Northern American Ensemble Forecast System (NAEFS)-Bias Correction

Yuejian Zhu,
Bo Cui and Zoltan Toth

Environmental Modeling Center
NOAA/NWS/NCEP

Acknowledgements:
DingChen Hou EMC

NAEFS Background Information

- First of a kind project
  - Operational multi-center ensemble system
  - Bias correction, climate percentiles never computed on such a scale operationally
- Timetable
  - Mar 2003 Project started
  - Oct 2003 Draft Research, Development and Implementation Plan
  - Sep 2004 Initial Operational Capability – Operational data exchange
  - May 2006 First Operational Implementation
  - Mar 2007 NAEFS upgrade
- Challenges
  - Developed joint plan with MSC personnel
  - Arranged operational data exchange
  - Coordinated GEFS development with international NAEFS developments
  - Coordinated software development & operational implementation with MSC
  - Worked with less THORPEX resources than planned originally
- Future expansion
  - Develop sustainable plans
    - Coordinate with partners
    - Rename NAEFS and position it as prototype GIFS system
  - Resource concerns
    - Computational (telecommunication, disc, etc)

First Implementation of NAEFS – Summary

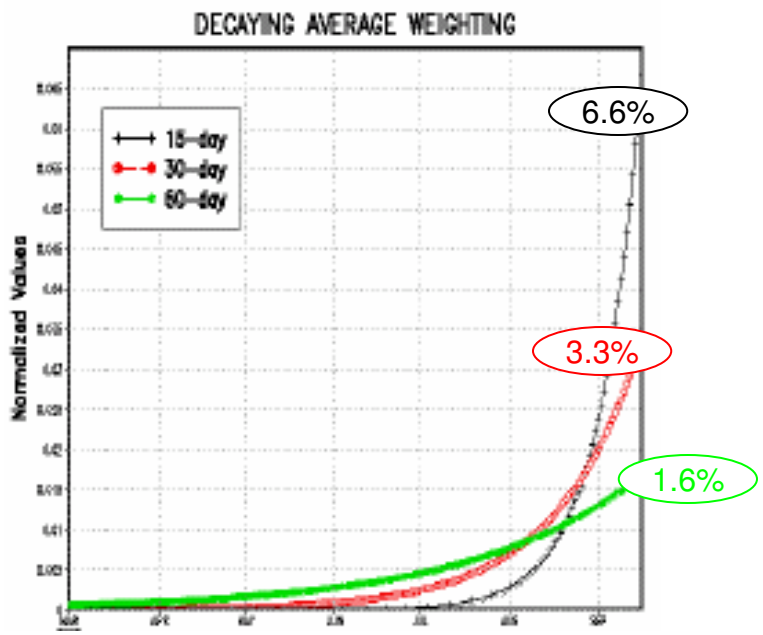
1. Bias corrected members of joint MSC-NCEP ensemble
   - Decaying accumulated bias (~past 50 days) for each var. for each grid point
   - For selected 35 of 50 NAEFS variables
   - 32(00Z), 15(06Z), 32(12Z) and 15(18Z) joint ensemble members
   - Bias correction against each center's own operational analysis
2. Weights for each member for creating joint ensemble
   (equal weights now – unequal weights to be added later)
   - Weights don't depend on the variables
   - Weights depend on geographical location (low precision packing)
   - Weights depend on the lead time
3. Climate anomaly percentiles for each member
   - Based on NCEP/NCAR 40-year reanalysis
     - Used first 4 Fourier modes for daily mean,
     - Estimated climate pdf distribution (standard deviation) from daily mean
   - For selected 19 of 50 NAEFS variables
   - 32(00Z), 15(06Z), 32(12Z) and 15(18Z) joint ensemble members
   - Adjustment made to account for difference between oper. & re-analysis
   - Provides basis for downscaling if local climatology available
     – Non-dimensional unit

Bias Correction Method & Application

- Bias Assessment: adaptive (Kalman Filter type) algorithm

decaying averaging mean error = (1-*w*) * prior t.m.e + *w* * (f – a)

For separated cycles, each lead time and individual grid point, t.m.e = time mean error



Toth, Z., and Y. Zhu, 2001

- Test different decaying weights. 0.25%, 0.5%, 1%, 2%, 5% and 10%, respectively
- Decide to use 2% (~ 50 days) decaying accumulation bias estimation

- Bias Correction: application to NCEP operational ensemble 15 members

List of Variables for Bias Correction, Weights and Forecast Anomalies for CMC & NCEP Ensemble

	CMC & NCEP
Ensemble	CMC (8 SEF, 8 GEM), NCEP (14 GFS)
GRID	1x1 deg (360x180 lat-lon)
DOMAIN	Global
FORMAT	WMO Grib Format
HOURS	6 hourly out of 384 hours (current 240 hours for CMC Ensemble)
GZ	200, 250, 500,700, 850 ,925,1000
TT	200, 250, 500,700, 850 ,925,1000
U,V	200, 250, 500,700, 850 ,925,1000
TT	2m
U, V	10m
MSLP	Sea Level Pressure
Sfc Pres	Surface Pressure
Tmax	2m
Tmin	2m

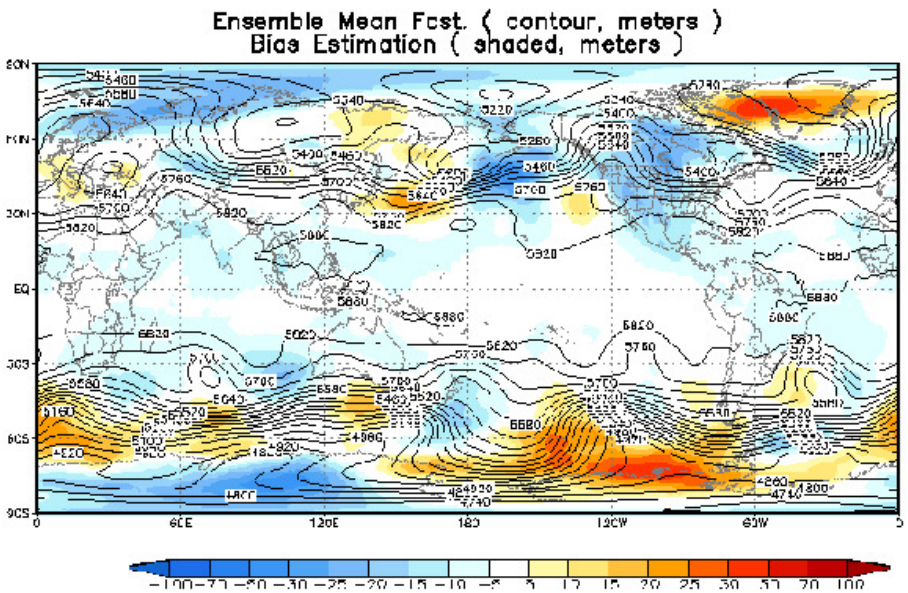
Note: 35 Variables in total, red variables are for climate anomalies only

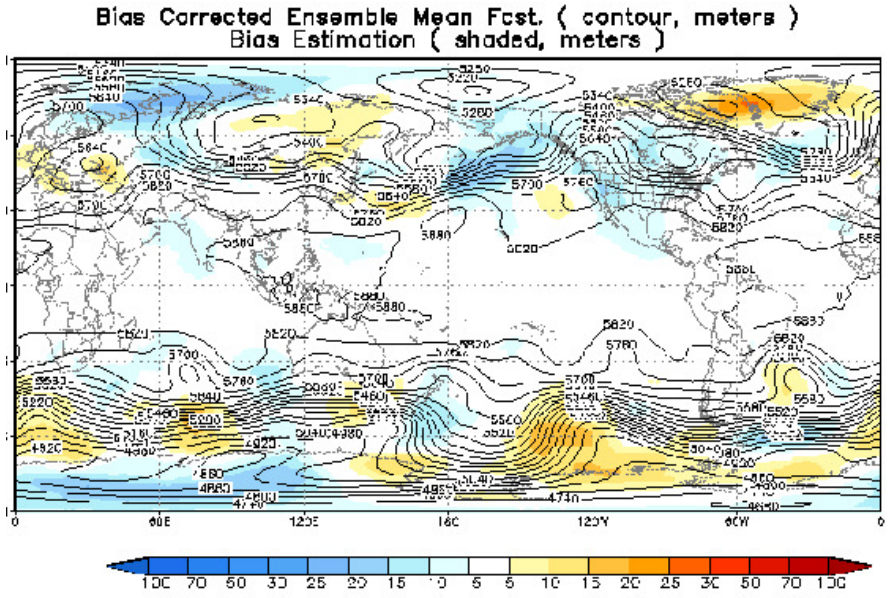
Summary of NAEFS First Implementation

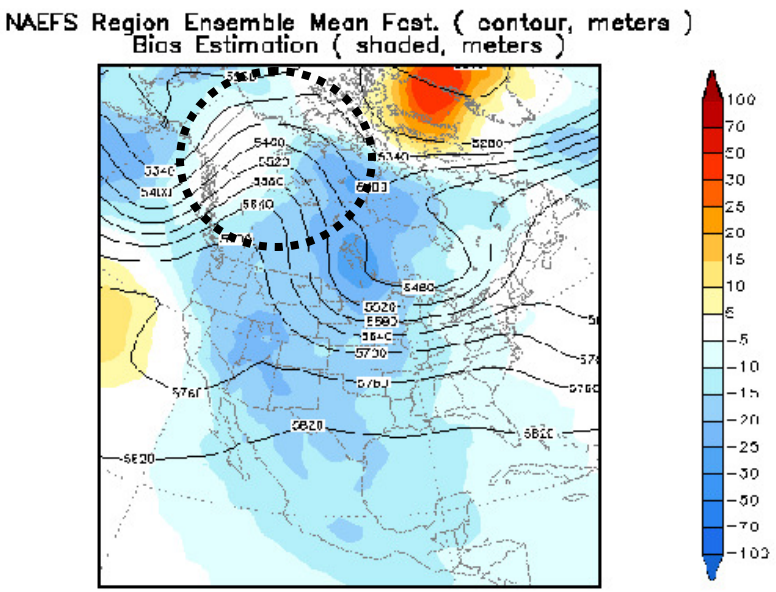
- Period:
  - 04/10/2006 – Current (NCO real time parallel)
- Maps comparison for bias (before and after)
  - 500hPa height, 2m temperature
- Statistics for
  - Bias reduction in percentage
    - Height, temperature, winds
  - RMS errors
  - Probabilistic verifications (ROC)
    - NH, SH and tropic
- Conclusions
  - Bias reduced (approximately 50% at early lead time)
  - RMS errors improved by 9% for d0-d3
  - Probabilistic forecast
    - Improved for all area, all lead time
    - Typically for NH, 20-24 hours improvement from d7

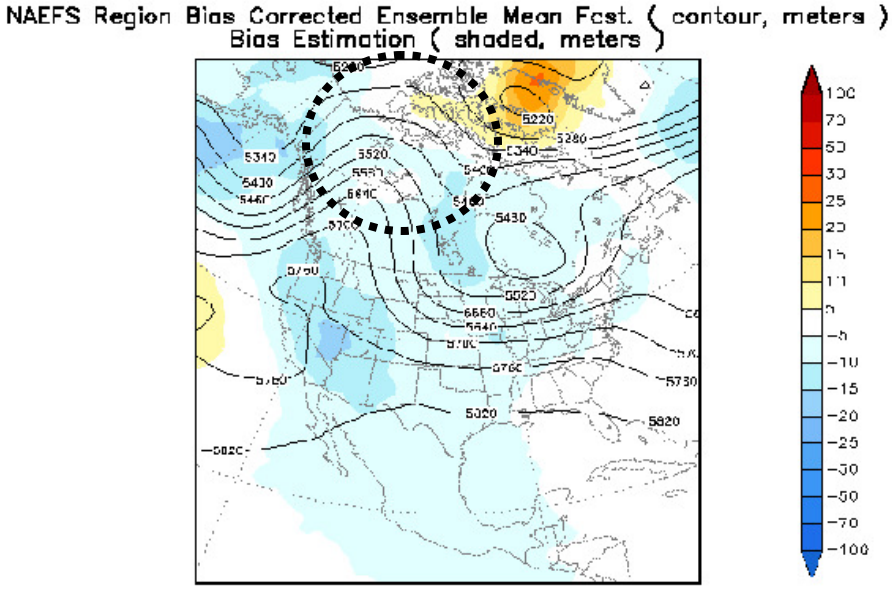
500hPa height: 120 hours forecast (ini: 2006043000)

Shaded: left – raw bias right – bias after correction



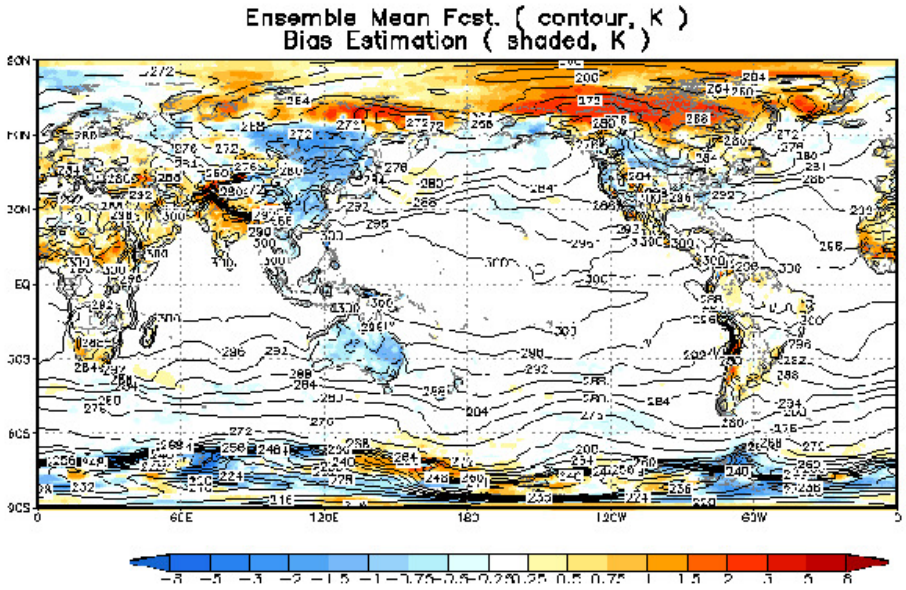


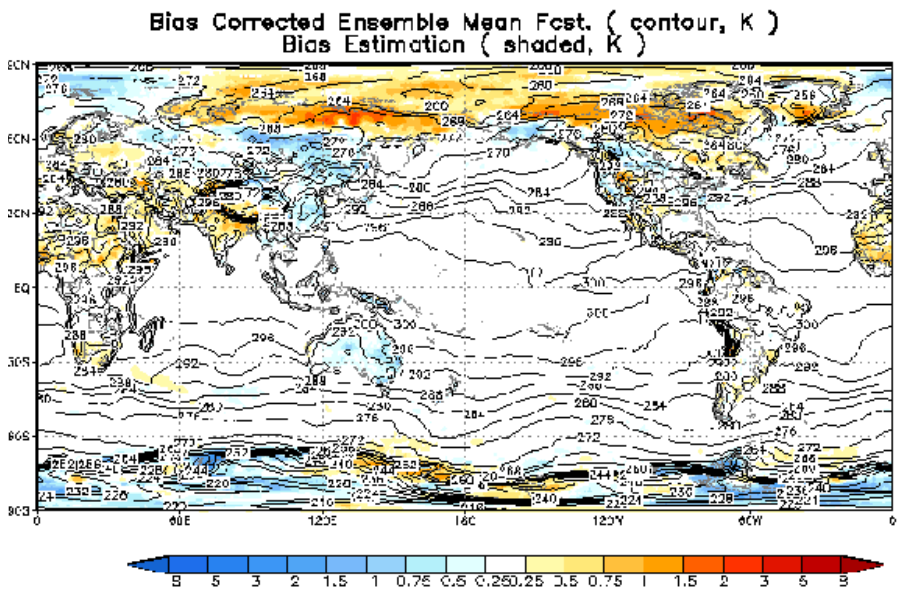


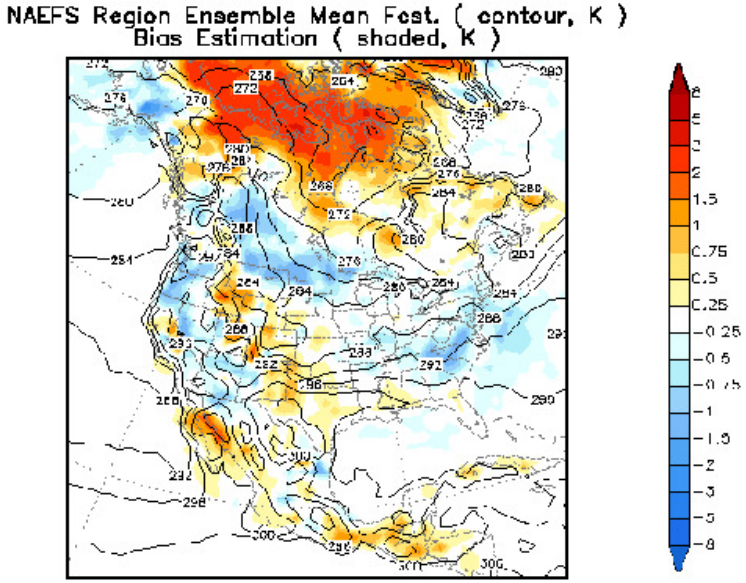


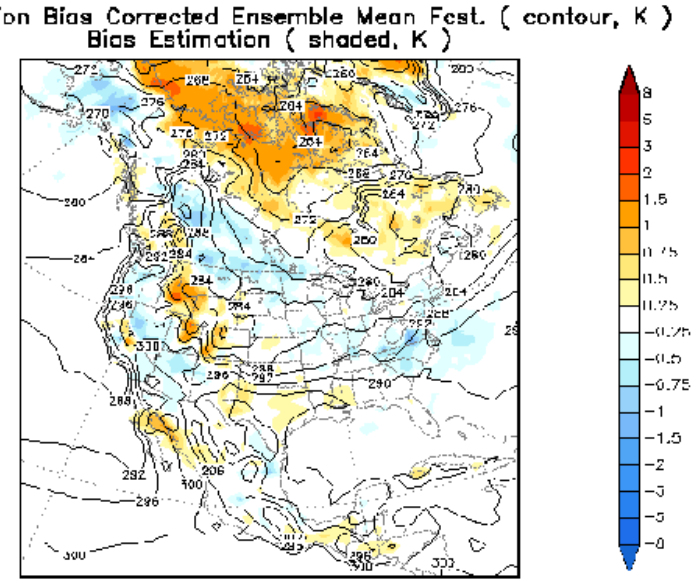
2 meter temperature: 120 hours forecast (ini: 2006043000)

Shaded: left – raw bias right – bias after correction

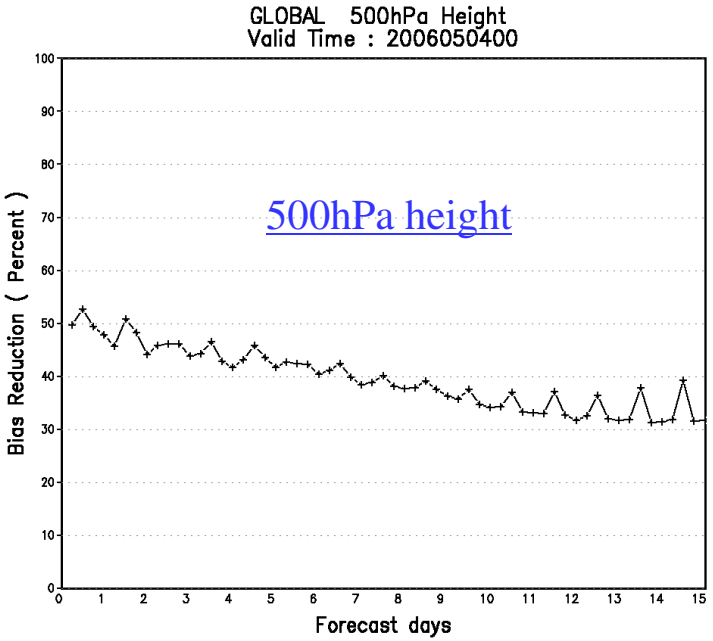








Bias Improvement (absolute value) after Bias correction



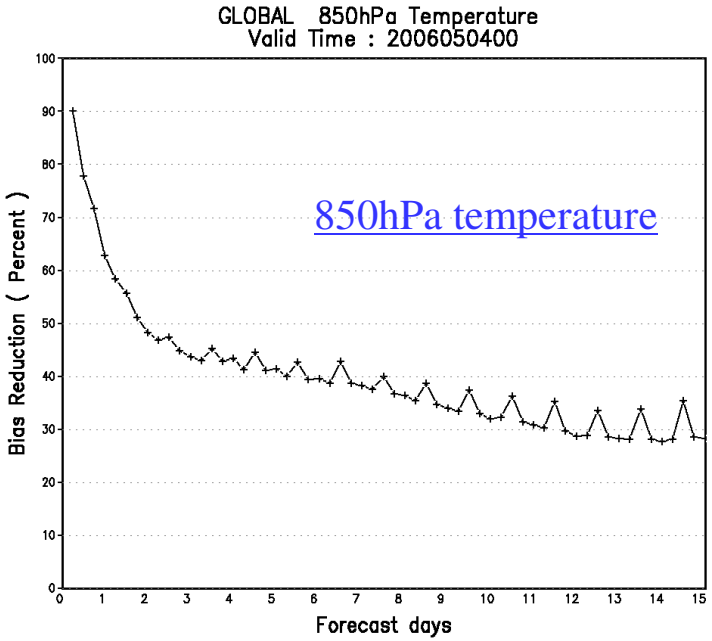
Overall bias reduction:

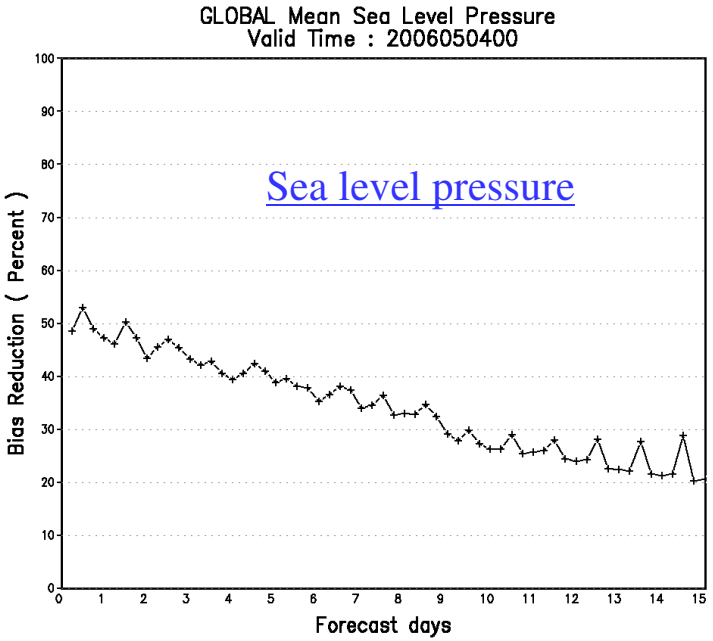
(globally)

D0-3: 50%

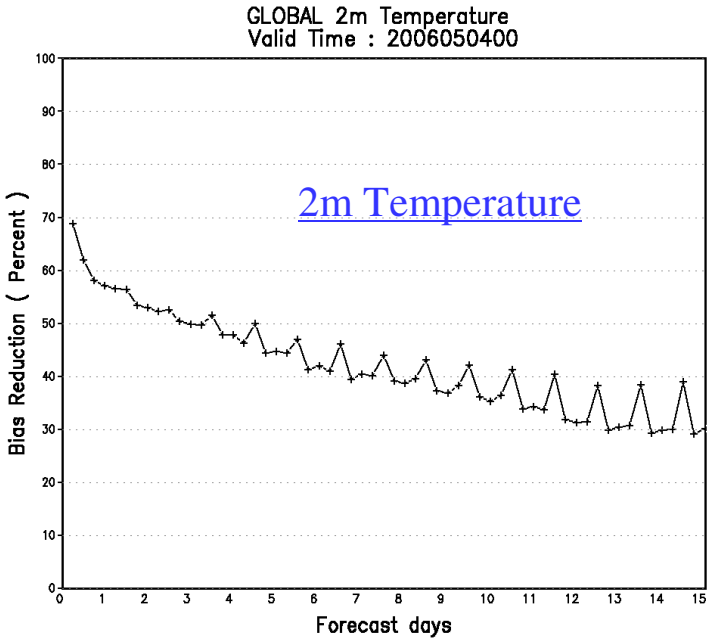
D3-8: 40%

D8-15: 30%

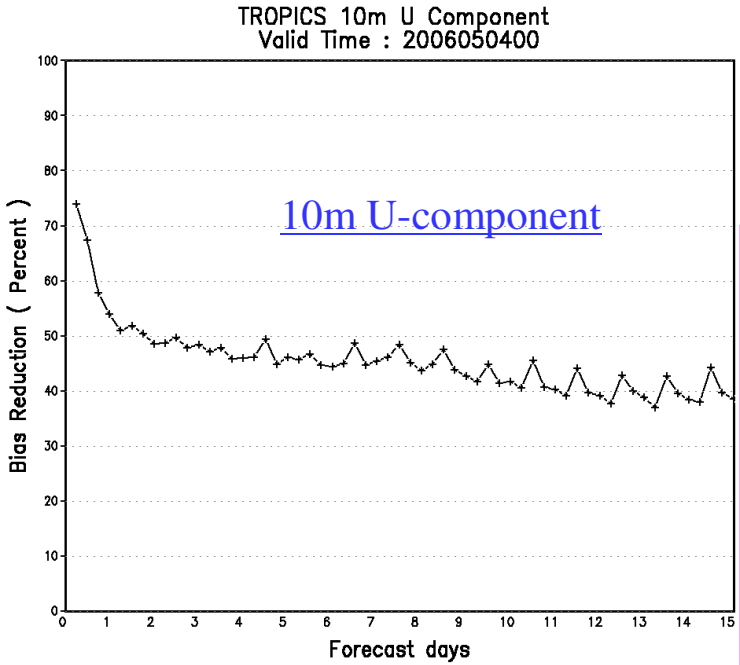




There is daily variation after bias correction, more bias reduced for valid 12Z cycle



Bias Improvement (absolute value) after Bias correction



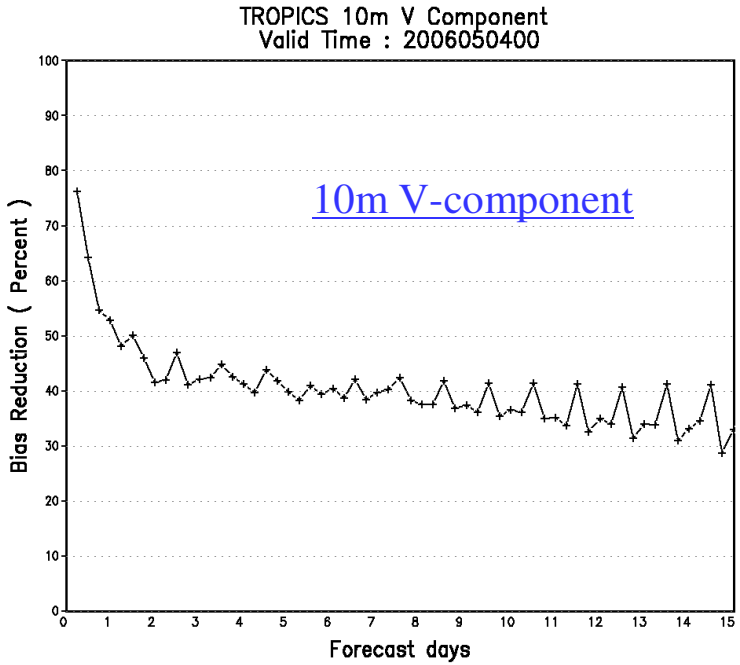
Overall bias reduction:

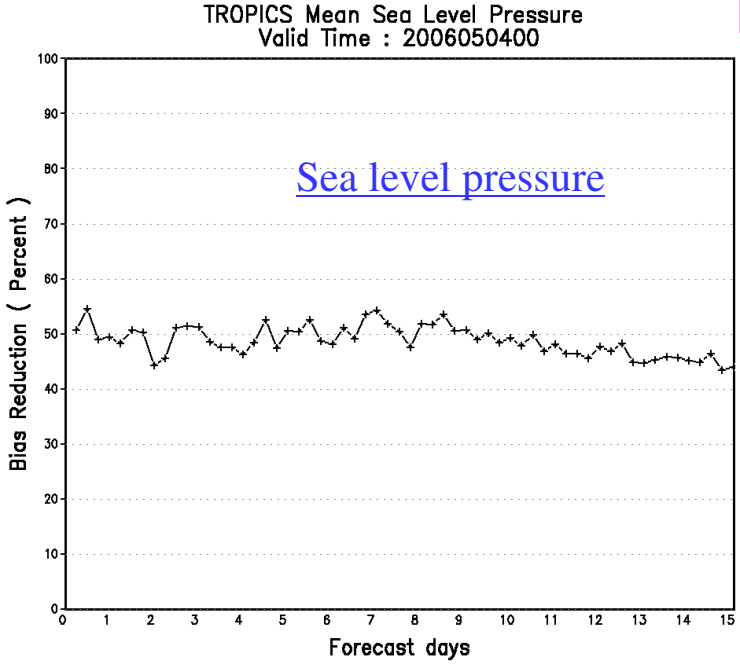
(Tropic)

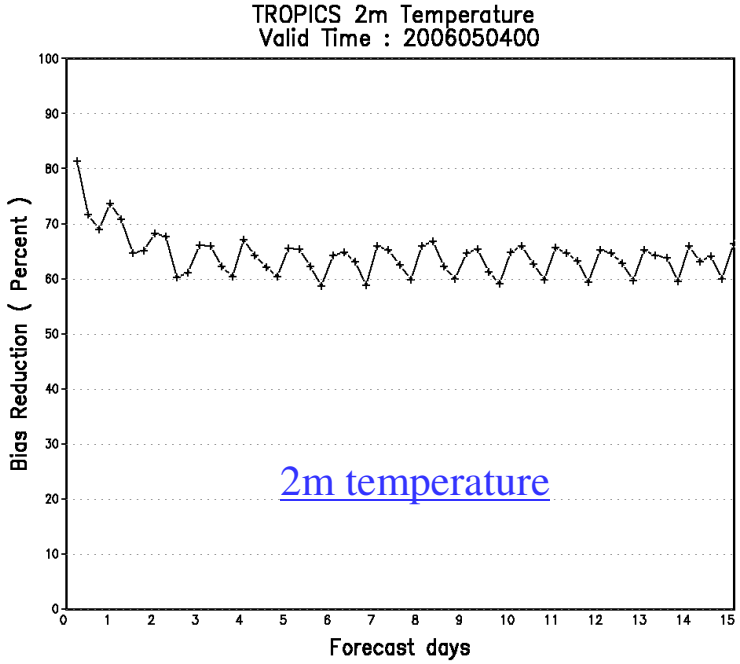
D0-3: 50%

D3-8: 45%

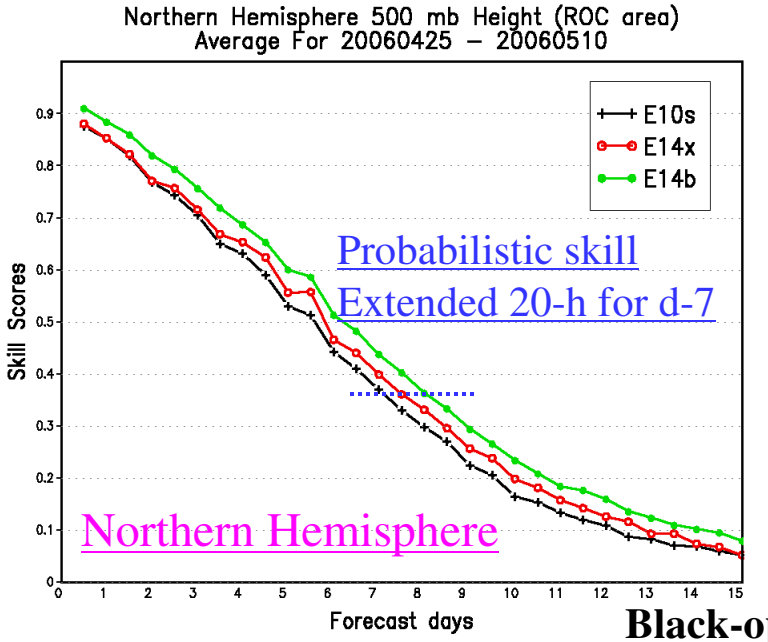
D8-15: 40%

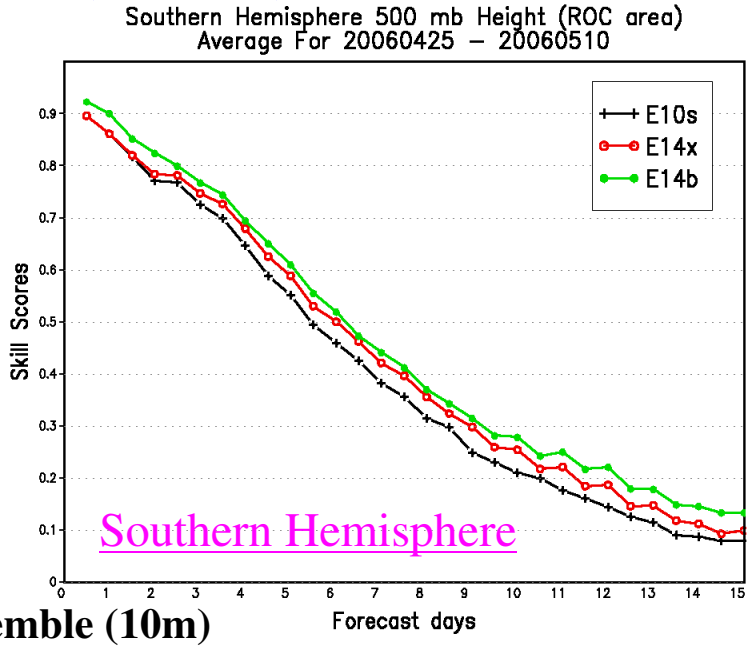




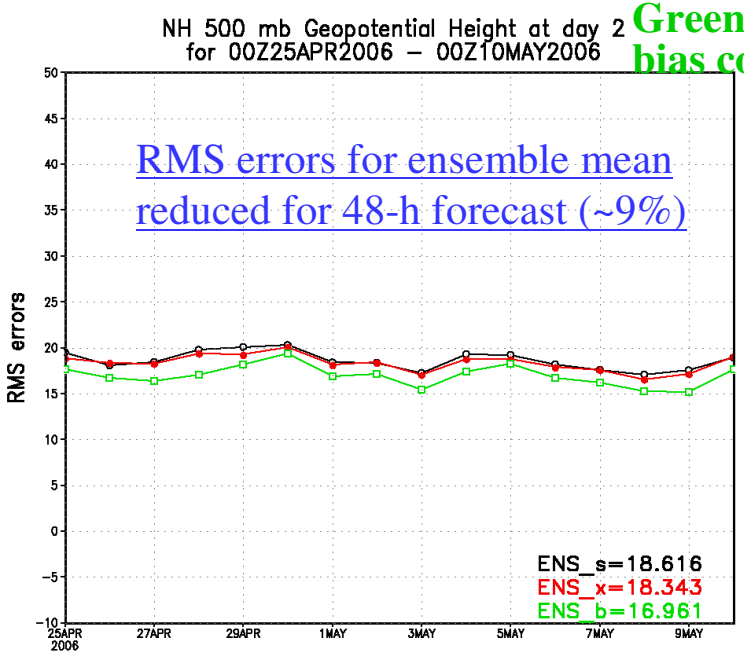


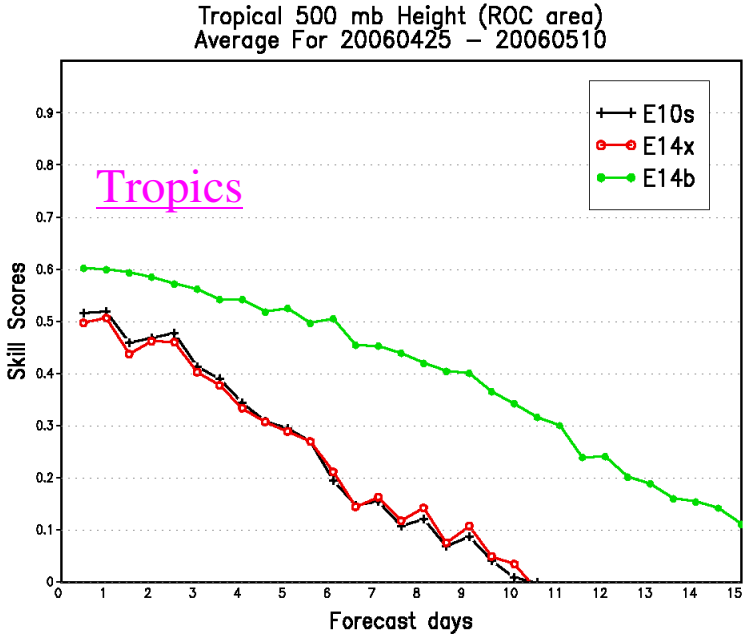
Evaluation after bias correction (16 cases)





Black-operational ensemble (10m)
Red-real time parallel ensemble (14m)
Green-real time parallel ensemble after bias correction (14m)





NAEFS verification

- Web-site:
  - http://wwwt.emc.ncep.noaa.gov/gmb/yzhu/html/opr/naefs.html
- Reference: NCEP/NCAR 40y reanalysis (next slide)
- Variables:
  - 1000hPa, 500hPa heights, 850hPa, 2m temperature, 10m u and v
- Verified for ensemble mean:
  - RMS errors, spread, mean error (bias) and absolute error
- Verified for ensemble distribution:
  - Histogram (Talagrand)
- Verified for ensemble probabilistic forecast
  - ROC, RPSS, CRPS, BSS (Resolution and Reliability), EV
- Regions:
  - NH, SH, Tropical, Asia, Europe and Northern American
- Statistics from seasonal average

Climatological Data

- NCEP/NCAR 40 years (1958-1997) reanalysis
- Monthly Sampling
  - For example: 40*30=1200
- Generating10 equally-a-likely, based on monthly sampling
- Projected to verify date
- All forecast skills will base on 10 equally-a-likely climatological bins.

Example of web-page setting:
http://wwwt.emc.ncep.noaa.gov/gmb/yzhu/html/opr/naefs.html

Global Ensemble Model Evaluation: (NCEP against NCEPb)

500 hPa Height Scores NCEP .vs NCEPb									
NH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
SH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
TROP	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM

1000 hPa Height Scores (NCEP .vs NCEPb)									
NH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
SH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
TROP	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM

850 hPa Temperature Scores (NCEP .vs NCEPb)									
NH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
SH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
TROP	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM

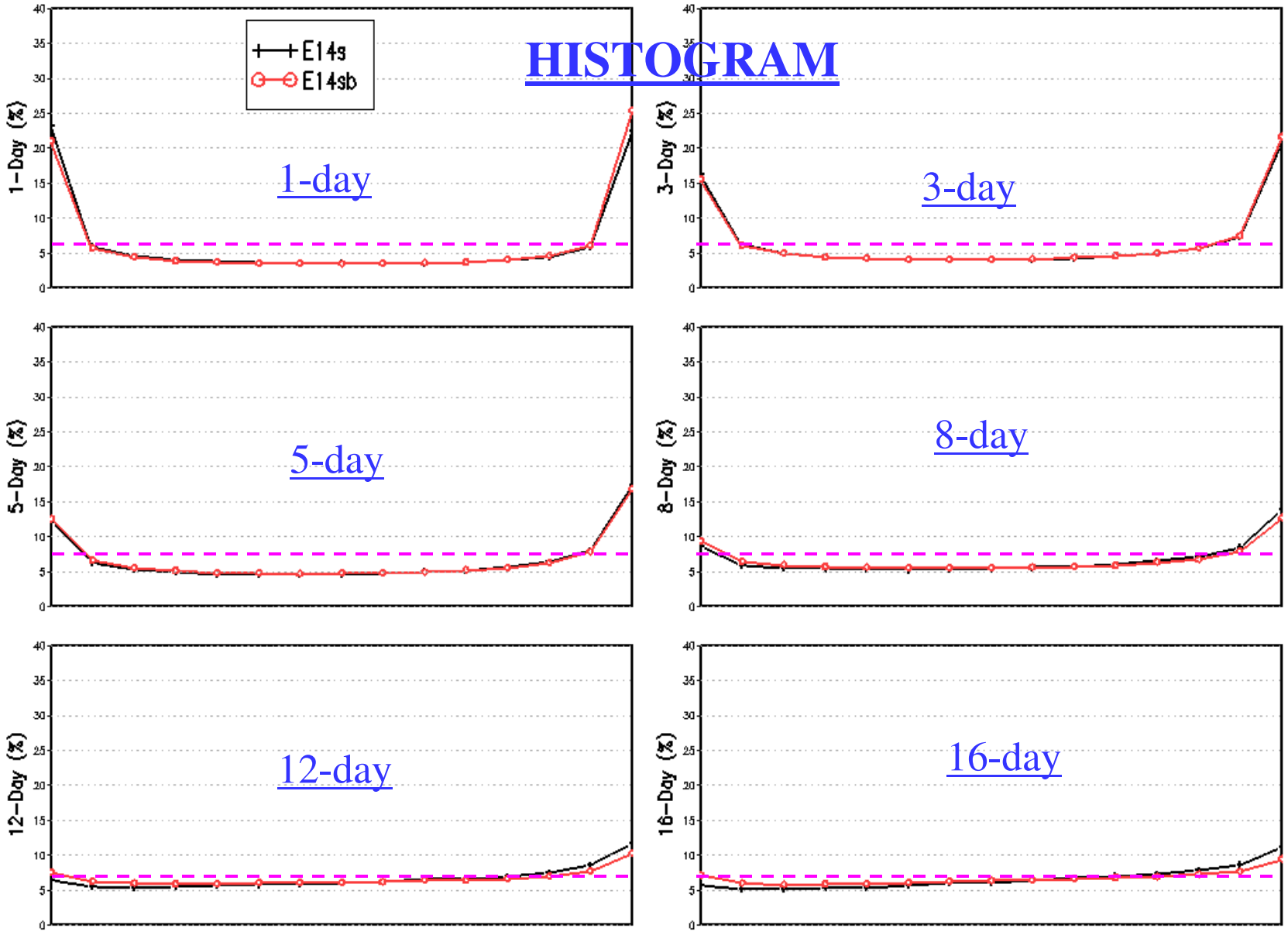
2 Meters Temperature Scores (NCEP .vs NCEPb)									
NH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
SH	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM
TROP	ROC	EV	RPSS	BSS	CRP	CRPS	RMS/SPRD	ERR/ABSE	HISTOGRAM

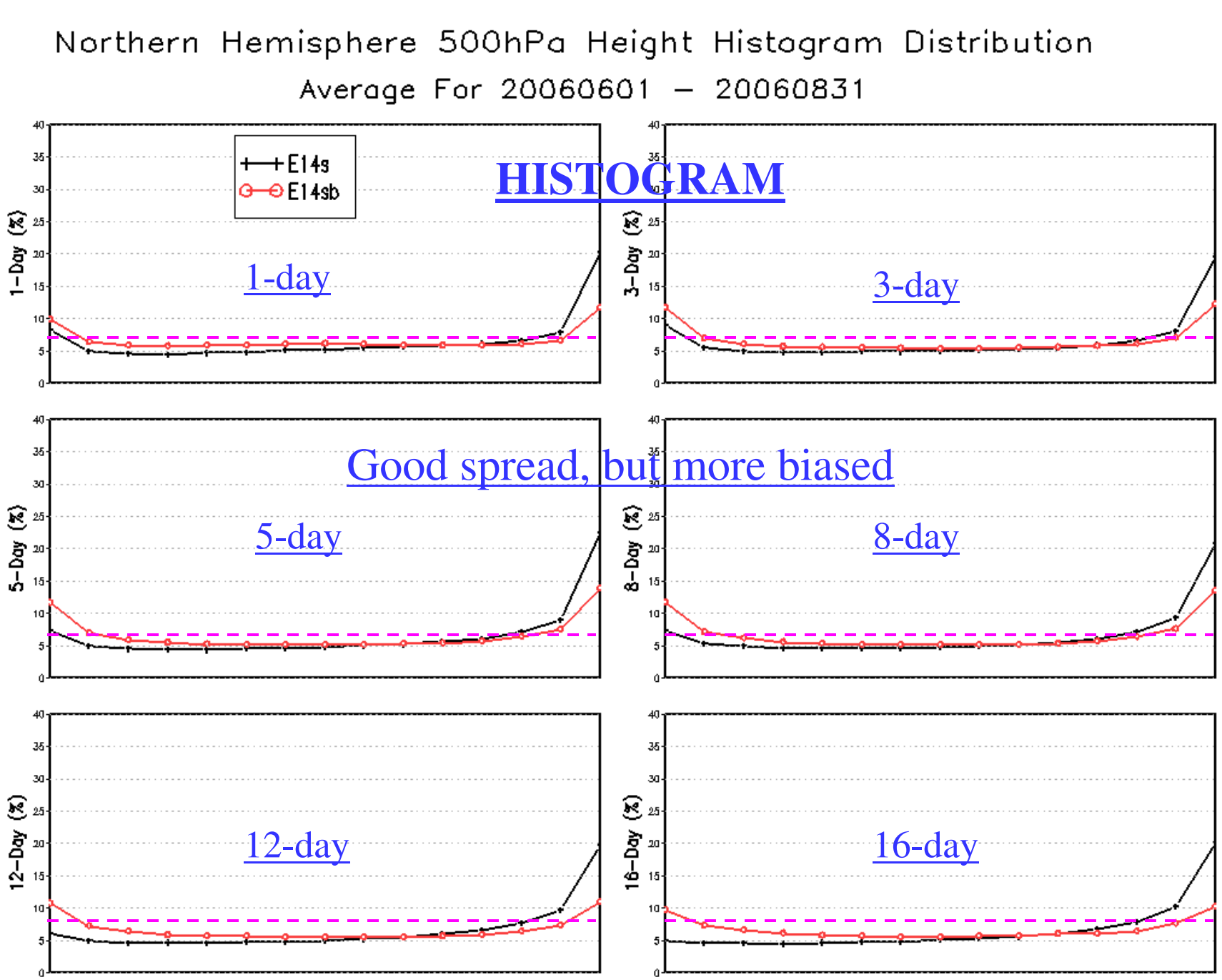
ISSUES ADDRESSED

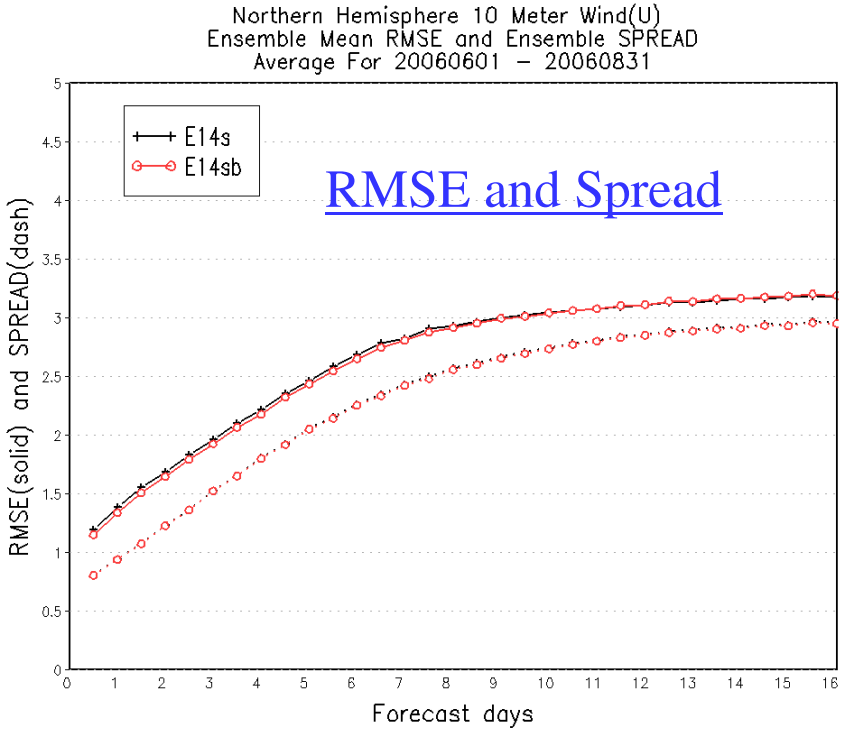
- Effect of bias-correction
  - Different variables
- Comparing of NCEP and CMC's forecasts
  - Before & after bias correction
- Impact of combined ensemble (NAEFS)
  - Before & after bias correction
  - Gains from bias correction + combination =
    - NAEFS advantage

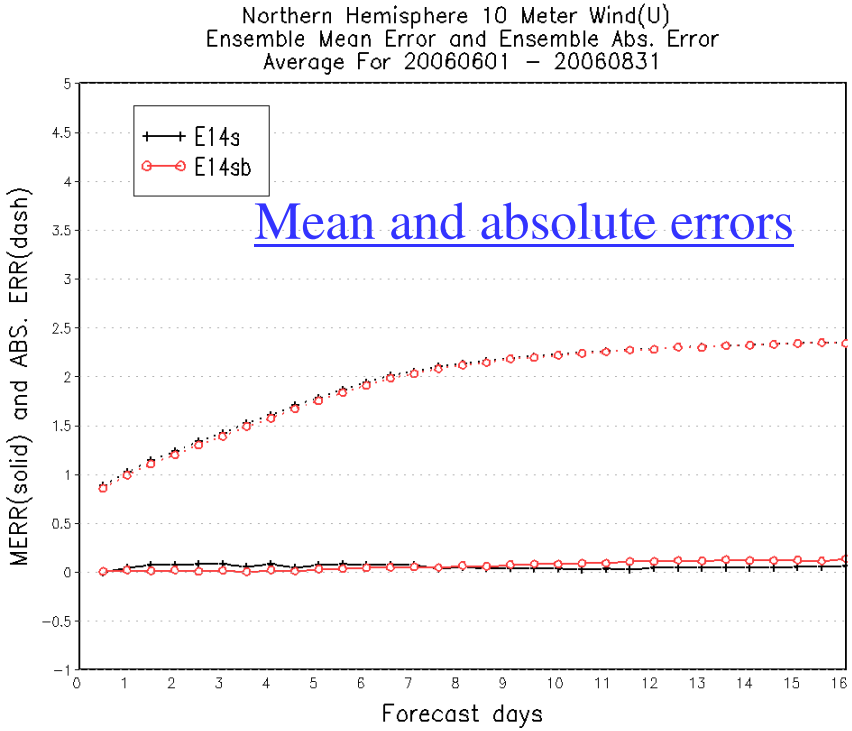
Northern Hemisphere 2 Meter Temp. Histogram Distribution

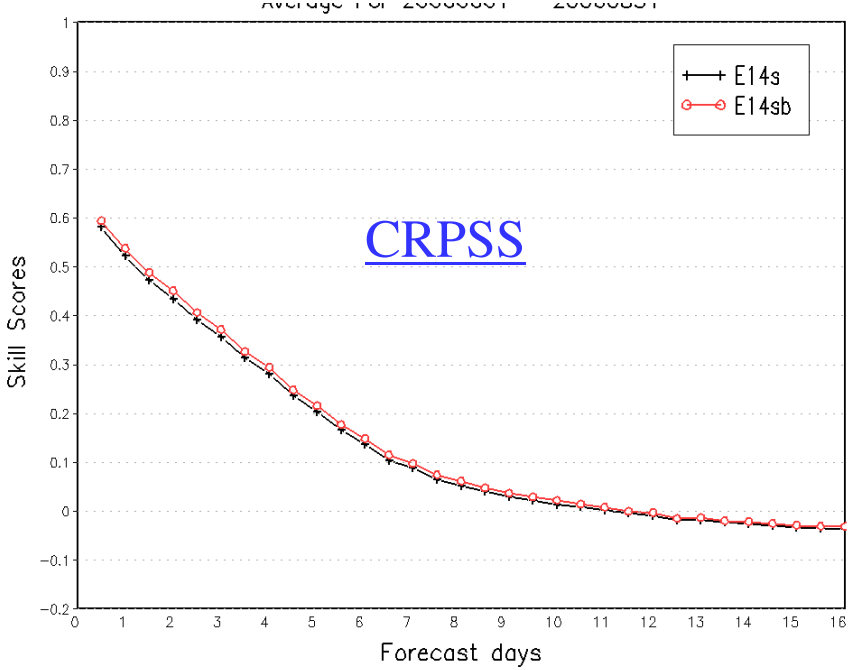
Average For 20061201 – 20070228




Northern Hemisphere 500hPa Height Histogram Distribution
Average For 20060601 – 20060831
E14s
E14sb
HISTOGRAM
1-Day (%)
1-day
3-Day (%)
3-day
Good spread, but more biased
5-Day (%)
5-day
8-Day (%)
8-day
12-Day (%)
12-day
16-Day (%)
16-day







10 meter wind (u-component)

Less biased,

There is less room to improve the skill by bias-correction only

ISSUES ADDRESSED

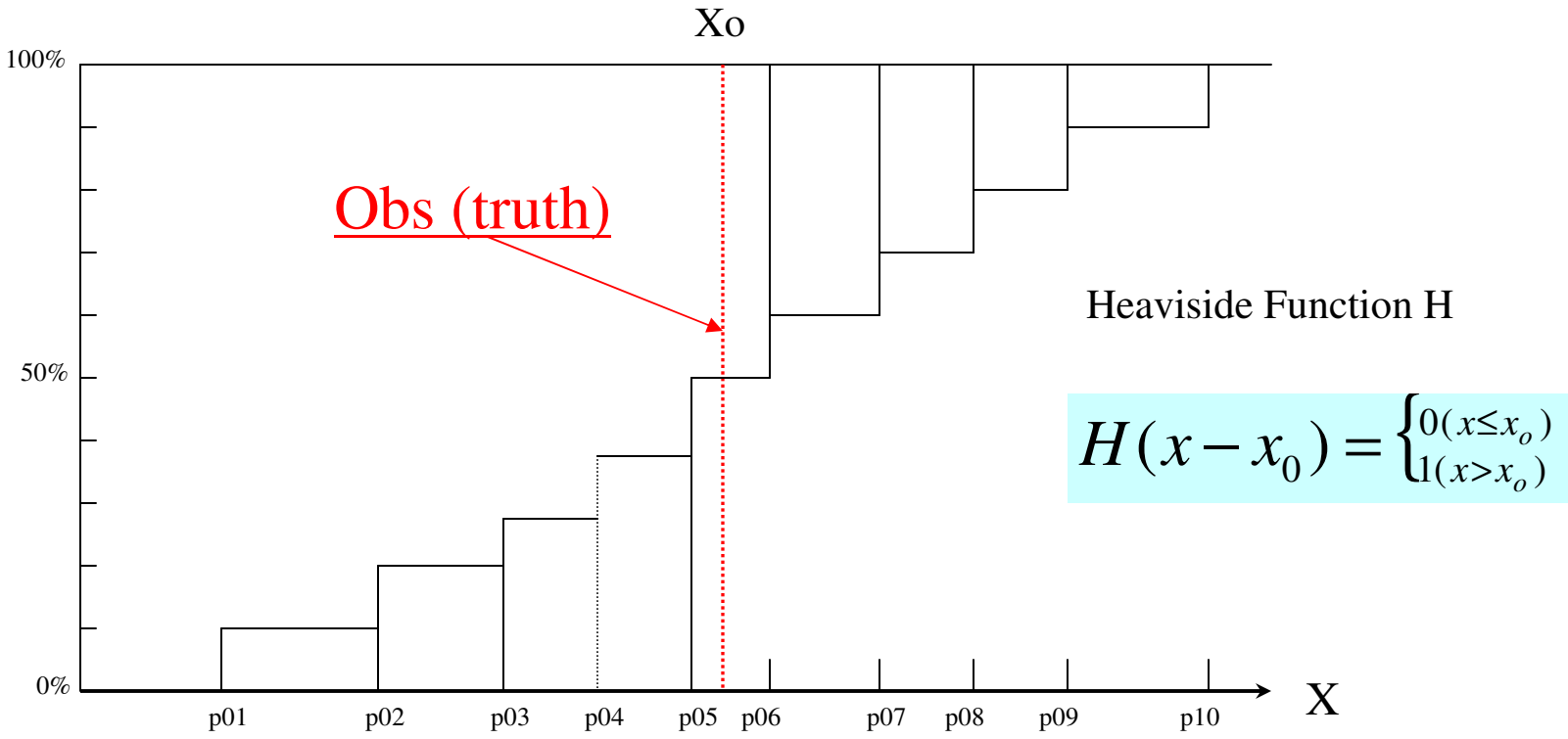
- Effect of bias-correction
  - Different variables
- Comparing of NCEP and CMC's forecasts
  - Before & after bias correction
- Impact of combined ensemble (NAEFS)
  - Before & after bias correction
  - Gains from bias correction + combination =
    - NAEFS advantage

Continuous Rank Probability Score

$$CRPS = \int_{-\infty}^{+\infty} [F(x) - H(x - x_0)]^2 \, dx$$

CRP Skill Score is

$$CRPSS = \frac{CRPS_c - CRPS_f}{CRPS_c}$$

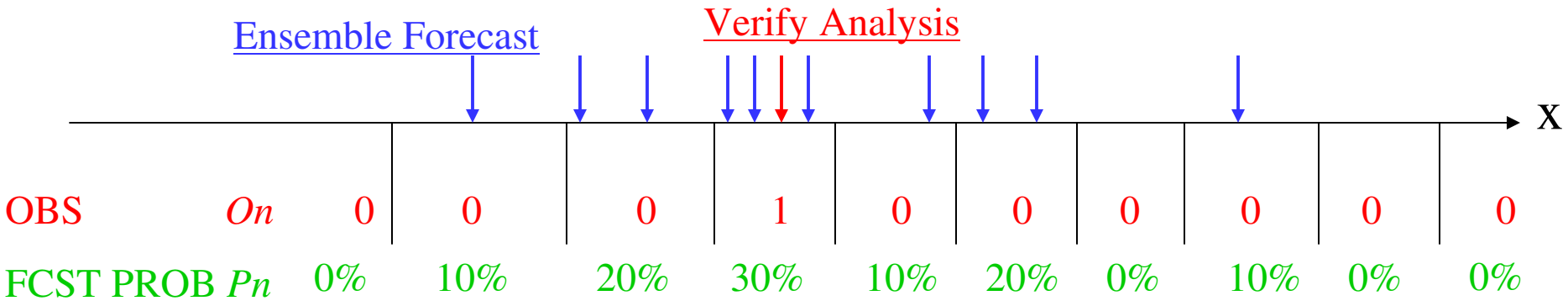


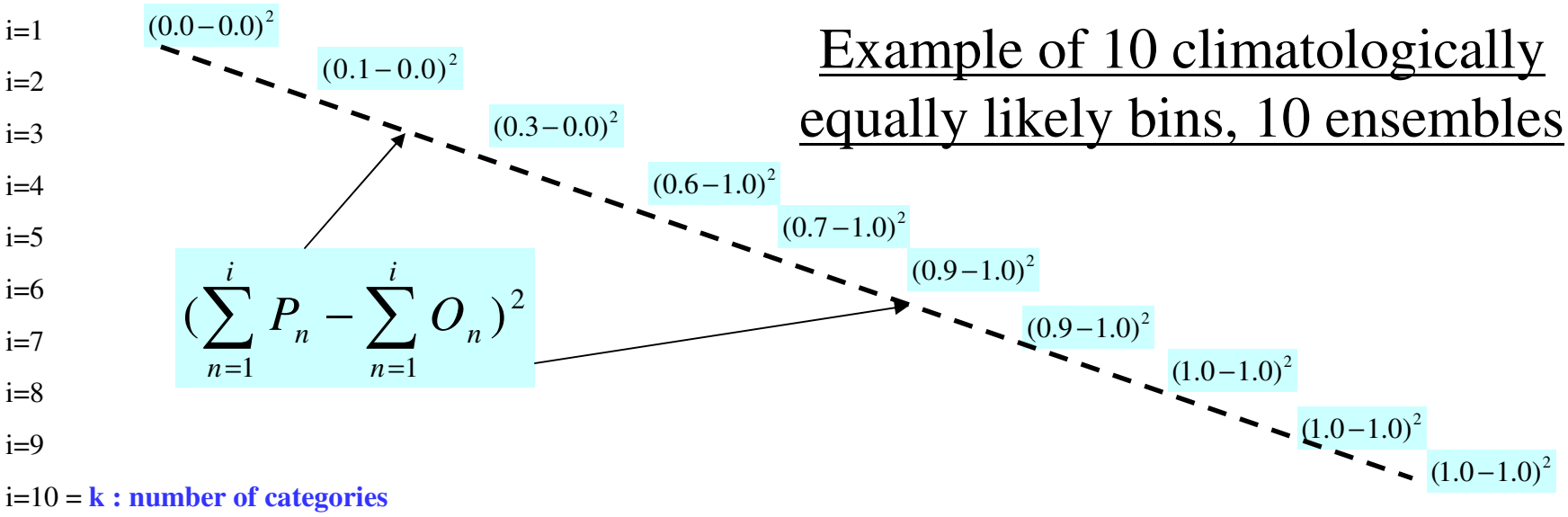
Order of 10 ensemble members (p01, p02,…,p10)

Ranked Probabilistic Score

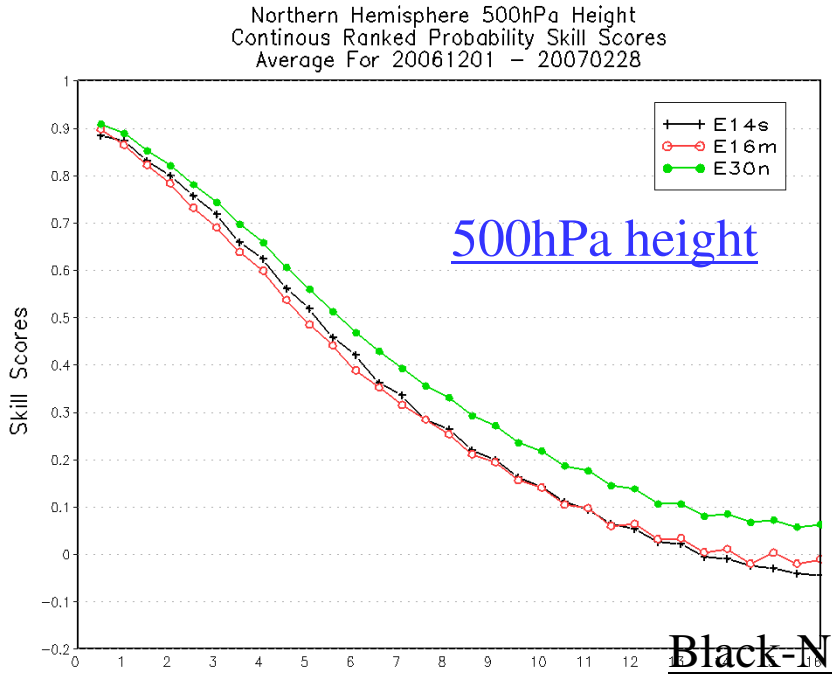
Ranked (ordered) Probability Score (RPS) is to verify multi-category probability forecasts, to measure both reliability and resolution which based on climatologically equally likely bins

$$RPS = 1 - \frac{1}{k-1}\left[\sum_{i=1}^{k}\left(\sum_{n=1}^{i} P_n - \sum_{n=1}^{i} O_n\right)^2\right] \quad \text{and} \quad RPSS = \frac{RPS_f - RPS_c}{1 - RPS_c}$$

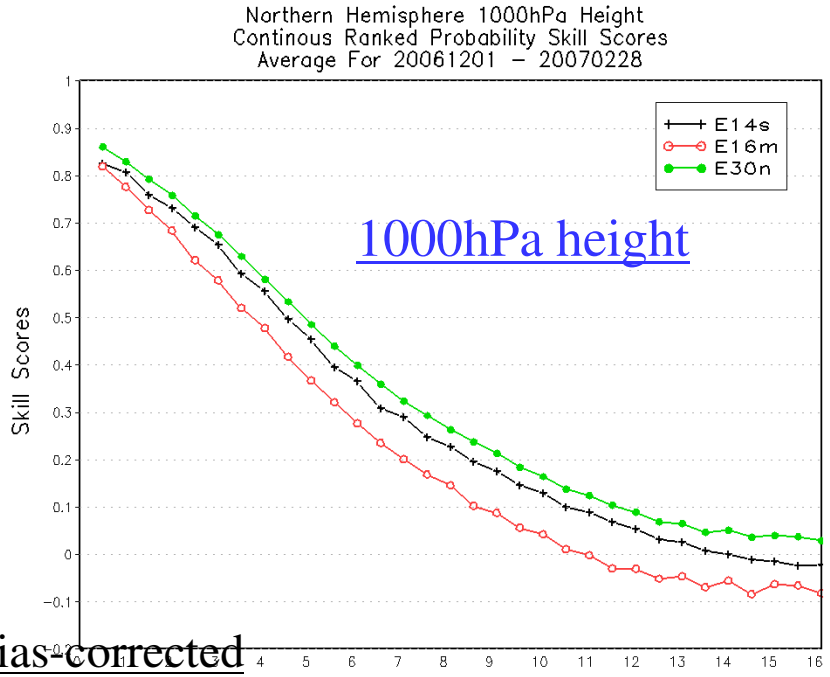




Example of 10 climatologically equally likely bins, 10 ensembles



500hPa height

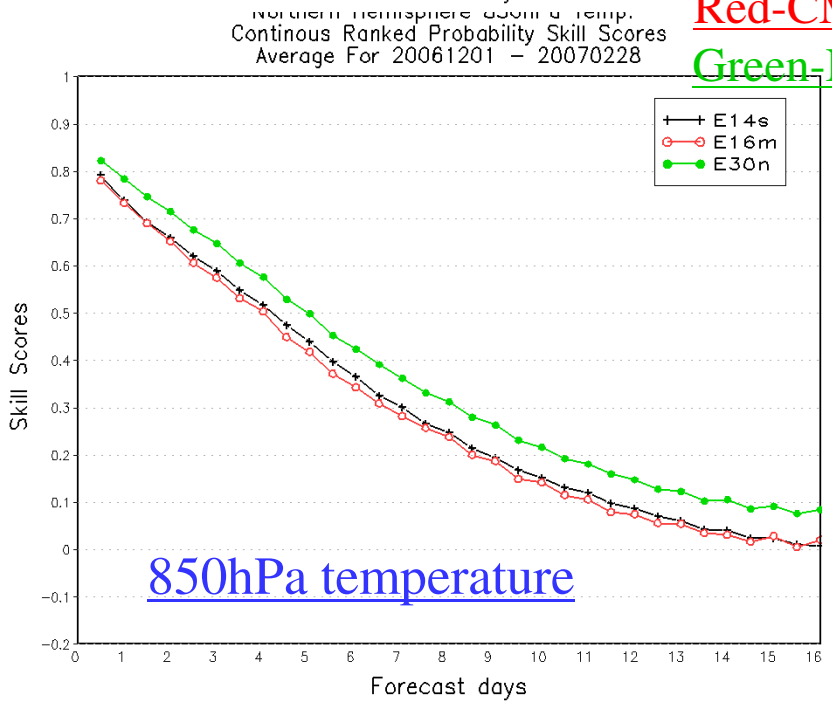


1000hPa height

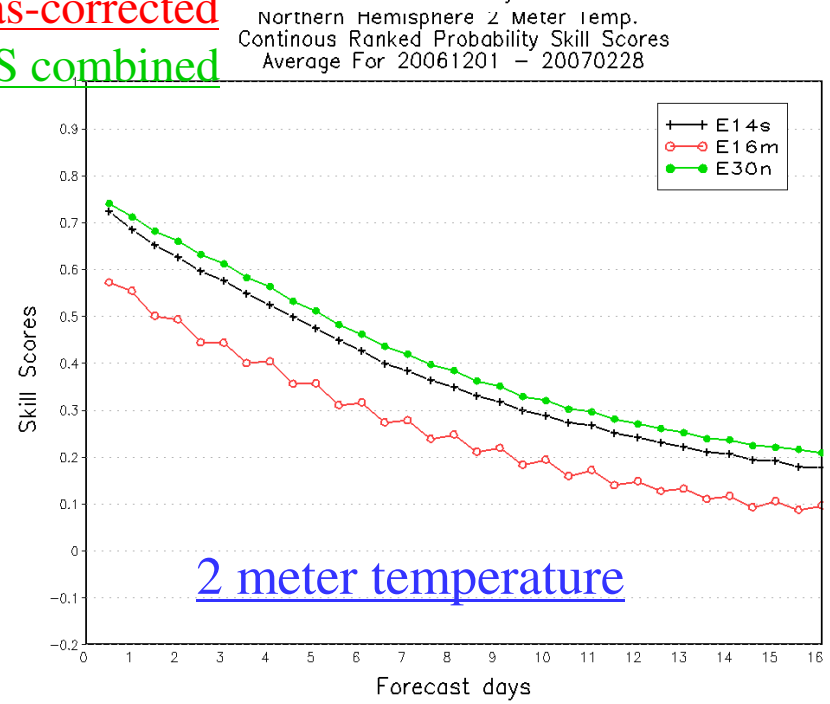
Black-NCEP bias-corrected

Red-CMC bias-corrected

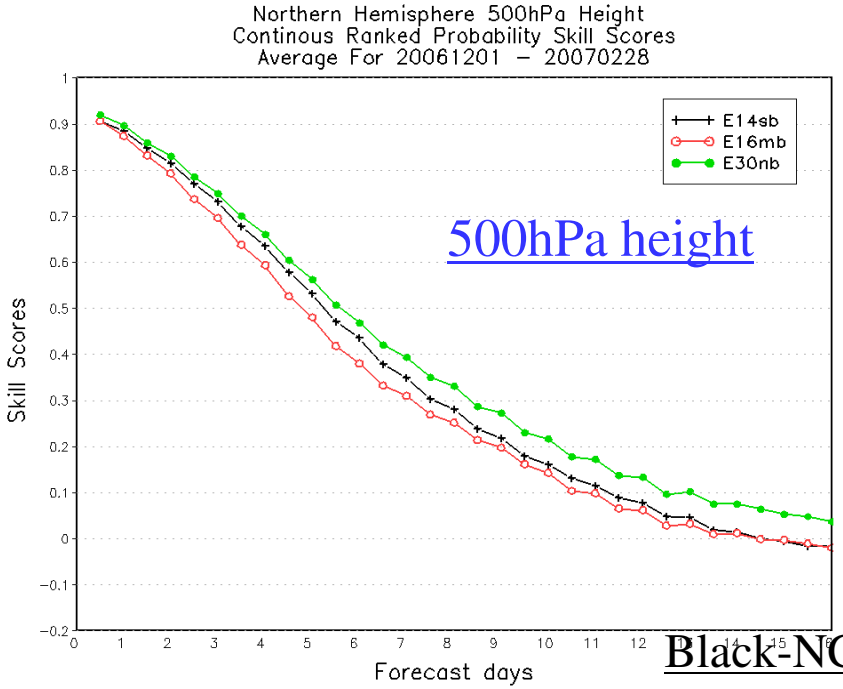
Green-NAEFS combined



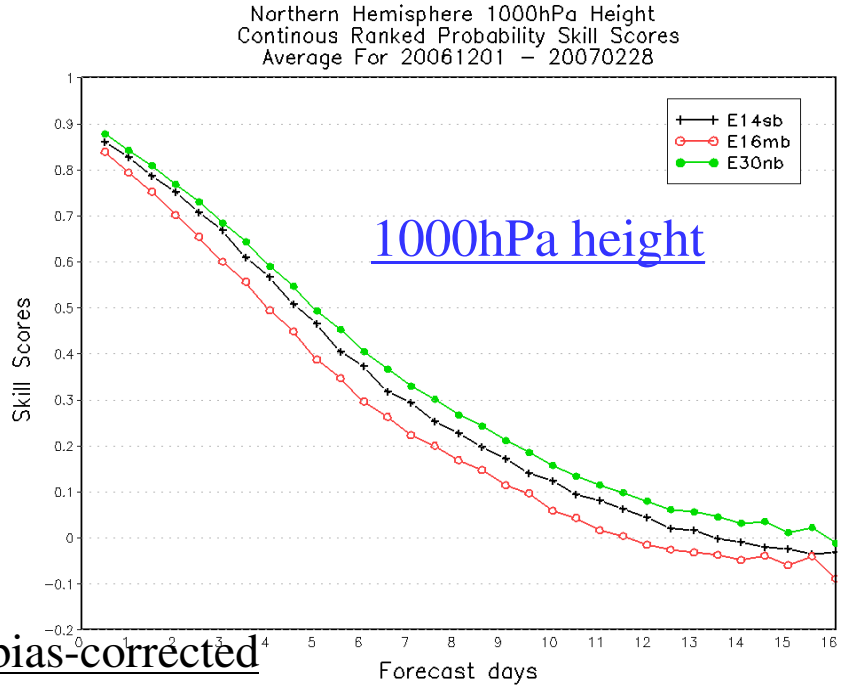
850hPa temperature



2 meter temperature



500hPa height

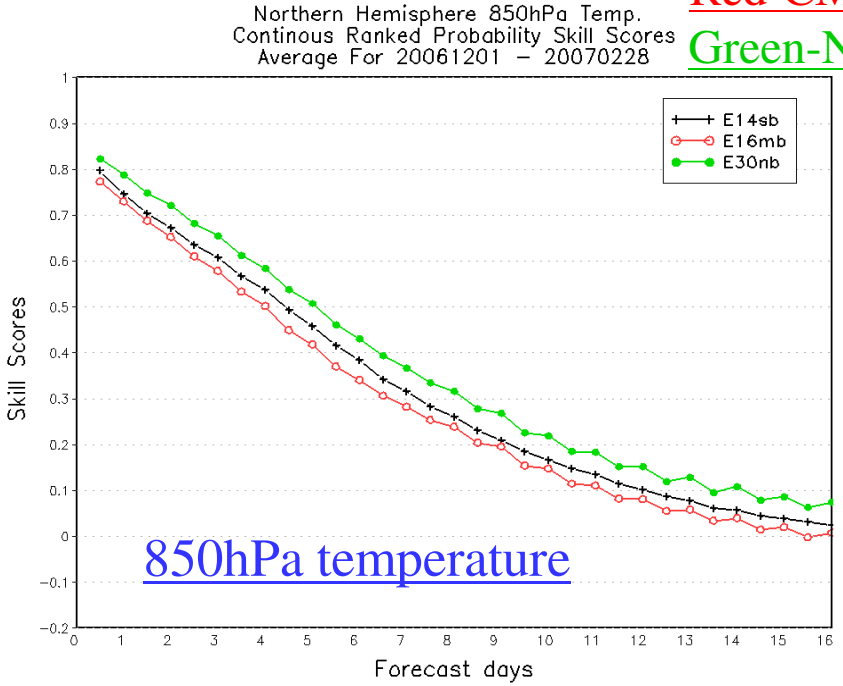


1000hPa height

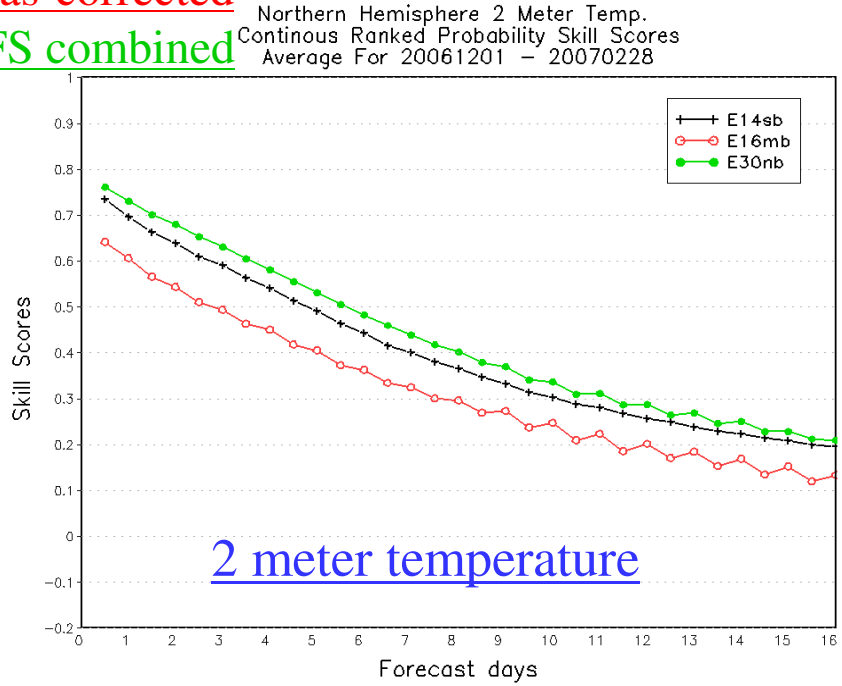
Black-NCEP bias-corrected

Red-CMC bias-corrected

Green-NAEFS combined



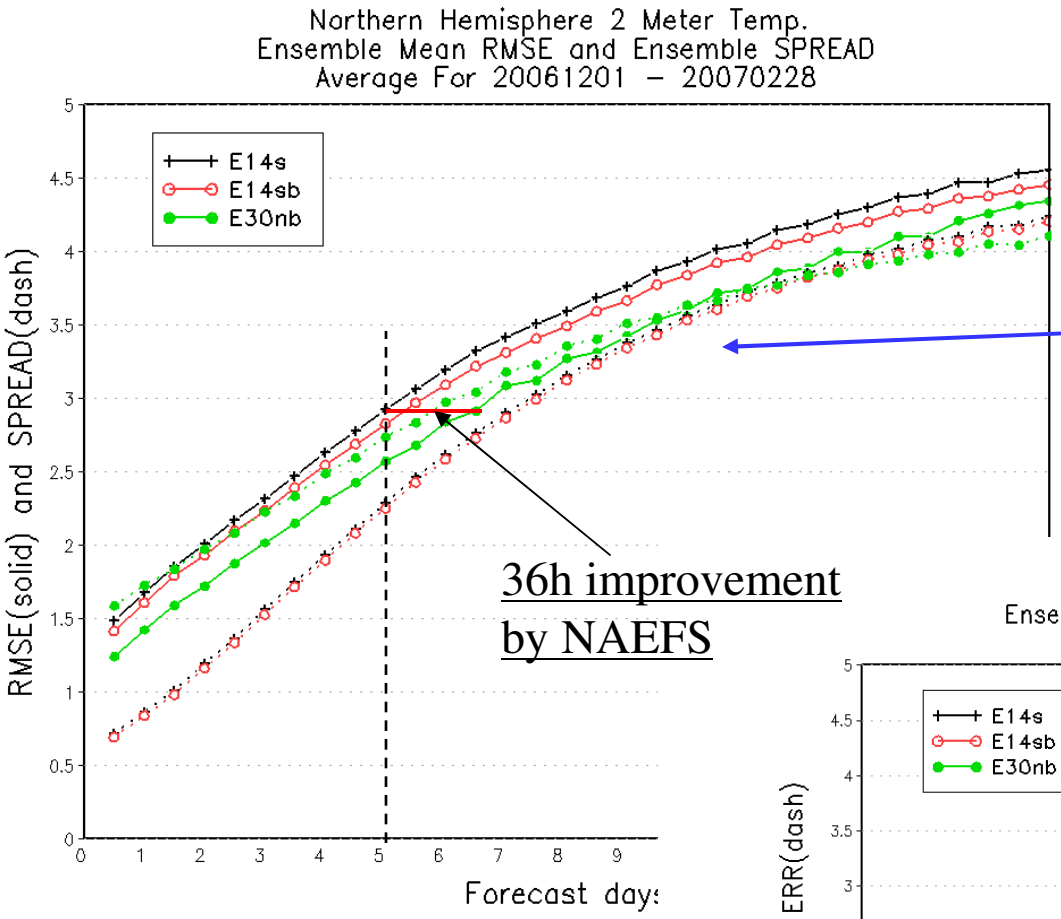
850hPa temperature



2 meter temperature

ISSUES ADDRESSED

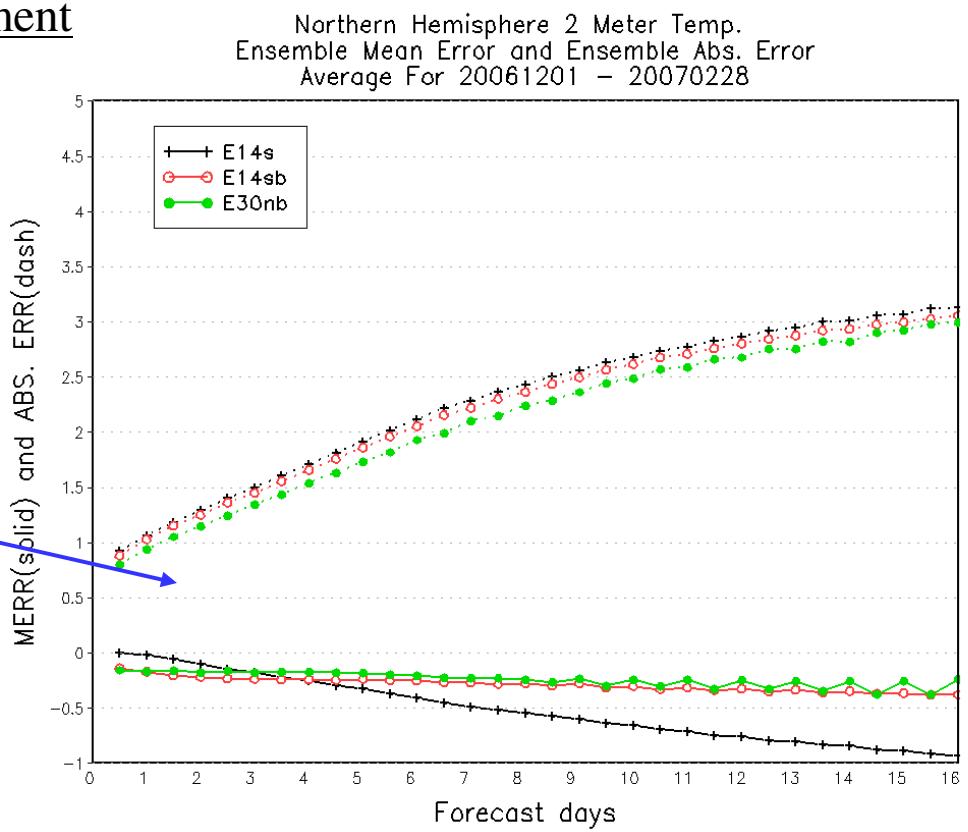
- Effect of bias-correction
  - Different variables
- Comparing of NCEP and CMC's forecasts
  - Before & after bias correction
- Impact of combined ensemble (NAEFS)
  - Before & after bias correction
  - Gains from bias correction + combination =
    - NAEFS advantage



Solid: RMS error

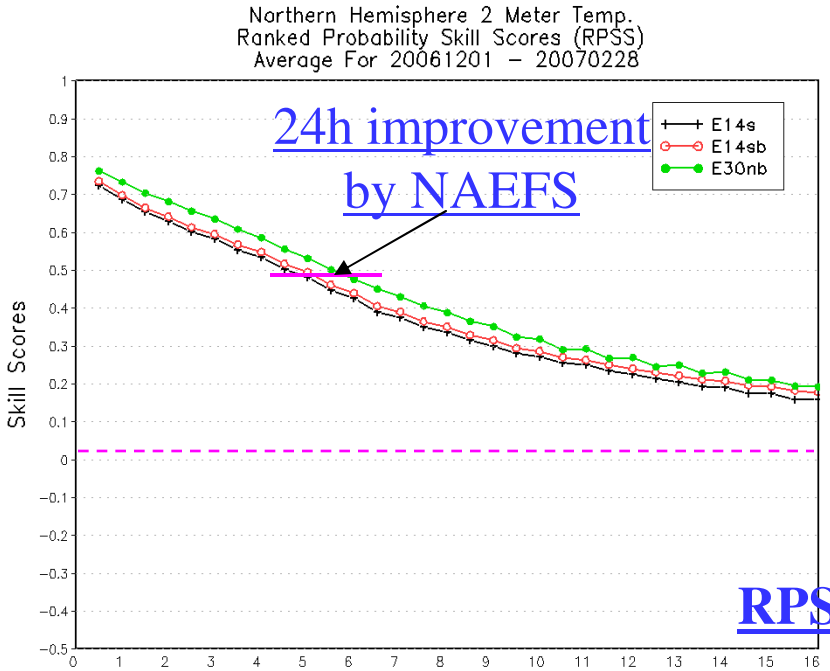
Dash: Spread

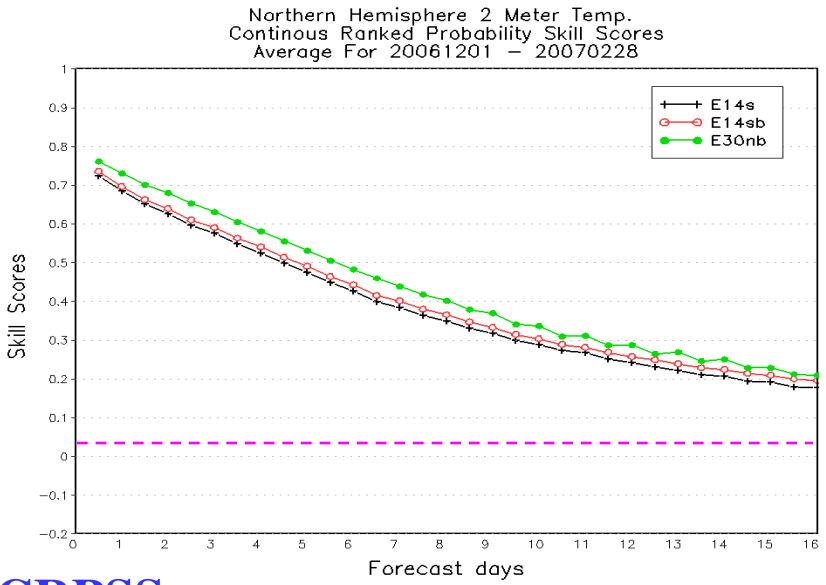
36h improvement by NAEFS



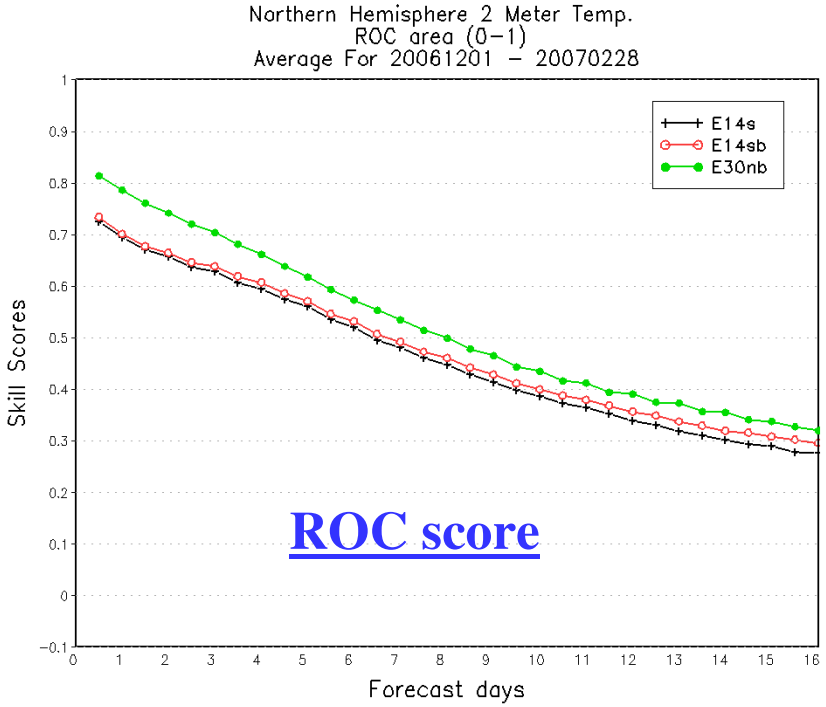
Solid: Mean error (bias)

Dash: Mean absolute error





RPSS .vs CRPSS



ROC score

Winter 2006-2007

NH 2m temperature

For

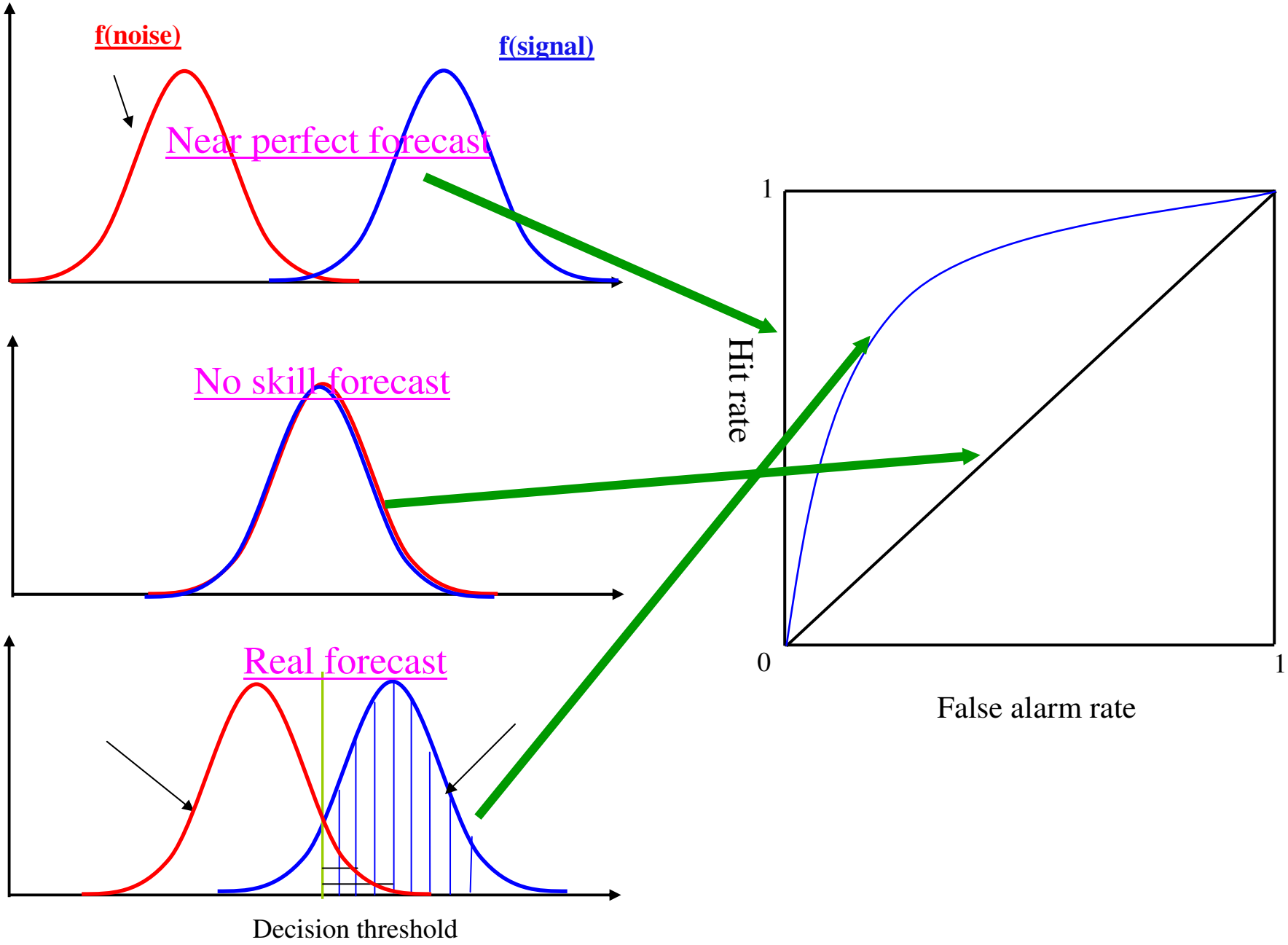
NCEP raw forecast (black)

NCEP bias corrected forecast (red)

NAEFS forecast (pink)

Background !!!!!

Relative Operating Characteristics area (ROC area)



NAEFS Performance Review

Appendix 6

KEY PERFORMANCE MEASURES

Improvement in Ensemble Forecasts				
Requirement		**Threshold**	**Actual 25Apr-10May06**	**Variance**
Ensemble Mean 3-14 Day Lead Time	Bias Reduction (%)	50%	30-70%	Met or exceeded in Tropics & up to D3 elsewhere; slightly below otherwise
	RMS Error Reduction (%)	10%	Up to 10%	Met up to D3, below expected D4 and beyond
Improvement in Ensemble-based Probabilistic Forecasts	3 Day	6 Hours	12 hrs	Exceeded
	7 Day	12 Hours	16 hrs	Exceeded
	10 – 14 Days	24 Hours	48 hrs	Exceeded

NAEFS Configuration Review (NCEP)

Appendix 8

MINIMAL (PREFERRED) CONFIGURATION FOR THE GLOBAL ENSEMBLE FORECAST SYSTEMS OPERATIONAL AT CMC AND NCEP

FEATURE	2005 Plan	2008 Plan	May 2006 Actual / Feb 2007 Plan (NCEP)
Forecast lead time (days)	16	16 (35)	16
Number of cycles per day	2 (4)	4	4
Number of ensemble members	10 (20)	20 (50)	14 / 20
Model resolution (km)	120 (90)	80 (60)	120 / ?
Number of vertical levels	28 (42)	42 (64)	28 / ?