Probabilistic Forecasting of Mixed Discrete-Continuous Weather Quantities Using Bayesian Model Averaging

J. McLean Sloughter

Based on research being conducted under Adrian E. Raftery and Tilmann Gneiting

This work was supported by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research under Grant N00014-01-10745.

Bayesian Model Averaging

- Weighted average of multiple models
- Weights determined by posterior probabilities of models
- Posterior probabilities given by how well each member fits the training data
- Weights, then, give an indication of the relative usefulness of ensemble members

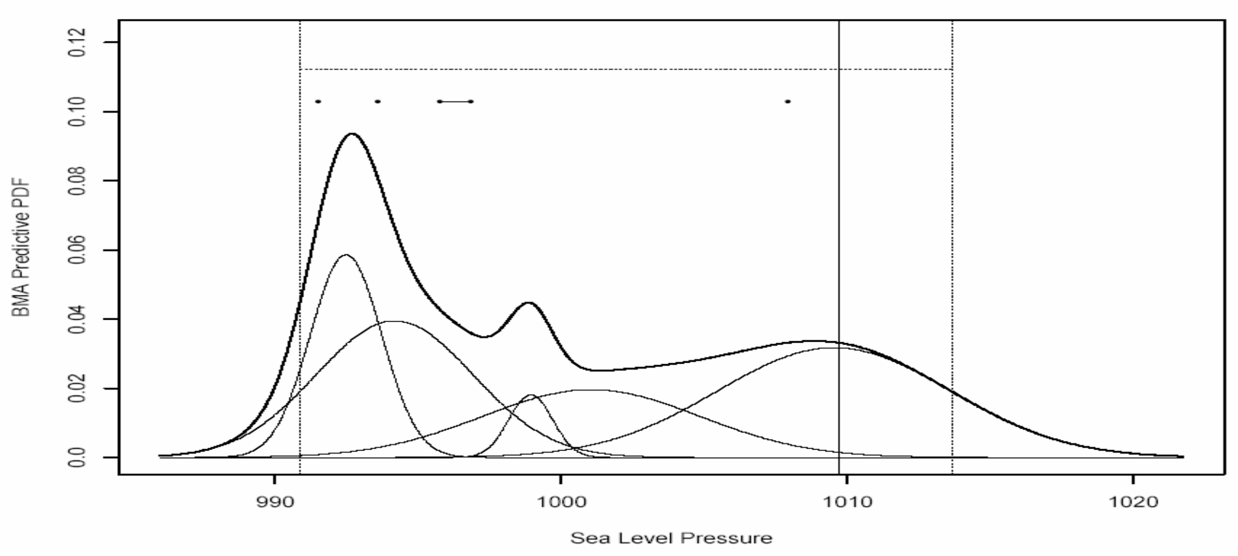
BMA for ensembles

$$p(Y \mid \tilde{f}_1, \ldots, \tilde{f}_k) = \sum_{i=1}^{k} w_i\, p(Y \mid \tilde{f}_i)$$

where $\tilde{f}_i$ is the forecast from member *i,*

w_i is the weight associated with member *i*, and

$p(Y \mid \tilde{f}_i)$ is the estimated distribution function for Y given member *i*



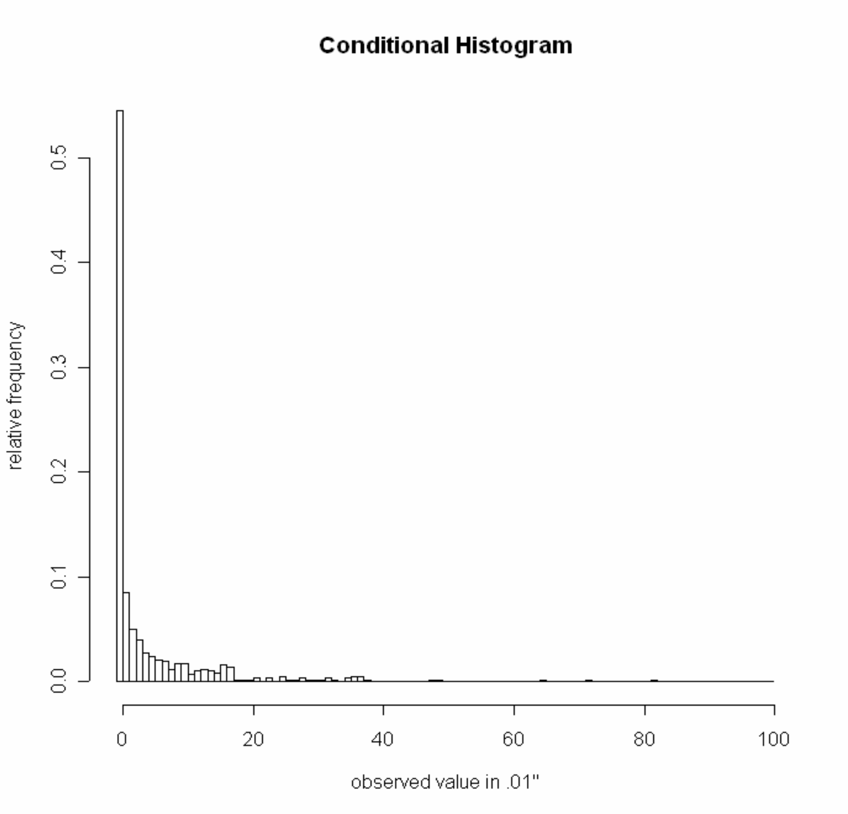
Picture taken from Raftery, Balabdaoui, Gneiting, and Polakowski (2003),
"Calibrated MesoscaleShort-Range Ensemble Forecasting Using Bayesian Model Averaging."

The Problem

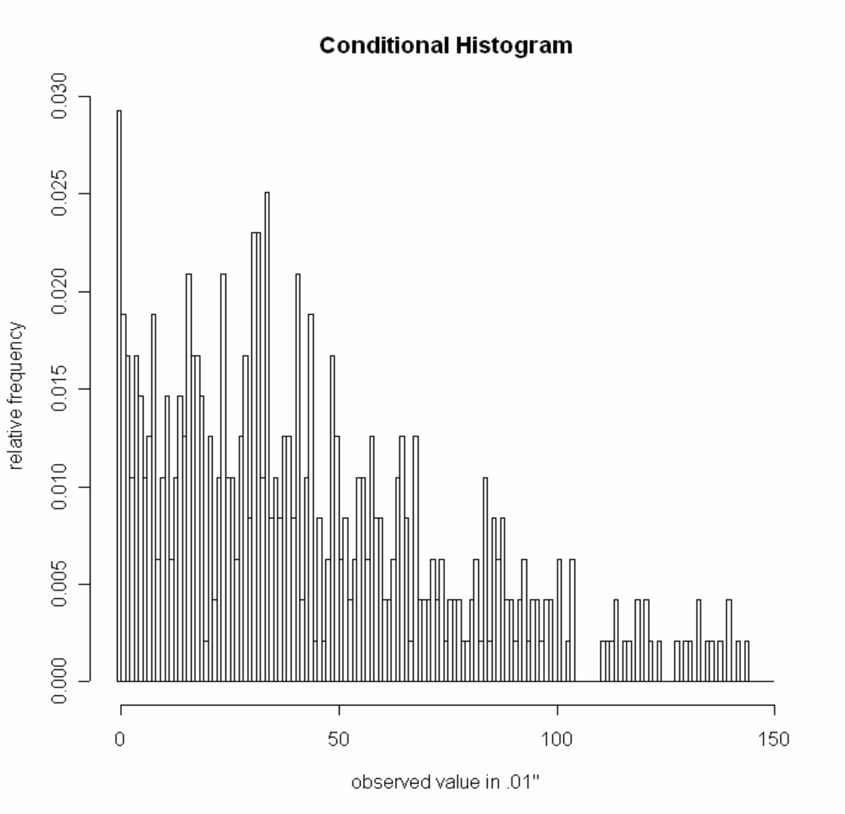
- Methods exist for using Bayesian Model Averaging to create probabilistic forecasts for weather quantities that can be expressed as a mixture of normals (Raftery et. al., 2005, MWR), such as temperature and pressure
- For quantities such as wind speed and precipitation, distributions are not only non-normal, but not purely continuous – there are point masses at zero

What Does a Precipitation Distribution Look Like?

Conditional Histograms



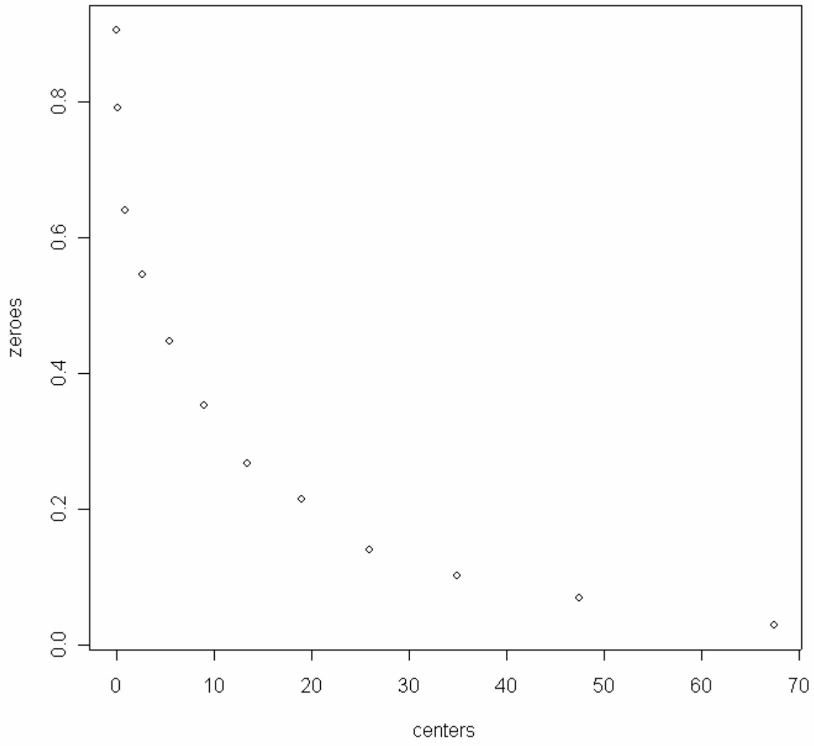
Observed given forecast from 1.5 to 4



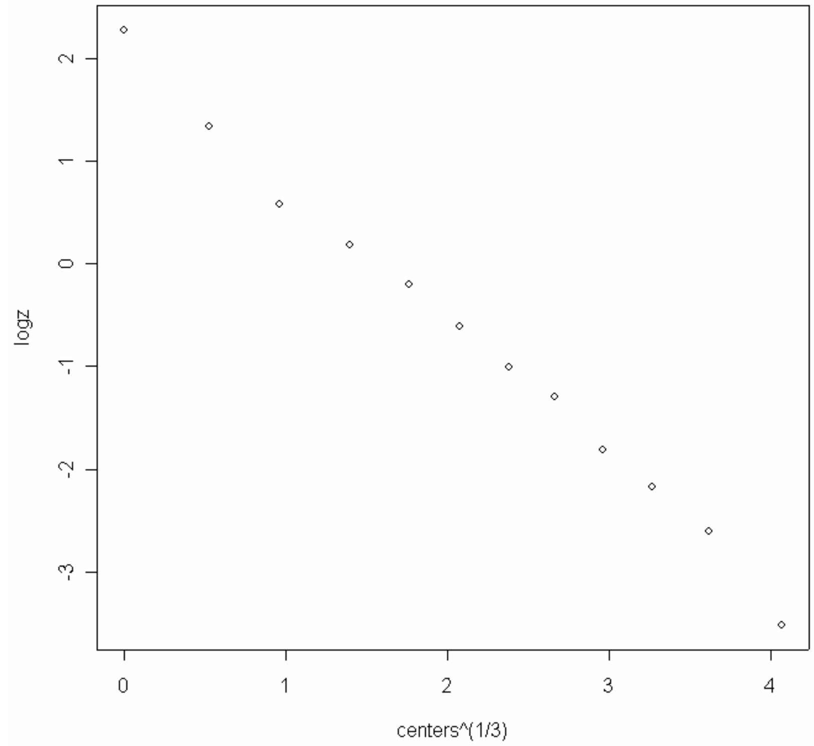
Observed given forecast from 55 to 80

in .01”

How to Model Zeroes

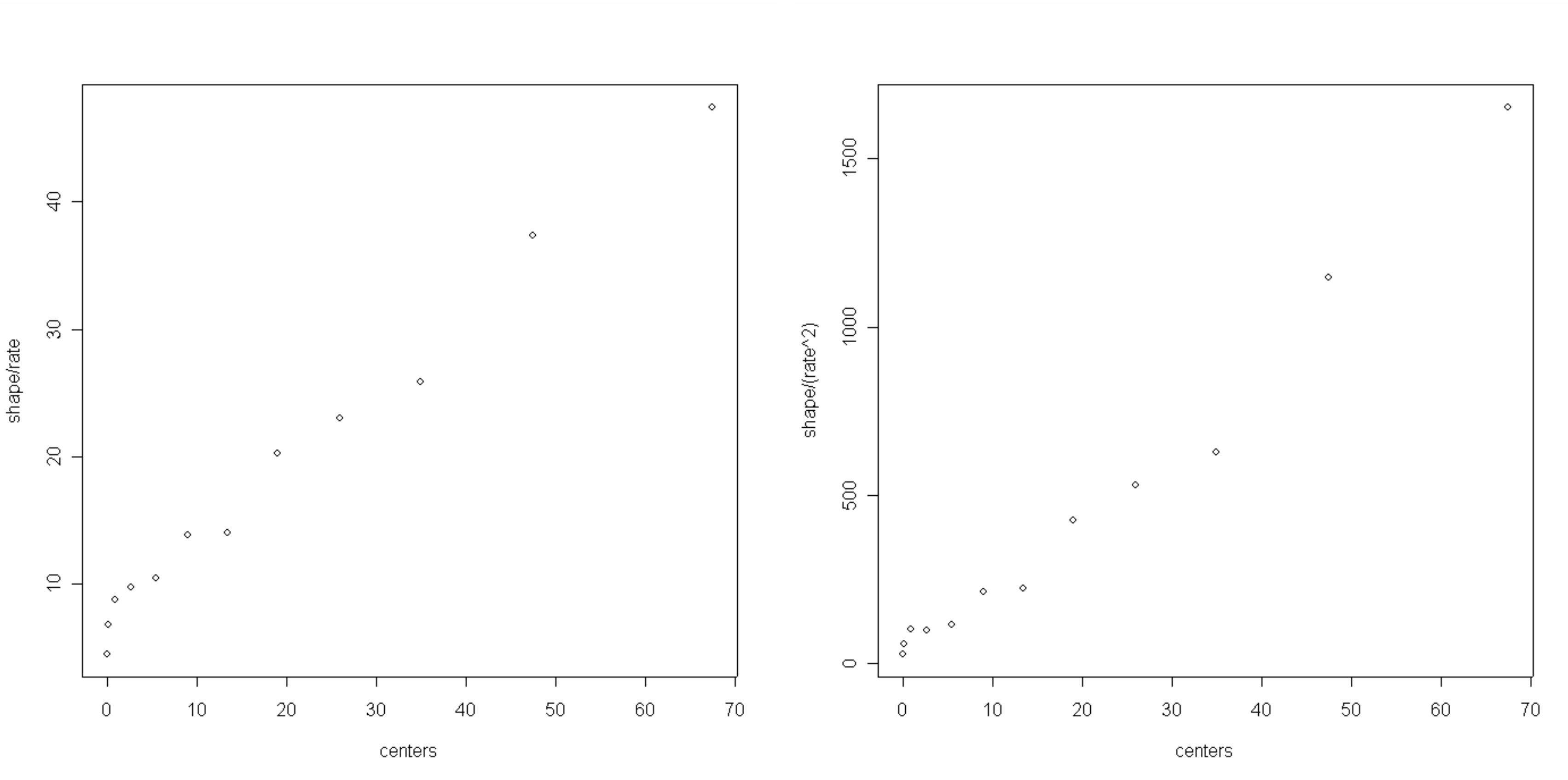


proportion of zeroes per bin



logit of proportion of zeroes versus cubed root of bin center

How to Model Non-Zeroes



mean (left) and variance (right) of fitted gammas on each bin

Complete Model

This suggests the following model:

$$p(Y \mid \tilde{f}_1, \ldots, \tilde{f}_k) = \sum_{i=1}^{k} w_i \left(\left[p_i(Y \neq 0 \mid \tilde{f}_i) g_i(Y \mid \tilde{f}_i, Y \neq 0) \right] I[Y \neq 0] + \left[p_i(Y = 0 \mid \tilde{f}_i) \right] I[Y = 0] \right)$$

where

$$\operatorname{logit}(p_i(Y = 0 \mid \tilde{f}_i)) = \alpha_i + \beta_i \sqrt[3]{\tilde{f}_i}$$

$$g_i(Y \mid \tilde{f}_i, Y \neq 0) \sim Gamma(\mu_{\tilde{f}_i}, \sigma^2_{\tilde{f}_i})$$

$$\mu_{\tilde{f}_i} = a_i + b_i(\tilde{f}_i)$$

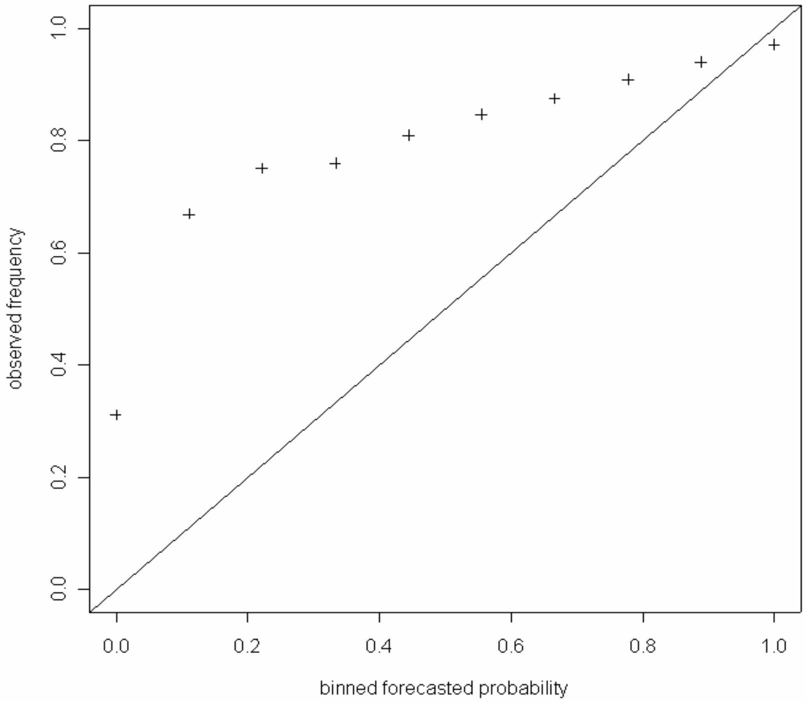
$$\sigma^2_{\tilde{f}_i} = c + d(\tilde{f}_i)$$

Current Implementation

- Fitting P(0) by logistic regression
- Fitting means as a linear bias correction
- Fitting variances and weights by EM algorithm

Results

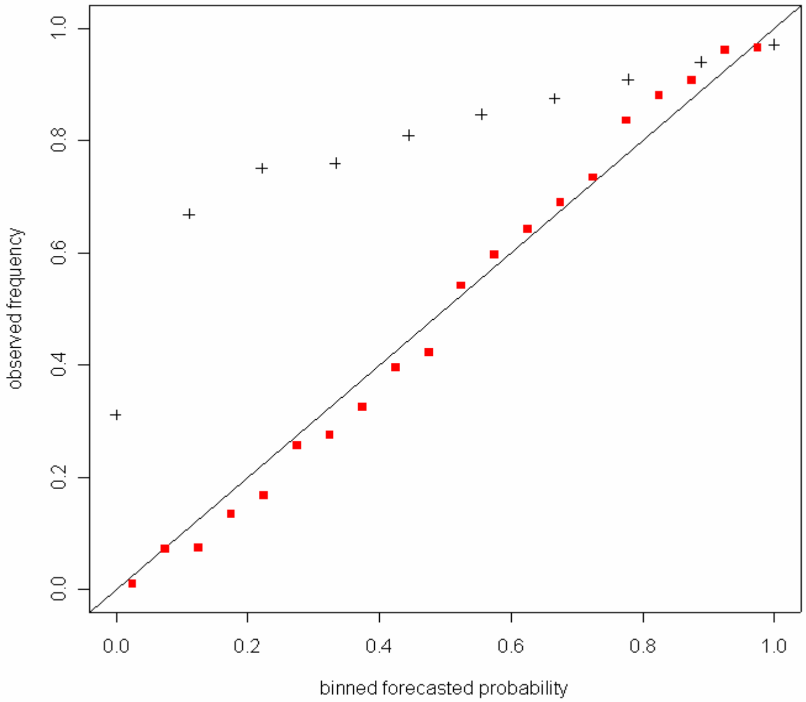
Results for December 12, 2002 through October 7, 2003 24-hour accumulation precip forecasts, with 40 Julian day training.



Consensus-voting forecasted probability of rain versus observed frequency of rain

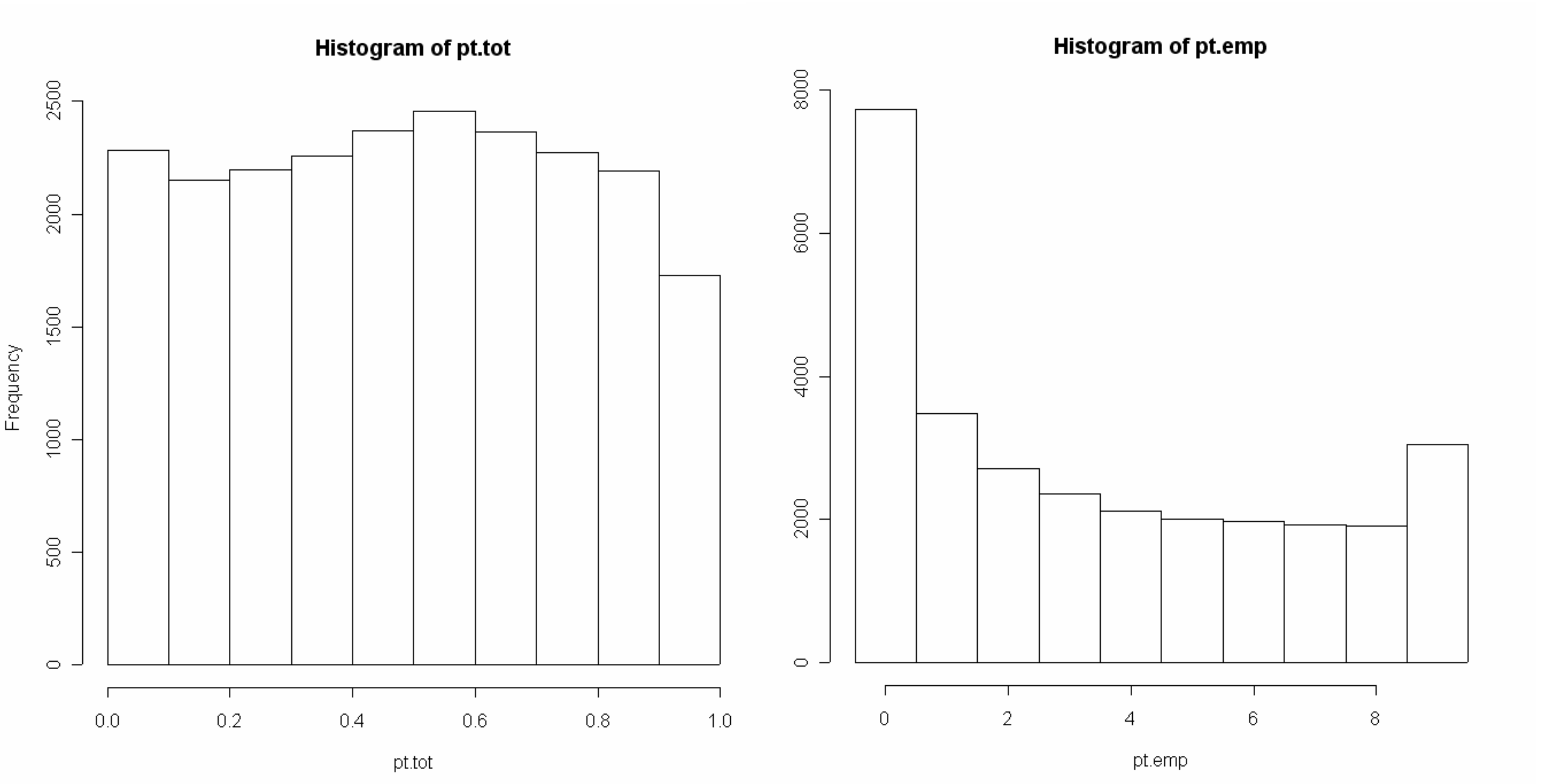
Results

Results for December 12, 2002 through October 7, 2003 24-hour accumulation precip forecasts, with 40 Julian day training.



Binned forecasted probability of rain versus observed frequency of rain – consensus voting as crosses, our model as red dots

PIT Histogram versus Verification Rank Histogram



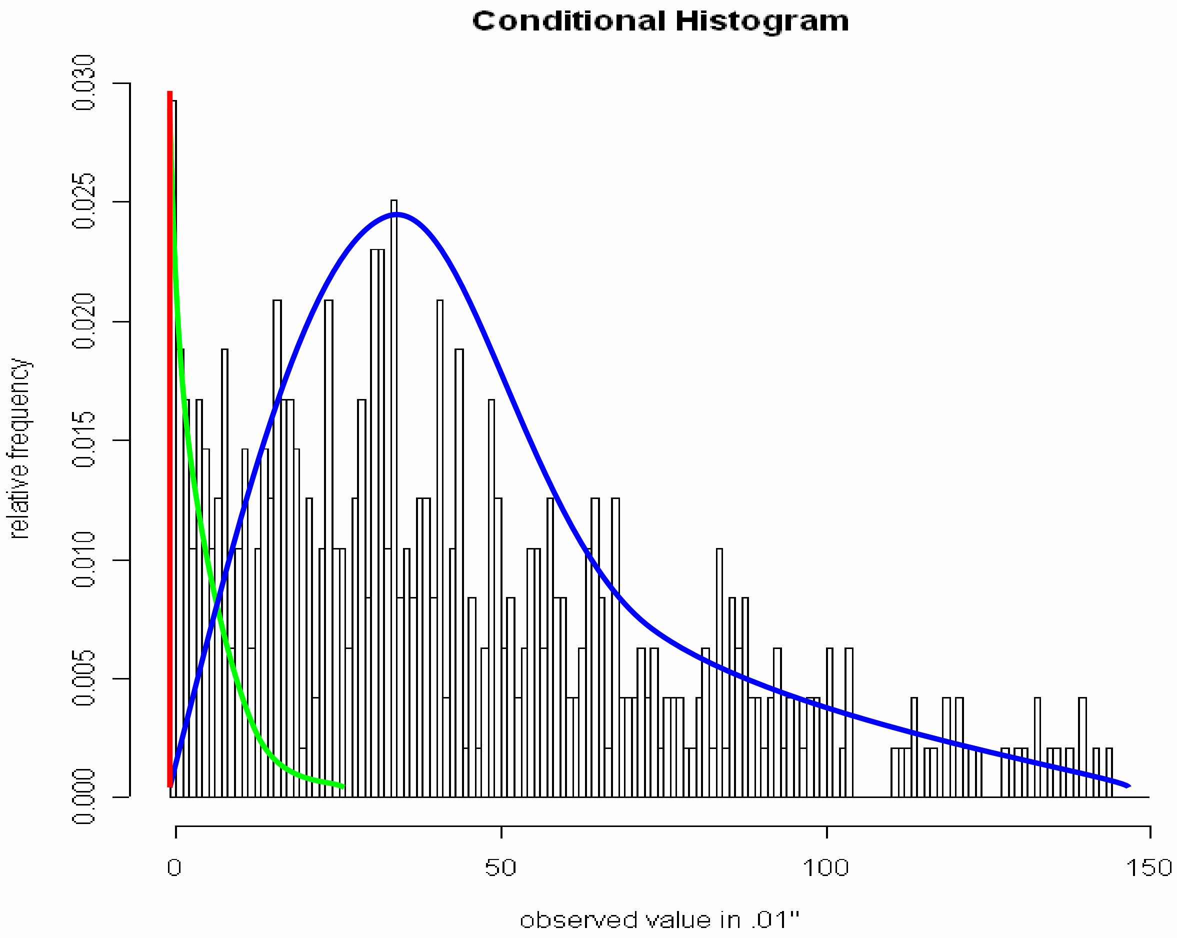
Interval Coverage

- 90% lower intervals had 95.9% coverage, with an average width of .18”
- 90% interval from climatology had a width of .24”
- 90% interval from raw ensemble had 92.9% coverage with an average width of .20”

What’s Next

- Try a more complicated model, fitting a point mass at zero, an exponential for “drizzle,” and a gamma for true rain around each member forecast

Proposed Model



Red: no rain, Green: drizzle, Blue: rain

Problems with this approach

- Potential of over-fitting
- Difficult to determine from inspection how the exponential should relate to the forecast
- Exponential could prove difficult to fit