STATISTICAL POST-PROCESSING –

A FRAMEWORK & A TECHNICAL SOLUTION

Zoltan Toth and Roman Krzysztofowicz[1]

Global Systems Division, ESRL/OAR/NOAA

[1] University of Virginia

Acknowledgements:
Yuejian Zhu, NCEP Ensemble Team, NAEFS project

March 27, 2014 – NCEP Ensemble User Workshop

OUTLINE / SUMMARY

- User needs
  - Various weather parameters at different times / locations

- Forecast formats
  - Probabilities, joint probabilities – limited choices
  - Ensemble scenarios ultimate “carrier” of full information

- Post-processing framework
  - Pre-processing, fusion, output

- Bayesian Processor of Ensembles (BPE)
  - Theoretically sound, practical/flexible, high performance

- Great R2O opportunity
  - NOAA can leverage 10 yrs ($1M+) of development
  - Must take last steps

WHAT USERS NEED?

- **Sensitivity to weather** - each user affected by
  - *Various weather elements*
  - *At different times*
  - *At different locations*

- What users need for **optimal decision making**
  - Probability distributions for single variables?
    - Incomplete - lack of information on cross-correlations
  - Covariances needed across
    - Variables, space, and time

DATABASE CHOICES

- Will **joint probability distributions** work?
  - Provision of all joint distributions possibly needed by users – Intractable

- Proposed solution – **calibrated ensemble members**
  - Forecast based on NWP ensemble
  - Encapsulate all forecast info into calibrated ensemble members
    - Possible weather scenarios

6-D DATA CUBE (6DDC)

- **Format**
  - Space (3D), lead time (1D), variables (1D), ens. members (1D)
    - Global, with embedded finer scale regional guidance at shorter leads

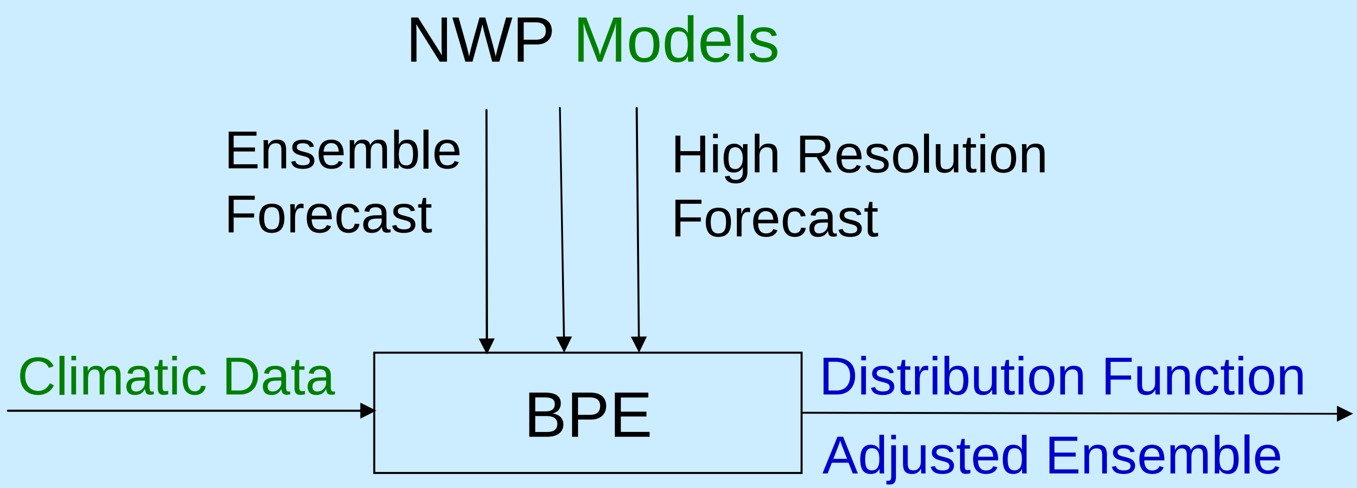
- **Quality requirements**
  - Consolidated
    - Single, authoritative source
  - Comprehensive
    - Addresses all needs, answers all questions
      - Aviation, other uses
  - Quality
    - Resolution, reliability

- **Interrogation toolkit** to
  - Derive any meteorological info
  - Answer any questions

MODULAR FRAMEWORK

COMPONENT	***PRE-PROSESSING***	***MERGING***	***ENSEMBLE OUTPUT***	***DOWNSCALING***
NEED	Reduce / homogenize bias	Combine all forecast & climate information	-	Connect model variables with user variables
NAEFS TOOLS / STEPS	Unconditional 1st moment bias correction	Equally weighed ensembles, ad hoc merge with high resolution control	-	Unconditional downscaling (GFS vs RTMA analyses)
PROPOSED TECHNOLOGY	-	**Bayesian Processor of Ensemble (BPE)**	Map ensemble members into posterior cdf	
POTENTIAL ADVANTAGE	-	**Objective, performance-based, higher moment bias correction, fuses climatology**	Preserves spatial / temporal / cross-variate covariance rank correlations	

CONCEPT OF BAYESIAN PROCESSOR



1. Extracts and Fuses Information
   maximizes *informativeness*
2. Quantifies Total Uncertainty
   guarantees *calibration*

Estimated for: predictand, grid point, day, lead time

Versions for: binary, multi-category, continuous predictands

MODELING CHALLENGES

Statistical properties of meteorological variates

1. Asymmetric samples from two sources
2. Non-stationary daily time series (seasonality)
3. Non-Gaussian marginal distributions (of many forms)
4. Non-linear, Heteroscedastic dependence structures
5. Non-random ensembles

Why to worry?

Want to: maximize INFORMATIVENESS

guarantee CALIBRATION

for all events (extremes as well)

at every grid point

on every day of year

$\Rightarrow$ for every real user (not “average user”)

*BPE offers theoretical and numerical solutions

METHODOLOGY to Implement BPE

Comprehensive approach to post-processing
(formulated jointly with *Zoltan Toth*)

Mathematical models and statistical procedures
(tested partially: on temperature, precipitation
at representative points)

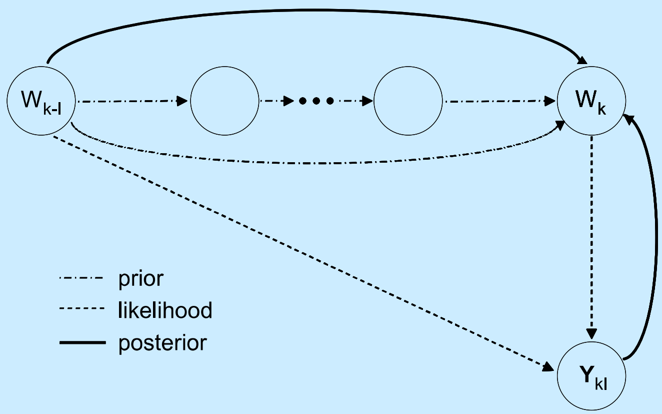
1. STANDARDIZATION of Time Series
2. NORMAL QUANTILE TRANSFORM (NQT) of Variate
3. SUFFICIENT STATISTICS of Ensemble
4. Basic GAUSSIAN-GAMMA BPE

BAYESIAN THEORY: General

W_k — predictand on day k $(k = 1, \ldots, 365)$

W_{k-l} — antecedent observation on day $k-l$ (forecast day)

$\mathbf{Y}_{kl}$ — ensemble forecast (vector of estimators) for day k with lead time l days



$h_{kl}(w_k \mid w_{k-l})$ — climatic l-step transition density function (prior)

$f_{kl}(\mathbf{y}_{kl} \mid w_k, w_{k-l})$ — conditional density function (likelihood function)

Bayes Theorem: Posterior density function

$$\phi_{kl}(w_k \mid \mathbf{y}_{kl}, w_{k-l}) = \frac{f_{kl}(\mathbf{y}_{kl} \mid w_k, w_{k-l})\, h_{kl}(w_k \mid w_{k-l})}{\kappa_{kl}(\mathbf{y}_{kl} \mid w_{k-l})}$$

Total Probability Law: Expected density function

$$\kappa_{kl}(\mathbf{y}_{kl} \mid w_{k-l}) = \int_{-\infty}^{\infty} f_{kl}(\mathbf{y}_{kl} \mid w_k, w_{k-l})\, h_{kl}(w_k \mid w_{k-l})\, dw_k$$

BAYESIAN FORECASTING THEORY

"Why Should a Forecaster and a Decision Maker Use Bayes' Theorem?" R. Krzysztofowicz *Water Resources Research*, 19(2), 327–336, 1983.

⋮

"Probabilistic Forecasts from the National Digital Forecast Data Base", R. Krzysztofowicz and W.B. Evans, *Weather and Forecasting*, 23(2), 270–289, 2008.

"The Role of Climatic Autocorrelation in Probabilistic Forecasting", R. Krzysztofowicz and W.B. Evans, *Monthly Weather Review*, 136(12), 4572–4592, 2008.

INFORMATION FUSION

Example

Predictand: 2m temp. at 12 UTC — Forecast time: 00 UTC

Location: Savannah, Georgia — Lead time: 108 h

Data Samples for Estimation

Asymmetric samples — W — predictand

Two sources — $\mathbf{Y}$ — ensemble (vector)

- $\boxed{\text{Climatic sample of } W \text{ — long}}$

  NCEP/NCAR re-analysis

  Jan. 1959 – Dec. 1998 (40 years)

  Sample size: $M = 600$ per day ($40\ y \times 15\ d$)

  → estimate prior density function

- $\boxed{\text{Joint sample of } (\mathbf{Y}, W) \text{ — short}}$

  NCEP ensemble forecasts and analyses

  Mar. 2007 – Feb. 2009 (2 years)

  Sample size $N \simeq 360$ per season $(2\ y \times 180\ d)$

  → estimate likelihood function

Reduced need for hindcast sample

FORECAST for Savannah, GA, 1 May 2008, 00 UTC

For: 5 May 2008, 12 UTC

LT: 108 h [day 5]

Ensemble Forecast:

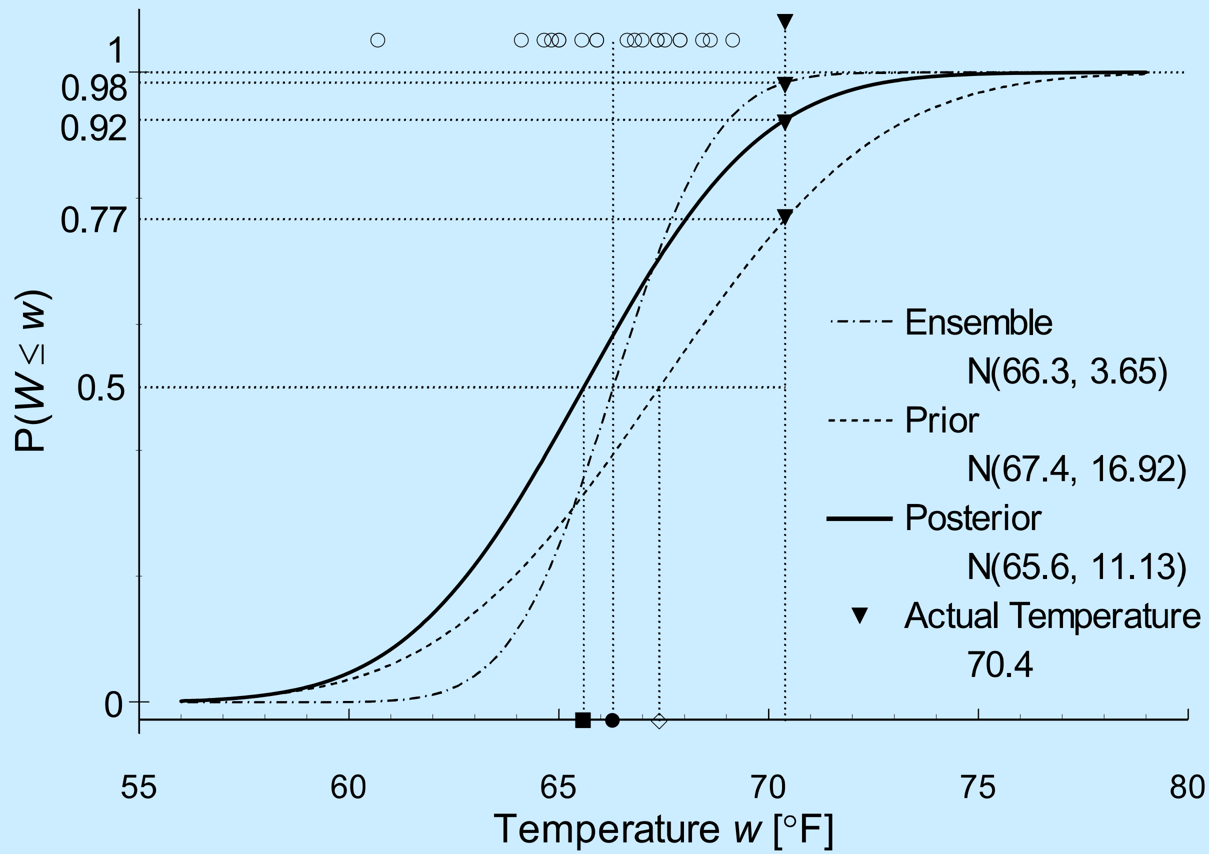
mean 66.3 °F

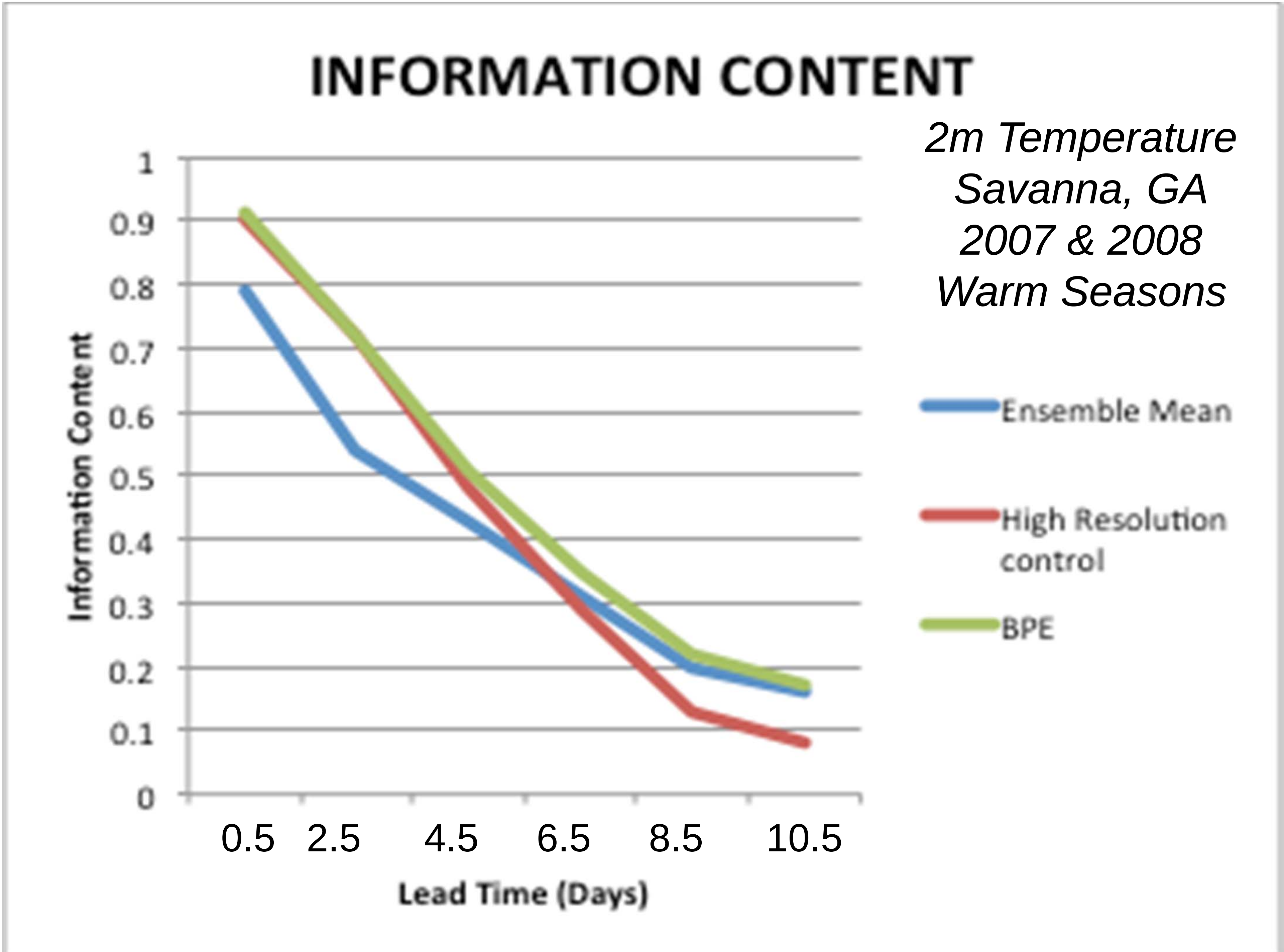
std. dev. 1.91 °F

Sufficient Statistics:

$x = -0.261$

$t = 2.057$




INFORMATION CONTENT
2m Temperature
Savanna, GA
2007 & 2008
Warm Seasons
Information Content
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.5
2.5
4.5
6.5
8.5
10.5
Lead Time (Days)
Ensemble Mean
High Resolution control
BPE

BPE — Basic Properties

Theoretically-based optimal fusion of ensemble forecast with climatic data

Revises prior (climatic) distribution given ensemble forecast
based on comparison of past forecasts with observations

1. CORRECT THEORETIC STRUCTURE
   - Always valid
   - Modular: Framework for different – modeling assumptions
     – estimation procedures

2. FLEXIBLE ANALYTIC MODELS
   - Handle distributions of any form (not only normal)
   - Handle non-linear, heteroscedastic dependence structures
   - Parametric (easy to estimate and manipulate)
   - Robust when joint sample is small (lesser need for "freeze" or "re-forecast ")

3. UNIQUE PERFORMANCE ATTRIBUTES
   - Removes bias in distribution
   - Guarantees calibration of the adjusted ensemble
     - Stable calibration (against climatic distribution | antecedent, regime)
     - Stationary calibration (equally good for all lead times)
     - User-specific calibration (point-specific, time-specific)
     - When predictability vanishes:
       adjusted ensemble ➞ climatic ensemble
   - Preserves spatial / temporal / cross-variate rank correlations in ensemble

BPE – GREAT R2O OPPORTUNITY

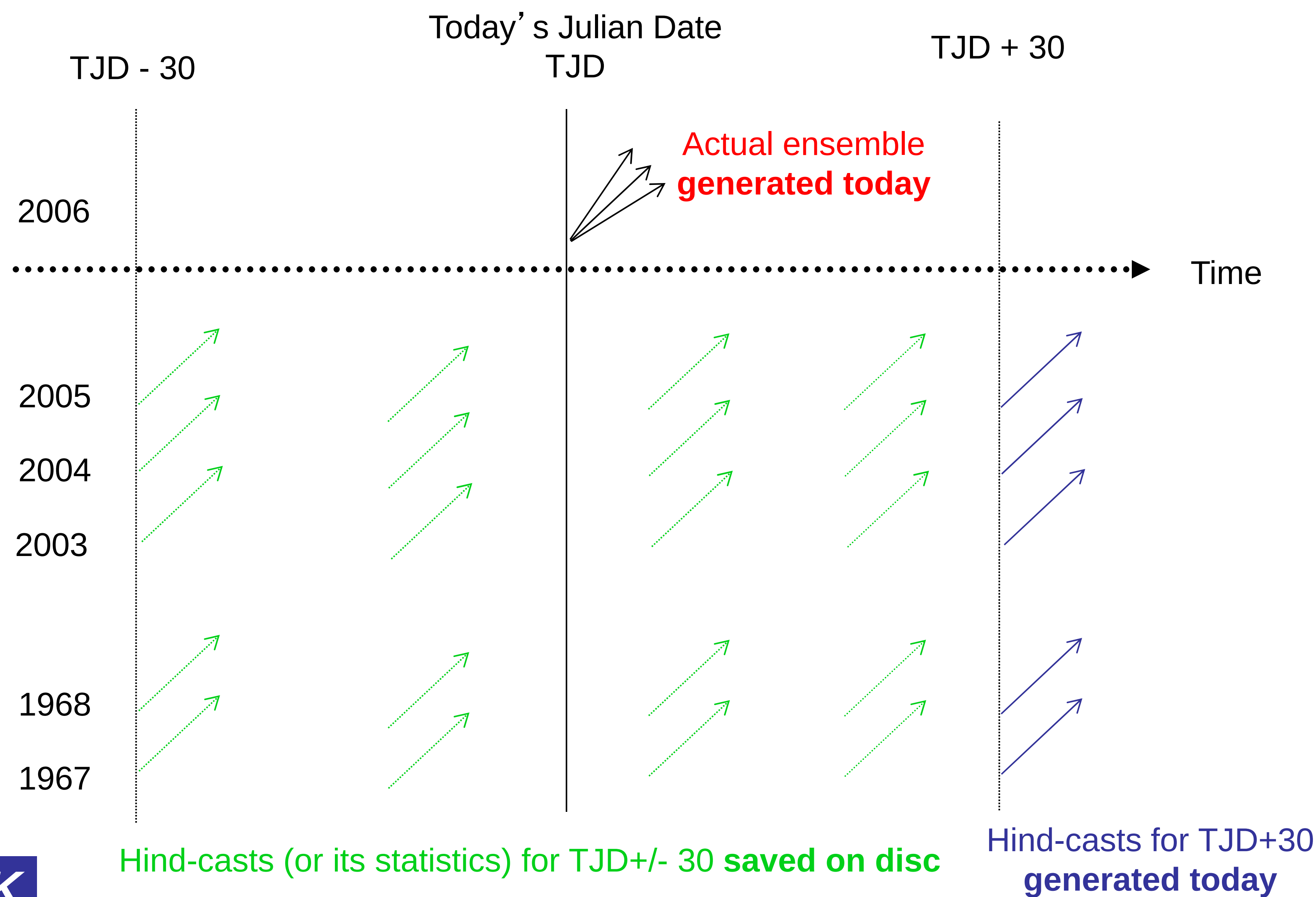
- **Leverage** significant **NSF & UV investments**
  - 10-year research/development effort by R. Krzysztofowicz & his team
  - $1M+ effort
- **Successful collaboration** between
  - Academia & Application – Operations
  - Roman Krzysztofowicz, theory Z.T., forecasting
  - Theoretical soundness Practically solutions
- All components developed & **thoroughly tested** in research environment
  - Solid results
  - 30+ related peer-reviewed papers
  - 3 Research Theses
- **Unique** technology
  - Bayesian Model Averaging (BMA) does NOT use climatological info
  - "B" refers to Bayesian estimation of statistical parameters
- **Next steps**
  - ***Assemble components and test in operational environment***

OUTLINE / SUMMARY

- User needs
  - Various weather parameters at different times / locations
- Forecast formats
  - Probabilities, joint probabilities – limited choices
  - Ensemble scenarios ultimate “carrier” of full information
- Post-processing framework
  - Pre-processing, fusion, output
- Bayesian Processor of Ensembles (BPE)
  - Theoretically sound, practical/flexible, high performance
- Great R2O opportunity
  - NOAA can leverage Leverage Bias correction, merging, downscaling, derivation of variables
  - Bayesian methods

BACKGROUND

REAL-TIME GENERATION OF HIND-CAST DATASET



BACK

POST-PROCERSSING DESIGN PRINCIPLES

- **Quality – utility**
  - Highest quality info required by users
- **Computational efficiency**
  - Affordable processing of large amounts of data
- **Scientific soundness**
  - Evaluate/test alternatives as statistical model is built
- **Parsimony**
  - Use simplest method that can solve problem
- **Modular design**
  - Stand-alone components for distinct functionalities
    - Aid collaborative development

BPE — Outputs

Input: Model ensemble (for predictand, grid point, lead time)

Output: (1) Posterior density function

(2) Posterior distribution function

(3) Posterior quantile function

(4) Posterior ensemble (calibrated ensemble)

Each member is mapped into a posterior quantile
via the inverse of the posterior distribution function

Usage

(1) *Solve* a decision problem analytically

(2) Find *probability* of non-exceedance of given ***threshold***

(3) Find *quantile* corresponding to given non-exceedance ***probability***

(4) *Simulate* operation of a dynamic system

HOW TO GENERATE HINDCASTS?

- Generate all hindcasts prior to implementation
  - Cons
    - Great one-time human effort
    - Frozen forecast system
      - Less frequent upgrades?
    - Possibly less cpu used if system upgraded less than once a year
- Real time generation
  - 30 days ahead of season
  - Pros
    - Flexibility
      - Upgrades at any time, subject to 2-3 mos parallel testing
        Statistics from parallel moved to operations at impl.
    - Run as part of operations
    - Can feed back to continuous developments/upgrades

WHAT USERS NEED?

- **Sensitivity to weather** - each user affected by
  - *Various weather elements*
  - *At different times*
  - *At different locations*

- What users need for **optimal decision making**
  - Probability distributions for single variables?
    - Incomplete - lack of information on cross-correlations
  - Covariances needed across
    - Variables, space, and time

- What **forecast format** may work?
  - Joint probability distributions?
    - Provision of all joint distributions possibly needed by users - Intractable
  - Encapsulate best forecast info into **calibrated ensemble members**
    - Possible *weather scenarios*

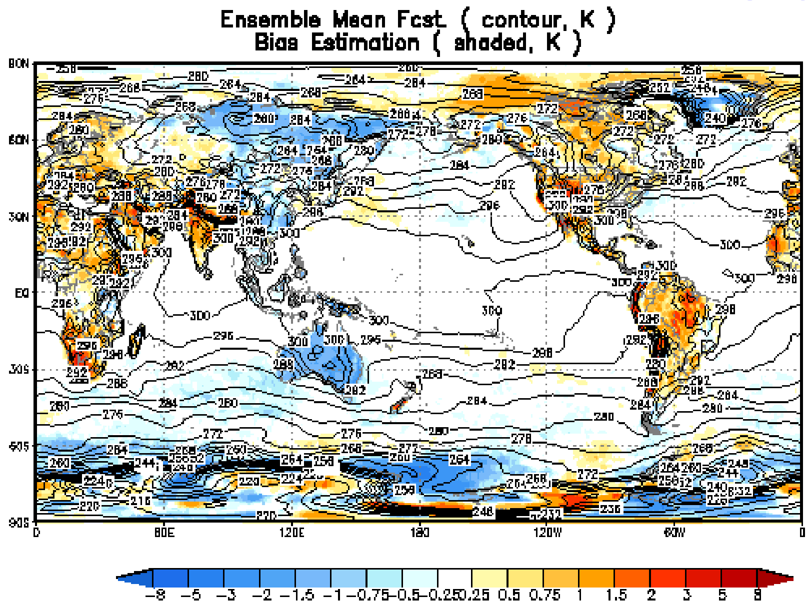
BAYESIAN PRE-PROCESSOR (BPP)

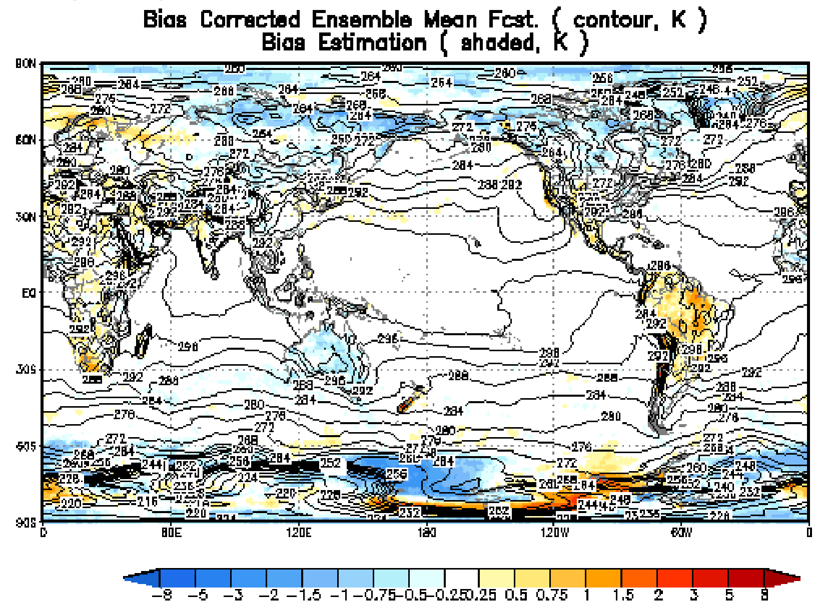
- Problem
  - Forecasts potentially generated with different models
    - Lead-time dependent bias different in each member
- Functionality
  - Unify systematic behavior of all members
- Method
  - Estimate and remove unconditional systematic error
    - Bayesian Pre-Processor on model grid
- Current status
  - Recursively estimated time mean removed
    - North American Ensemble Forecast System operations

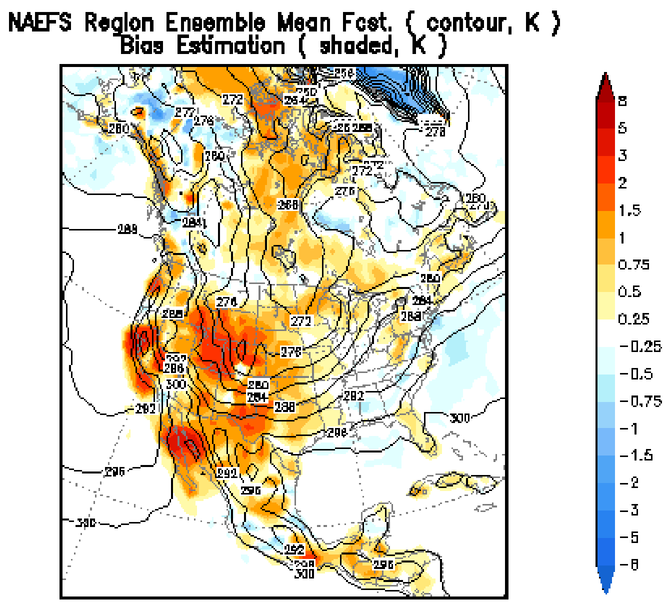
ESTIMATED BIAS – 2m Temperature, 5-d forecast

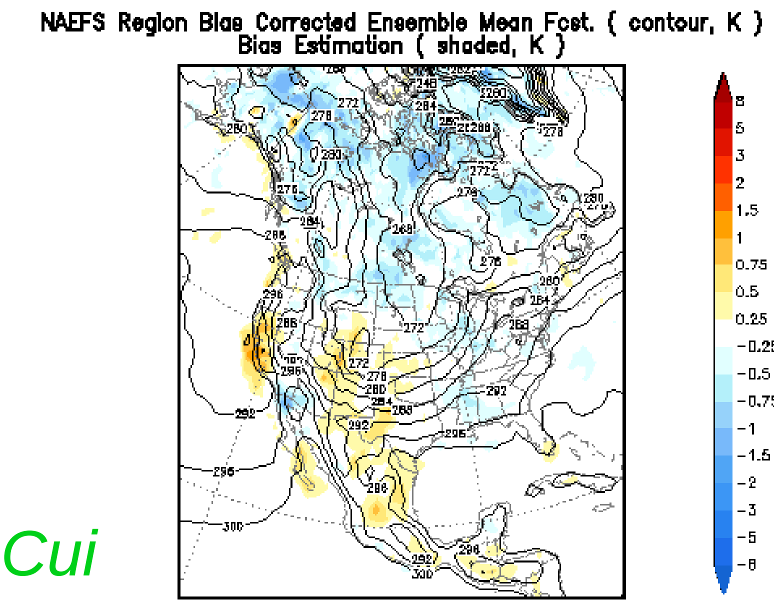
BEFORE AFTER

BIAS CORRECTION





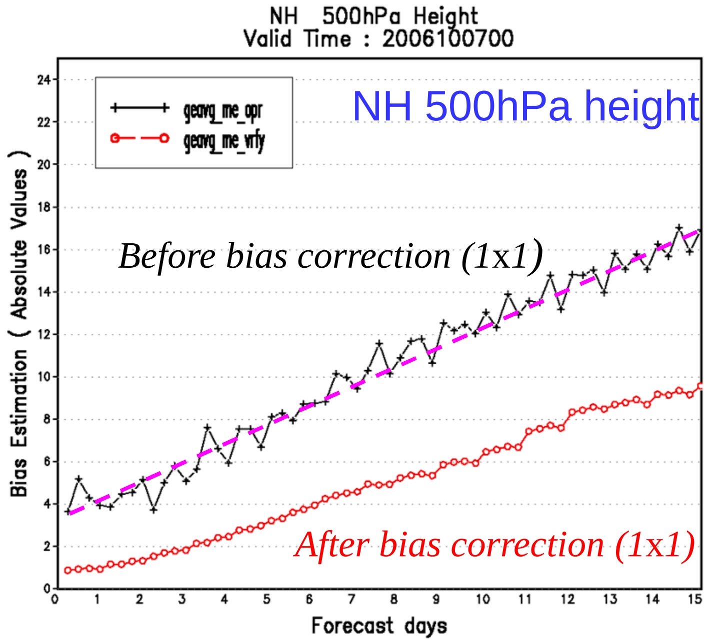


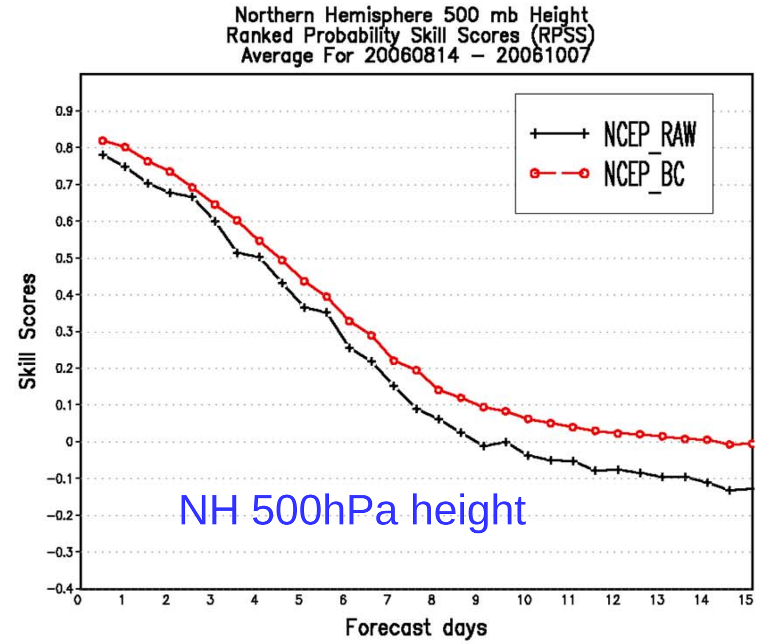


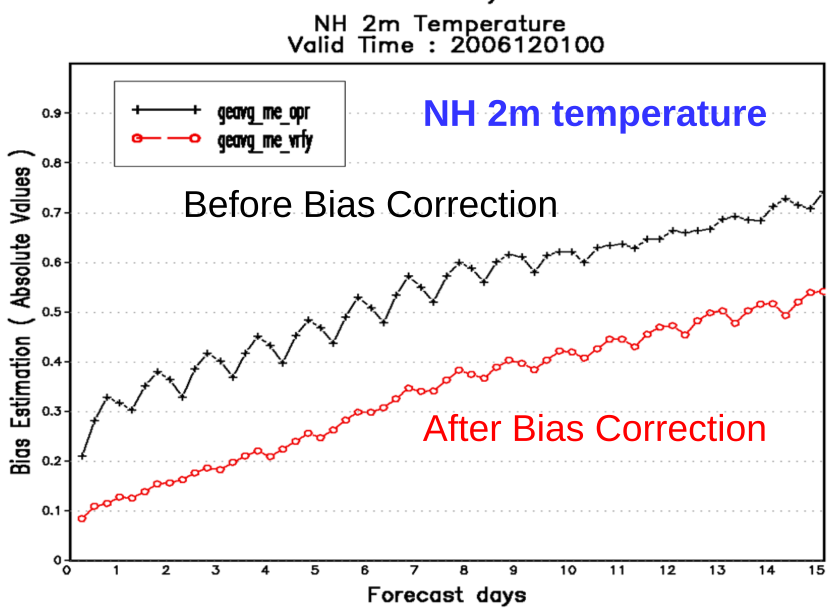
Bo Cui

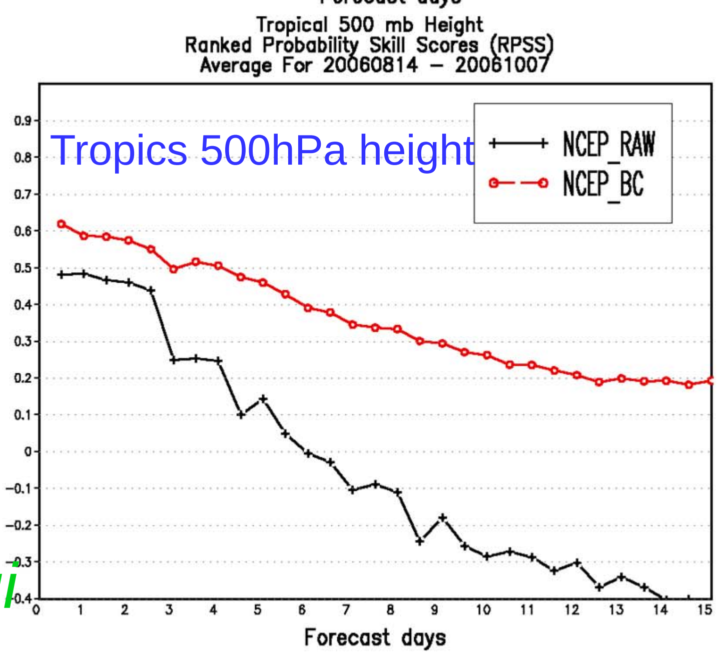
IMPACT OF BIAS CORRECTION ON

ESTIMATED SYSTEAMTIC ERROR   PROBABILISTIC SCORES









Bo Cui

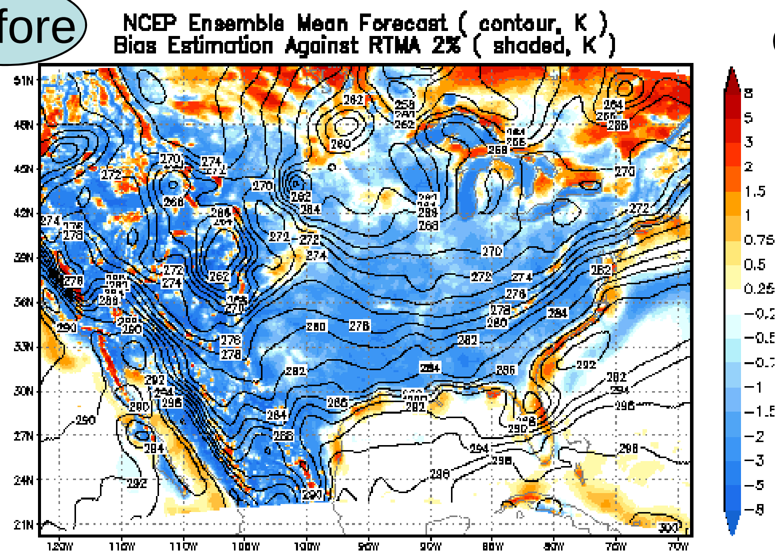
DOWNSCALING

- Problems
  - NWP forecasts do not produce all required variables
    - Missing variables (fog, wind shear, etc)
    - Spatial / temporal resolution too coarse
- Functionality
  - Derive user variables statistically etc consistent with calibrated NWP ensemble members
- Methods
  - Statistically downscale ensemble members
    - Ensure long-term mean of downscaled forecasts matches climatology
    - Ensure spatial / temporal variability on fine scales match observed variability
  - Derive additional variables using statistical & other relationships
    - CIN, CAPE, etc
- Current status
  - Recursively estimate time mean difference between
    - NWP analysis interpolated onto fine grid & obs-based mesoscale analysis
  - Add estimate of downscaling vector to each ensemble member
    - North American Ensemble Forecast System operations

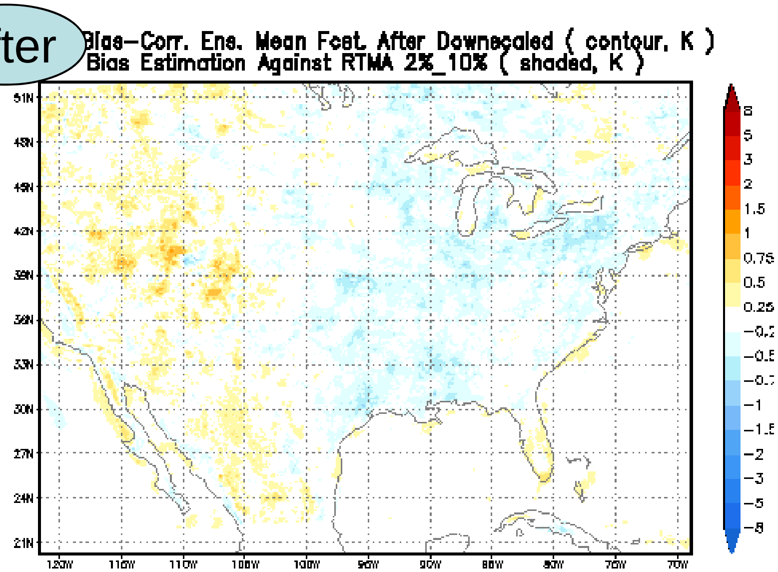
00hr GEFS Ensemble Mean & Bias Before/After Downscaling 10%

2m Temperature

Before



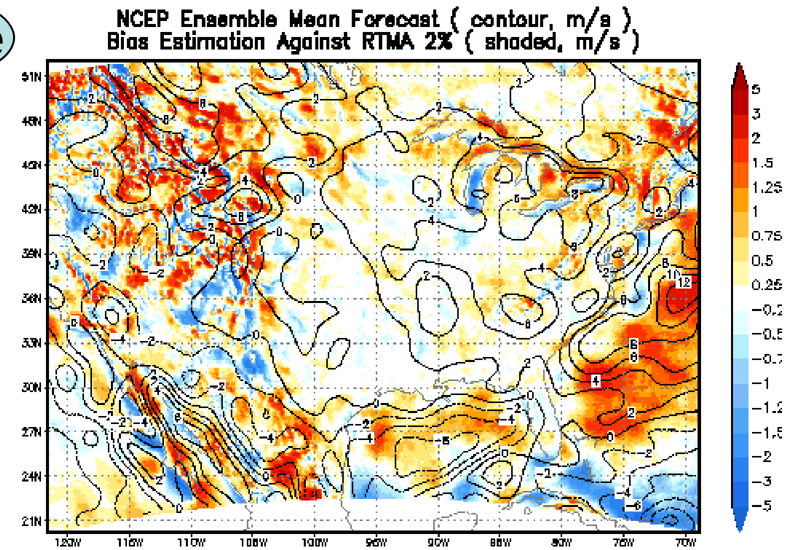
After



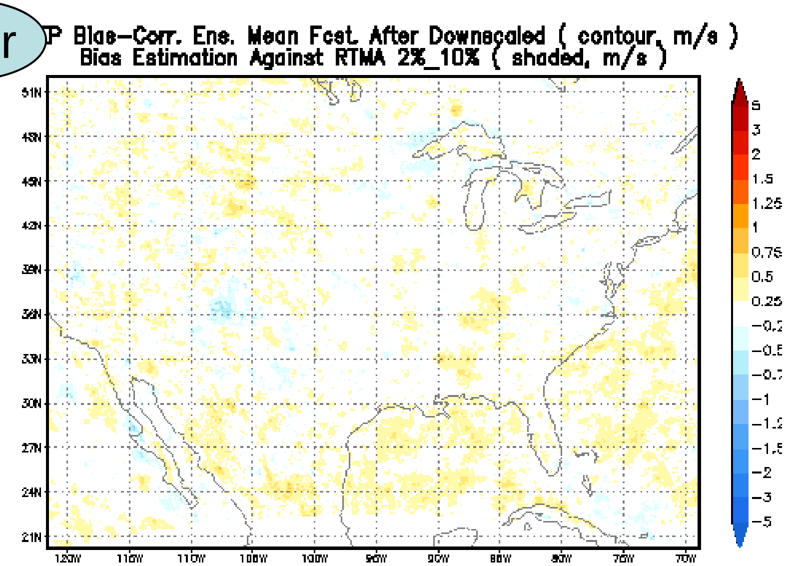
BO CUI, GCWMB/EMC/NCEP/NOAA

10m U Wind

Before

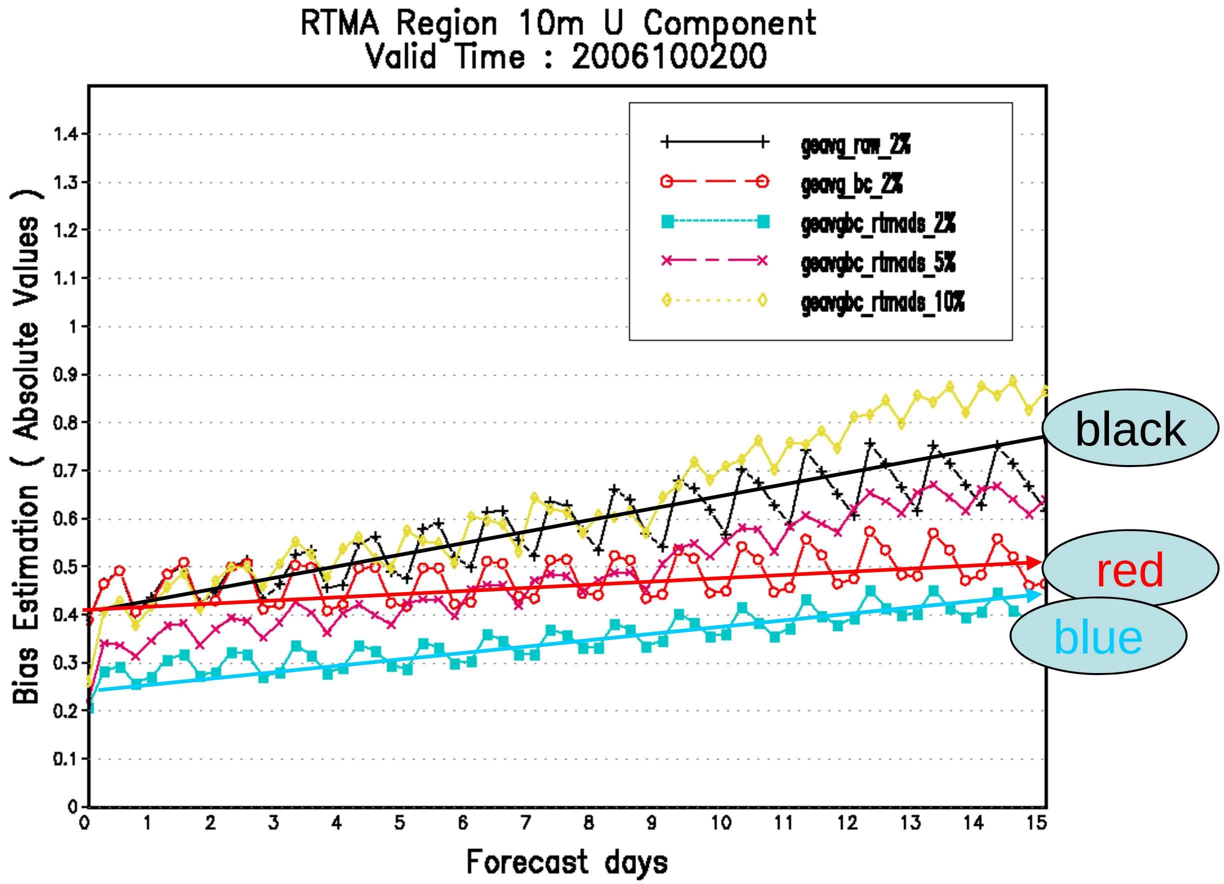


After



BO CUI, GCWMB/EMC/NCEP/NOAA

Accumulated Bias Before/After RTMA Downscaling



Black- operational ensemble mean, 2%

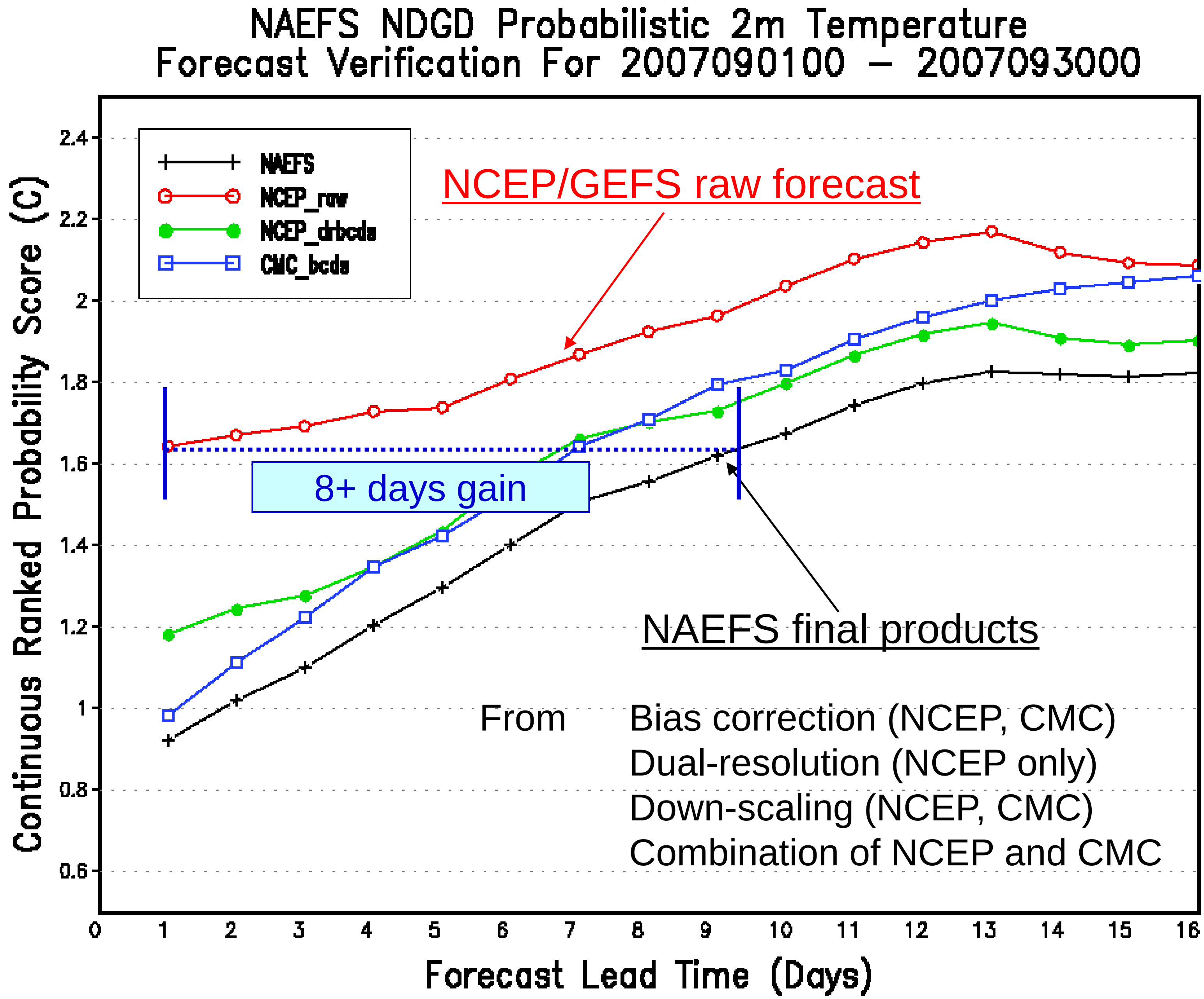
Pink- bias corrected ens. mean after downscaling, 5%

Red- NAEFS bias corrected ensemble mean, 2%

Blue-bias corrected ens. mean after downscaling, 2%

Yellow-bias corrected ens. mean after downscaling, 10%

CONTINUOUS RANKED PROBABILITY SCORE
RAW / BIAS CORR. & DOWNSCALED & HIRES MERGED / NAEFS



High resolution control & Canadian ensemble adds significant value

=>

8-day total gain in skill

BACKGROUND