NAEFS post-processing

Yuejian Zhu

Environmental Modeling Center
NOAA/NWS/NCEP

Acknowledgments:
Bo Cui, Dingchen Hou and Yan Lou

Presentation for EMC/OHD THORPEX-HYDRO meeting
September 17 2010

Overview

- Postprocessing project
- Bias correction and statistical downscaling
  - Bias correction globally (NAEFS 50 variables) and regional (SREF)
  - Downscaling for CONUS (4 variables) and Alaska (8)
- Probabilistic verifications
  - Include bias (mean error) and absolute error from ensemble mean
- Calibration of precipitation forecast
  - Working in progress for improved version
- Future plan (THORPEX proposal)
  - Improving the methods
  - Improving extreme events forecast
  - Using reforecast information
- References

Post processing project

- Focused on Multi-Model processing
- Supported by THORPEX program with interagency and international contributions
- Generating community-based software
  - Public access
  - Managed by Subversion
- Implementing many improvements in FY10-11

Bias correction and downscaling

- Bias correction at 1*1 degree resolution (weight=0.02 for Kalman filter algorithm)
  - Bias corrected NCEP/GEFS, GFS (out to 180 hours) and CMC/GEFS forecasts
    - Consider the same bias for NCEP all ensemble members
    - Consider the different bias for each model (member)
  - Combine bias corrected high resolution GFS and low resolution ensembles
    - Dual resolution ensemble approach for short lead time
    - GFS has higher weights at short lead time
  - NAEFS products based on all bias corrected forecasts
    - Produce Ensemble mean, spread, mode, 10% 50%(median) and 90% probability forecast
    - Climate anomaly (percentile) forecasts also generated for ensemble mean

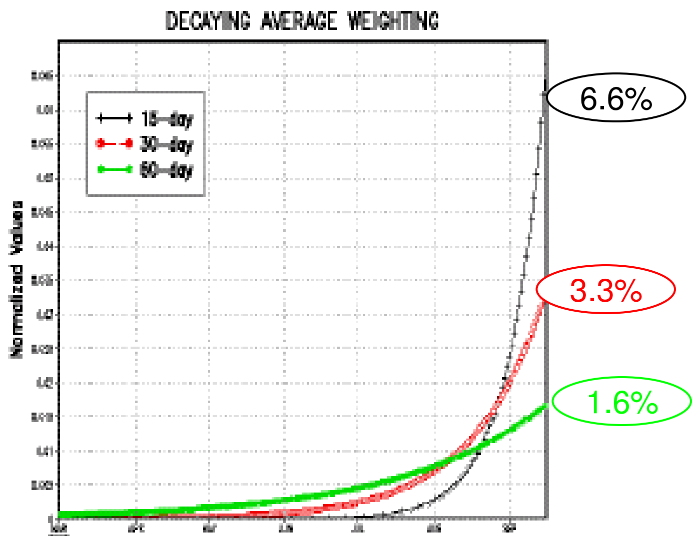
- Statistical downscaling to NDGD grids (weight=0.2 for Kalman filter algorithm)
  - Proxy for truth - RTMA at 5km/6km resolutions
    - Variables (surface pressure, 2-m temp (and Max/min), and 10-m wind (and speed/direction)
  - Downscaling vector
    - Interpolate GDAS analysis to 5km/6km resolutions
    - Compare difference between interpolated GDAS and RTMA
    - Apply decaying weight to accumulate this difference – downscaling vector
  - Downscaled forecast
    - Interpolate bias corrected 1*1 degree NAEFS to 5km/6km resolutions
    - Add the downscaling vector to interpolated NAEFS forecast
  - NAEFS products from downscaling
    - CONUS – NDGD grid/resolution (5km)
      - 4 variables for Ensemble spread, mean, mode, 10%, 50%(median) and 90% forecasts
    - Alaska – NDGD grid/resolution (6km)
      - 8 variables for Ensemble spread, mean, mode, 10%, 50%(median) and 90% forecasts

Bias Correction Method & Application

- **Bias Correction Techniques** – array of methods
  - Estimate/correct bias moment by moment
    - Simple approach, implemented partially
    - May be less applicable for extreme cases
- **Moment-based method at NCEP**: apply adaptive (Kalman Filter type) algorithm

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

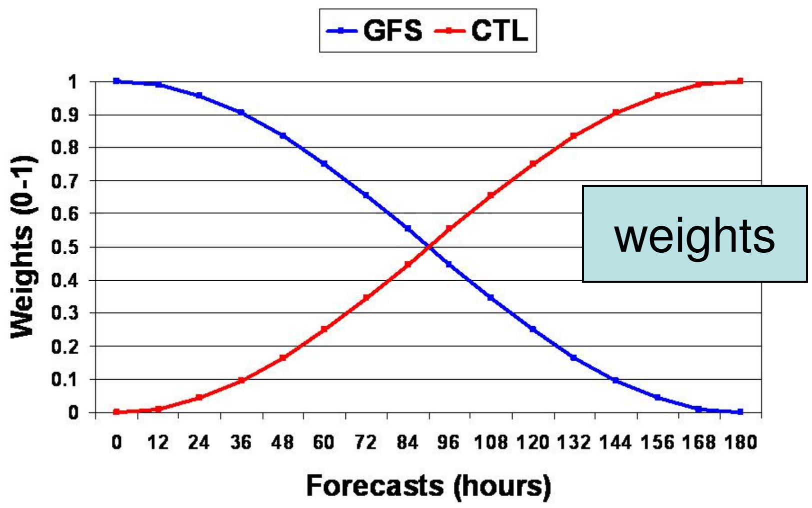
For separated cycles, each lead time and individual grid point, a.m.e = averaging mean error

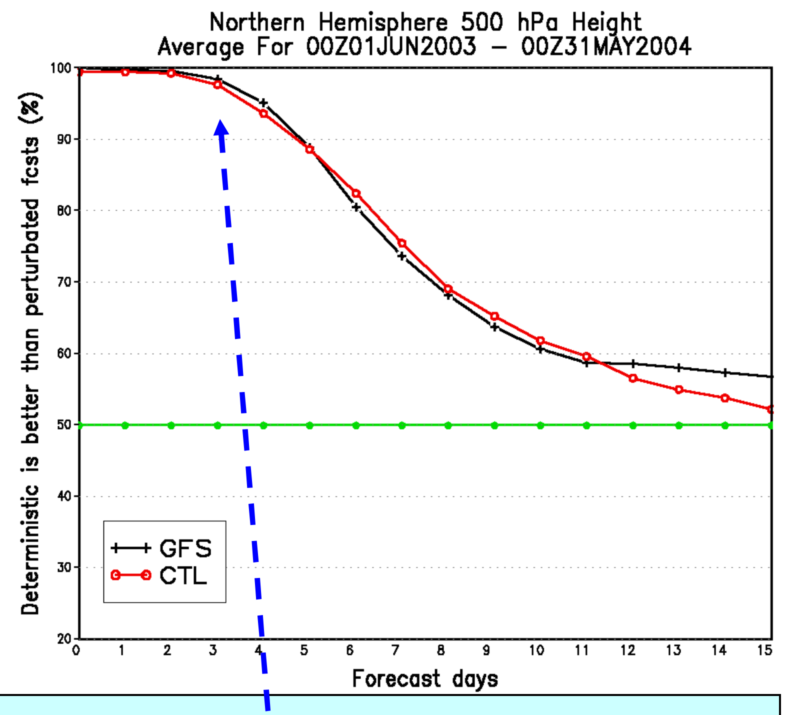


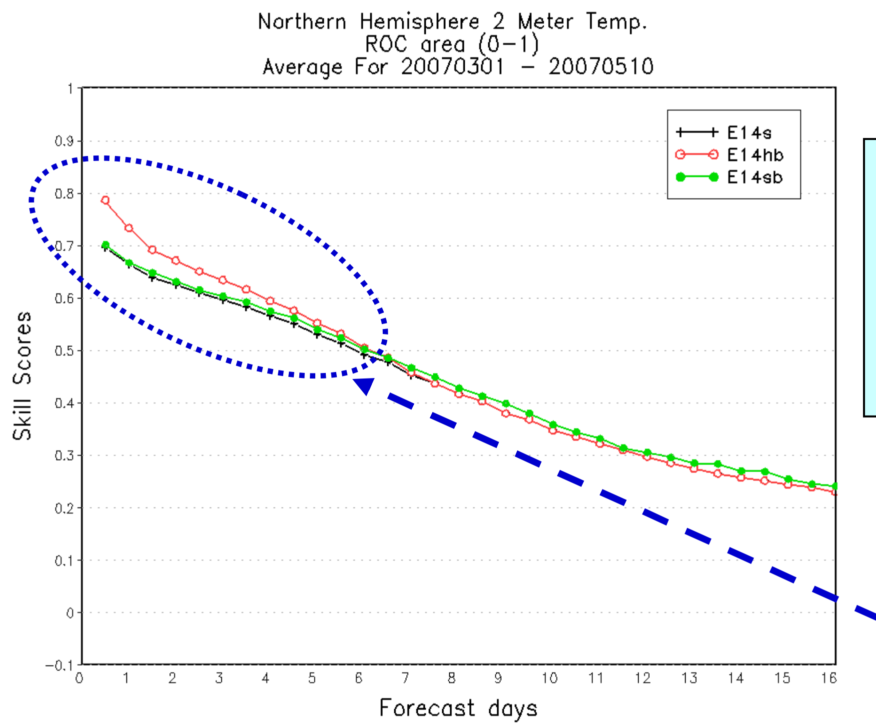
Toth, Z., and Y. Zhu, 2001

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

Hybrid GFS and GEFS

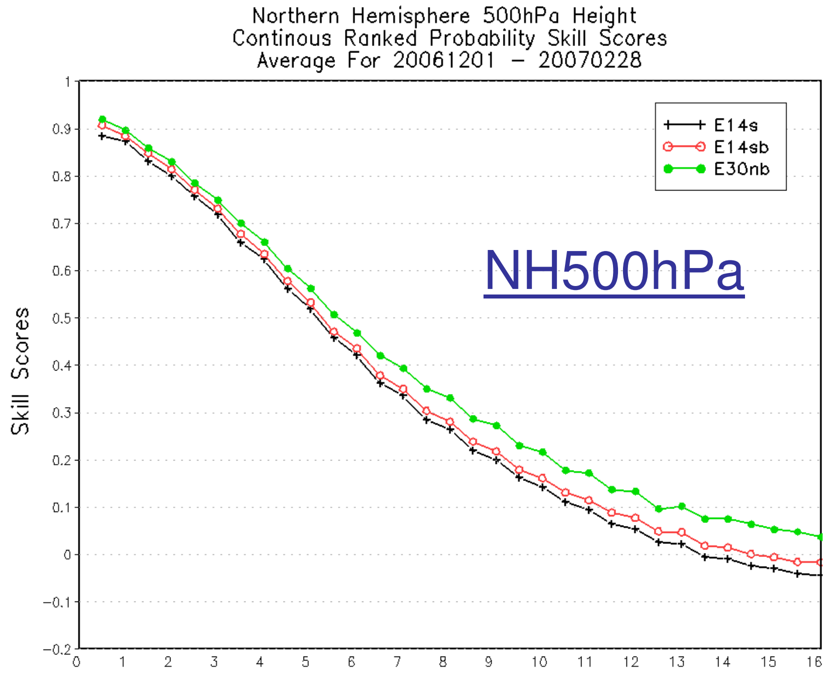


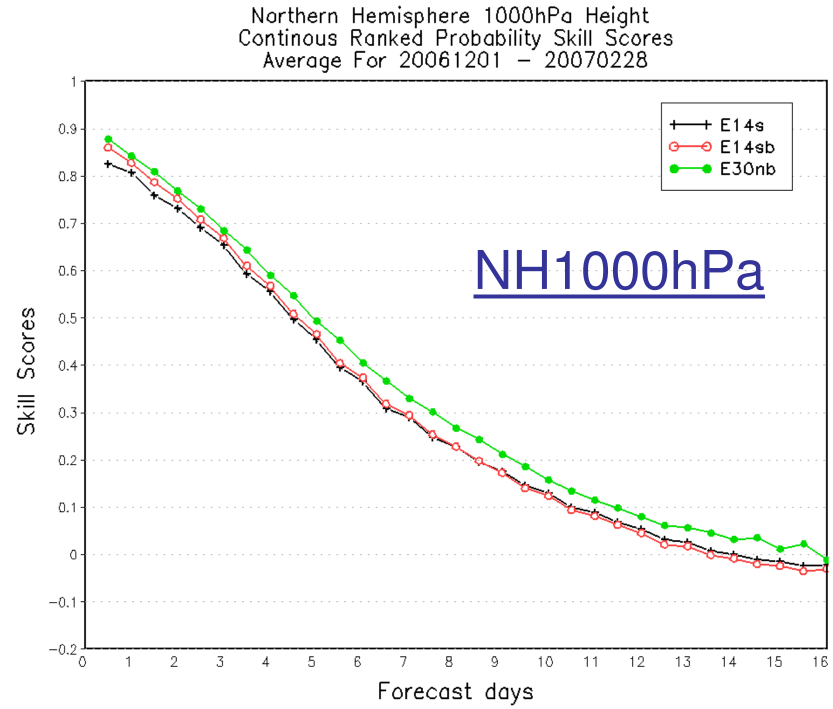


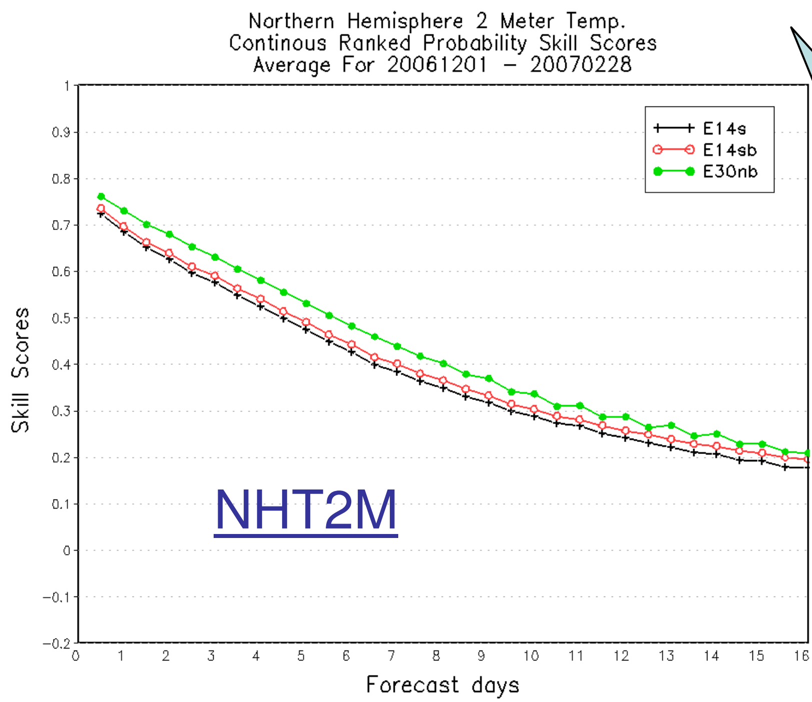


<u>Statistically demonstrate that GFS has more skill than ensemble control (lower resolution) for short lead time</u>

Combined GFS and GEFS bias Corrected forecast for first 180 hours improves the skill (red) than GEFS only (black)

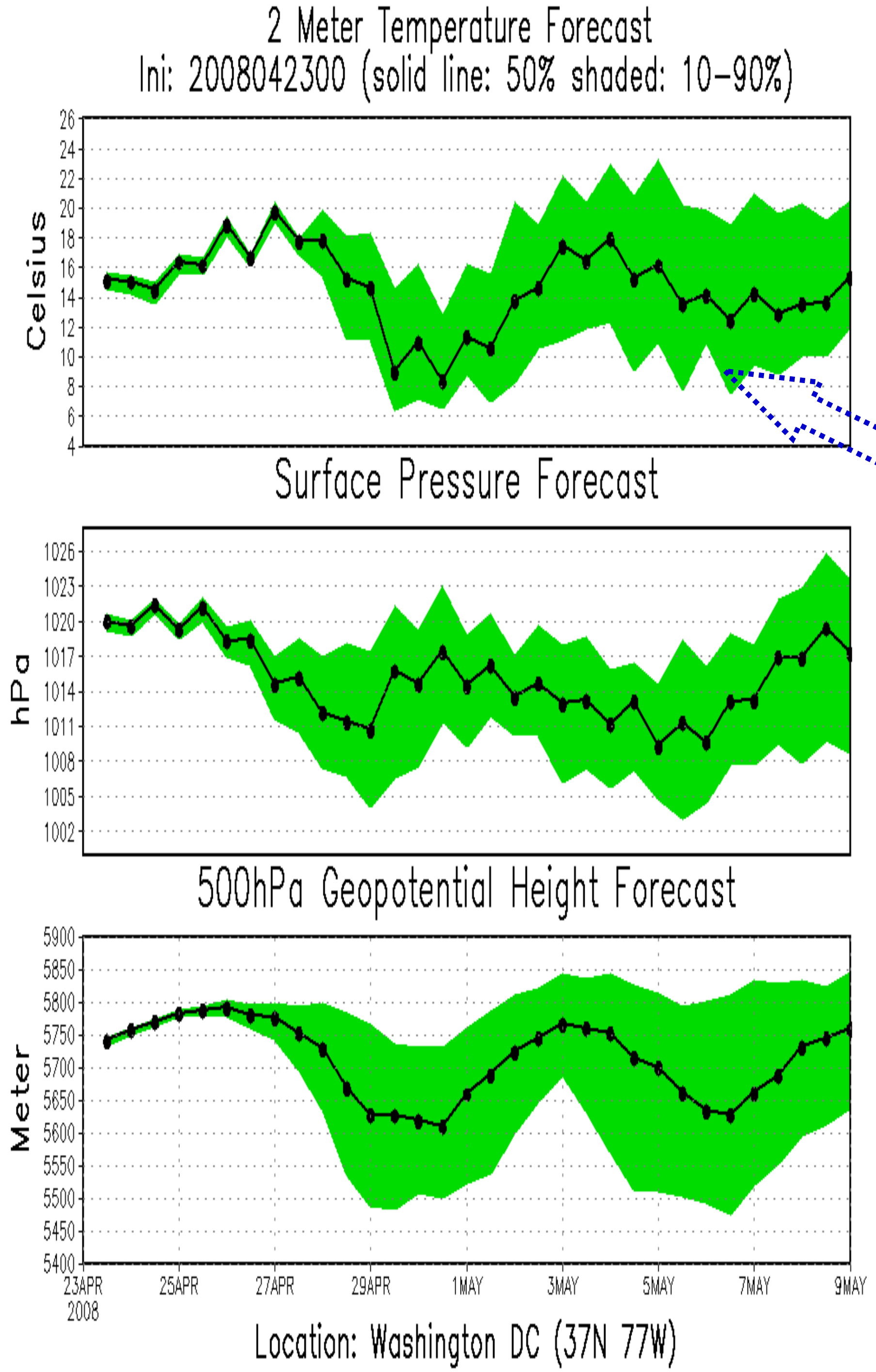


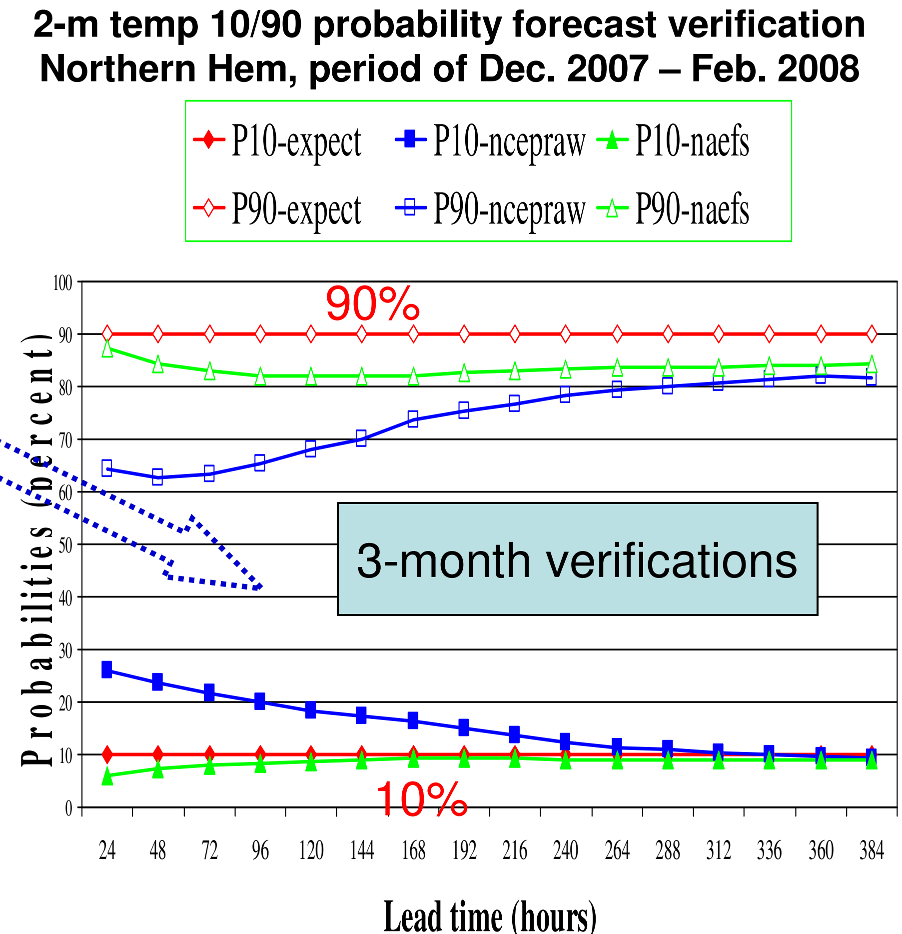




All these stats show the positive impact for probabilistic forecast by apply bias correction and multi-model ensemble (NAEFS) for upper atmosphere (500hPa) and near surface (2-m temperature).

Green line is for NAEFS.





Top: 2-m temperature probabilistic forecast (10% and 90%) verification red: perfect, blue: raw, green: NAEFS

Left: example of probabilistic forecasts (meteogram) for Washington DC, every 6-hr out to 16 days from 2008042300

Downscaling Method with Decaying Averaging Algorithm

- True = high resolution analysis
  - Operational North American Real-Time Mesoscale Analysis (RTMA)
    - 5x5 km National Digital Forecast Database (NDFD) grid (e.g. G. DiMego et al.)
    - 4 variables available: surface pressure, T2m, 10m U and V
  - Other data can also be used

- Downscaling method: apply decaying averaging algorithm

$$\text{Downscaling Vector}^{5km}(t_0) = (1-w) * \text{prior DV}^{5km}(t_{-1}) + w * (\text{GDAS}^{5km}(t_0) - \text{RTMA}^{5km}(t_0))$$

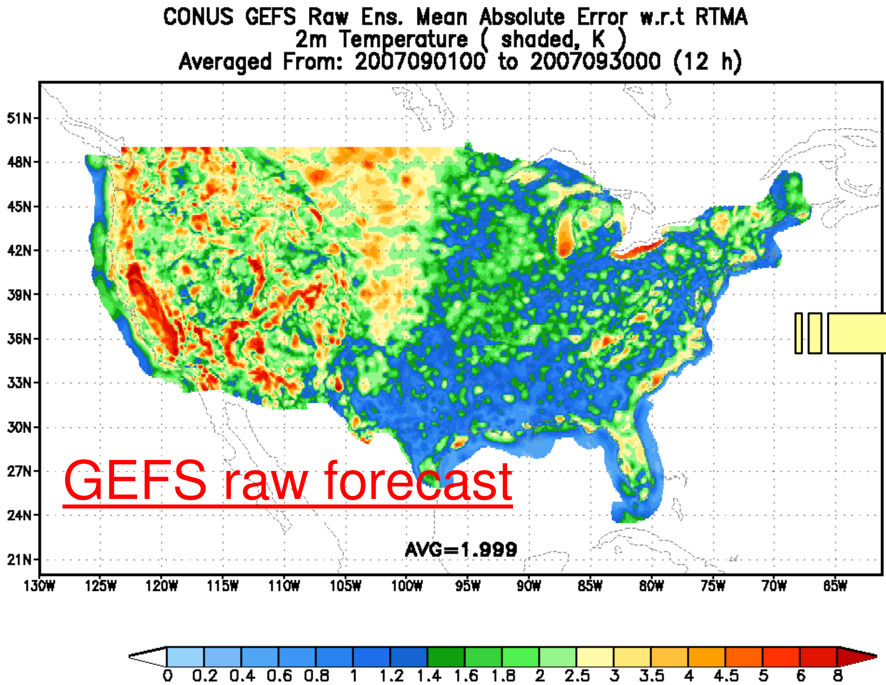
- *$GDAS^{5km}$: GDAS 1x1 analysis interpolated to $RTMA^{5km}$ grids by bilinear interpolation*
- *4 cycles, individual grid point, DV^{5km} = Downscaling Vector on 5km grids*
- *choose different weight: 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 and 0.5*
- ***weight = 0.2 is best and used for weight to calculate downscaling vector***

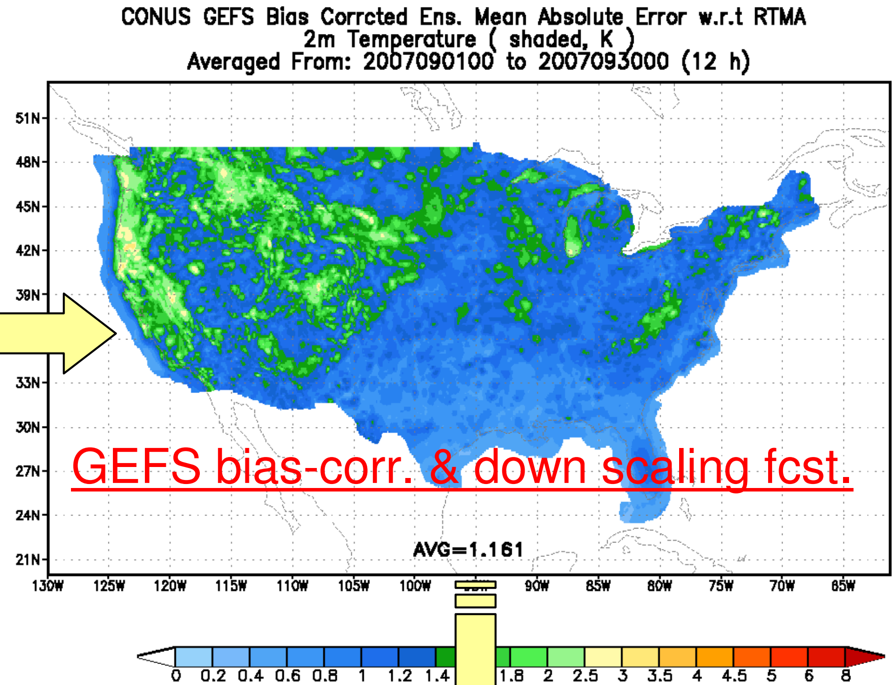
- Downscaling forecast:

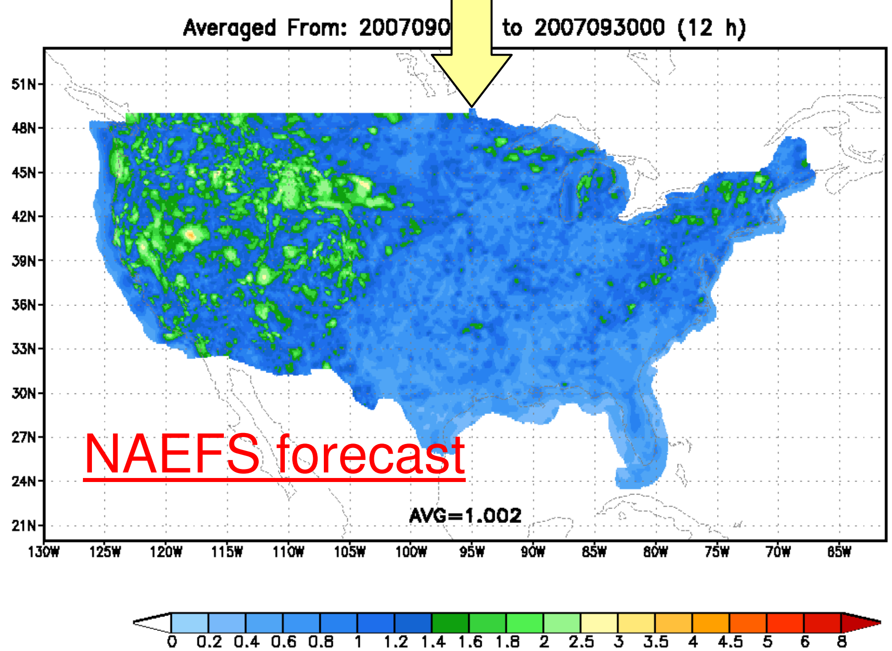
$$\text{Downscaled Forecast}^{5km}(t) = \text{Bias-corrected Forecast}^{5km}(t) - DV^{5km}(t_0)$$

- *Bias-corrected Forecast$^{5km:}$ interpolated to $RTMA^{5km}$ grids by bilinear interpolation*
- *subtract DV^{5km} from bias-corrected forecast^{5km} valid at analysis time*

One month average of mean absolute error for 2-m temp (against RTMA)

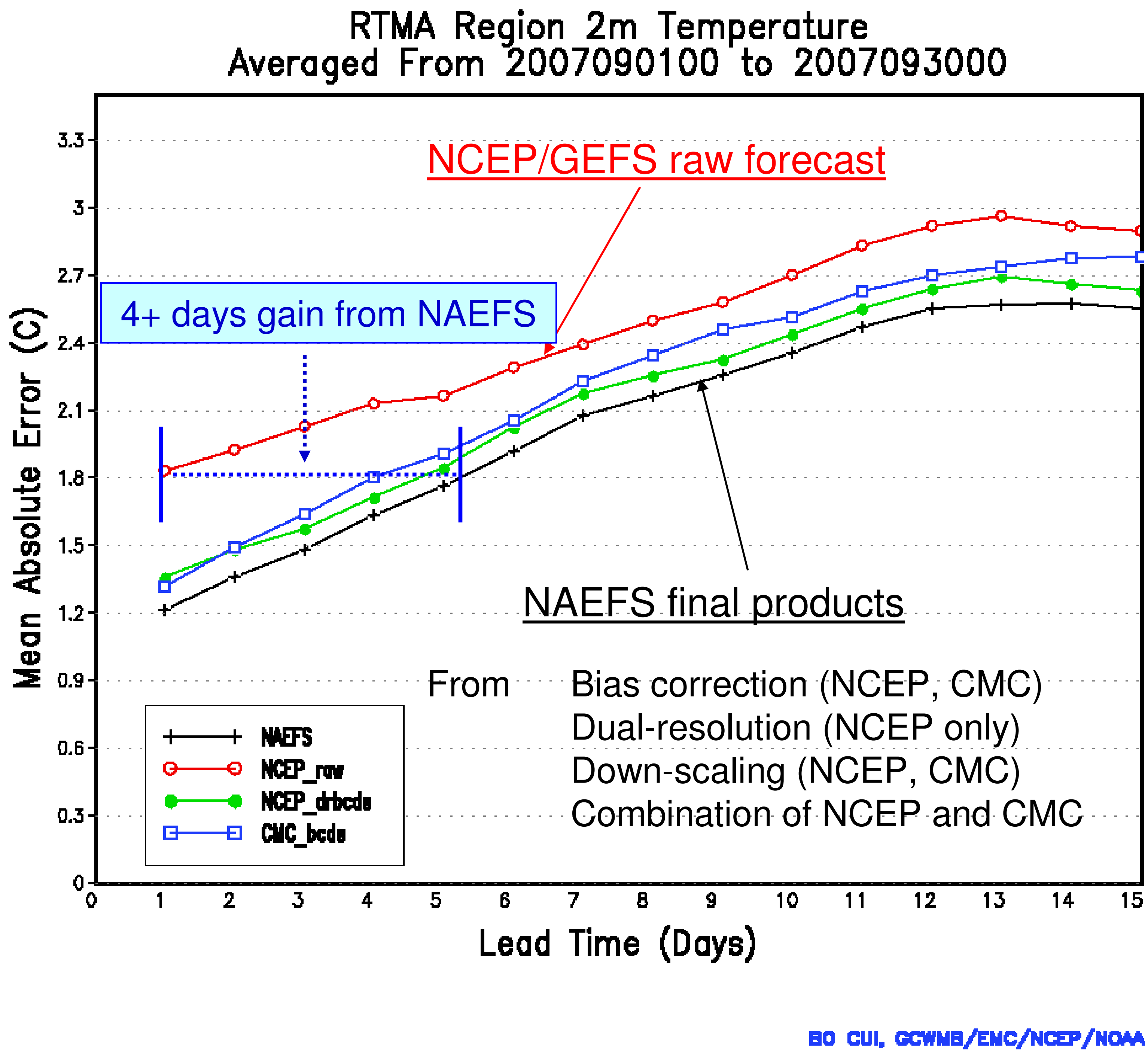


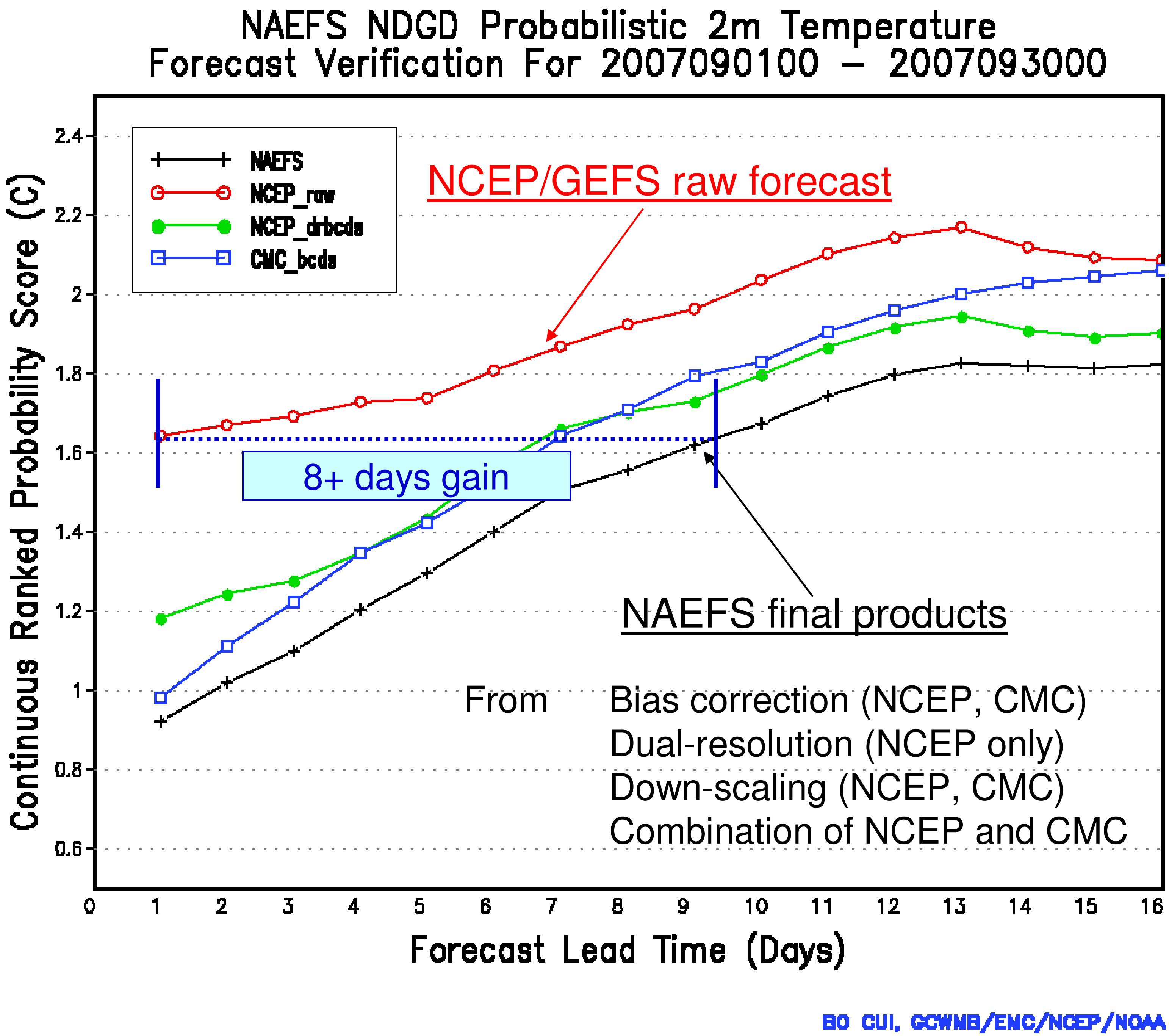




12hr 2m temp forecast
Mean Absolute Error
w.r.t RTMA for CONUS
average for Sept 2007

MAE (raw) =1.999
MAE (bc_GEFS) = 1.161
MAE (NAEFS) = 1.002


RTMA Region 2m Temperature
Averaged From 2007090100 to 2007093000
NCEP/GEFS raw forecast
4+ days gain from NAEFS
NAEFS final products
From
Bias correction (NCEP, CMC)
Dual-resolution (NCEP only)
Down-scaling (NCEP, CMC)
Combination of NCEP and CMC
NAEFS
NCEP_raw
NCEP_drbcds
CMC_bcds
Mean Absolute Error (C)
Lead Time (Days)
BO CUI, GCWMB/EMC/NCEP/NOAA


NAEFS NDGD Probabilistic 2m Temperature
Forecast Verification For 2007090100 – 2007093000
NAEFS
NCEP_raw
NCEP_drbcds
CMC_bcds
NCEP/GEFS raw forecast
8+ days gain
NAEFS final products
From
Bias correction (NCEP, CMC)
Dual-resolution (NCEP only)
Down-scaling (NCEP, CMC)
Combination of NCEP and CMC
Continuous Ranked Probability Score (C)
Forecast Lead Time (Days)
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
BO CUI, GCWMB/EMC/NCEP/NOAA

The evaluations from WFO (State College) forecaster for NAEFS mean minimum temperature

Minimum temperature forecast: Average over past 30 days: (20080929-20081028)

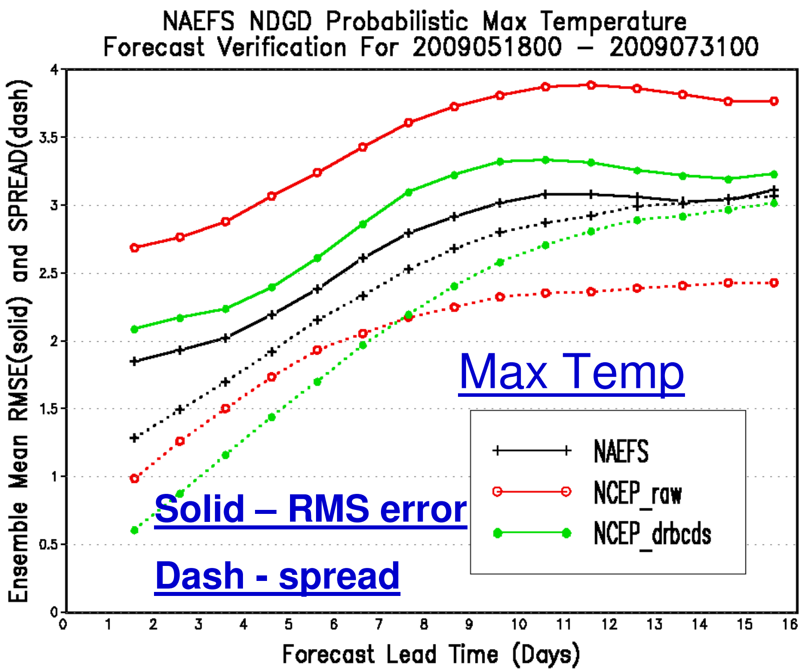
		MAE	Bias	>10 err	<3 err	off. rank	Best G.		2nd G.		Worst G.	
1	12-hr	3.17	-1.2	1.0%	53.4%	3 out of 7	NAEFS	59.7%	SREF	57.1%	NGM80	21.8%
2	24-hr	3.03	-0.9	0.6%	55.5%	2 out of 7	SREF	57.2%	NAEFS	54.2%	NGM80	24.9%
3	36-hr	3.25	-0.8	0.9%	51.6%	3 out of 7	NAEFS	54.2%	SREF	53.9%	NGM80	23.2%
4	48-hr	3.94	-1.1	2.9%	43.2%	3 out of 7	NAEFS	51.9%	SREF	45.8%	NGM80	6.2%
5	60-hr	4.30	-0.4	4.4%	39.1%	4 out of 6	NAEFS	49.2%	SREF	43.0%	NAM40	8.9%
6	72-hr	4.76	0.1	6.4%	33.7%	5 out of 5	NAEFS	42.9%	SREF	40.1%	NAM12	35.2%
7	84-hr	4.85	0.3	7.5%	34.7%	2 out of 6	NAEFS	40.0%	MOSGd	33.4%	NAM12	8.9%
8	96-hr	5.24	0.4	13.0%	33.1%	1 out of 3	NAEFS	32.7%	MOSGd	29.9%	MOSGd	29.9%
9	108-hr	5.11	0.8	12.8%	35.4%	1 out of 4	HPCGd	34.5%	NAEFS	32.1%	MOSGd	30.5%
10	120-hr	5.31	0.7	12.0%	31.9%	1 out of 3	MOSGd	31.6%	NAEFS	24.8%	NAEFS	24.8%
11	132-hr	4.97	0.7	9.9%	35.1%	2 out of 4	HPCGd	38.0%	MOSGd	30.9%	NAEFS	27.2%
12	144-hr	5.42	0.6	15.0%	35.0%	1 out of 3	MOSGd	31.3%	NAEFS	29.0%	NAEFS	29.0%
13	156-hr	5.40	0.5	14.9%	35.7%	1 out of 4	HPCGd	32.9%	MOSGd	32.7%	NAEFS	23.4%
14	168-hr	5.46	1.1	17.7%	38.1%	1 out of 3	MOSGd	35.6%	NAEFS	28.4%	NAEFS	28.4%

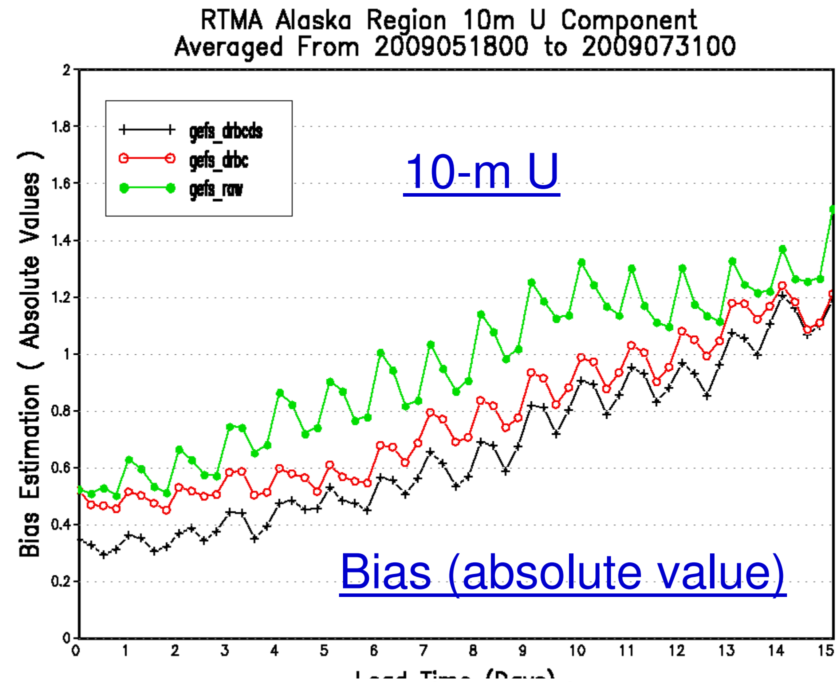
Official Guidance: NGM80, NAM40, SREF, NAM12, MOSGd, HPCGd, NAEFS

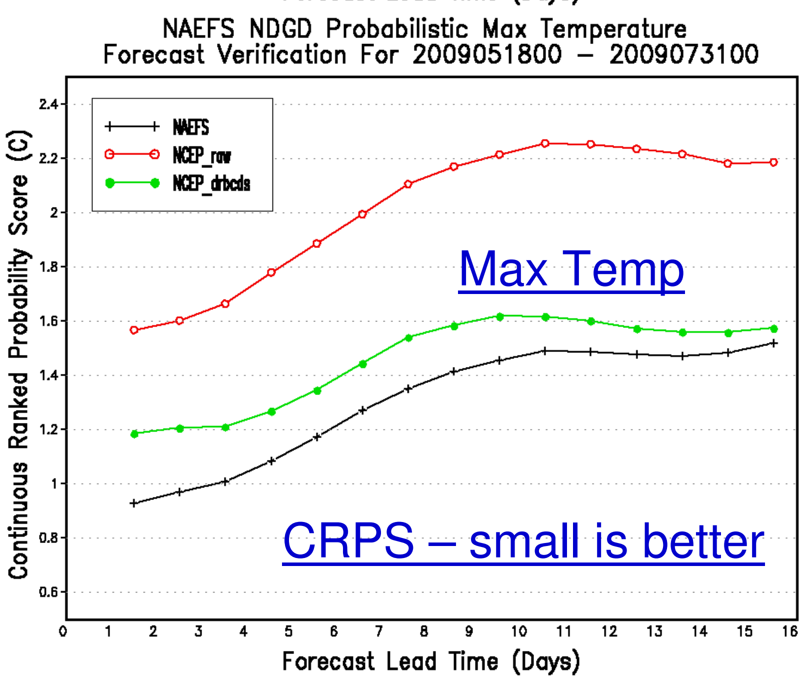
NAEFS downscaled 5km (NDGD) minimum temperature (mean) is the best guidance for first 96hr forecasts from 7 different guidance

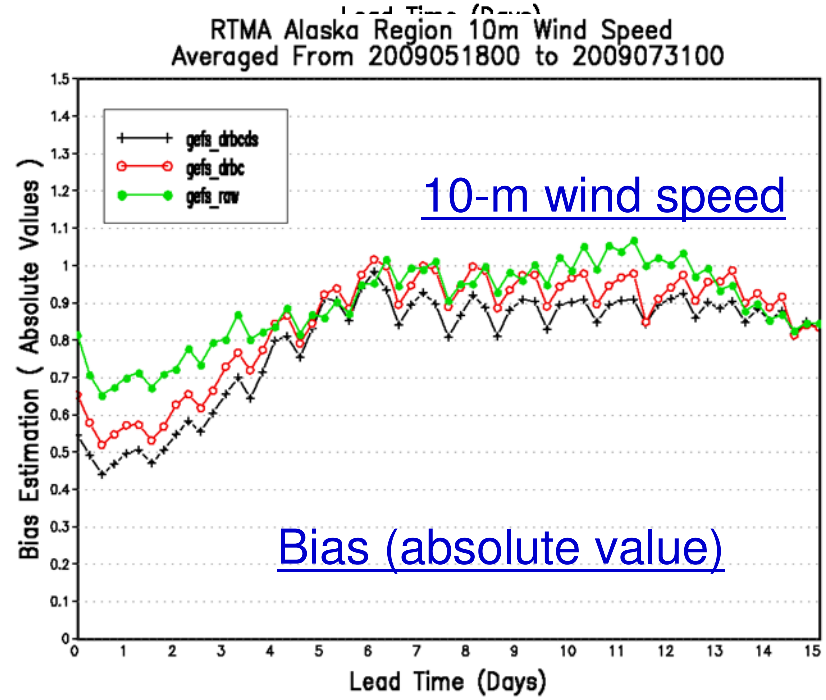
Contributed by Richard Grumm (WFO)

Application for Alaska region and HPC Alaska desk





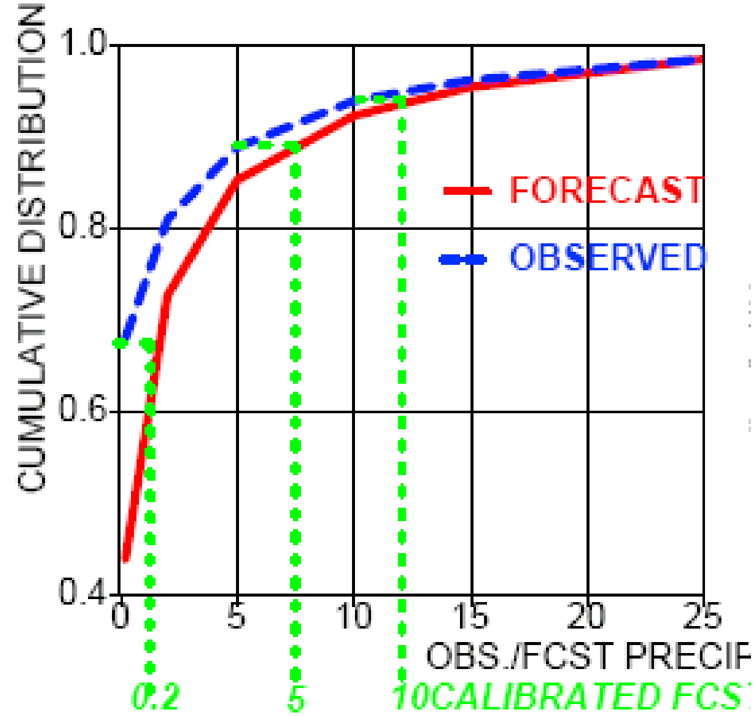




Calibration of precipitation forecast

Implemented May 2004 (HPC and CPC endorsed)

Latest experiments for every 6hr instead of 24hr



METHOD

1) Construct cumulative frequency distributions for forecast QPF & corresponding observed values
2) For each forecast value, find the observed value that has the same frequency as forecast value
3) Re-label forecast value with corresponding observed value

DETAILS

Observations used:

CCPA – climatological calibrated precipitation analysis

Adaptive method, training data accumulated over:

Most recent ~30-day period – *Decaying averaging*

More weight on most recent data

Continental US

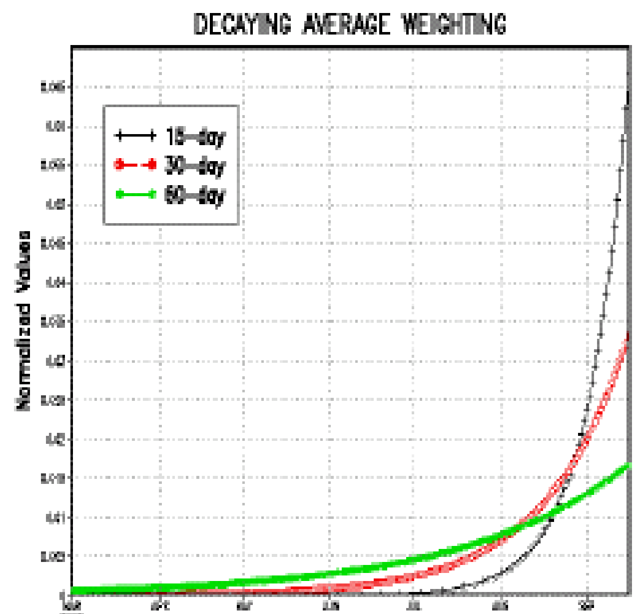
Linear inter/extrapolation

Corrections applied CONUS (and globally) on model grid

Correction is function of forecast value

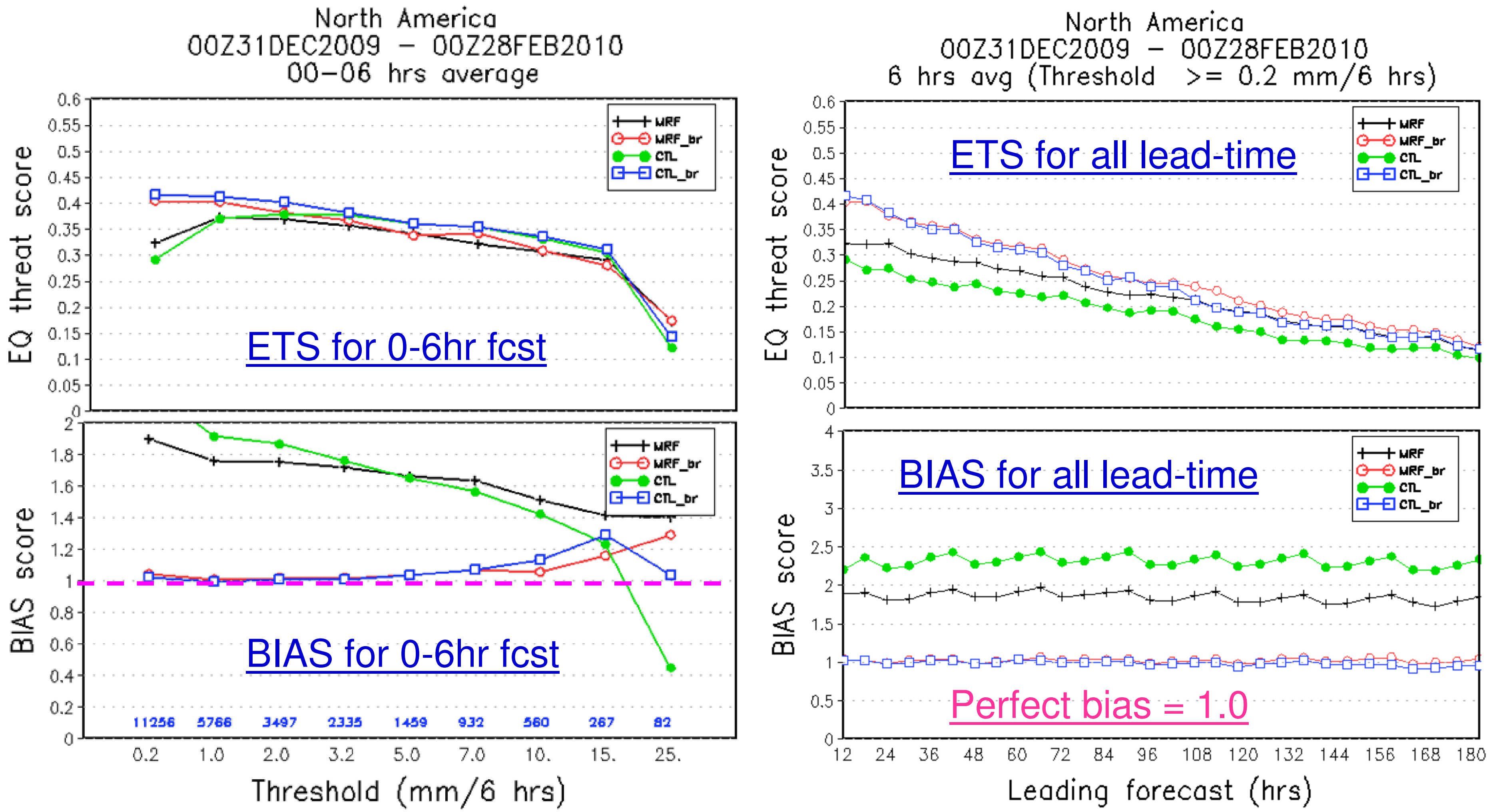
1*1 degree (and 5km – downscaled) spatial resolution

Every 6-hour forecast interval (3-hour later)



Precipitation calibration for 2009-2010 winter season (CONUS only)

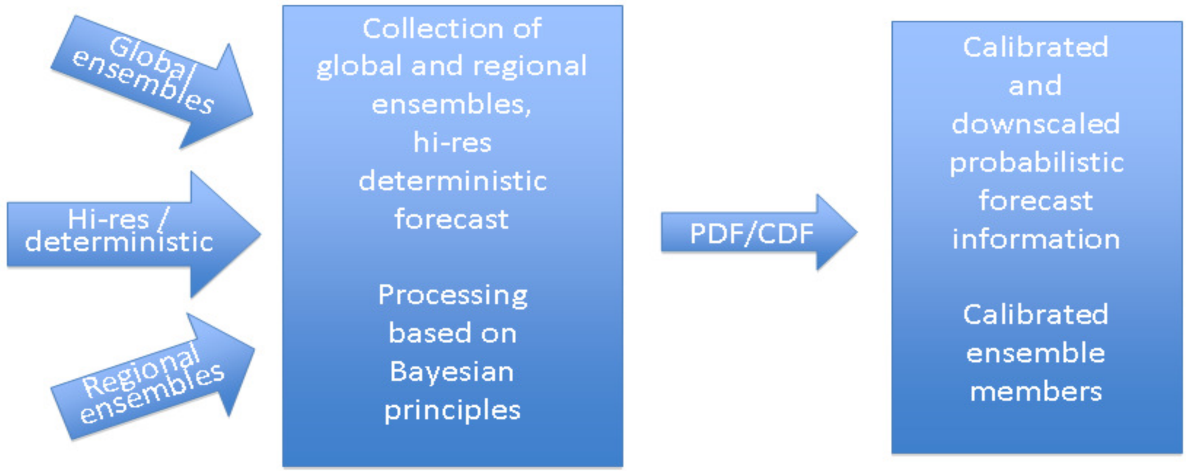
Comparison for GFS and ensemble control (raw and bias corrected)



The probabilistic scores (CRPS -not show here) is much improved as well.
We are still working on the different weights, different RFC regions, downscaled to 5km as well. More results will come in soon. Plan for implementation: Q2FY11

Development Plan of Statistical Post-Processing for NAEFS

Cui/Yuan's THORPEX proposal (2010-2013)



- Opportunities for improving the post-processor
  - Utilization of additional input information
    - More ensembles - high resolution control forecasts, SREF, GEFS …
    - Using reforecast information to improve week-2 and precipitation
    - Improving analysis fields (such as RTMA and etc..)
  - Improving calibration technique
    - Calibration of higher moments (especially spread)
    - Use of objective weighting in input fields combination
    - Processing of additional variables with non-Gaussian distribution
  - Improve downscaling methods
  - Future plan (overlook beyond 2-3 years)
    - Bias correct all model output variables (>200 which include precipitation)

Software – public access through subversion

- Available software – already in NWS operation
  - Bias correction for all near Gaussian distribution variables
  - Downscaling for surface variables
    - Include maximum/minimum temperature (derived variables)
    - Include wind speed/direction (derived variables)
  - Precipitation calibration (2004 version)
    - New version will be implemented soon
  - Ensemble verification packages
    - Verification for ensemble mean (such as RMS error, bias, et al.)
    - Verification for probabilistic forecast
- Advantage – shared the same algorithm
  - MSC (Meteorological Service of Canada) uses it in operation
  - FNMOC – already receive it, will use it later this year
  - ESRL/GSD (Toth) – in testing
  - ESRL/PSD (Whitaker) – use EMC's verification package
  - OHD (DJ Seo) – shared verification package
  - Many institutions use EMC's verification package
- Public access through subversion
  - The same as GFS, GSI, HWRF and et al.

References:

- December 14 2007 implementation: http://www.emc.ncep.noaa.gov/gmb/yzhu/html/imp/200711_imp.html
- February 23 2010 implementation: http://www.emc.ncep.noaa.gov/gmb/yzhu/html/imp/201002_imp.html
- Q4 FY10 implementation: http://www.emc.ncep.noaa.gov/gmb/yzhu/html/imp/201004_imp.html
- Zoltan and et al. 2005: http://www.emc.ncep.noaa.gov/gmb/ens/papers/toth_naefs_thorpex_montreal.pdf
- Cui and et al. 2006: http://www.emc.ncep.noaa.gov/gmb/ens/papers/manuscript_thorpex_bocui.pdf
- Zhu and Toth, 2008: http://www.emc.ncep.noaa.gov/gmb/yzhu/gif/pub/AMS_Zhu_2008.pdf
- Son and et al. 2008: http://www.emc.ncep.noaa.gov/gmb/ens/papers/npg-15-1013-2008.pdf
- Cui and et al. 2010: http://www.emc.ncep.noaa.gov/gmb/yzhu/gif/pub/manscript_bocui_bias_correction_20100709.pdf
- Cui and et al. 2010 (draft for downscaling):

Background !!!

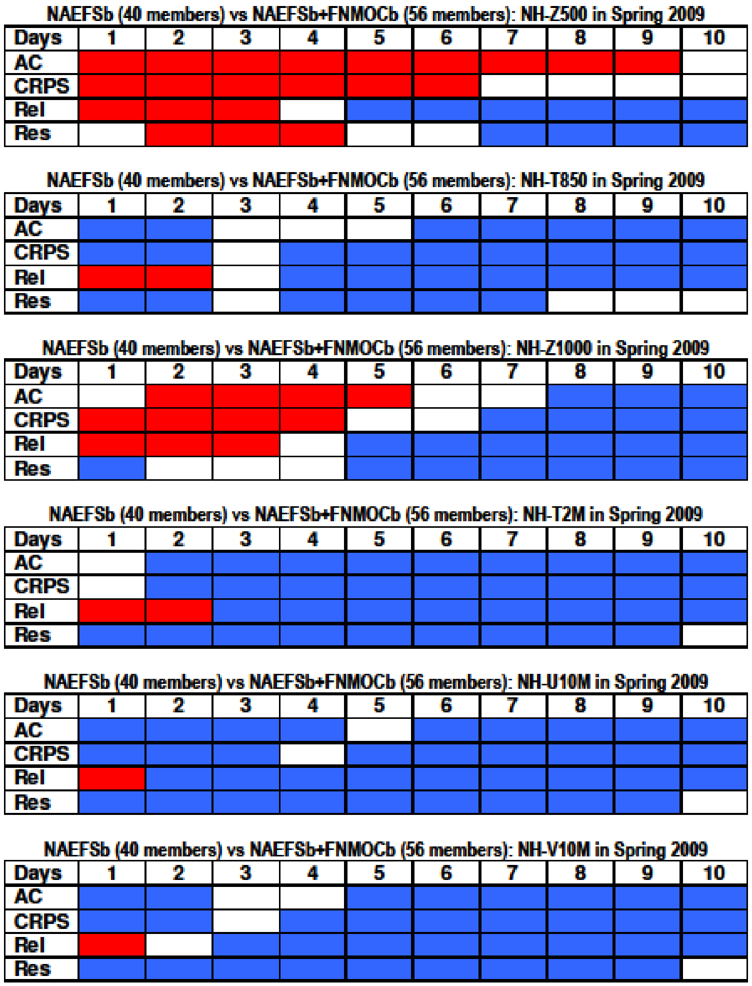
NAEFS inclusion of FNMOC ensembles

Yuejian Zhu & Bo Cui

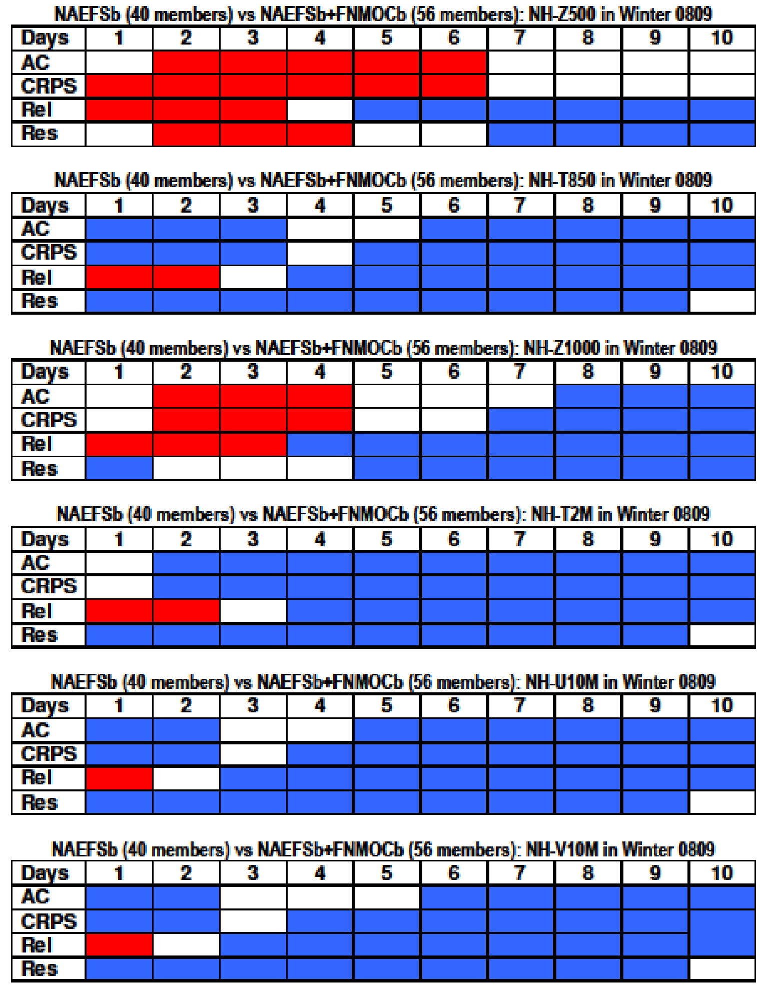
December 2010

Example of score cards for ensembles evaluation

Comparison for NAEFS with/without FNMOC ensembles



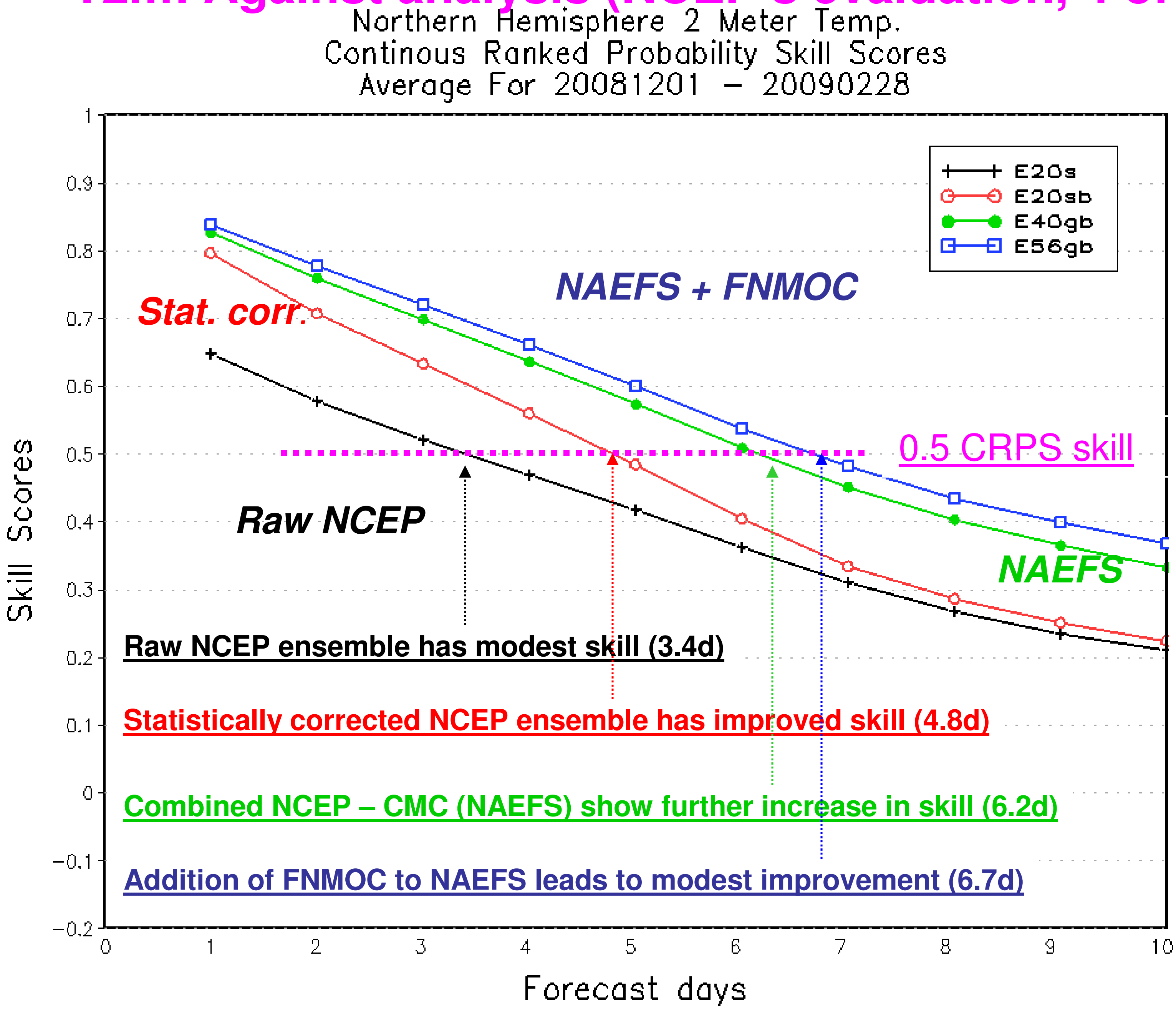
- Using 95% confidence interval (2.5%-97.5%), **BLUE** means NAEFSb+FNMOCb is significantly better than NAEFSb, **RED** means otherwise.
- The reliability (Rel) and resolution (Res) are from Brier Score decomposition.



- Using 95% confidence interval (2.5%-97.5%), **BLUE** means NAEFSb+FNMOCb is significantly better than NAEFSb, **RED** means otherwise.
- The reliability (Rel) and resolution (Res) are from Brier Score decomposition.

Blue means better to have FNMOC ensemble in NAEFS, red is not

Value-added by including FNMOC ensemble into NAEFS
T2m: Against analysis (NCEP's evaluation, 4 of 4)



Preliminary Conclusions and Plans

- **Individual ensemble systems (individual Centers' forecasts)**
  - NCEP and CMC have similar performance
  - FNMOC performance similar to NCEP & FNMOC for near surface variables, including precipitation
  - FNMOC is less skillful than NCEP and CMC for upper atmosphere variable (500hPa)

- **Combined ensemble system (without bias correction)**
  - Multi-model ensembles have higher skill than single system
  - Adding FNMOC ensemble to current NAEFS (NCEP+CMC) adds value for most forecast variables
    - Noticeable improvement for surface variables
    - Minimal improvement for upper atmosphere

- **Combined ensemble system (with operational NAEFS bias correction)**
  - Improved near surface variables with FNMOC ensemble
    - NCEPbc + CMCbc + FNMOCbc
  - Less improvement for upper atmosphere (e.g. 500hPa height))
    - Some degradation for short lead times (related to large spread in FNMOC ensemble)

- **Plan to NAEFS upgrade (NUOPC IOC Q1FY11)**
  - Based on score card for past season
  - Include the variables/parameters to current NAEFS if it adds values