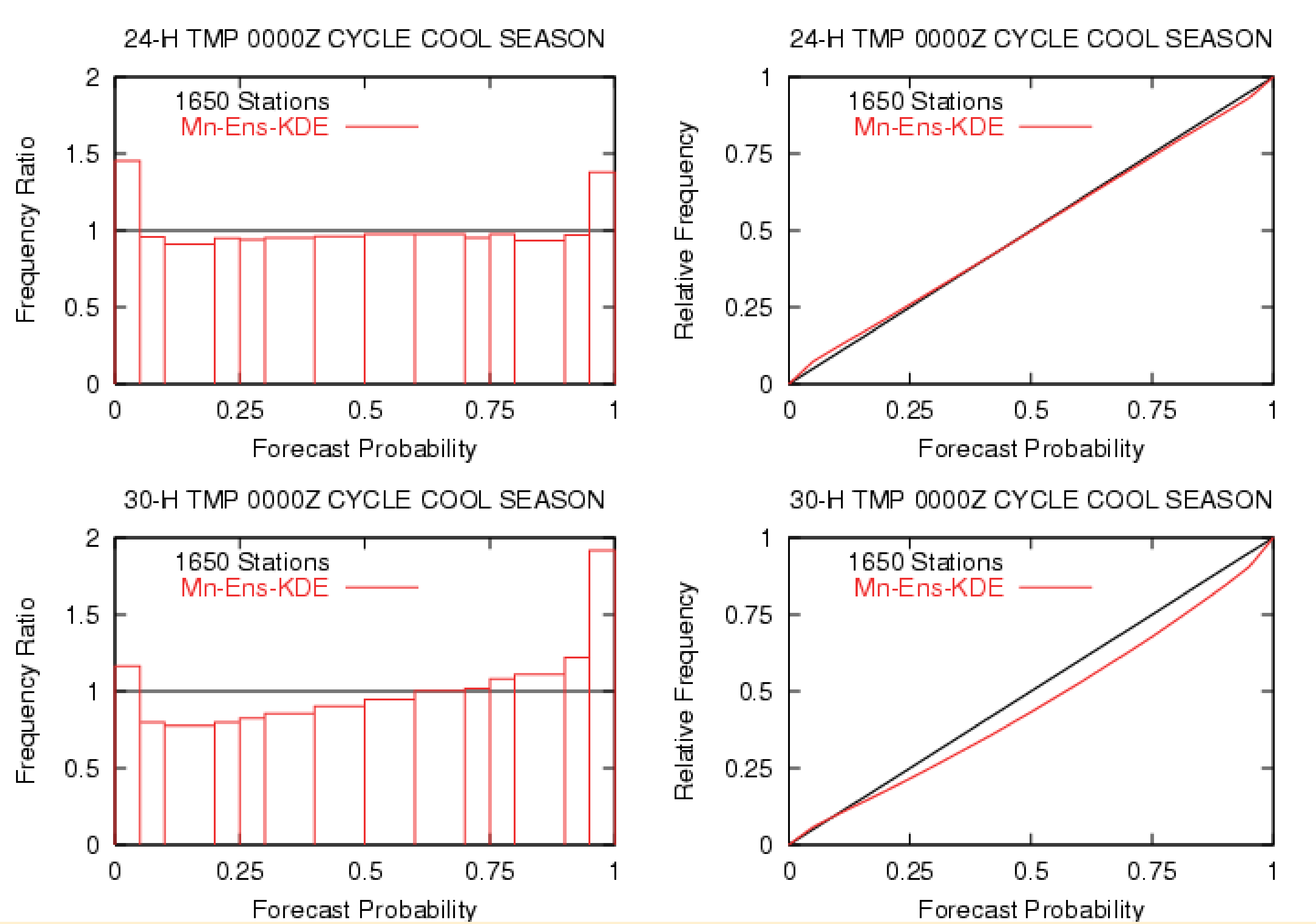


Cumulative Reliability Diagram (CRD)

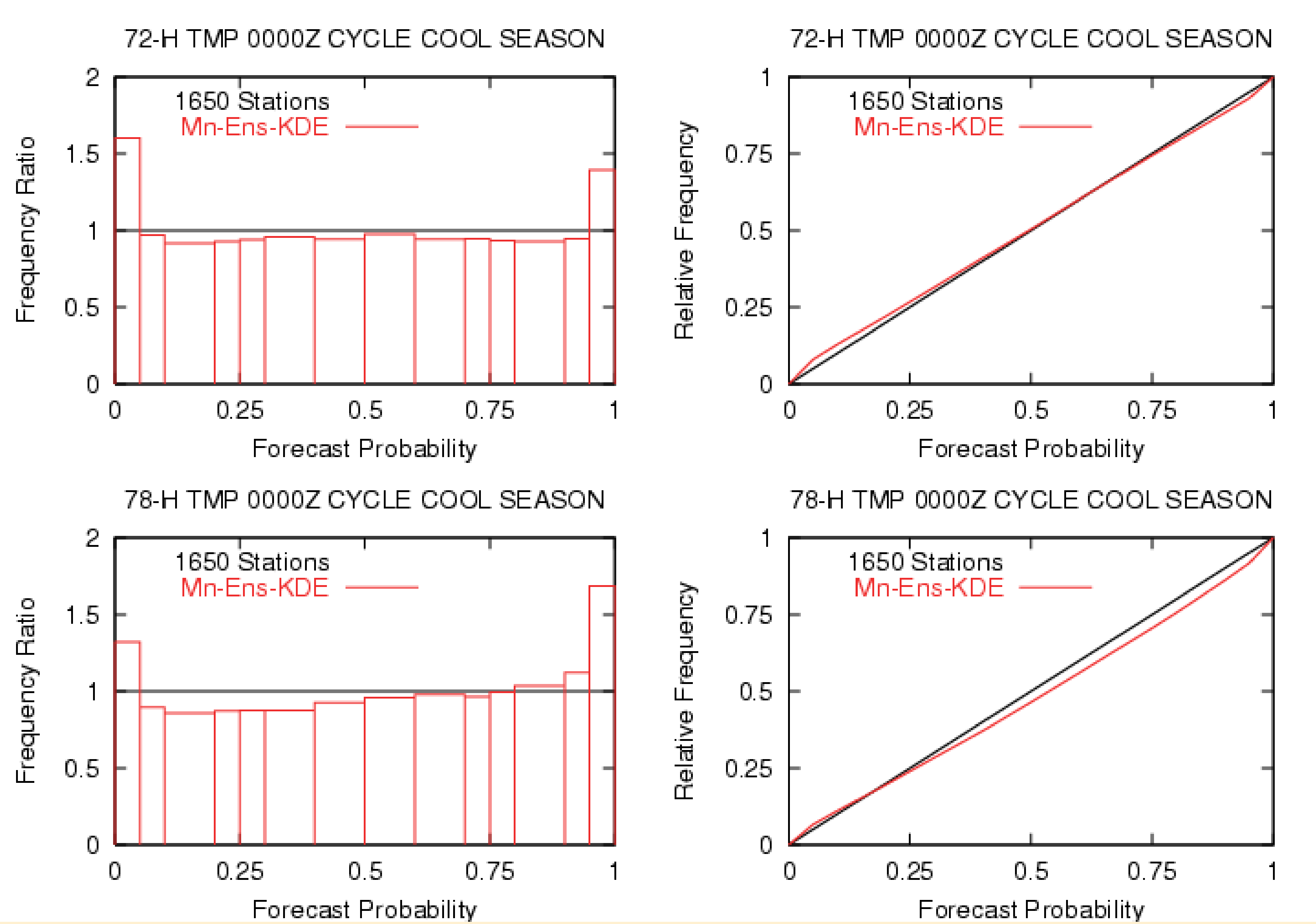
- Graphically assesses reliability for a set of probabilistic forecasts. Visually similar to reliability diagrams for event-based probability forecasts.
- Method
  - For each forecast-observation pair, probability associated with observed event is computed.
  - Cumulative distribution of verifying probabilities is plotted against the cumulative distribution of forecasts.



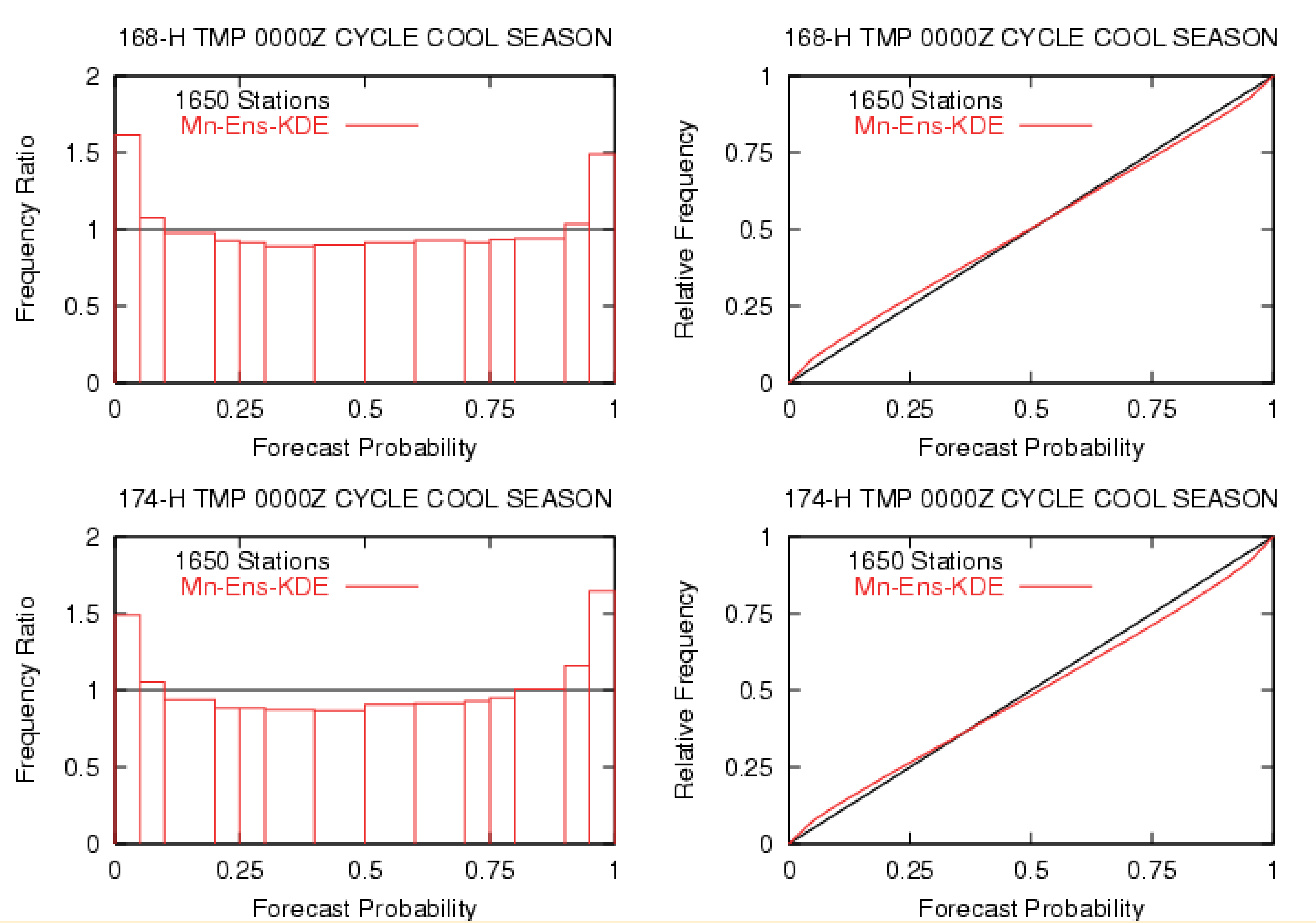
Day 1 Reliability



Day 3 Reliability



Day 7 Reliability



Continuous Ranked Probability Score

The formula for CRPS is

$$CRPS = CRPS\left(P, x_a\right) = \int_{-\infty}^{\infty} \left[P(x) - P_a(x)\right]^2 dx$$

where *P(x)* and $P_a(x)$ are both CDFs

$$P(x) = \int_{-\infty}^{x} \rho(y)dy$$

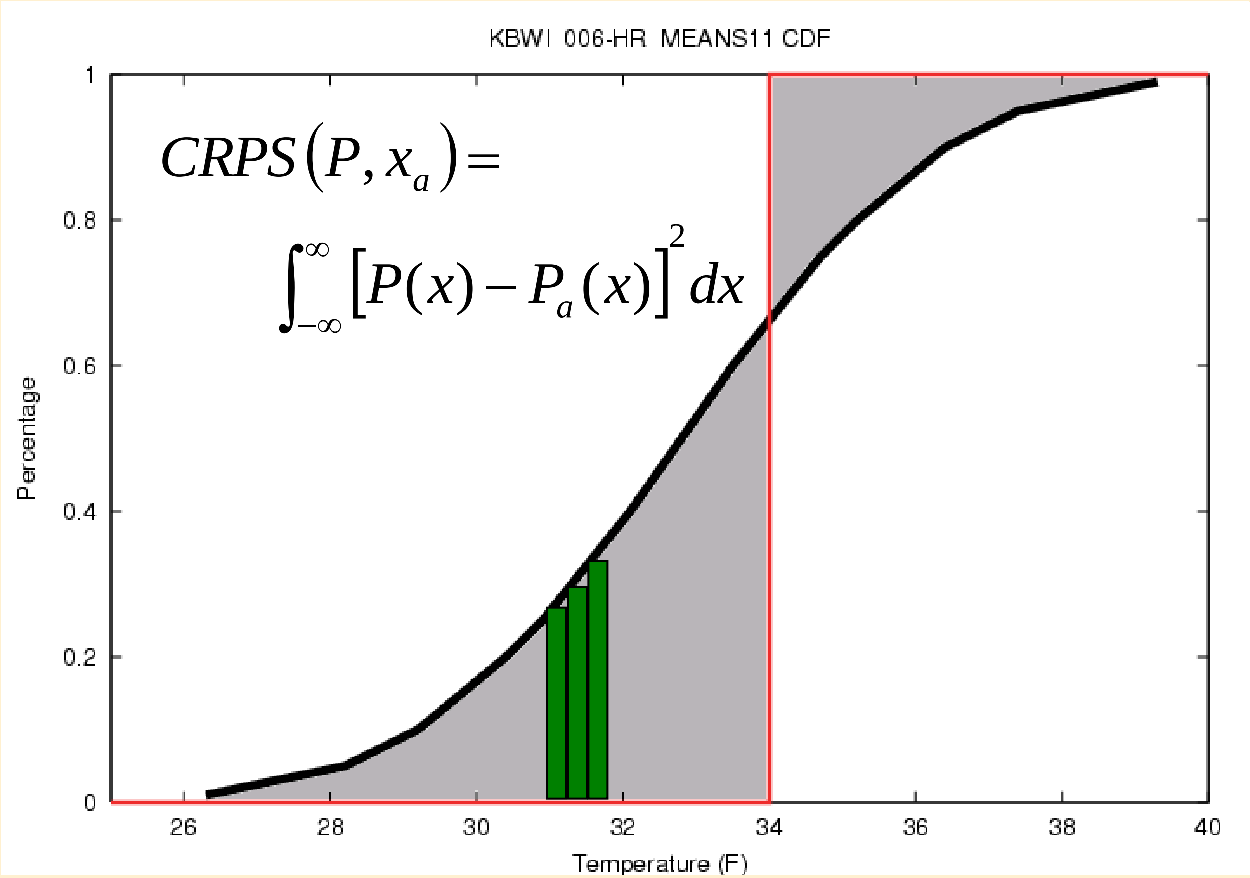
$$P_a(x) = H(x - x_a)$$

and

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$$

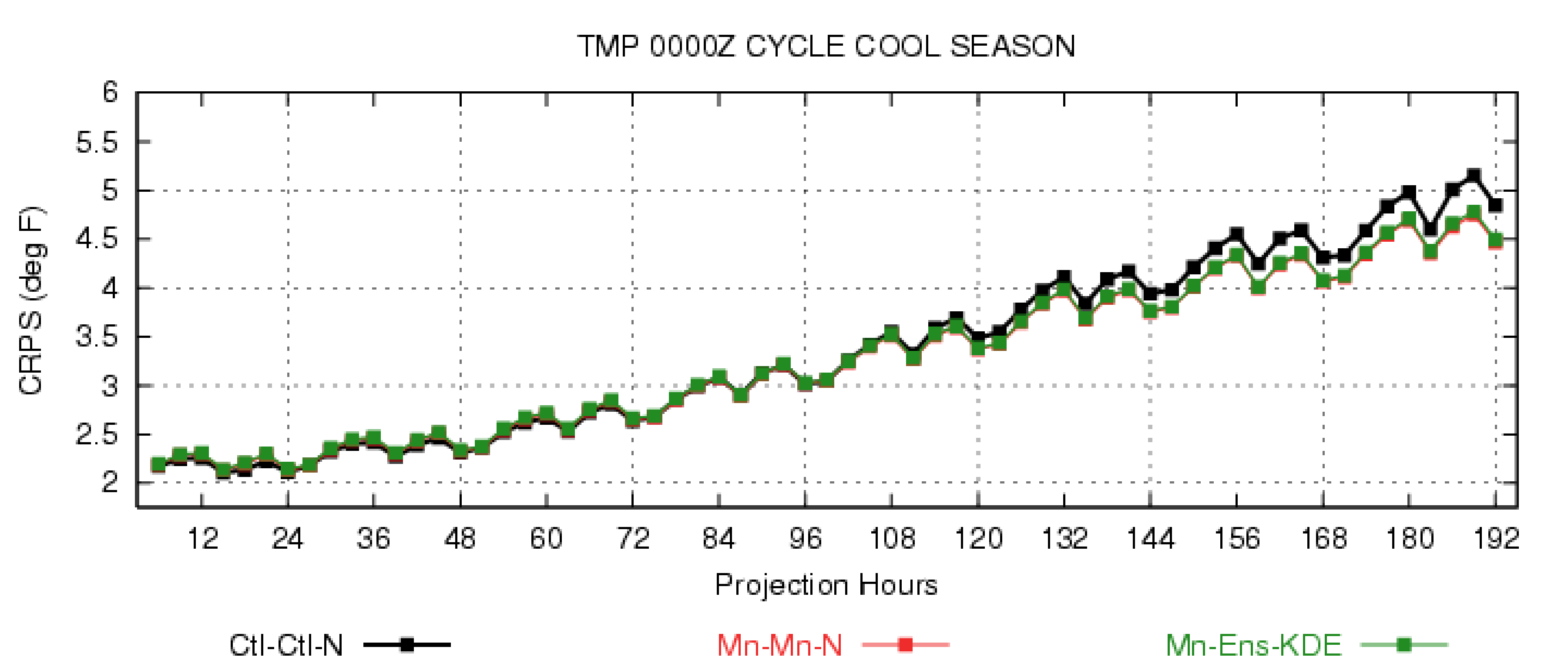
Continuous Ranked Probability Score

- Proper score that measures the accuracy of a set of probabilistic forecasts.
- Squared difference between the forecast CDF and a perfect single value forecast, integrated over all possible values of the variable.



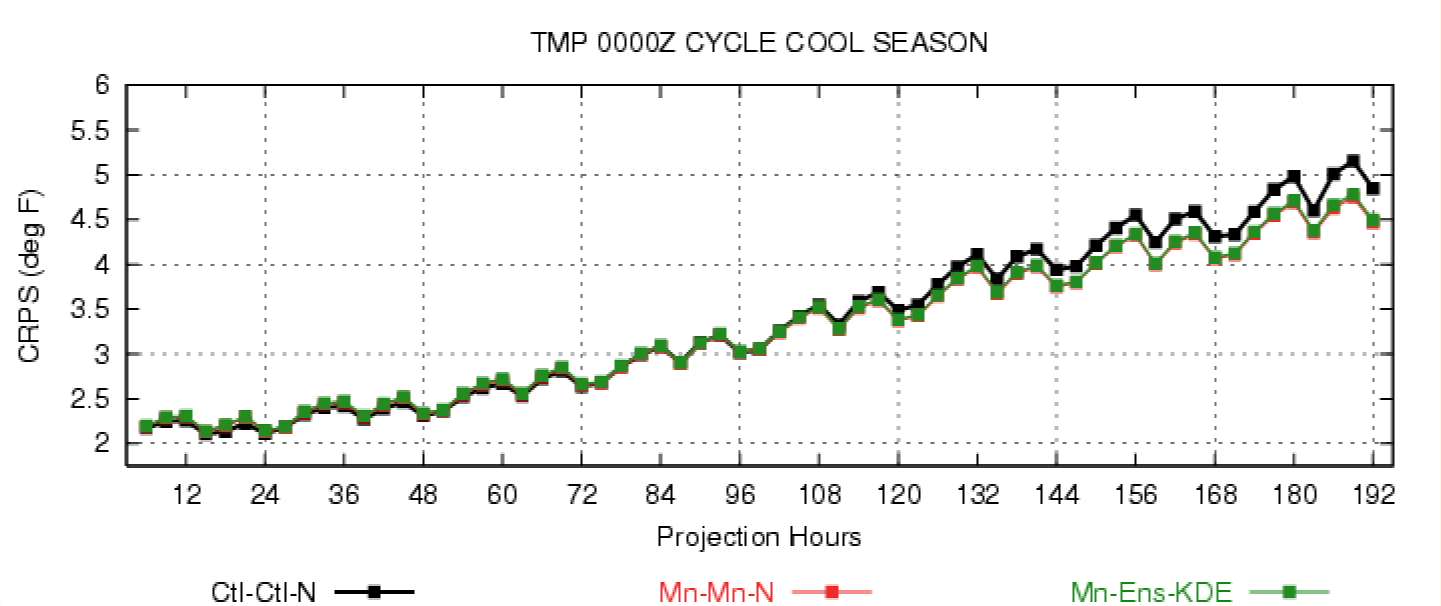
  Units are those of the variable.
- Zero indicates perfect accuracy. No upper bound.

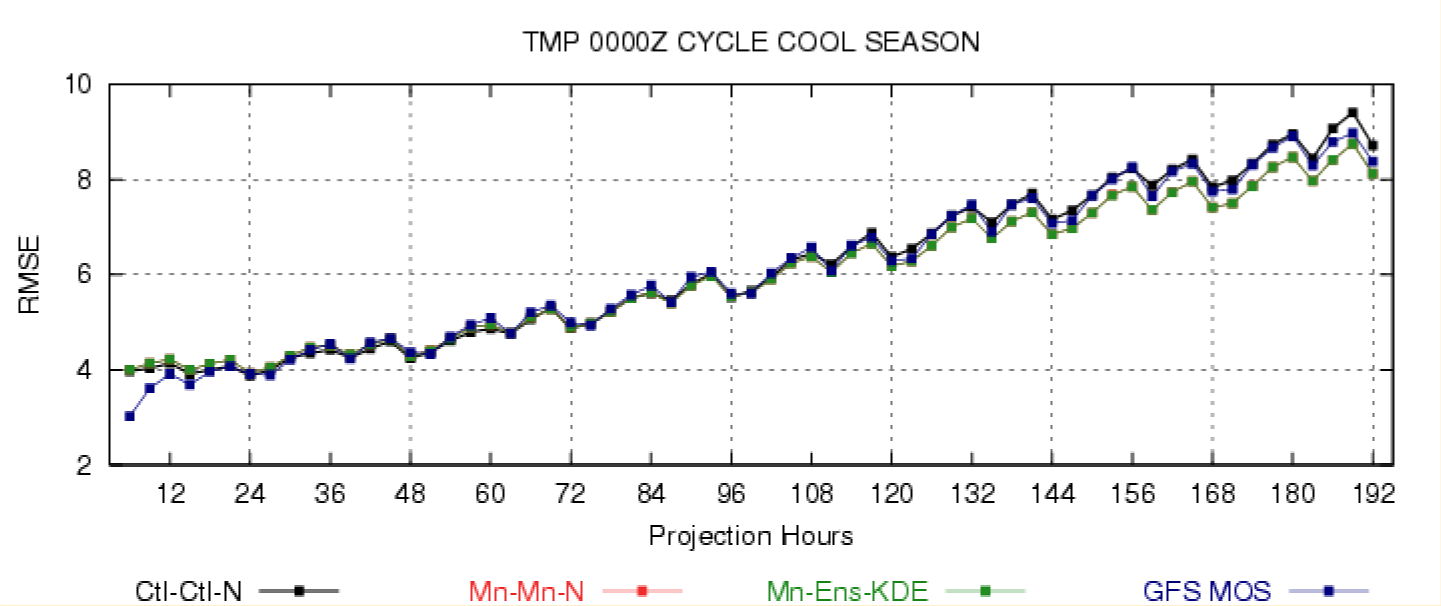
Continuous Rank Probability Score



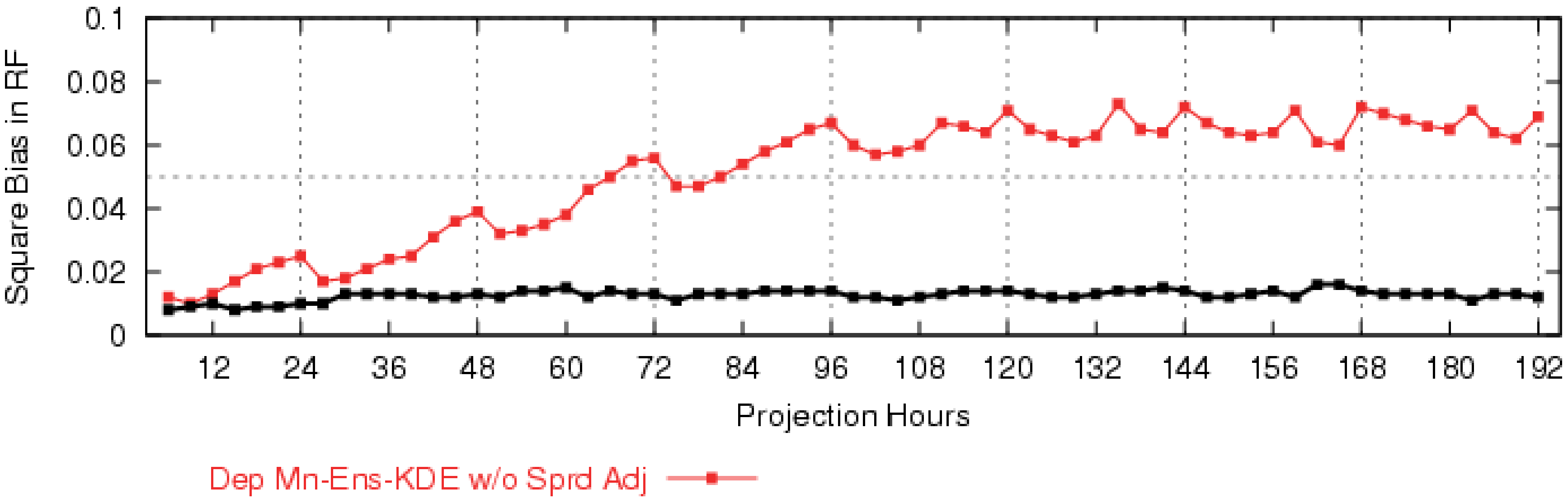
- All techniques show considerable accuracy.
- After Day 5 the 2 techniques that use ensembles show ~0.5 deg F improvement (~12 h).

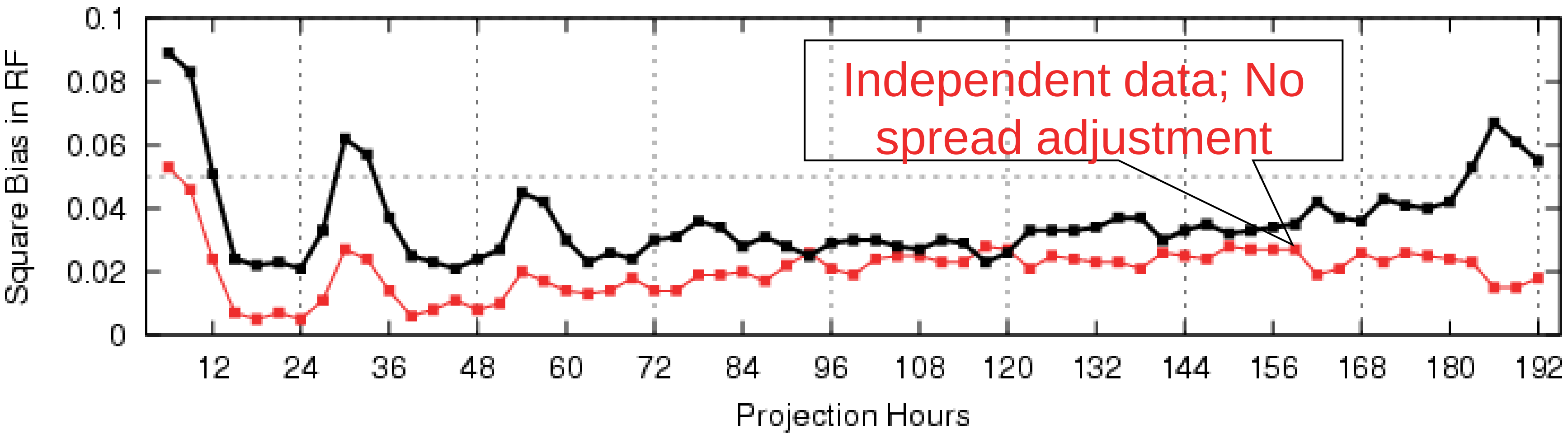
Accuracy Comparison





Effects of Spread Adjustment

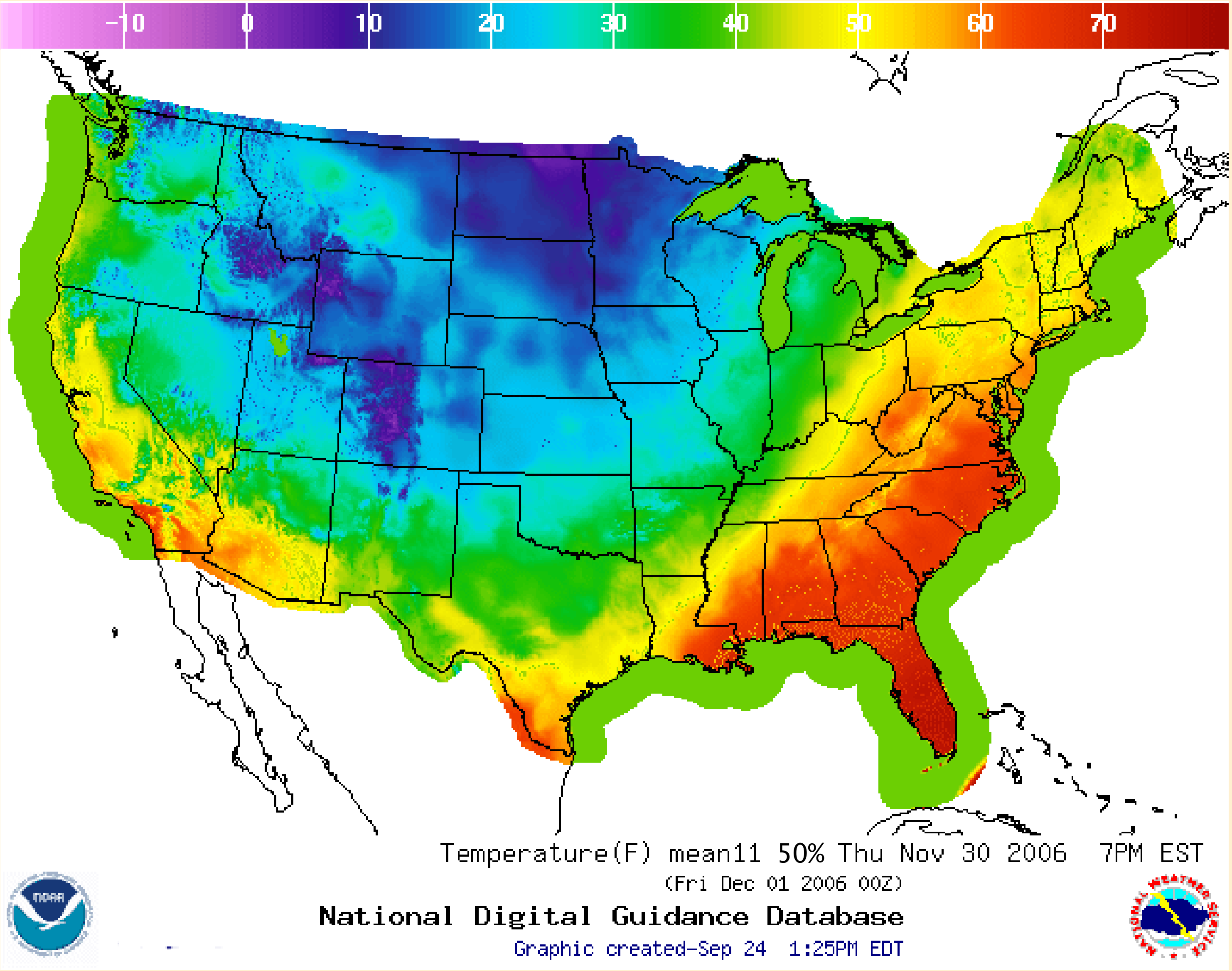




Grids

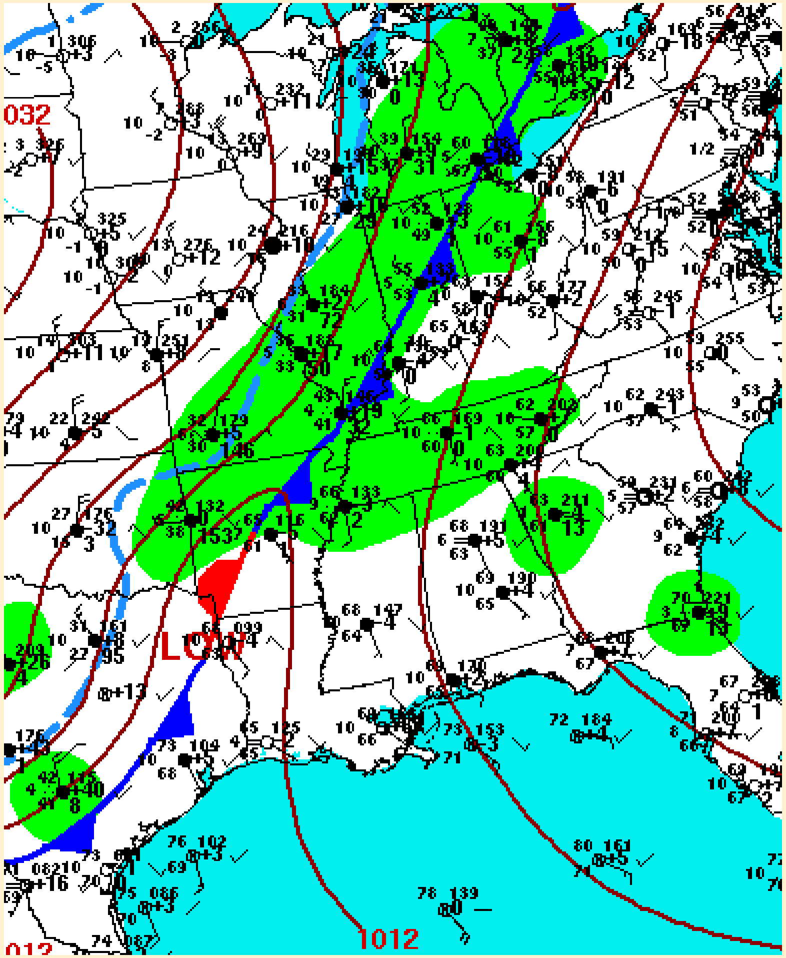
- Temperature forecasts for 1650 stations can be used to generate grids.
  - Technique is identical to that used currently for gridded MOS.
- Each grid is associated with an exceedence probability.

Gridded [.05, .95] Temperatures

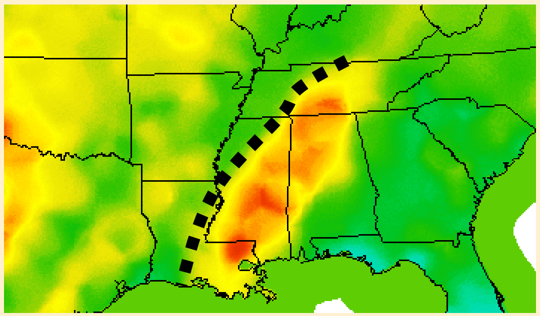
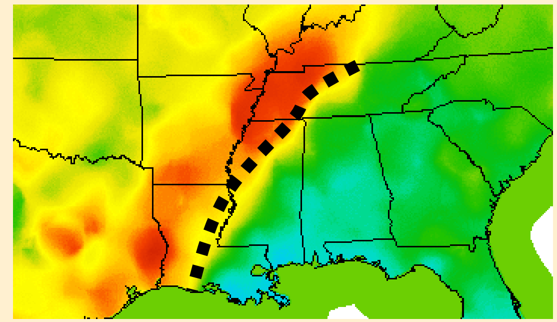
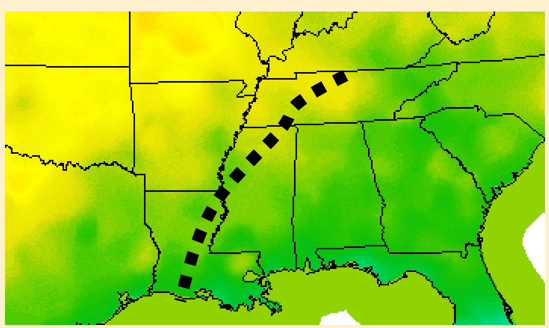
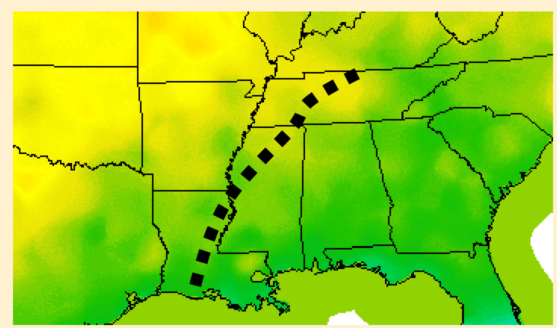


Case Study

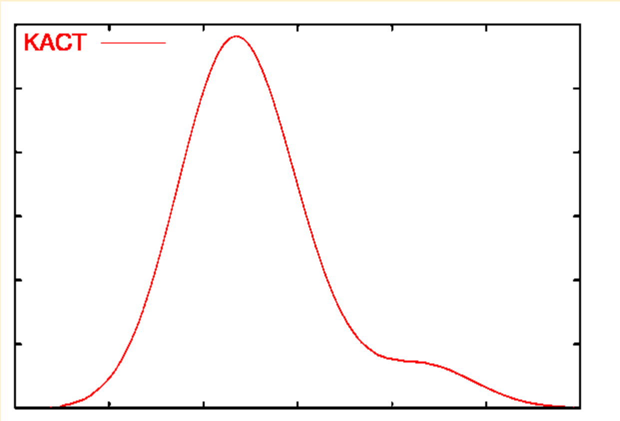
- 120-h Temperature forecast based on 0000 UTC 11/26/2006, valid 0000 UTC 12/1/2006.
- Daily Weather Map at right is valid 12 h before verification time.
- Cold front, inverted trough suggests a tricky forecast, especially for Day 5.
- Ensembles showed considerable divergence.



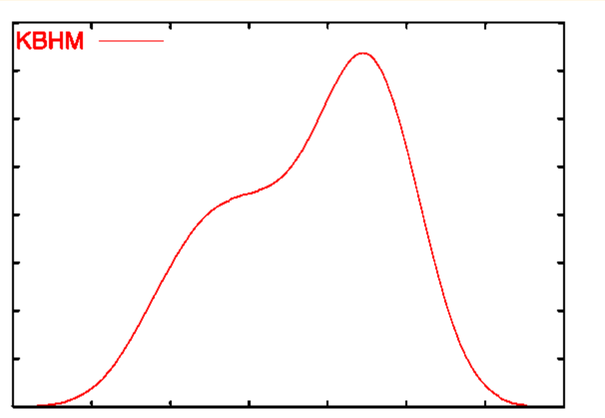
Skew in Forecast Distributions

	$(T_{50}-T_{10})$; Cold Tail	$(T_{90}-T_{50})$; Warm Tail
Mn-Ens-KDE		
	 0 5 10° F	
Mn-Mn-N		

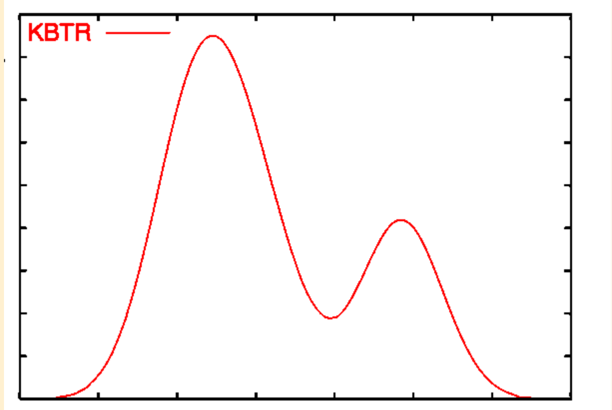
A "Rogue's Gallery" of Forecast PDFs



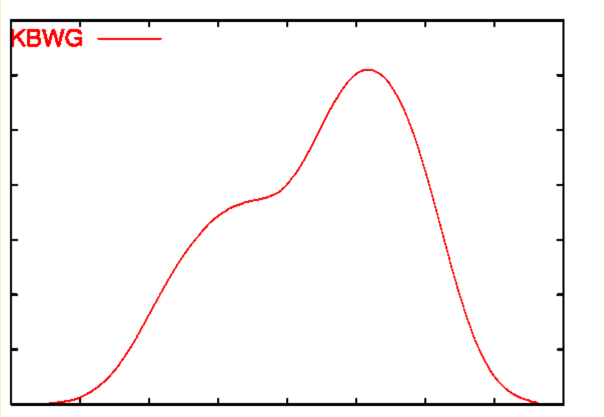
Waco, Texas



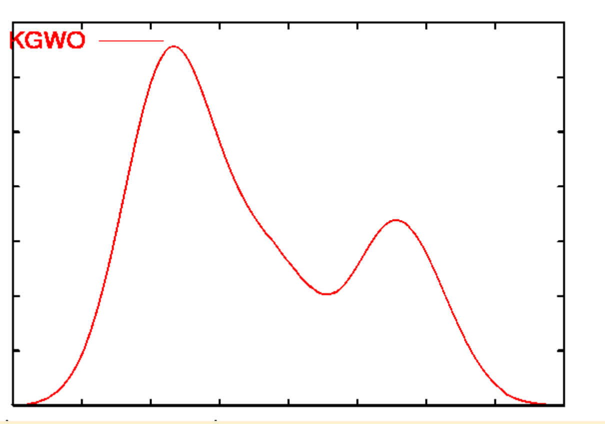
Birmingham, Alabama



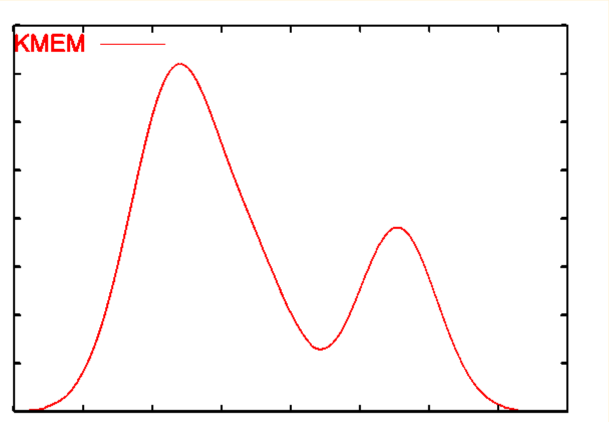
Baton Rouge, Louisiana



Bowling Green, Kentucky



Greenwood, Mississippi



Memphis, Tennessee

Conclusions

- These techniques can capture the uncertainty in temperature forecasts and routinely forecast probability distributions.
- Linear regression alone can be used to generate probability distributions from a single model run.
- Means of ensemble output variables are useful predictors.
- The Mn-Ens-KDE technique shows considerable promise, and it would be relatively easy to implement within the current MOS framework.
- Enumerating the points of the quantile function is an effective way to disseminate probability distributions.

Future Work

- Improve spread adjustment technique.
- Examine characteristics of forecast distributions and their variation.
- Verify individual stations.
- Extend temperature, dew point, maximum/ minimum temperature development to four forecast cycles and two seasons.
- Consider forecast sharpness and convergence as well as reliability and accuracy.
- Create forecast distributions of QPF and wind speed.
- Explore dissemination avenues.