EMC Implementation Briefing of SREF.v7.0 (Q4FY15)

Jun Du, Geoff DiMego, Dusan Javic, Brad Ferrier, Bo Yang

(NCWCP, September 25, 2015)
Changes

1. 3 model core system becomes 2 model core system (NMMB, WRF_ARW, WRF_NMM)

2. Vertical resolution is increased from 35 to 40 layers (horizontal resolution remains the same of 16km)

3. Ensemble membership is increased from 21 to 26 members

4. IC diversity is enhanced: (a) mix use of multi analyses (NDAS, GFS and RAP) for each model core, and (b) blending of GEFS and SREF IC perturbations for all members

5. Physics diversity is enhanced: (a) more variety of physics schemes, and (b) stochastic flavor in physics parameters (GWD and soil moisture)

6. Others: name change from em to ARW; individual member ID in pgrb files; addition of 138 new stations in bufr/sounding output, unit change of ceiling height (from AGL to ASL) etc.
Vertical level increase

• As for HiResWindow displayed here (35-40), adding vertical levels to SREF members will improve models’ ability to distinguish ceiling heights at boundaries between critical flight categories
# 13 NMMB members

<table>
<thead>
<tr>
<th>Mod-Mem</th>
<th>IC</th>
<th>IC pert</th>
<th>LBCs</th>
<th>Physics 1</th>
<th>Physics 2</th>
<th>GWD</th>
<th>Land Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conv</td>
<td>PBL</td>
<td>Sfc layer</td>
<td>Microphys</td>
</tr>
<tr>
<td>nmbb_ctl</td>
<td>NDAS</td>
<td>Blending (GEFS + SREF)</td>
<td>GFS</td>
<td>BMJ old shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Fer_hires</td>
</tr>
<tr>
<td>nmbb_n1</td>
<td>GEFS 2</td>
<td>SAS</td>
<td>GFS</td>
<td>MYJ</td>
<td>WSM6</td>
<td>GFDL</td>
<td>cleffamp=0.5</td>
</tr>
<tr>
<td>nmbb_p1</td>
<td>GEFS 1</td>
<td>BMJ new shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Fer_hires</td>
<td>RRTM</td>
<td>cleffamp=2</td>
</tr>
<tr>
<td>nmbb_n2</td>
<td>GEFS 4</td>
<td>SAS</td>
<td>GFS</td>
<td>MYJ</td>
<td>WSM6</td>
<td>GFDL</td>
<td>cleffamp=1</td>
</tr>
<tr>
<td>nmbb_p2</td>
<td>GEFS 3</td>
<td>BMJ old shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>WSM6</td>
<td>RRTM</td>
<td>cleffamp=0.5</td>
</tr>
<tr>
<td>nmbb_n3</td>
<td>GFS</td>
<td>Blending (GEFS + SREF)</td>
<td>GEFS 6</td>
<td>SAS</td>
<td>GFS</td>
<td>MYJ</td>
<td>WSM6</td>
</tr>
<tr>
<td>nmbb_p3</td>
<td>GEFS 5</td>
<td>BMJ new shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>WSM6</td>
<td>RRTM</td>
<td>cleffamp=1</td>
</tr>
<tr>
<td>nmbb_n4</td>
<td>GEFS 8</td>
<td>SAS</td>
<td>GFS</td>
<td>MYJ</td>
<td>WSM6</td>
<td>RRTM</td>
<td>cleffamp=0.5</td>
</tr>
<tr>
<td>nmbb_p4</td>
<td>GEFS 7</td>
<td>BMJ old shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Fer_hires</td>
<td>GFDL</td>
<td>cleffamp=2</td>
</tr>
<tr>
<td>nmbb_n5</td>
<td>RAP</td>
<td>Blending (GEFS + SREF)</td>
<td>GEFS 10</td>
<td>SAS</td>
<td>GFS</td>
<td>MYJ</td>
<td>WSM6</td>
</tr>
<tr>
<td>nmbb_p5</td>
<td>GEFS 9</td>
<td>BMJ new shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Fer_hires</td>
<td>RRTM</td>
<td>cleffamp=0.5</td>
</tr>
<tr>
<td>nmbb_n6</td>
<td>GEFS 12</td>
<td>SAS</td>
<td>GFS</td>
<td>MYJ</td>
<td>Fer_hires</td>
<td>GFDL</td>
<td>cleffamp=2</td>
</tr>
<tr>
<td>nmbb_p6</td>
<td>GEFS 11</td>
<td>BMJ old shal</td>
<td>MYJ</td>
<td>MYJ</td>
<td>WSM6</td>
<td>GFDL</td>
<td>cleffamp=1</td>
</tr>
</tbody>
</table>
## 13 ARW members

<table>
<thead>
<tr>
<th>Mod-Mem</th>
<th>IC</th>
<th>IC pert</th>
<th>LBCs</th>
<th>Physics 1</th>
<th>Physics 2</th>
<th>Land surface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conv</td>
<td>PBL</td>
<td>Sfc layer</td>
</tr>
<tr>
<td>arw_ctl</td>
<td>RAP</td>
<td>Blending (GEFS + SREF)</td>
<td>GFS</td>
<td>KF</td>
<td>YSU</td>
<td>MM5</td>
</tr>
<tr>
<td>arw_n1</td>
<td>GEFS 14</td>
<td>BMJ</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Fer</td>
<td>GFDL</td>
</tr>
<tr>
<td>arw_p1</td>
<td>GEFS 13</td>
<td>Grell</td>
<td>MYNN</td>
<td>MYNN</td>
<td>Thompson</td>
<td>Old RRTM</td>
</tr>
<tr>
<td>arw_n2</td>
<td>GEFS 16</td>
<td>KF</td>
<td>YSU</td>
<td>MM5</td>
<td>Fer</td>
<td>GFDL</td>
</tr>
<tr>
<td>arw_p2</td>
<td>GEFS 15</td>
<td>BMJ</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Thompson</td>
<td>RRTMG</td>
</tr>
<tr>
<td>arw_n3</td>
<td>GFS</td>
<td>Blending (GEFS + SREF)</td>
<td>GEFS 18</td>
<td>Grell</td>
<td>MYNN</td>
<td>MYNN</td>
</tr>
<tr>
<td>arw_p3</td>
<td>GEFS 17</td>
<td>KF</td>
<td>YSU</td>
<td>MM5</td>
<td>Thompson</td>
<td>Old RRTM</td>
</tr>
<tr>
<td>arw_n4</td>
<td>GEFS 20</td>
<td>BMJ</td>
<td>MYJ</td>
<td>MYJ</td>
<td>WSM6</td>
<td>RRTMG</td>
</tr>
<tr>
<td>arw_p4</td>
<td>GEFS 19</td>
<td>KF</td>
<td>YSU</td>
<td>MM5</td>
<td>Fer</td>
<td>GFDL</td>
</tr>
<tr>
<td>arw_n5</td>
<td>NDAS</td>
<td>Blending (GEFS + SREF)</td>
<td>GEFS 2</td>
<td>Grell</td>
<td>MYNN</td>
<td>MYNN</td>
</tr>
<tr>
<td>arw_p5</td>
<td>GEFS 1</td>
<td>KF</td>
<td>YSU</td>
<td>MM5</td>
<td>WSM6</td>
<td>RRTMG</td>
</tr>
<tr>
<td>arw_n6</td>
<td>GEFS 4</td>
<td>BMJ</td>
<td>MYJ</td>
<td>MYJ</td>
<td>Thompson</td>
<td>Old RRTM</td>
</tr>
<tr>
<td>arw_p6</td>
<td>GEFS 3</td>
<td>Grell</td>
<td>MYNN</td>
<td>MYNN</td>
<td>Thompson</td>
<td>RRTMG</td>
</tr>
</tbody>
</table>
Expected Benefits

1. Reduce **cold bias** in surface temperature (2m T)

2. Reduce **wet bias** in surface moisture field (2m RH, not precipitation)

3. Increase ensemble **spread** (diversity) especially for ARW members

4. Improve overall **skill of probabilistic** forecasts in general

5. Improve **visibility and cloud ceiling etc.** due to increased vertical resolution (verified by AWC)
Improvement in forecast performance
Reduced cold bias in 2mT to a lesser degree
Reduced cold bias in 2mT to a lesser degree (warm season: April – Sept. 2015)
Reduced wet bias in 2mRH to a lesser degree (cold season: Oct. 2014 – March 2015)

Talagrand Distribution
Almost eliminate wet bias in 2mRH (warm season: April – Sept. 2015)
More skillful probabilistic forecasts (measured by RPSS) of 2mT and 2mRH cold season (Oct. 2014 – Mar. 2015) and warm season (Apr. – Sept. 2015).
More reliable probabilistic forecasts (measured by Reliability) of 2mT and 2mRH.

2m T
Reliability 2mT hr75

- cold season (Oct. 2014 – Mar. 2015)
- warm season (Apr. – Sept. 2015)

2m RH
Reliability 2mRH hr75

- cold season (Oct. 2014 – Mar. 2015)
- warm season (Apr. – Sept. 2015)
Improvements in various other measures
(SREF vs. SREFx, cold season, Binbin’s g2g)

- **ROC**
  - FAR (forecast time = ALL hr)

- **CRPSS**
  - Forecast hours

- **INFO**
  - Forecast hours

- **ECON-V**
  - Cost/Loss (forecast time = ALL hr)
Improvements in various other measures
(SREF vs. SREFx, warm season, Binbin’s g2g)

- ROC
- CRPSS
- INFO
- ECON-V

warm season
(May – Sept. 2015)
Precipitation forecasts of ensemble mean in cold season: similar in position and improvement in amount
Precipitation forecasts of probabilistic information in cold season: an improvement in Brier Skill Score
Precipitation forecasts of ensemble mean in warm season: similar in both position and amount
Precipitation forecasts of probabilistic information in warm season: similar in Brier Skill Score
Equal-likelihood of ensemble members (24h-apcp forecasts)

Cold season (Dec. 2014 - Mar. 2015)

Warm season (Apr. – July 2015)
Improvement in ensemble spread
Increased IC-perturbation size and more mixed members in forecast projection (Texas flooding case)

- Little initial IC spread of ARW members
- Larger initial IC spread of ARW members
- Grouping
- Less grouping
Forecaster’s complain: all ARW members are grouped to be too wet: March 5, 2015 DC snow storm
Spread closer to ensemble mean forecast error and less outlier in forecasts in the new SREF than the old SREF (cold season: Oct. 2014 – March 2015)
Spread closer to ensemble mean forecast error and less outlier in forecasts in the new SREF than the old SREF (warm season: April – July 2015)
Too high local peak value of 2m Td from a couple of ARW members (p1, n3, n5 and p6 in particular)
2-m Td Plume for DSM (Des Moines, Iowa): 20150711/15Z (Israel Jirak of SPC)
© (>=90°F)

Para SREFx: F81h of Td_2m
(init: 15z July 11, 2015,
verif at 00z July 15, 2015)
It’s only in 2m diagnostic Td but not in atmospheric lower level Td (Andy Dean of SPC)
Slightly reduced total error in 2m-Td (warm season: April – Sept. 2015)

No improvement in bias magnitude

But reduced bias frequency
Improved spread: better spread-error relation and less chance to miss truth
(warm season: April – Sept. 2015)

5% outlier reduction
Probabilistic forecasts: more skillful, slightly more reliable in low prob end and less reliable in higher prob (but overall reliability is good) (warm season: April – Sept. 2015)
Tests of three methods to fix local too high 2m-Td values of a few ARW members (particularly the 4 using MYNN scheme)

(1) Simple cap of 82F

(2) Using model’s lowest layer q to calculate 2m-Td over the entire model domain (sophisticated but degraded the overall performance by being too dry)

(3) Using model-lowest layer q to calculate 2m-Td only over the area where Td > 82F (kept the overall performance un-degraded but destroyed the spatial structure of high-impact area)

It turns out that the method 1 is the best method to keep both the overall performance not being degraded and spatial structure not being distorted.
Method 2: severely degraded overall performance of 2m-Td by dramatically increasing dry bias

Increased RMSE and dry bias over all forecast hours from 00 to 87h

RMSE and BIAS of 2m Td based on ARW-ensemble mean (09z, 07/26/2015)

RMSE and BIAS of 2m Td based on ARW-ensemble mean (09z, 07/28/2015)
Reason

• First layer in the SREF is 7mb deep (1.000 – 0.9930) which is roughly 52m thick and its mid-point would be at 26m.
• 26m level is usually drier than 2m level

So, how about apply this method only over the area where 2m-Td >82F and leave other areas unchanged (method 3)?
Method 3: inadvertently changed the maximum to local minimum

2m-Td

MUCAPE
SPC comment

• After looking at this case, it seems that applying the LML specific humidity value where the 2-m Td exceeds the cap has too strong of a drying effect. While it may not be an elegant solution, it might be better and more physically consistent to just use a constant cap value of 82F (i.e., to keep a maximum plateau rather than introduce a relative minimum).

Method 1 is, therefore, implemented
The impact is limited

• The cap applies only to the dew point at the 2-m level and that it's purely a diagnostic value.
• The primary products that were affected by the problem are the SPC SREF plumes and MUCAPE with a dew members.
• GSD has a new way to calculate 2m-Td to deal with this issue in newer version of ARW which will be used for the next SREF upgrade.
Summary note about the 2m Td cap

• A few ARW members especially p01, n03, n05 and p06 occasionally produce too high dew-point temperature value at 2m level (2m-Td) such as greater than 90F at local locations where 2m-Td is generally expected to be high (greater than 80F). Therefore, a cap of 28C or 82.4F has been added. Based on the tests done by both EMC and SPC, this fix is the best one among the three approaches being tested at this time. It preserves good domain-averaged performance (i.e. no impact on the overall performance of 2m Td) and correct spatial structure of individual cases besides taking care of these occasional high value spikes.

• Note that 2m Td is strictly an alone diagnostic parameter done at model post and does not impact any other forecast parameters. All model-produced variables are good and nothing to do with this issue. In other words, the issue is only pertinent to this particular variable 2m Td itself. Impact on the ensemble products of 2m Td such as its mean and probability is expected to be minimal since majority of 26 members have no this issue. No impact is expected on the overall performance of this field. The newest version of ARW model has a new way to calculate 2m Td to take care of this issue, which will be used in the next SREF upgrade.
### Activities of involving users and managers during the development

<table>
<thead>
<tr>
<th>Step leading to implementation</th>
<th>SREF.v7.0</th>
</tr>
</thead>
</table>
| Brief staff, team and/or management via **EMC Branch** and/or quarterly science briefings | 07/10/2014 (WC OSS Science Quarterly)  
10/07/2014 (WC OSS Science Quarterly)  
11/18/2014 (NAEFS conference)  
12/03/2014 (NCEP Production Suite Review)  
04/20/2015 (EMC-WPC meeting about winter weather exp)  
06/29/2015 (NWP/WAF conference)  
07/16/2015 (Model Evaluation Group, MEG, meeting)  
08/03/2015 (EMC-SPC meeting about 2m Td)  
10/05/2015 (EMC-WPC meeting reviewing winter storm cases) |
| 1-year parallel data to forecasters to use (such as WPC winter weather experiment in Jan.-Feb. 2015) | Oct. 2014 – April 2015 (by EMC)  
April 2015 – Sept. 2015 (by NCO) |
| Specifically requested 2-month retrospective run for SPC | April 25 – June 30, 2014 |
| Hold initial coordination discussions with NCO (aka EE or kickoff meeting) | 12/3/2014 |
| Draft/Issue/Amend/Final-issuance of Technical Information Notice (TIN) | 5/20/2015D  
TIN 15-32  
6/25/2015 I  
7/07/2015 A  
8/20/2015A |
| Hold Change Control Board briefing for EMC & NCO prior to code delivery | 3/23/2015 |
| Code frozen - begin 30-day pre-implementation test | 6/22/2015 -- 7/24/2015  
8/17/2015 – 9/15/2015 (restart 30 day for SPC due to 2m Td) |
| Brief NCEP Director to obtain authorization to implement | 9/25/2015 (primary for overall) and 10/13/2015 (supplementary for winter storms) |
| NCO implements into NCEP Production Suite | 10/20/2015 |
Summary

Highlight: Unified the models by eliminating NMM: one step closer to NCEP strategic unified modeling system, increased ensemble membership from 21 to 26

1. Reduced cold bias in surface temperature (2m T)
2. Reduced wet bias in surface moisture field (2m RH, not precipitation)
3. Increased ensemble spread (diversity) especially for ARW members
4. Precipitation: improved in winter and similar in summer due probably to the removal of NMM model (WPC concern)
5. Improved overall skill of probabilistic forecasts for most variables.
6. Improved visibility and cloud ceiling etc. aviation products due to the increased vertical resolution (verified by AWC)

Future: To improve precipitation forecasts by adding probability-matched mean
To implement the new way of calculating 2m Td in ARW model (GSD)
To add reliability score in precipitation verification
Supplementary material for

EMC Implementation Briefing of SREF.v7.0 (Q1FY16):

Performance in winter storms

(NCWCP, October 13, 2015)
## High-impact winter storms investigated (Mar. 2014 – Feb. 2015)

<table>
<thead>
<tr>
<th>Case</th>
<th>Model cycle</th>
<th>Verification time</th>
<th>Parameter</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Northeast Blizzard</td>
<td>09z, 01/25/15</td>
<td>00z, 01/28/15</td>
<td>24h-apcp</td>
<td>Improvement</td>
</tr>
<tr>
<td>2. Mid-west holiday storm</td>
<td>09z, 12/22/14</td>
<td>00z, 12/25/14</td>
<td>SLP/cyclone position and 12h-apcp</td>
<td>Improvement</td>
</tr>
<tr>
<td>3. False alarm mid-Atlantic clipper</td>
<td>09z, 12/28/14</td>
<td>00z, 12/31/14</td>
<td>24h-apcp</td>
<td>Improvement</td>
</tr>
<tr>
<td>4. Mid-Atlantic heavy rain</td>
<td>09z, 02/24/15</td>
<td>00z, 02/27/15</td>
<td>24h-apcp</td>
<td>Improvement</td>
</tr>
<tr>
<td>5. Lake effect snow</td>
<td>09z, 11/26/14</td>
<td>00z, 11/29/14</td>
<td>24h-apcp</td>
<td>Both excellent, SREF even better in max amount</td>
</tr>
<tr>
<td>6. East Cost snow</td>
<td>09z, 03/01/14</td>
<td>12z, 03/03/14</td>
<td>T2m and 24h-apcp</td>
<td>No improve in T2m and improve in precipitation</td>
</tr>
<tr>
<td>7. Midwest/Western Great Lakes snow</td>
<td>09z, 03/16/14</td>
<td>00z, 03/19/14</td>
<td>24h-apcp</td>
<td>Improvement</td>
</tr>
<tr>
<td>8. South New England snow</td>
<td>09z, 02/07/15</td>
<td>00z, 02/10/15</td>
<td>24h-apcp</td>
<td>Mixed</td>
</tr>
<tr>
<td>9. Midwest and western Great Lakes snow</td>
<td>21z, 11/08/14</td>
<td>00z, 11/11/14</td>
<td>24h-apcp</td>
<td>Both excellent, SREFP better in north boundary</td>
</tr>
</tbody>
</table>
Case 1: Northeast blizzard, 24h-apcp over 00z 1/27 – 00z 1/28, 2015: SREF extended heavy precip too much to the southwest, while SREFP corrected most of this error.
Case 2: Mid-west holiday storm, SLP/cyclone position at 00z, 12/25, 2015: SREFP has much smaller error than SREF in cyclone position.
Case 2: Mid-west holiday storm, 24h-apcp over 00z 12/24 - 00z 12/25, 2015: as a result of better cyclone position, a large false alarm area of heavy precipitation southwest of Chicago is correctly removed in SREFP.
Case 3: False alarm mid-Atlantic clipper, 24h-apcp over 00z 12/30 - 00z 12/31, 2014: SREFP has smaller error in precip forecast

DC: <0.01”

DC: 0.25-0.5”

DC: 0.1-0.25”
Case 4: Mid-Atlantic heavy precipitation, 24h-apcp over 00z 2/26 - 00z 2/27, 2015: SREFP has smaller error and better heavy rain band structure.
Case 5: Lake effect snow events, 24h-apcp over 00z 11/18 - 00z 11/19, 2014: both SREF and SREFP did an excellent job but SREF is even better in max amount.

COM_US 24h-apcp (in) 63H fcst from 09Z 16 nov 2014 verified time: 00z, 11/19/2014

COM_US 24h-apcp (in) 63H fcst from 09Z 16 nov 2014 verified time: 00z, 11/19/2014
Case 5 (zooming to storm#5): Lake effect snow, 24h-apcp over 00z 11/18 - 00z 11/19, 2014: both SREF and SREFP did an excellent job but SREF is even better in max amount.
Case 6: East coast wintery precipitation, 2m T at 12z 3/3, 2014: no improvement in warm and dry bias
Case 6: East coast wintery precipitation, 24h-apcp over 00z 3/3 - 00z 3/4, 2014: but SREFP has more accurate heavy precipitation band location.
Case 7: Midwest/Western Great Lakes snow event, 24h-apcp over 00z 03/18 - 00z 03/19, 2014: SREFP has slightly better south storm track
Case 8: South New England snow, 24h-apcp over 00z 2/9 – 00z 2/10, 2015: too large area of heavy precip has been partially corrected, improvement in the north edge but degraded in the south edge, a mixed result.
Case 9: similar and both excellent jobs maybe better in defining the north boundary for SREFP
Summary

In overall, SREFP noticeably outperformed SREF for major high-impact winter storms, which should greatly help WPC and WFOs in daily winter storm prediction.
Backup
South Carolina historical flooding events: 24h-apcp of 21z 10/1/15 run
SREFP improved in position and structure of heavy rain area