Report of Inclusion of FNMOC Ensemble into NAEFS

S. Lord (NCEP/EMC)
Andre Methot (MSC)
Yuejian Zhu, and Zoltan Toth (NOAA)

Acknowledgements
Bo Cui (EMC), Stephane Beauregard (MSC), Mike Sestak (FNMOC), Rebecca Cosgrove (NCEP/NCO)

October 20, 2009
Overview

• Background & Testing Procedure
• Results
• Conclusions
• Issues
• Recommendation and outlook
Background & Testing Procedure

• North American Ensemble Forecast System (NAEFS)
  – Collaboration between NCEP, Meteorological Service of Canada (MSC), FNMOC and Mexico Weather Service

  – Elements:
    • Demonstrate value of Multi-Model Ensemble (MME)
    • Engage in collaborative software development, focused on postprocessing products from an arbitrary number of forecast systems
    • Establish operational data transfer
    • Application to operational products with shared software
    • Continue to monitor value-added with MME strategy

• Global ensemble products
  – NCEP – operational
    • 20 members -16 days
  – CMC – operational
    • 20 members - 16 days
  – FNMOC – experimental
    • 16 members – 10 days
Background & Testing Procedure (cont)

- **Forecast data**
  - 9 months of data collected (off line)
  - Communications pathway established with FNMOC
  - Raw forecasts
    - Fall 2008 (September 1\(^{st}\) – November 30\(^{th}\) 2008)
    - Winter 2008/2009 (December 1\(^{st}\) 2008 – February 28\(^{th}\) 2009)
    - Spring 2009 (March 1\(^{st}\) – May 31\(^{st}\) 2009)
  - Bias corrected forecasts – All ensembles bias corrected against NCEP analysis
    - Winter 2008/2009 (December 1\(^{st}\) 2008 – February 28\(^{th}\) 2009)
    - Spring 2009 (March 1\(^{st}\) – May 31\(^{st}\) 2009)

- **Verification methods**
  - Reference analysis
    - Individual ensembles – Each center’s own
    - Combined ensembles – NCEP analysis
  - Scores
    - NCEP standard probabilistic verification package
      - AC and RMS for ensemble mean, spread, histogram
      - CRPS, RPSS, ROC, BSS (resolution and reliability)
  - Variables
    - 500 hPa and 1000 hPa height
    - 850 hPa and 2-meter temperature
    - 10-m U and V
    - Precipitation (limited scores, CONUS only)
2 meter temperature: 120 hours forecast (ini: 2006043000)

Shaded: left – uncorrected right – after bias correction

Bias reduced approximately 50% at early lead time

RMS errors improved by 9% for d0-d3
NAEFS NDGD Probabilistic 2m Temperature Forecast Verification For 2007090100 – 2007093000

From Bias correction (NCEP, CMC)

Dual-resolution (NCEP only)

Down-scaling (NCEP, CMC)

Combination of NCEP and CMC

NAEFS final products

NCEP/GEFS raw forecast

8+ days gain
Northern Hemisphere 500hPa Height
Ensemble Mean Anomaly Correlation
Average For 20080901 – 20081130

NH 500hPa Height
Fall 2008 (AC)

FNMOC is about 12h behind CMC and NCEP

E20s – NCEP 20 members raw ensemble mean
E20m – CMC 20 members raw ensemble mean
E16f – FNMOC 16 members raw ensemble mean

Forecast days
Value-added by including FNMOC ensemble into NAEFS
T2m: Against analysis (NCEP’s evaluation, 1 of 4)

Northern Hemisphere 2 Meter Temp.
Continuous Ranked Probability Skill Scores
Average For 20081201 - 20090228

Raw NCEP ensemble has modest skill (3.4d)

0.5 CRPS skill

Raw NCEP
Value-added by including FNMOC ensemble into NAEFS
T2m: Against analysis (NCEP’s evaluation, 2 of 4)

Northern Hemisphere 2 Meter Temp.
Continuous Ranked Probability Skill Scores
Average For 20081201 - 20090228

Raw NCEP ensemble has modest skill (3.4d)
Statistically corrected NCEP ensemble has improved skill (4.8d)
Value-added by including FNMOC ensemble into NAEFS

T2m: Against analysis (NCEP’s evaluation, 3 of 4)

Northern Hemisphere 2 Meter Temp.
Continuous Ranked Probability Skill Scores
Average For 20081201 – 20090228

- Raw NCEP ensemble has modest skill (3.4d)
- Statistically corrected NCEP ensemble has improved skill (4.8d)
- Combined NCEP – CMC (NAEFS) show further increase in skill (6.2d)
Value-added by including FNMOC ensemble into NAEFS T2m: Against analysis (NCEP’s evaluation, 4 of 4)

Northern Hemisphere 2 Meter Temp.
Continuous Ranked Probability Skill Scores
Average For 20081201 – 20090228

Raw NCEP

Statistically corrected NCEP ensemble has improved skill (4.8d)

Combined NCEP – CMC (NAEFS) show further increase in skill (6.2d)

Addition of FNMOC to NAEFS leads to modest improvement (6.7d)

Raw NCEP ensemble has modest skill (3.4d)

Stat. corr.

0.5 CRPS skill

NAEFS + FNMOC

E20s

E20sb

E40gb

E56gb

Skill Scores

Forecast days

-0.2

-0.1

0

0.1

0.2

0.3

0.4

0.5

0.6

0.7

0.8

0.9

1

10
Preliminary Results from CMC (raw forecast)

Verification Against Observations

NAEFS (40 mb) vs NAEFS+FNMOC (56 mb): GZ500 in AUG08

Reliability
Resolution
Dispersion
Bias
CRPS

NAEFS (40 mb) vs NAEFS+FNMOC (56 mb): GZ500 in JAN09

Reliability
Resolution
Dispersion
Bias
CRPS

NAEFS (40 mb) vs NAEFS+FNMOC (56 mb): TT850 in AUG08

Reliability
Resolution
Dispersion
Bias
CRPS

NAEFS (40 mb) vs NAEFS+FNMOC (56 mb): TT850 in JAN09

Reliability
Resolution
Dispersion
Bias
CRPS
Preliminary Results from CMC (bias corrected forecast)

 Verification Against Observations

### NAEFSdb (40 mb) vs NAEFSdb+FNMOCdb (56 mb): GZ500 in AUG08

<table>
<thead>
<tr>
<th>Reliability</th>
<th>Resolution</th>
<th>Dispersion</th>
<th>Bias</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
</tr>
</tbody>
</table>

### NAEFSdb (40 mb) vs NAEFSdb+FNMOCdb (56 mb): TT850 in AUG08

<table>
<thead>
<tr>
<th>Reliability</th>
<th>Resolution</th>
<th>Dispersion</th>
<th>Bias</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
</tr>
</tbody>
</table>

### NAEFSdb (40 mb) vs NAEFSdb+FNMOCdb (56 mb): GZ500 in JAN09

<table>
<thead>
<tr>
<th>Reliability</th>
<th>Resolution</th>
<th>Dispersion</th>
<th>Bias</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
</tr>
</tbody>
</table>

### NAEFSdb (40 mb) vs NAEFSdb+FNMOCdb (56 mb): TT850 in JAN09

<table>
<thead>
<tr>
<th>Reliability</th>
<th>Resolution</th>
<th>Dispersion</th>
<th>Bias</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
<td>NAEFSdb (40 mb) scores better (95% c.i.)</td>
<td>NAEFSdb+FNMOCdb (56 mb) scores better (95% c.i.)</td>
</tr>
</tbody>
</table>
Preliminary Conclusions

• **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
    • Noticable 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)

• **CMC evaluation against observations**
  – Preliminary results combining raw ensembles are mixed
  – Results with bias corrected data still mixed
Issues

1. Data flow
   - FNMOC processing at NCEP must be completed by the time NAEFS processing begins
   - Currently
     • NAEFS processing begins at 0730 and 1930 Z
     • Processing of FNMOC data takes 30 minutes
     • FNMOC delivery to NCO is 0730 and 1930 Z
   - Require 30 minute overall gain for timely availability of FNMOC ensemble for NAEFS (0730 and 1930) processing
     • Processing time at NCEP can be reduced by ~10 minutes
     • Arrival at NCEP by 0710, 1910 required (if NCEP speedup is 10 minutes)
   - Data delivery needs to be accelerated by 20 minutes

2. FNMOC ensemble upgrades
   - Extend forecast from 10 days to 16, and add 4 members
   - Expand variables from 52 to 80
   - Reduce initial spread in ensemble generation
   - Receive in GRIB2 format

3. FNMOC use of MSC ensemble
   - Optional
   - May be security issues
Recommendation and Outlook

- **NCEP plans to include FNMOC ensemble in NAEFS based on**
  - Preliminary evaluations (shown here)
  - Future improvements
    - NOGAPS 4-D Var (recently implemented)
    - Ensemble system upgrade
      - Reduced initial ensemble spread for variables related to 500hPa height
    - Extended forecast from current 10d to 16d
    - 4 additional members (16 → 20)
    - Increase variables from 52 to 80
    - Upgrade exchange data format to GRIB2 for reduced data flow
  - Earlier data delivery from FNMOC
  - Final Real Time parallel evaluation (Q3FY10) with all partners (NCEP, FNMOC, MSC) for 3-months including above improvements
    - MSC reserves right to not include FNMOC data but no decision yet

- **Proposed data flow**
  - NCEP data: NCEP to FNMOC and CMC directly
  - FNMOC data: FNMOC to NCEP, then NCEP to CMC
  - CMC data: CMC to NCEP, then NCEP to FNMOC (?)

- **Anticipated implementation: Q4FY10**
  - Address new issues as they arise
Backup
Standard Probabilistic Scores

- **Continuous Ranked Probabilistic Skill Score (CRPSS)**
  - Ability of ensemble to forecast the observed (climatological) distribution of values
  - Maximum value is 1.0, >0 more skillful than climatology
- **Brier Skill Score (BSS)**
  - Ability of ensemble to predict spatial and temporal variability of observed events (e.g. T2>10 K) skillfully (relative to climatological probability)
  - BSS=1 for perfect, BSS=0 for no skill
- **Relative Operating Characteristic**
  - Ability of an ensemble membership to distinguish “hits” and “false alarms”
Bayesian Processor of Ensemble (BPE)

NWP Model Ensemble

Climatic Data → BPE → Distribution Function → Adjusted Ensemble

- extracts and fuses information
- quantifies total uncertainty
- calibrates (de-biases) ensemble

Prior d.f. $g(w)$
Posterior d.f. $\phi(w|x)$

Posterior D.F. $\Phi(w|x)$
Model ensemble
Posterior ensemble

Graphs showing precipitation distribution and quantile mapping.
Construction of Optimum Forecast Guidance from Multi-Model Ensembles

1. Multiple independent realizations
2. Historical “reforecast” data set
3. Optimal post-processing to produce “the best” forecast
4. Compact information dissemination

Potential Benefits of Using 9 Models
Lead 5 Nino34 forecast 1981-
Ensemble Spread
500hPa height
(example)

Anomaly correlation (45-day mean)

NCEP spread will be much increased after 2009 NCEP/GEFS implementation (due to introduction of stochastic scheme, higher resolution model & higher order horizontal diffusion)
500hPa Height Forecast (Ini: 2008090900; NCEP 20m)

Examples of plume

spread

500hPa Height Forecast (Ini: 2008091500; NCEP 20m)

NCEP/GEFS

FNMOC/GEFS

500hPa Height Forecast (Ini: 2008090900; FNMOC 16m)

500hPa Height Forecast (Ini: 2008091500; FNMOC 16m)

FNMOC/GEFS

Location: Washington DC (37N 77W)
Precipitation

Ensemble Precipitation Verification for CONUS
ETS and TSS for threshold >= 0.20mm/24hours
Average For 20090419 – 20090520

AVERAGE BOXES: TOTAL=979.13 OBS=382.03 BIAS(NCEP_01)=1.34
Solid=NCEP,Dash=CMC

ETS
TSS

Skill Scores (0-1)
Forecast days

Ensemble Precipitation Verification for CONUS
RMSE, ABSE, SPREAD and CRPS
Average For 20090419 – 20090520

1. NCEP and CMC
2. NCEP and FNMOC

Individual ensembles

Raw Fcst

Forecast Errors (mm)
Forecast days
### NEXT NAEFS exchange pgrba files

<table>
<thead>
<tr>
<th>Variables</th>
<th>pgrba file</th>
<th>Total 80 (28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHT</td>
<td><strong>Surface</strong>, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>11 (3)</td>
</tr>
<tr>
<td>TMP</td>
<td>2m, 2mMax, 2mMin, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>13 (3)</td>
</tr>
<tr>
<td>RH</td>
<td>2m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>11 (3)</td>
</tr>
<tr>
<td>UGRD</td>
<td>10m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>11 (3)</td>
</tr>
<tr>
<td>VGRD</td>
<td>10m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>11 (3)</td>
</tr>
<tr>
<td>VVEL</td>
<td>850hPa</td>
<td>1 (1)</td>
</tr>
<tr>
<td>PRES</td>
<td>Surface, PRMSL</td>
<td>2 (0)</td>
</tr>
<tr>
<td>PRCP (types)</td>
<td>APCP, CRAIN, CSNOW, CFRZR, CICEP</td>
<td>5 (0)</td>
</tr>
<tr>
<td>FLUX (surface)</td>
<td>LHTFL, SHTFL, DSWRF, DLWRF, USWRF, ULWRF</td>
<td>6 (6)</td>
</tr>
<tr>
<td>FLUX (top)</td>
<td>ULWRF (OLR)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>PWAT</td>
<td>Total precipitable water at atmospheric column</td>
<td>1 (0)</td>
</tr>
<tr>
<td>TCDC</td>
<td>Total cloud cover at atmospheric column</td>
<td>1 (0)</td>
</tr>
<tr>
<td>CAPE and CIN</td>
<td>Convective available potential energy, Convective Inhibition</td>
<td>2 (1)</td>
</tr>
<tr>
<td>SOIL</td>
<td>SOILW(0-10cm), WEASD(water equiv. of accum. snow depth), SNOD(surface), TMP(0-10cm down)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Notes</td>
<td>Surface GHT is only in analysis file and first pgrb file when the resolution changed. 25 of 28 new variables are from pgrbb files, 10, 50hPa RH and SNOD are new variables</td>
<td>28 new vars</td>
</tr>
</tbody>
</table>
**NEXT NAEFS pgrba_bc files**  
*(bias correction)*

<table>
<thead>
<tr>
<th>Variables</th>
<th>pgrba_bc file</th>
<th>Total 49 (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHT</td>
<td>10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>10 (3)</td>
</tr>
<tr>
<td>TMP</td>
<td>2m, 2mMax, 2mMin, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>13 (3)</td>
</tr>
<tr>
<td>UGRD</td>
<td>10m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>11 (3)</td>
</tr>
<tr>
<td>VGRD</td>
<td>10m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa</td>
<td>11 (3)</td>
</tr>
<tr>
<td>VVEL</td>
<td>850hPa</td>
<td>1(1)</td>
</tr>
<tr>
<td>PRES</td>
<td>Surface, PRMSL</td>
<td>2(0)</td>
</tr>
<tr>
<td>FLUX (top)</td>
<td>ULWRF (toa - OLR)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Notes</td>
<td></td>
<td>14 new vars</td>
</tr>
</tbody>
</table>
Data Flow

– NCEP receives 00 and 12Z cycle data
– Data path from FNMOC to the NWS/TOC then to the NCEP/CCS
– April 2009 requirements study
  • NCO, TOC, FNMOC examined data delivery
  • Offline delivery time (for evaluation) is 11Z and 23Z
  • For operations, NCO requires data here and packaged appropriately by 730Z (1930 for the 12Z cycle) to meet the current start time of the NAEFS processing
– NCO currently receives FNMOC ensemble data 720 to 740Z for the 00Z (1930 to 2000Z for the 12Z)
– Processing takes 30 minutes
– Delivery by 0710, 1910 required (if NCEP speedup is 10 minutes)