A Practical Model Blending Technique Based on Bayesian Model Averaging

Bruce Veenhuis
bruce.veenhuis@noaa.gov
NWS/MDL
February 6, 2014

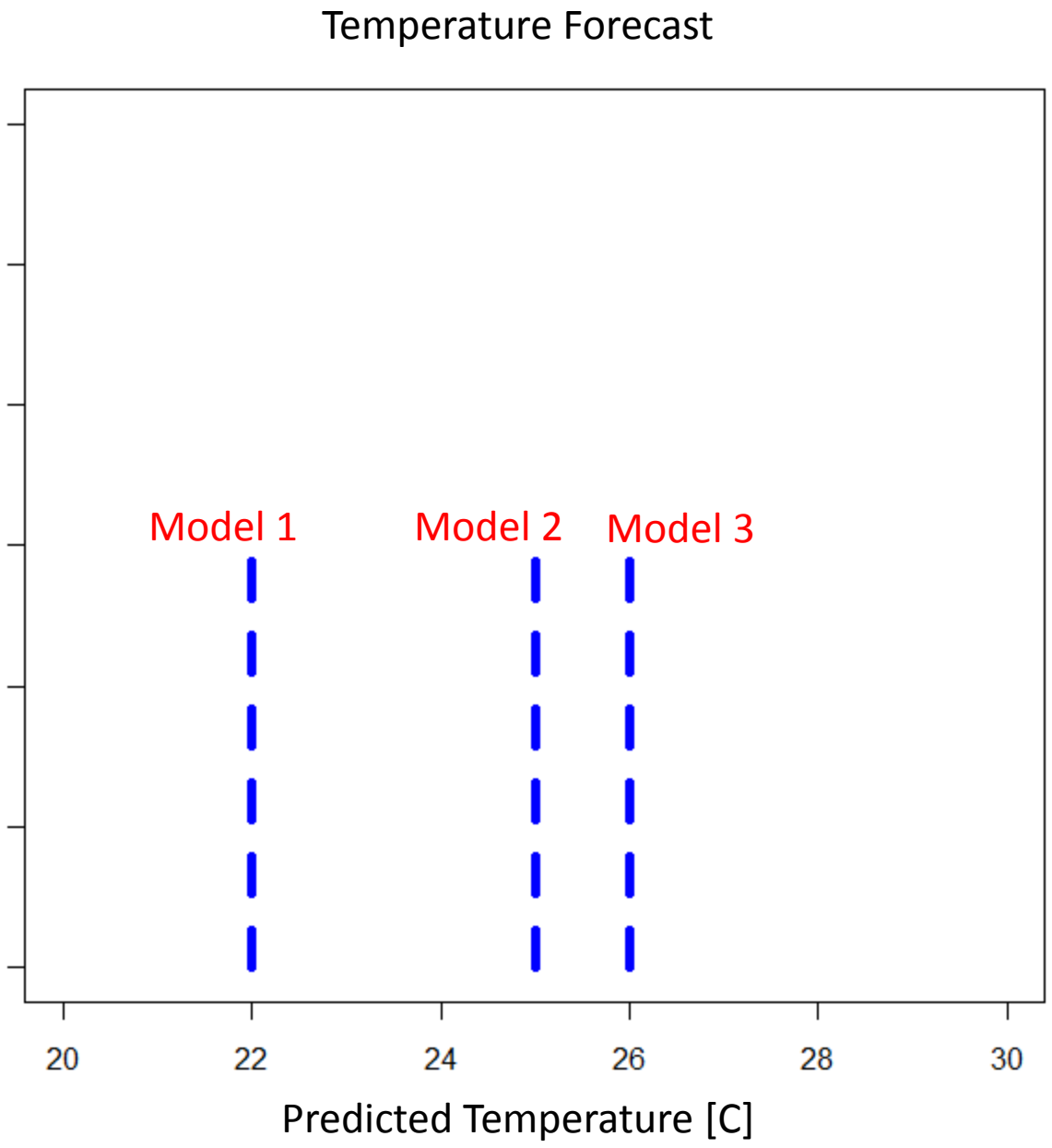
Acknowledgments: John L. Wagner, Michelle Cohen, Mark Oberfield

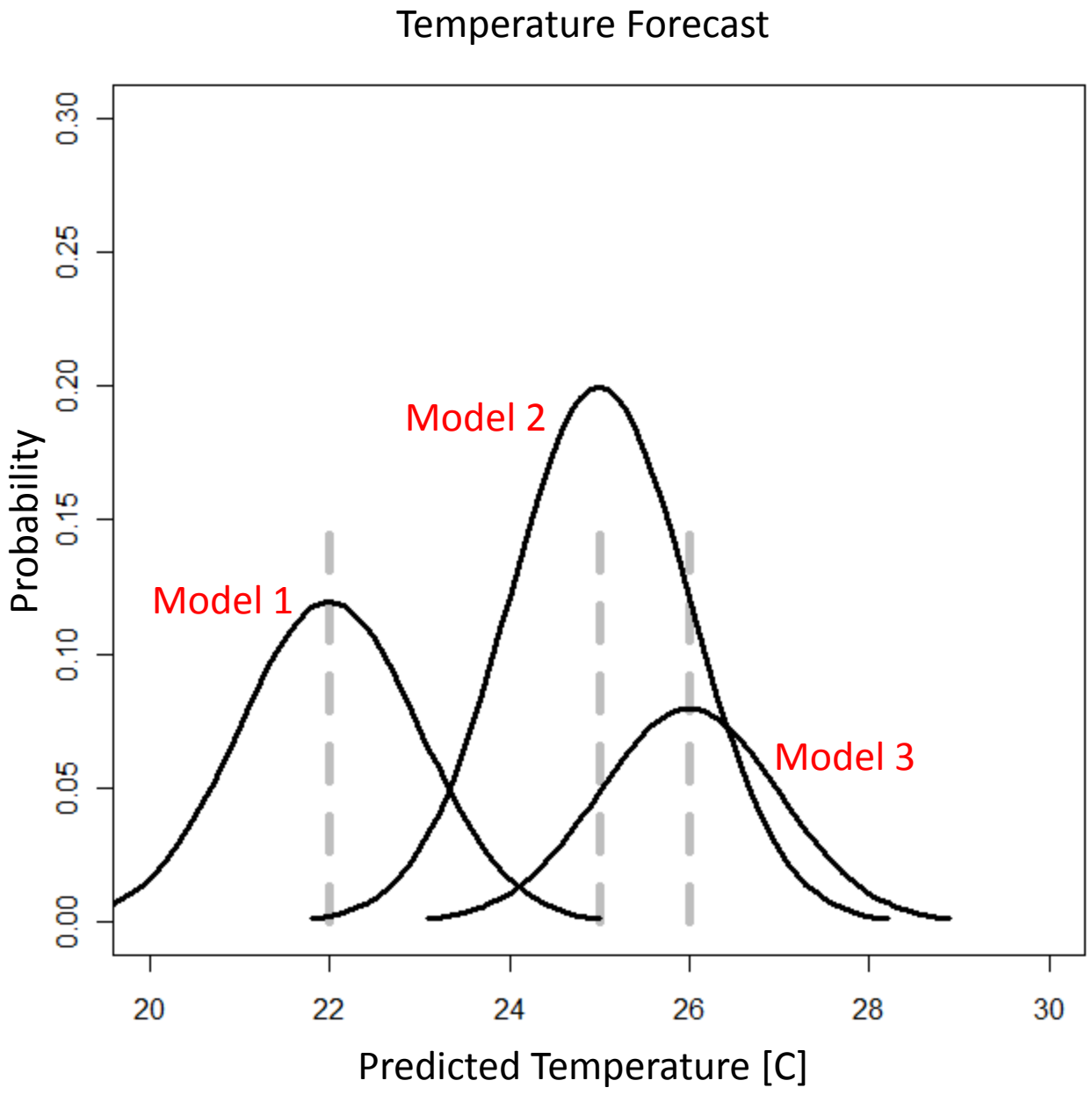
Motivation

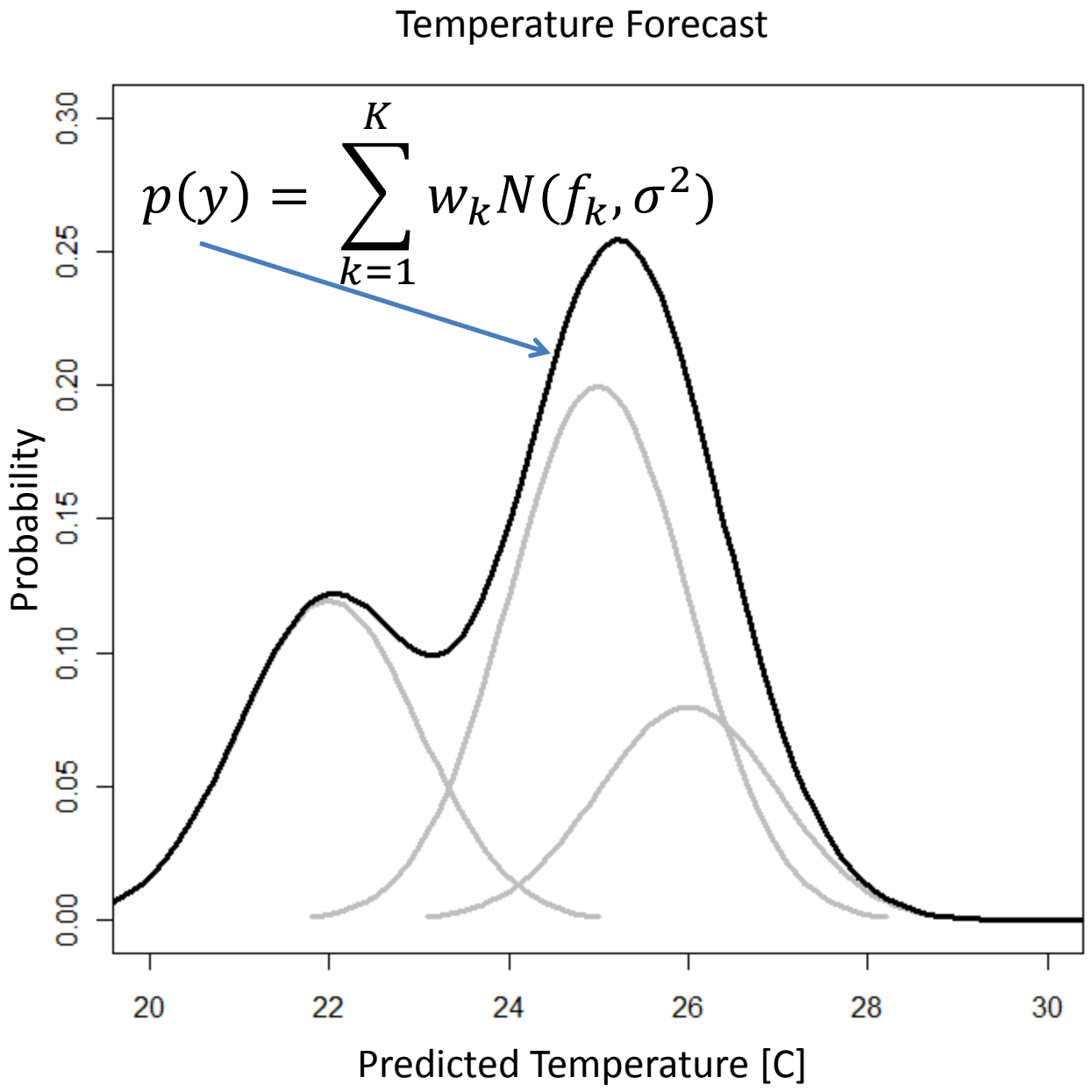
- Many operational meteorological centers run numerical weather prediction (NWP) models
  - Deterministic & Ensembles Forecasts
  - NCEP, Environment Canada, ECMWF
- We wish to create a single multi-model consensus
  - Optimally weight individual models
  - Create calibrated probability distributions
- Mainly concerned with sensible weather elements such as 2-m temperature, 10-m wind speed, etc.

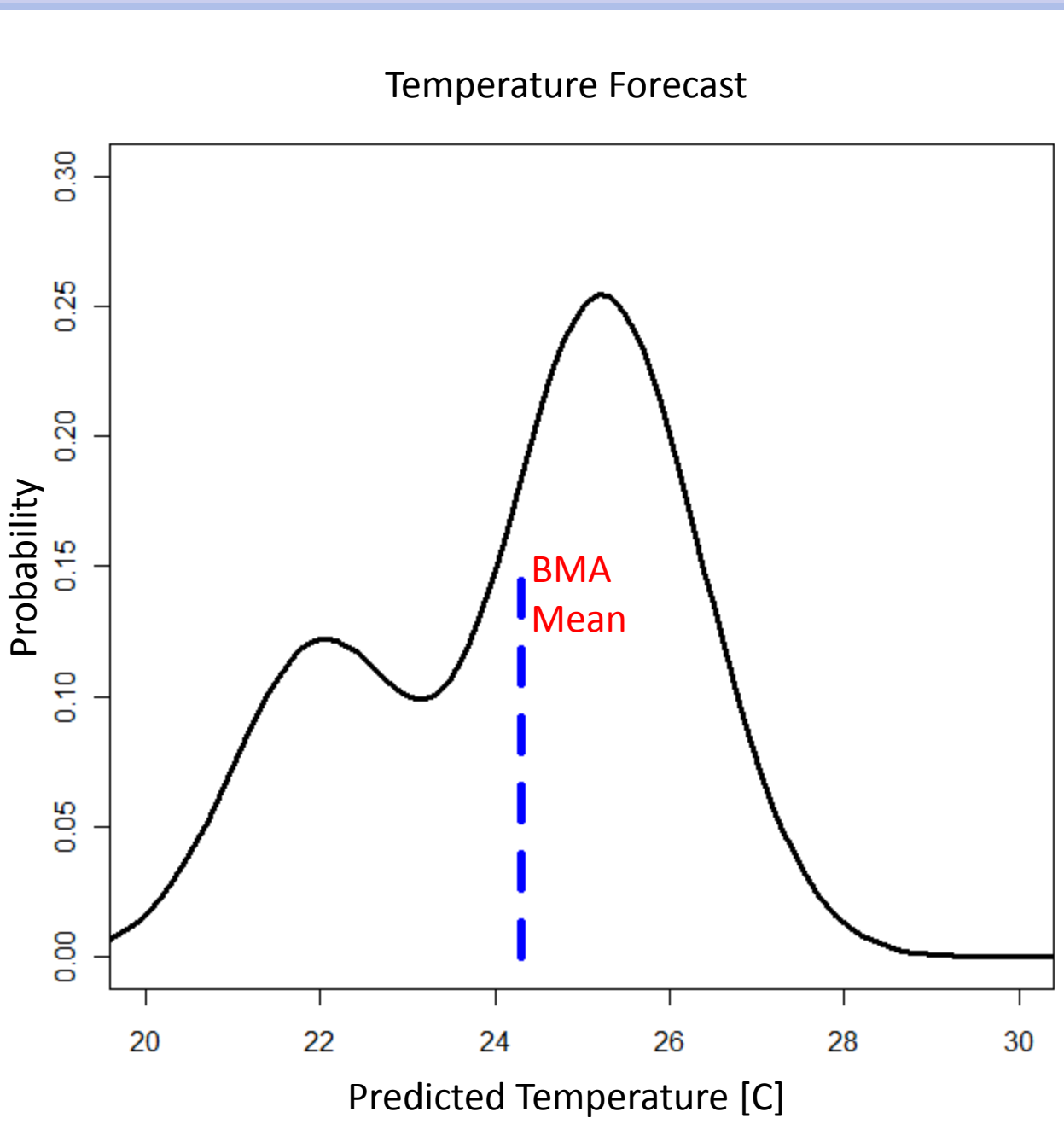
Bayesian Model Averaging (BMA)

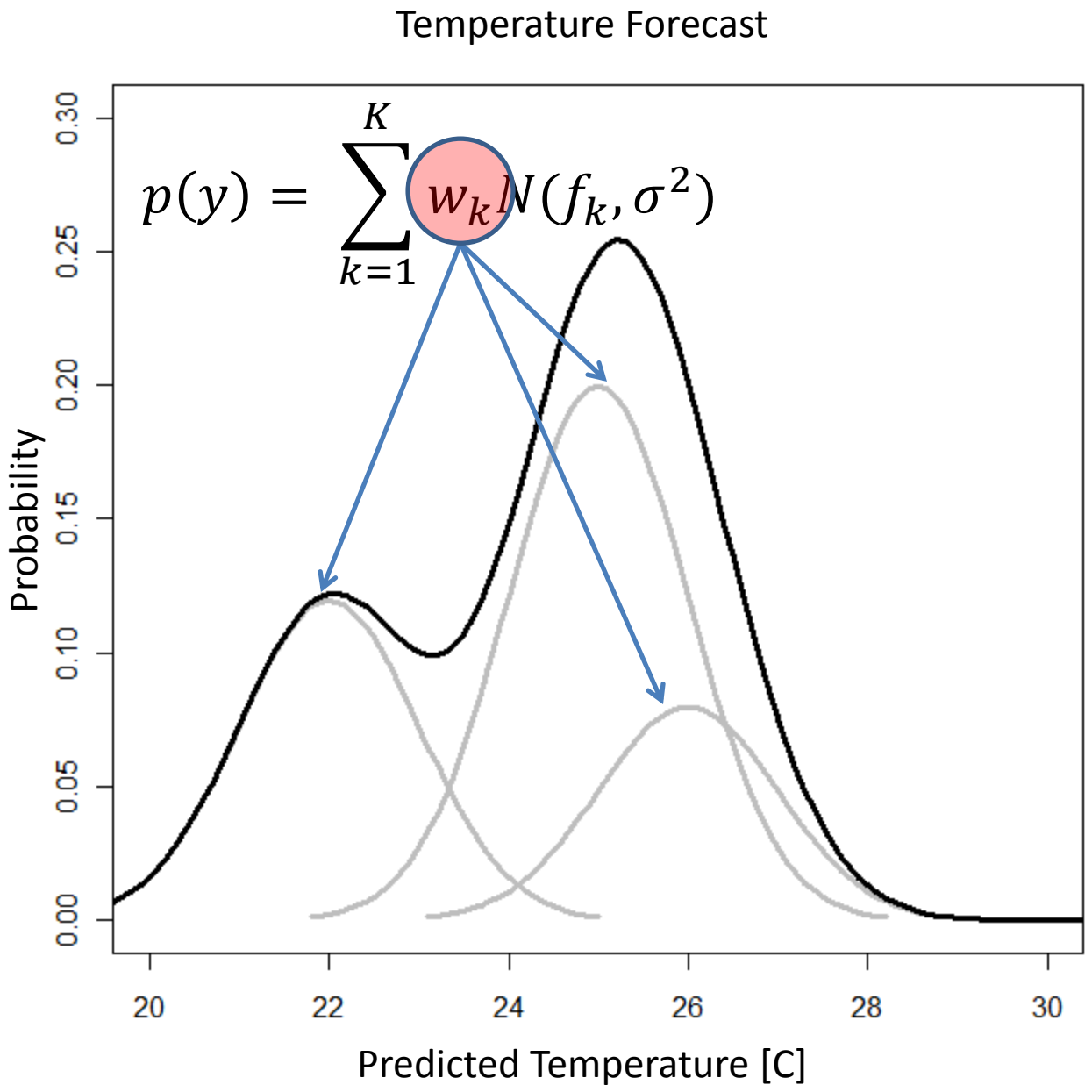
- A statistical postprocessing technique for ensembles (Raftery et al. 2005)
- BMA dresses each ensemble member with a probabilistic kernel
- Combine kernels to create a weighted mean forecast and a reliable probability distribution

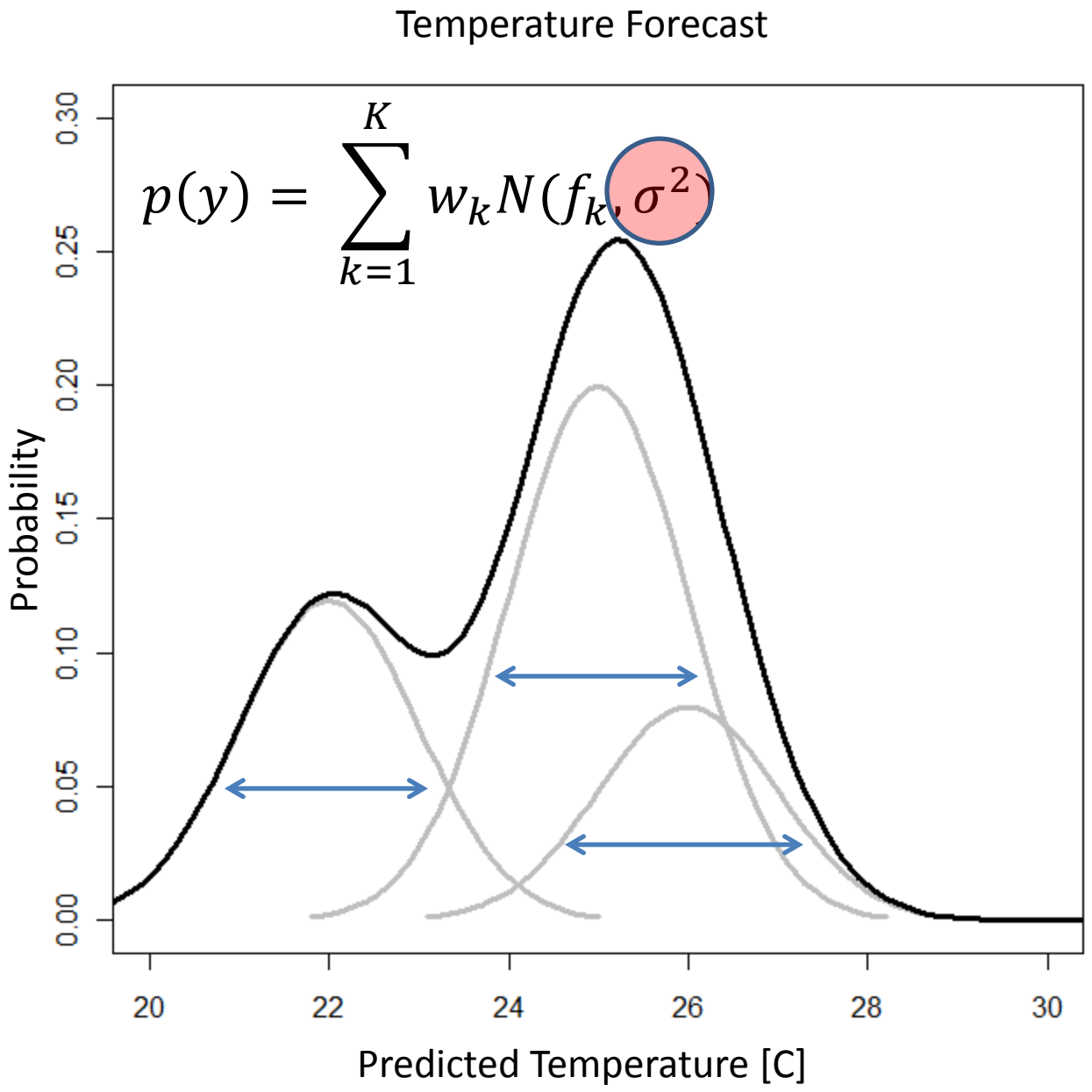

Temperature Forecast
Model 1
Model 2
Model 3
20
22
24
26
28
30
Predicted Temperature [C]


Temperature Forecast
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Model 1
Model 2
Model 3
20
22
24
26
28
30
Predicted Temperature [C]


Temperature Forecast
$p(y) = \sum_{k=1}^{K} w_k N(f_k, \sigma^2)$
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
20
22
24
26
28
30
Predicted Temperature [C]


Temperature Forecast
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
BMA
Mean
20
22
24
26
28
30
Predicted Temperature [C]

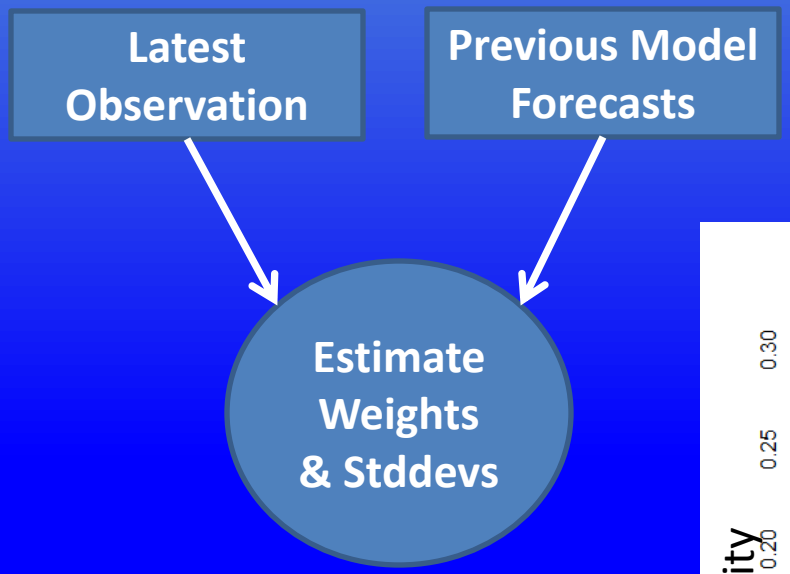

Temperature Forecast
$p(y) = \sum_{k=1}^{K} w_k N(f_k, \sigma^2)$
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
20
22
24
26
28
30
Predicted Temperature [C]

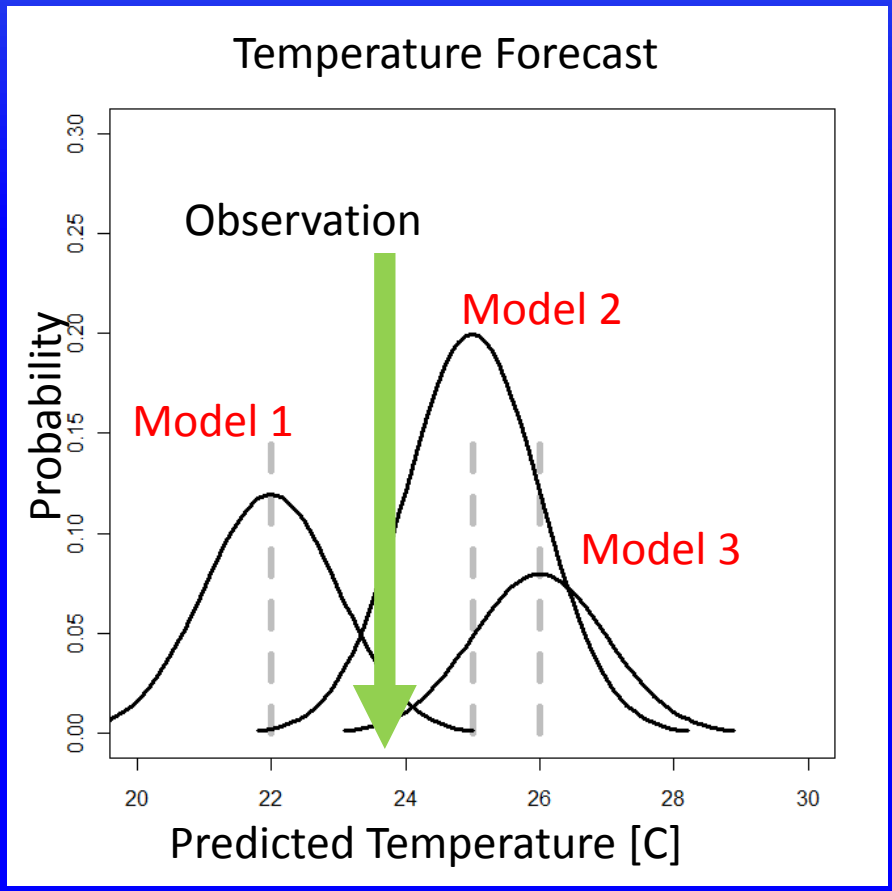

Temperature Forecast
$p(y) = \sum_{k=1}^{K} w_k N(f_k, \sigma^2)$
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
20
22
24
26
28
30
Predicted Temperature [C]

Decaying Average BMA (DABMA)

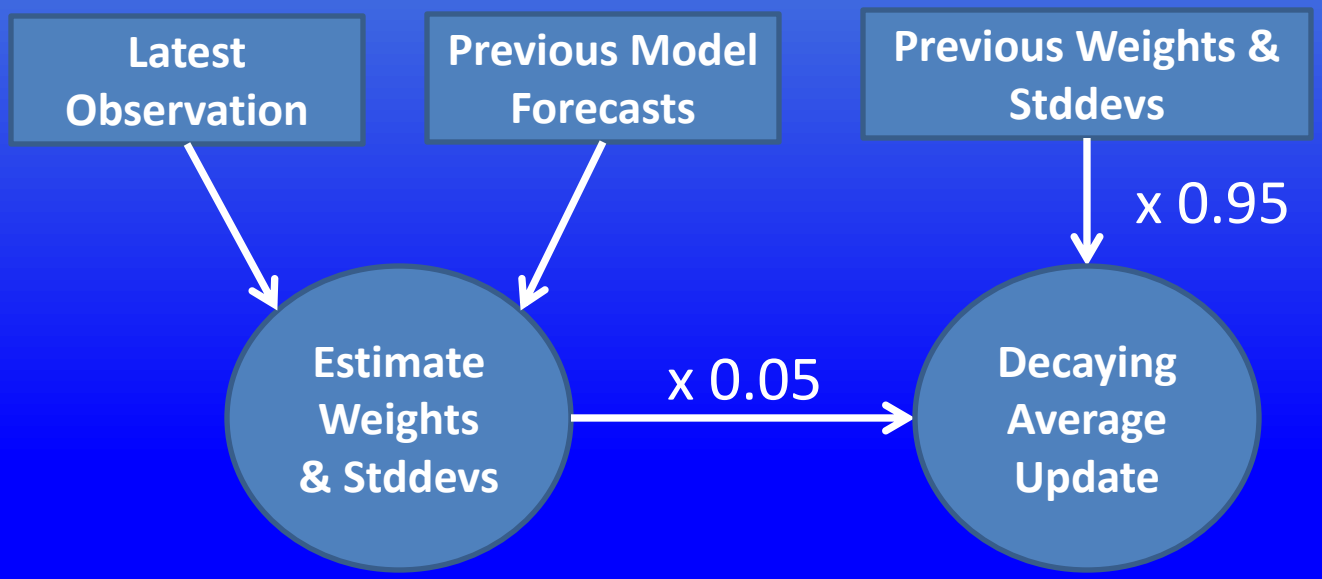
- MDL's implementation of BMA
  - First apply decaying average bias correction to each model
  - Continuously updates the weights and standard deviations with a decaying average algorithm
  - Training is based on recent performance (i.e. the last 60 days determines 95% of the weighting)
  - Simple to implement and computationally cheap

Decaying Average BMA Basics

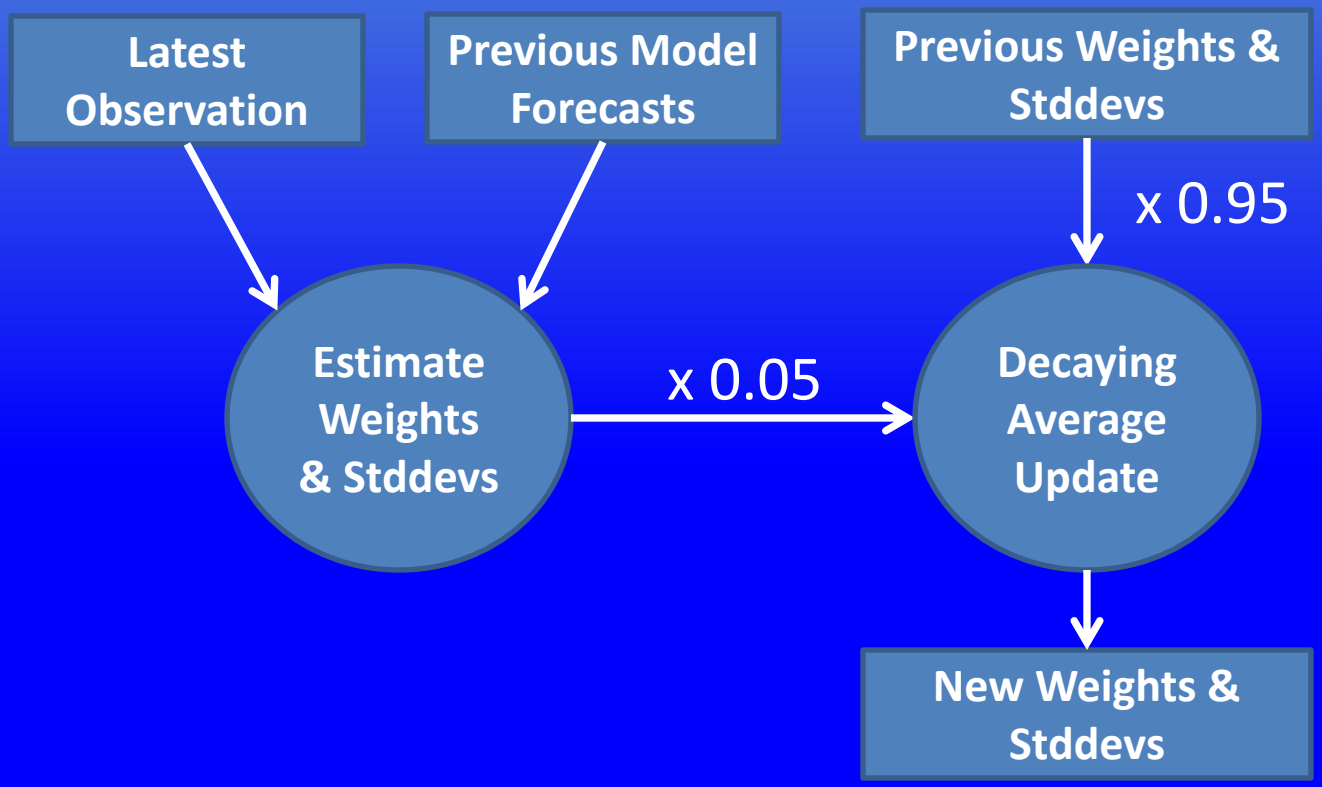




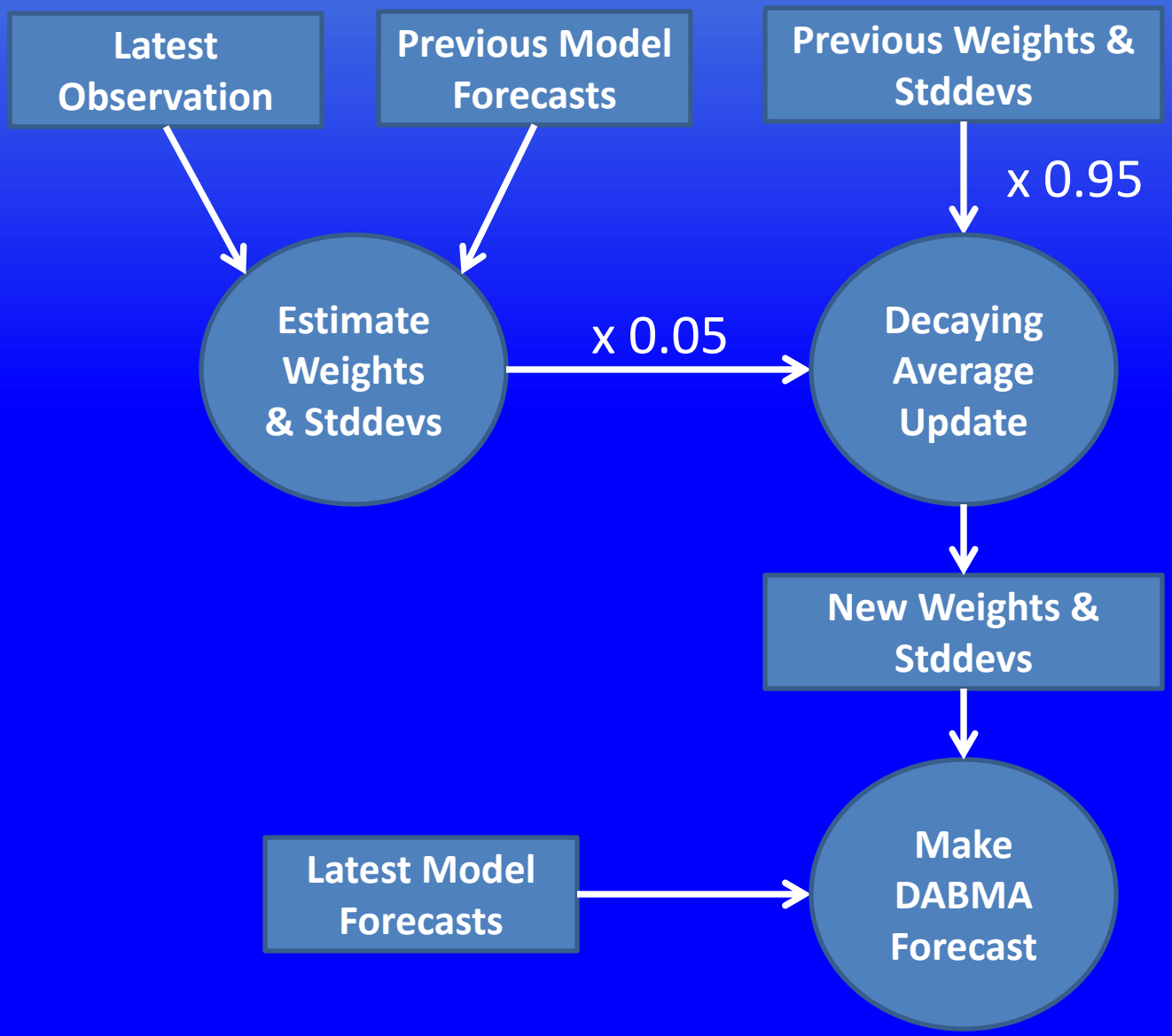
Decaying Average BMA Basics



Decaying Average BMA Basics

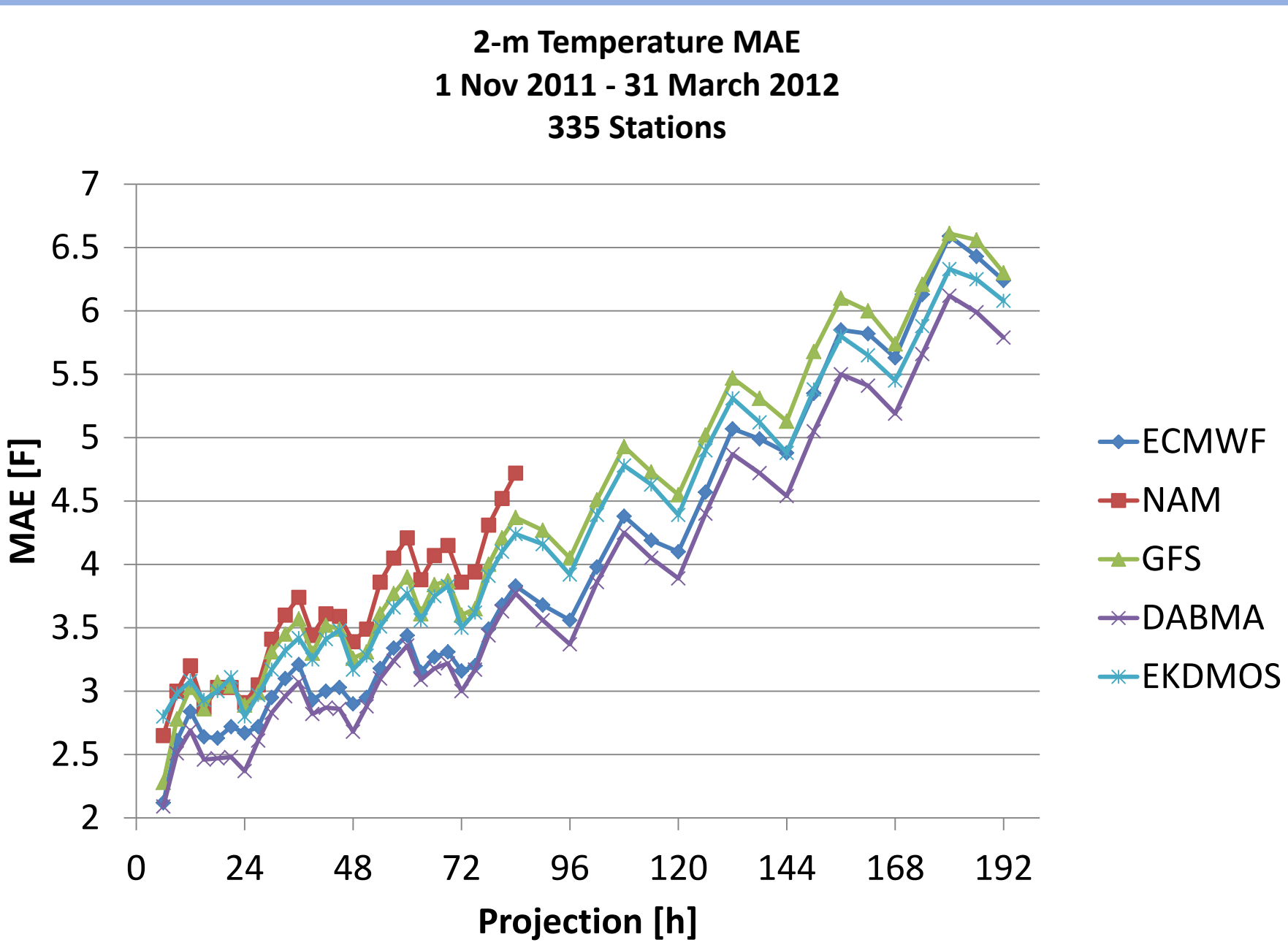


Decaying Average BMA Basics

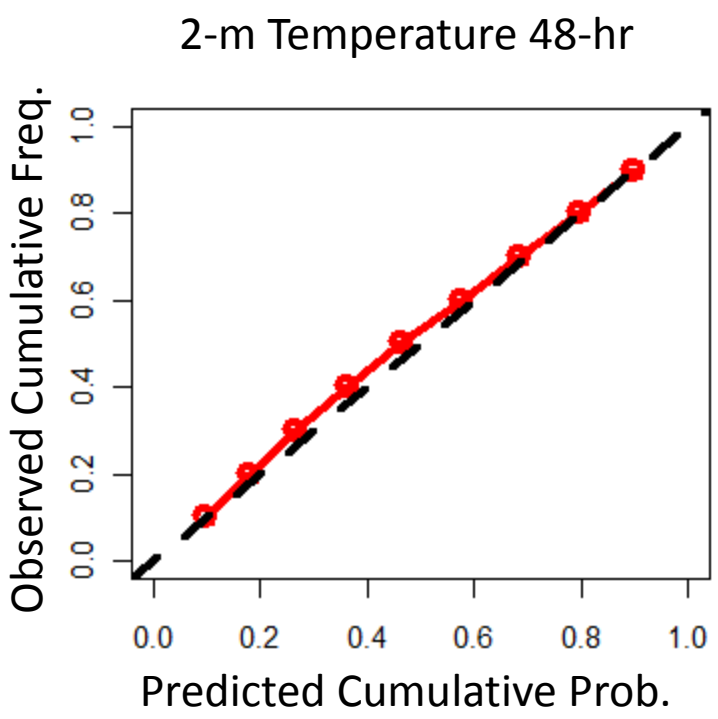


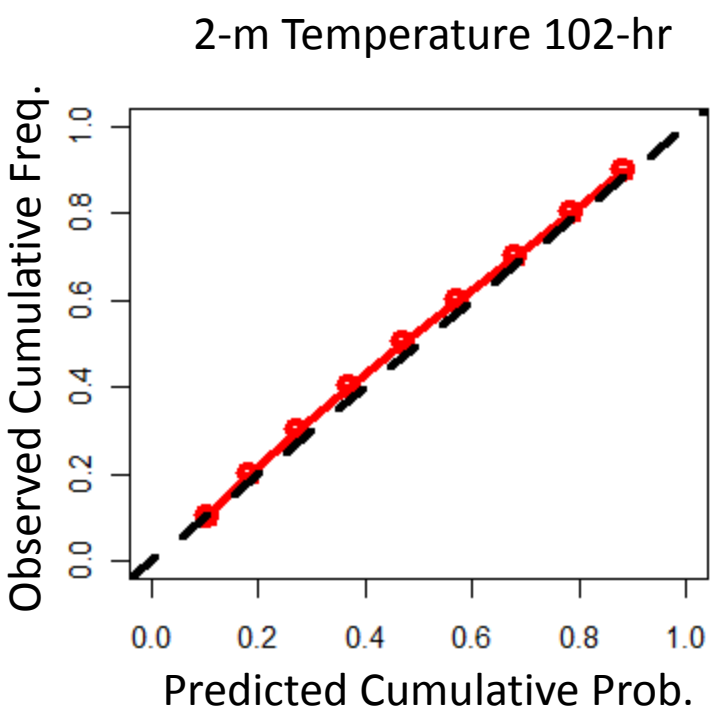
Example Application 1: Consensus MOS with DABMA

- MDL creates a variety of MOS guidance
  - GFS
  - NAM
  - ECMWF
  - EKDMOS Mean (GEFS and CMCE)
- 1 November 2011 – 31 March 2012
- 2-m Temperature, 335 Stations
- Accuracy and Reliability


2-m Temperature MAE
1 Nov 2011 - 31 March 2012
335 Stations
MAE [F]
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
0
24
48
72
96
120
144
168
192
Projection [h]
ECMWF
NAM
GFS
DABMA
EKDMOS

Cumulative Reliability Diagrams

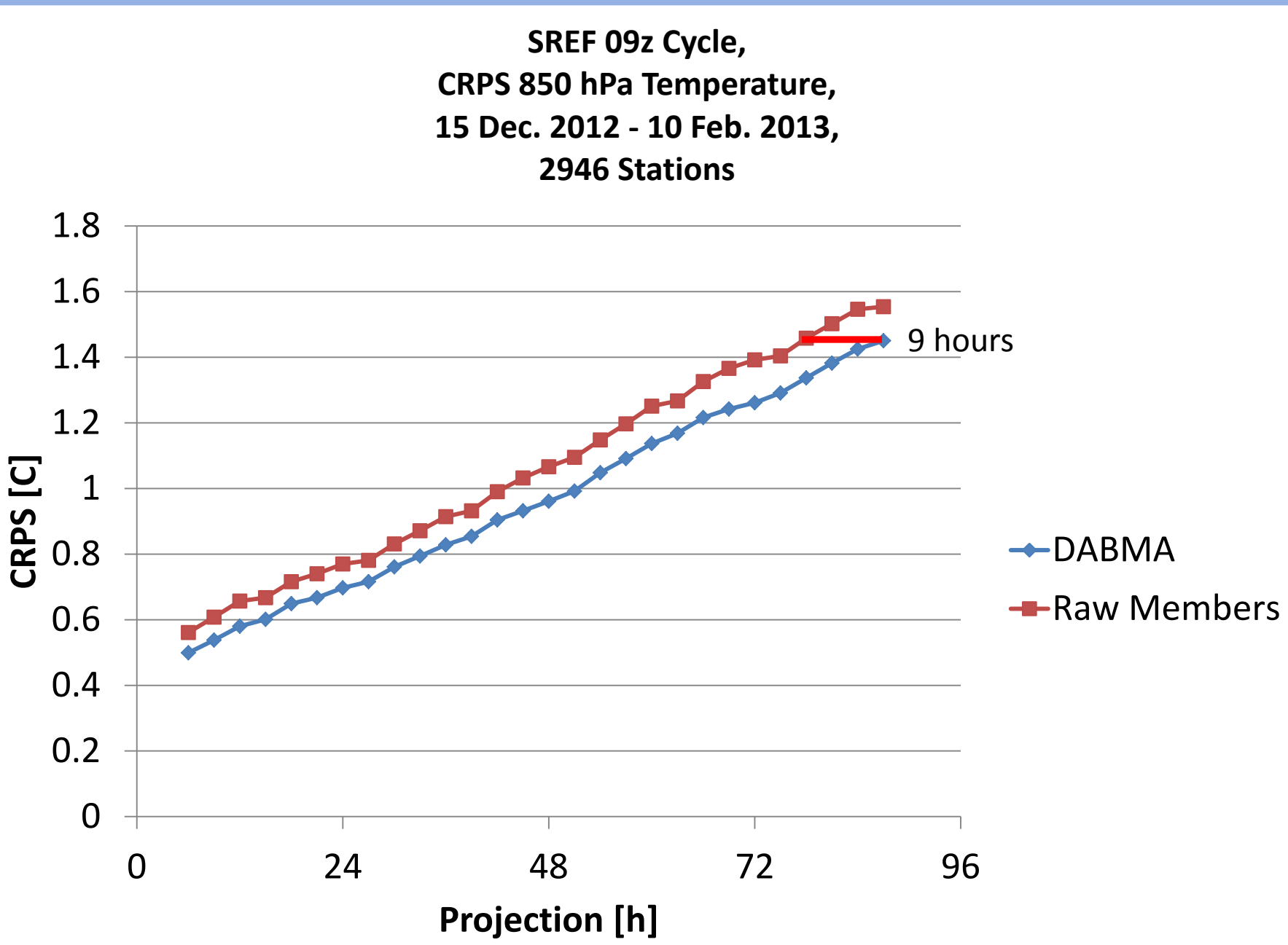




Example Application 2: DABMA Applied to SREF

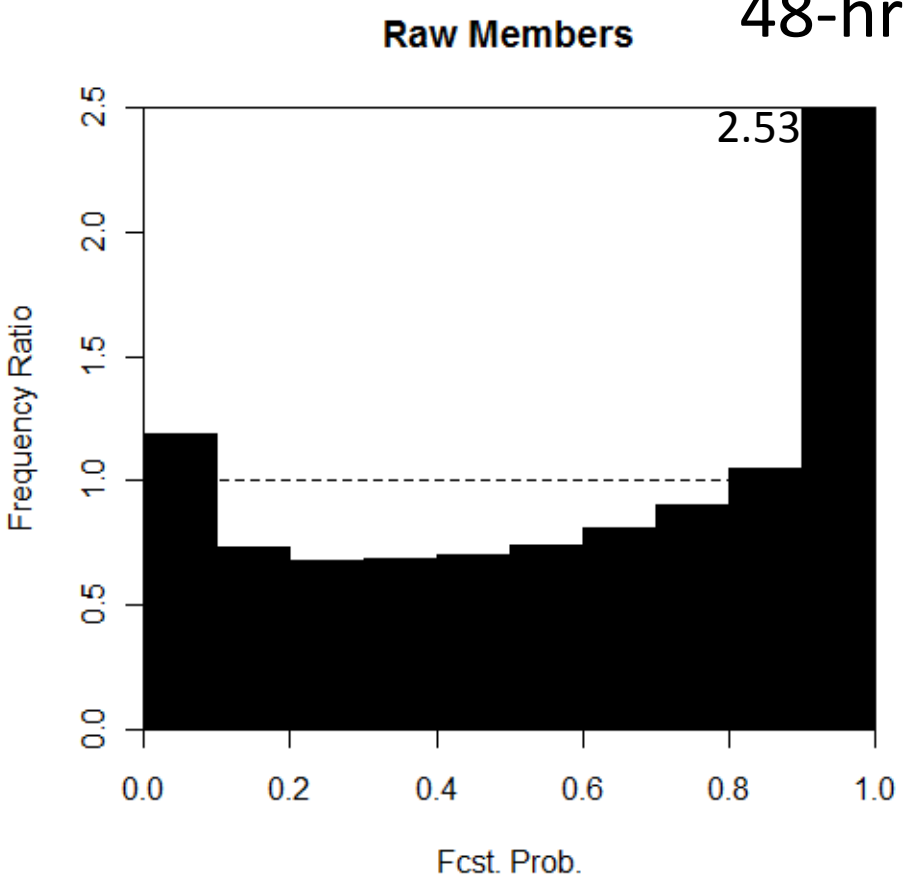
- Calibrated 850 hPa temperature forecasts from the Short Range Ensemble Forecast (SREF)
- To support precipitation type forecasting
- NDAS analysis: proxy for truth
- SREF and NDAS interpolated to stations
- Experimental Products Available:

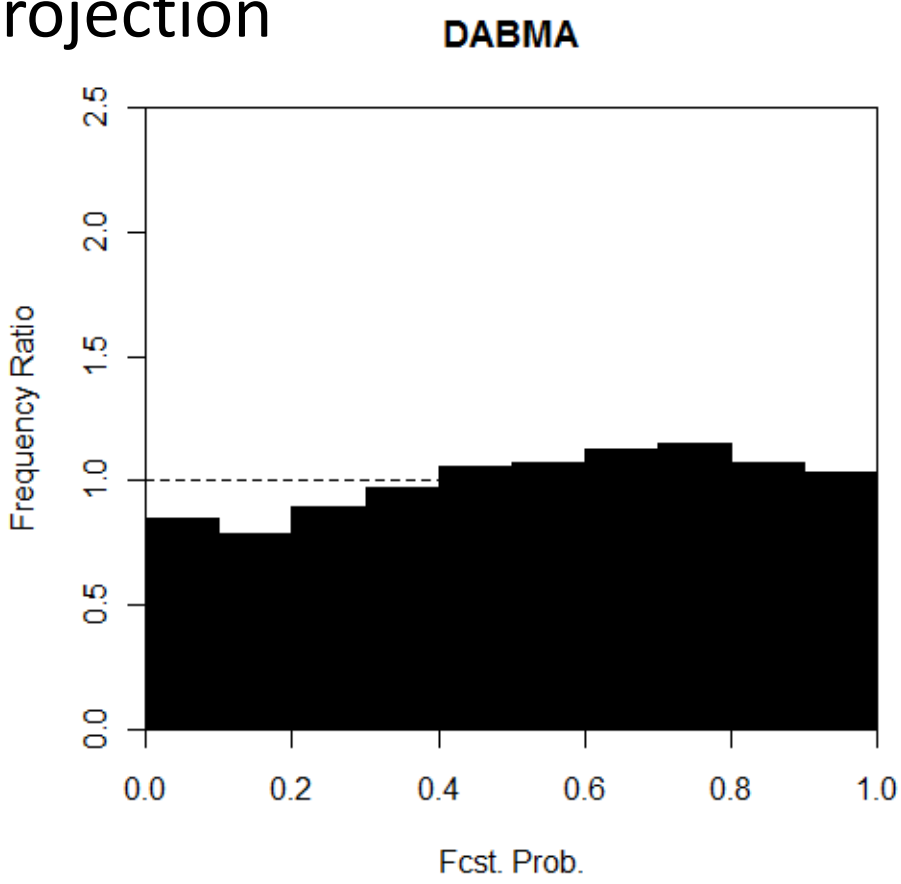
http://www.mdl.nws.noaa.gov/~BMA-SREF/BMAindex.php


SREF 09z Cycle,
CRPS 850 hPa Temperature,
15 Dec. 2012 - 10 Feb. 2013,
2946 Stations
CRPS [C]
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
24
48
72
96
Projection [h]
9 hours
DABMA
Raw Members

Probability Integral Transform (PIT) Histograms

48-hr Projection





Conclusions

- DABMA – MDL's implementation of BMA
  - Estimate weights and standard deviations with a decaying average algorithm
- Pros:
  - Computationally cheap
  - Simple to implement
  - Improves accuracy and reliability
- Cons:
  - Can only increase ensemble spread
  - May be problematic with an overdispersed ensemble
  - DABMA only tested for Gaussian elements