

HRRR SUMMER 2014 EVALUATION

Steve Weygandt, Curtis Alexander, Stan Benjamin

NOAA ESRL GSD for FAA AWRP MDE

1. Introduction

This is an assessment of the HRRR performance during the 2014 summer real-time evaluation. For this evaluation, the GSD real-time experimental HRRR and associated parent RAP model were frozen from April 2014 through the end of October 2014. The evaluation will focus primarily on the HRRR skill during the core of the summer convection season (June 1 – August 31, 2014). A summary of the changes made for these RAP and HRRR versions is provided in the appendix. On Sept. 30, 2014, a version of the HRRR that includes many, but not all, of the features in the GSD real-time experimental HRRR, was implemented at NCEP as an official NOAA operation model. The upgrades made in the 2014 GSD real-time HRRR, along with additional enhancements from retrospective and parallel testing, will form the basis of the code transferred to NCEP for the next RAP / HRRR. This next NCEP implementation is schedule for 2015.

2. HRRR reliability for 2014

Matched HRRR configurations were run on two different supercomputers (JET and ZEUS) for portions of the 2014 season, however, the cycle on ZEUS was also used for evaluation of further enhancements to HRRR. Overall, run reliability for the HRRR was quite good as indicated in Fig. 1

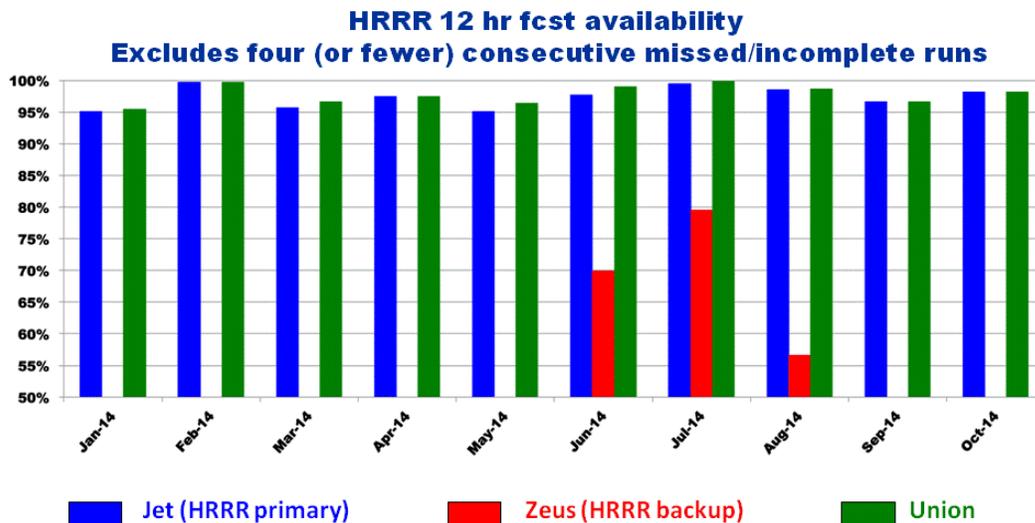


Fig. 1 HRRR reliability by month for 2014, showing JET reliability (blue), ZEUS reliability (red) and the combined reliability (green). Note, ZEUS reliability (red) is only shown for the 3 mos. – June, July, mad August.

3. HRRR forecast skill for 2014

Overall HRRR reflectivity forecast performance for 2014 shows a small improvement over 2013. In particular, similar CSI scores are obtained (slightly better for 2014 compared to 2013 for longer lead times) with a significant reduction in the high bias to values closer to 1. This can be seen in Figs. 2 and 3, which shows an un-matched, but long term (3 month) comparison of CSI and bias for the 2014 HRRR (red curve) vs. the 2013 HRRR (blue curve). Note that while this is un-matched comparison (same dates from different years), inclusion of corresponding dates from different years and a long verification period (> 90 days) makes this a more credible comparison. As shown in Fig. 2, CSI scores are quite similar between the two years; however, there is a slight increase in CSI for longer lead times for 2014. Most importantly however, is that there is a significant reduction in the high bias from 2013 to 2014. In 2013, bias scores peaked at > 1.5 for the 3-5 hour forecast range and stayed near or above 1.4 throughout the 15 hour forecast duration. In contrast, 2014 HRRR bias scores peak at less than 1.4 (also at the 3-5 forecast range) and settle at a value near 1.1 after about 8 hours. In day to day forecast plots, this reduction in bias translates into fewer instances of false alarm storms and less coverage of spurious convection. Both of these are desirable traits for use of the HRRR for aviation guidance.

Fig. 4 shows HRRR forecasts for a day with a complex pattern of active convection across the eastern U.S., July 14, 2014, illustrating general aspects of the HRRR forecast skill. Shown in the Fig. are HRRR 9-h and 3-h forecasts valid 21z. Examination of the details of the HRRR forecast and the radar verification reveals a number of aspects. The 9-h forecast (top panel) does a good job of capturing many general aspects of the convection, including 1) the scattered convective coverage across the mid-Atlantic region into the Northeast, 2) the existence of two lines of convective focus further south and west, one stretching from southern Michigan to southern Missouri and the second stretching from West Virginia through Arkansas, and 3) the general convective trend along these focus lines (more solid band of storms with the northern line and a more broken sequence of convective clusters along the southern line). Despite this overall skill, there are errors with specific details of the HRRR 9-h forecast, including 1) too much storm coverage from central Pennsylvania into Eastern West Virginia, 2) too continuous with the storm line from western Tennessee into Arkansas, and 3) too broad with the storm line from Indiana into Illinois. In the 3-h HRRR forecast (center panel), many of the smaller-scale convective details are forecast more accurately, including 1) a more solid, narrow convective line from Indiana into Illinois, 2) a series of 3 convective clusters along the southern line (southwest WV eastern TN, western TN, and central Arkansas, and 3) better depiction of the convection-free area and especially the gaps in the convective lines. Around the New York City area, there is some displacement in the location of HRRR convective clusters, but is accurate overall indication of convection in the vicinity, especially to the west. Overall, the example illustrates the ability of the HRRR to capture the overall pattern of convection and a fair number of convective details with a high degree of accuracy.

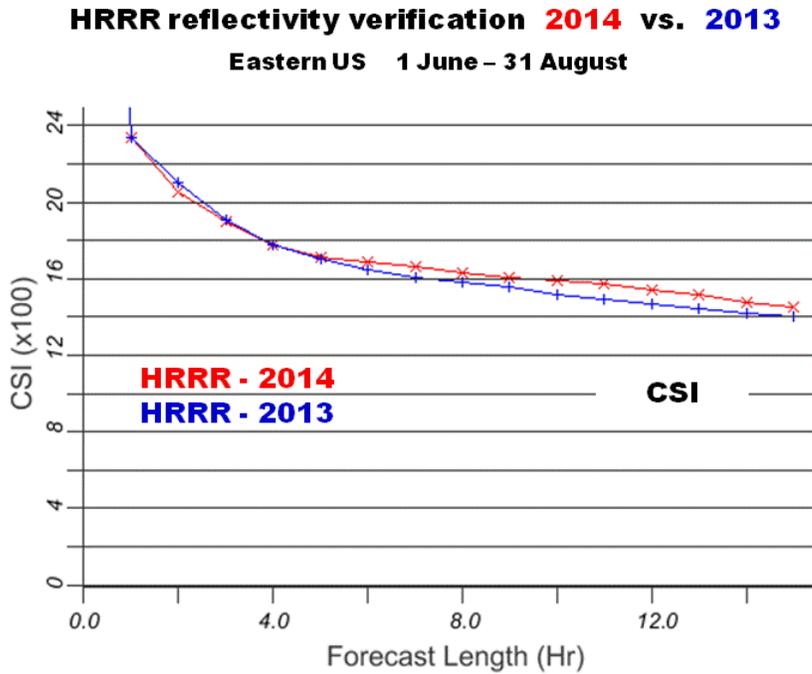


Fig 2. Comparison of HRRR warm season (JJA) forecast reflectivity CSI scores (25 dBZ, 13-km comparison grid) for 2014 vs. 2013

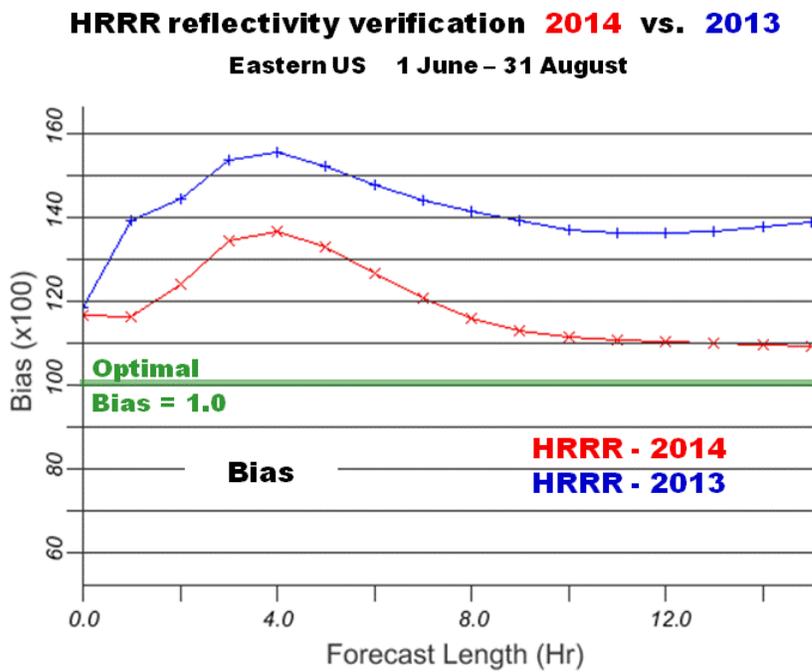


Fig 3. Comparison of HRRR warm season (JJA) forecast reflectivity bias scores (25 dBZ, 13-km comparison grid) for 2014 vs. 2013

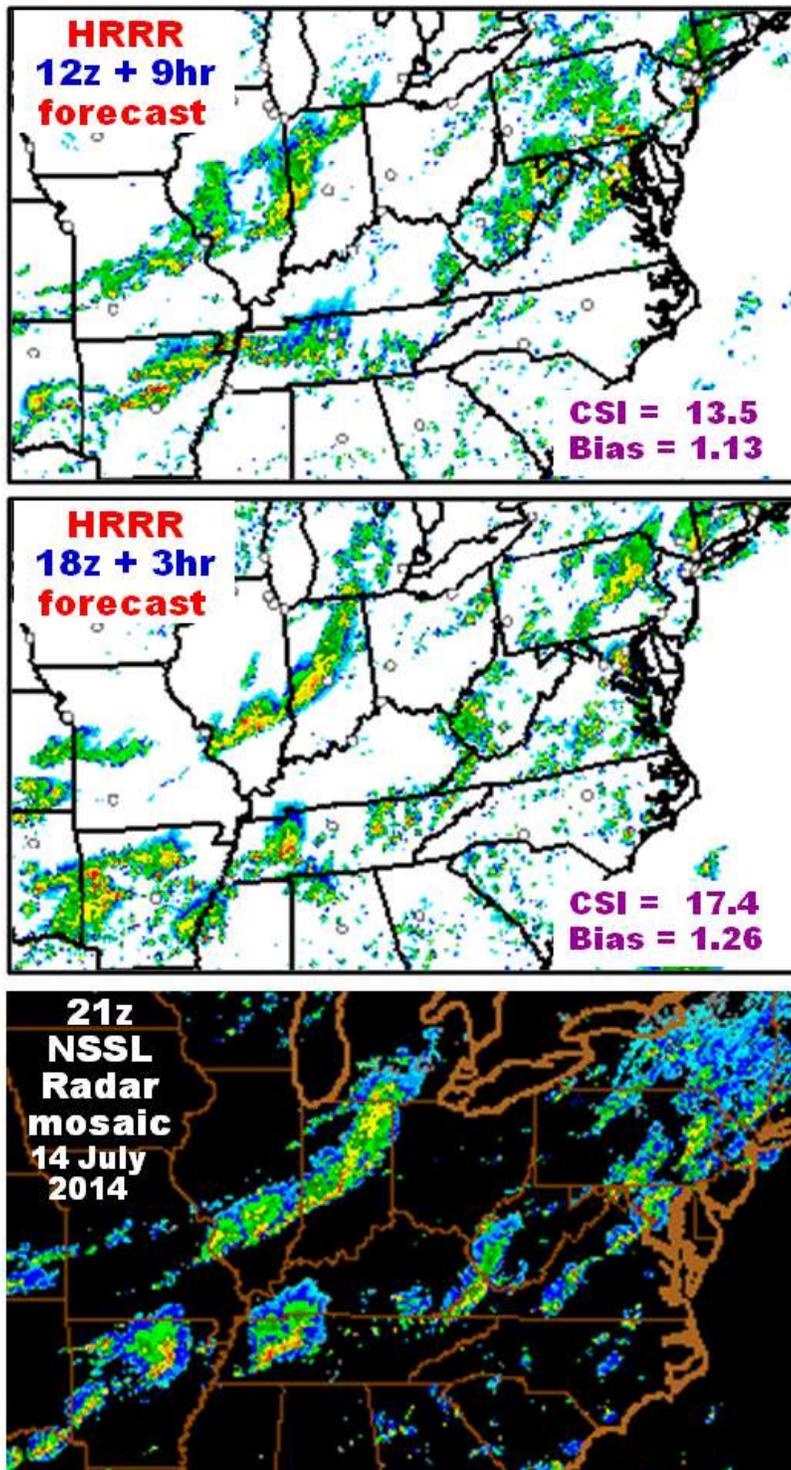


Fig. 4. Radar observed (bottom) and 9-h HRRR forecast reflectivity (top) and 3-h HRRR forecast reflectivity (middle) for a complex pattern of convection over the eastern U.S. at 21z July 14, 2014. Comparison of HRRR 3-h and 9-h forecasts with the radar observations reveals good HRRR depiction of overall storm mode and structure, with increased accuracy of small-scale details in the 3-h forecast. CSI, bias scores for 25 dBZ up-scaled to a 13-km grid.

Examination of many HRRR forecast plots and comparisons of skill scores indicates the forecast accuracy exhibited in Fig. 4 is fairly typically of the overall HRRR forecast performance.

Fig. 5 show another sample HRRR forecast example, a sequence of forecasts initialized at 16z 11 May, 2014.

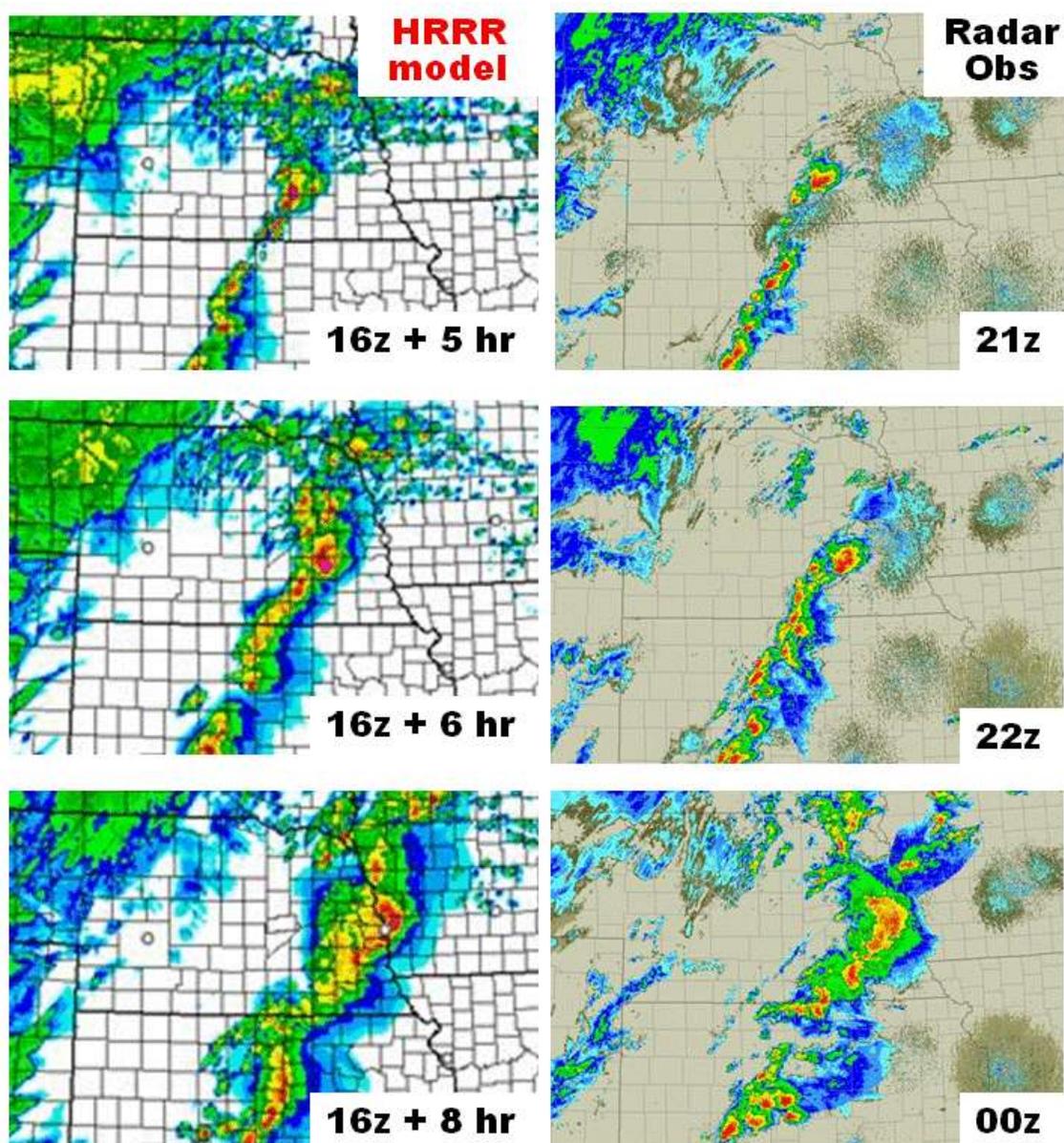


Fig. 5. Radar observed (left) and real-time HRRR forecast reflectivity (right) for broken line of storms on May 11, 2014. Comparison of HRRR 5, 6, and 8 hour forecasts with the radar observations reveals good HRRR depiction of overall storm mode and structure, as well weaknesses and gaps in the line, and even location of individual cells.

As can be seen in the 5, 6, and 8-h from the 16z initialized HRRR, the overall character and evolution of the convective system has well captured by the HRRR. In particular, the dominant supercell over Nebraska at 21z is well depicted, as is its evolution into a bow echo by 00z, with a weakness/gap in the line in north central Kansas.

Evaluation of surface fields from the HRRR has revealed a warm, dry bias in the 2014. This warm, dry bias is also evident in the RAP and is maximized in the warm season, during the afternoon through evening. Testing and evaluation of improvements within the boundary layer scheme, land surface model, and shallow cumulus scheme in the RAP and HRRR has yielded improvement in a parallel test versions of the RAP and HRRR. Fig. 6 illustrates this improvement in the warm, dry bias for three versions of the RAP.

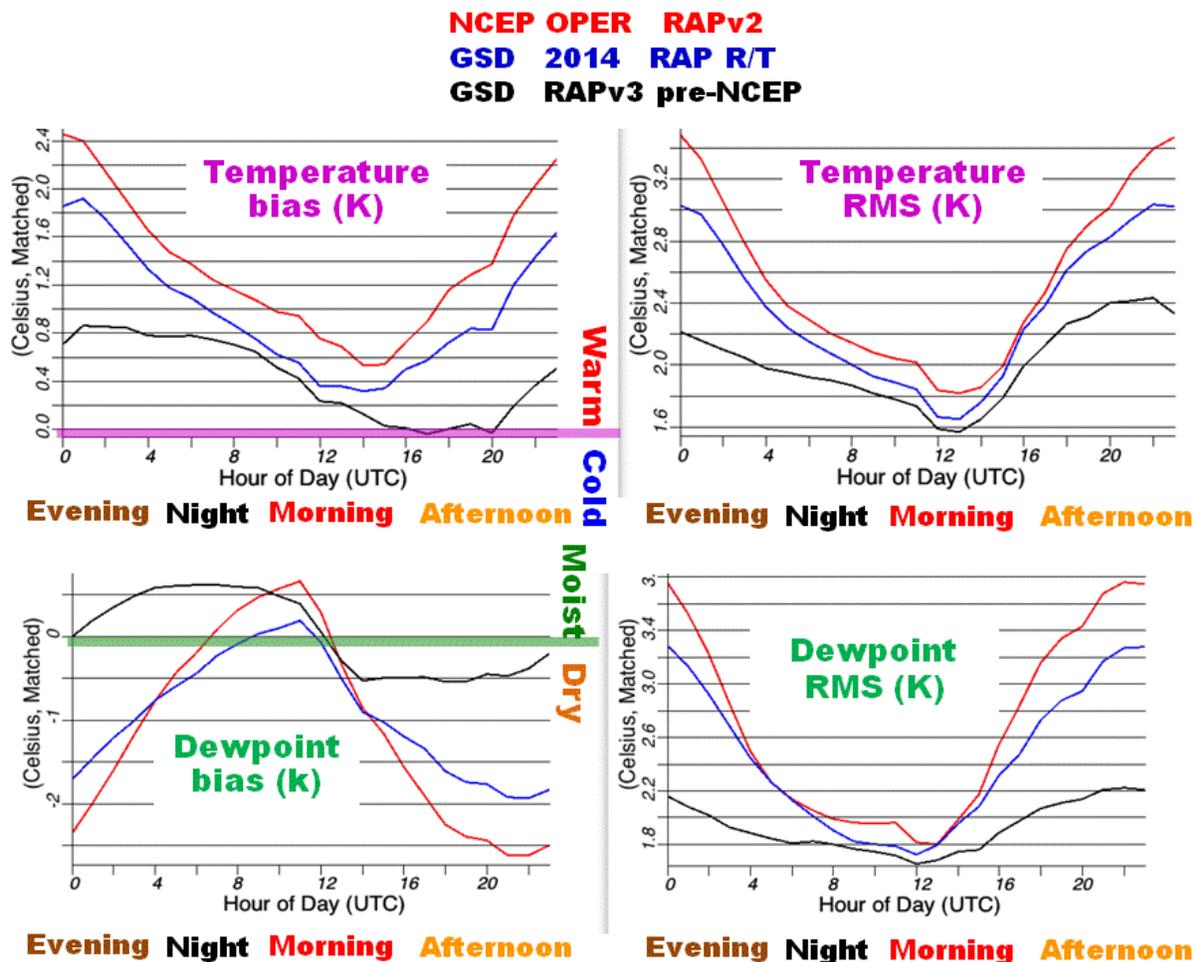


Fig. 6. Comparison of diurnal cycle of bias error (left) and root mean square error (right) for surface temperature (top) and surface dewpoint (bottom) for the following RAP real-time runs: NCEP oper RAPv2 (red), GSD 2014 R/T RAP (blue), GSD RAPv3 pre-NCEP implementation (black). Evaluation period is from 15 Aug. – 18 Sept. 2014.

As can be seen, the GSD 2014 RAP real-time runs (parent model to the GSD real-time HRRR being evaluated, blue) shows a reduction in both the bias and the root mean square errors for both surface temperature and surface dewpoint compared to the NCEP operational RAPv2. The GSD pre-implementation version (black), which incorporated the new changes listed above yields a further significant error reduction, especially for the warm, dry bias. These changes, along with additional enhancements, have been incorporated into GSD RAPv3 / HRRRv2 real-time experimental cycle and will be included in the next NCEP operational upgrade of the RAP / HRRR (slated of implementation in 2015).

APPENDIX:

RAP / HRRR changes for GSD 2014 real-time evaluation (code frozen April 2014).

(Additional pre-NCEP implementation changes made after Oct. 31)

ESRL RAP 13--km Data Assimilation Changes (in chronological order):

(1) Updated Gridpoint Statistical Interpolation (GSI) package to a recent source code trunk revision **(change effective 06 UTC 12 March 2014)**.

- Important for consistency with trunk code
- Change in forecast quality -- **small**

(2) Changed GSI hybrid data assimilation to increase the weight to 75% for model background error covariance derived from a GFS 80--member ensemble forecast at 60--km using the ensemble Kalman filter while decreasing the weight to 25% for static 3--D model background error covariance to further improve the assimilation of all observations. A slightly tighter fit to upper--level observations at the analysis time and throughout the forecast period can be expected **(change effective 06 UTC 12 March 2014)**.

- Change in forecast quality -- **high** for all applications and seasons

(3) Introduced cycled satellite radiance bias correction (for both angle and mass) in GSI during assimilation. Added AMSUA/MHS channels from METOP--B and GOES sounder data in radiance assimilation. Removed some upper level channels for AMSUA/GOES/HIRS4 data. A slightly tighter fit to upper--level observations at the analysis time and throughout the forecast period can be expected **(change effective 06 UTC 12 March 2014)**.

- Change in forecast quality -- **medium** for all applications and seasons

(4) Corrected soil temperature and moisture adjustments in GSI during data assimilation to avoid water points on the coarser analysis grid **(change effective 06 UTC 12 March 2014)**.

- Change in forecast quality -- **small** for all applications and seasons

(5) Corrected the non-variational cloud and precipitating hydrometeor analysis to (a) properly conserve virtual potential temperature during saturation of cloud bearing layers, (b) limit the saturation of cloud bearing layers to 100% relative humidity, and (c) properly scale rain number concentration in addition to rain-water mixing ratio when reducing model values to observed values at specific levels. An increase in low-level moisture and temperature can be expected in some cases where low-level (less than 1.2 km AGL) clouds are observed. A closer fit to lower-level observations in some cloudy areas can be expected at the analysis time **(changes effective 06 UTC 12 March 2014)**.

- Change in forecast quality -- **small** for all applications and seasons

(6) Changed the cycled snow cover fields through modifications in building of snow cover in the 00 UTC cycle based upon the Interactive Multisensor Snow and Ice Mapping System (IMS) snow cover analysis. A neighborhood gridpoint approach is now used to build snow based upon snow cover characteristics from surrounding gridpoint(s). The skin/snow temperature for points with built snow are now limited to no more than 272K **(change effective 06 UTC 12 March 2014)**.

- Change in forecast quality -- **medium** for 2m temperature in snow-cover or near-snow-cover areas, corrects large local surface temperature errors from erroneous snow cover, especially in spring and late winter (**small** overall)

(7) Enhanced the assimilation of surface dewpoint observations by accounting for the difference between the height of the lowest model level (~8 m AGL) and the height of the surface observation (2 m AGL). A reduction in moist bias of lower-level relative humidity through the forecast period can be expected **(change effective 17 UTC 05 April 2014)**.

- Change in forecast quality -- **medium-high** for 2m dewpoint forecasts and related moisture fields for all seasons

ESRL RAP 13-km Model Changes (in chronological order):

(1) Updated Advanced Weather and Research Forecast model (WRF-ARW) from a version 3.4.1 code base to a version 3.5.1 code base including an updated Thompson microphysics version **(changes effective 06 UTC 12 March 2014)**.

- Important for consistency with community code
- Change in forecast quality -- **small** overall

(2) Changed to the Grell-Freitas (GF) convective parameterization scheme (from an older Grell scheme). The scheme is designed to become less active as the grid size reduces to cloud-resolving scales with enhanced shallow cumulus parameterization. A tighter fit to upper-level observations can be expected at the analysis time and throughout the forecast period from improved convective forecasts. Improved precipitation forecasts can also be expected with a reduction in high bias of lower precipitation amounts (less than a half inch in a six hour period) and an improved bias of higher precipitation amounts (more than a half inch in a six hour period) **(change effective 06 UTC 12 March 2014)**.

- Change in forecast quality -- **high** for all applications and seasons, more so in warm season.

(3) Changed to the Rapid Radiative Transfer Model -- Global (RRTMG) scheme for both shortwave and longwave radiation (from the previous Goddard/RRTM radiation schemes). The

RRTMG scheme is designed to use a statistical method to resolve sub-grid scale cloud variability and includes the potential for improved interaction with the Thompson microphysics scheme and aerosols. The period between radiation calls has been increased from 10 min to 20 min during the forecast period, but an additional option has been enabled to interpolate shortwave radiation based on the updated solar zenith angle between radiation calls (**changes effective 06 UTC 12 March 2014**).

- Change in forecast quality -- **medium** for surface solar radiation forecasts, **small** overall

(4) Enhanced RUC Land Surface Model (RUCLSM) including an increased thickness of the top snow layer and increased value of exchange coefficient for stable stratification in 2-m diagnostics when 2-m temperature is set equal to the first atmospheric level. These changes will help reduce the 2-m cold temperature bias over existing snow cover at night (**changes effective 06 UTC 12 March 2014**).

- Change in forecast quality -- **high** for 2 m temperature forecasts over snowpack at night, **small** overall

(5) Enhanced Mellor-Yamada-Nakanishi-Niino (MYNN) planetary boundary layer (PBL) scheme including improved coupling of the PBL scheme with radiation feedback in the GF shallow cumulus convective parameterization scheme and a reduced thermal roughness length over snow. The thermal roughness change will help reduce the 2-m cold temperature bias over existing snow cover at night in concert with changes in (4) (**changes effective 06 UTC 12 March 2014**).

- Change in forecast quality -- **medium-high** for 2 m temperature forecasts over snowpack at night, **medium** overall

(6) Relaxed the restriction for diagnosis of ice pellets (sleet) as a surface precipitation type and enforced diagnosis of 2-m dewpoint to be equal or less than the 2-m temperature (**changes effective 06 UTC 12 March 2014**).

- Change in forecast quality -- **medium** for 2 m temperature/dewpoint and precipitation type forecasts, **small** overall

(7) Increased the surface roughness length values to 1 m for urban and 20 cm for cropland land-use categories to help reduce the high wind speed bias near the surface (**changes effective 17 UTC 05 April 2014**).

- Change in forecast quality -- **medium** for near-surface wind forecasts, **small** overall

(8) Added seasonally varying MODIS vegetation fraction and fractional leaf area index for improved surface roughness, and sensible and latent heat fluxes (**change effective 17 UTC 05 April 2014**).

- Change in forecast quality -- **medium** especially for 2 m temperature/dewpoint and 10-m wind forecasts

(9) Added model terrain elevation blending near the lateral boundaries of the RAP domain for a smoother transition to the coarser resolution GFS grid and included a correction in vertical interpolation of GFS data for lateral boundary conditions. These changes enhance the numerical

stability in the lateral boundary regions (**changes effective 06 UTC 12 March 2014 and 17 UTC 05 April 2014**).

- Change in forecast quality -- **small**

HRRR 3--km Data Assimilation Changes (in chronological order):

(Changes 1--5 below for HRRR assimilation match those also implemented for RAP)

(6) Enhanced Mellor--Yamada--Nakanishi--Niino (MYNN) planetary boundary layer (PBL) scheme including a reduced thermal roughness length over snow. The thermal roughness change will help reduce the 2--m cold temperature bias over existing snow cover at night in concert with changes in (5) (**changes effective 03 UTC 10 April 2014**).

- Change in forecast quality -- **medium--high** for 2 m temperature forecasts over snowpack at night, **medium** overall

(7) Increased the surface roughness length values to 1 m for urban and 20 cm for cropland land-use to help reduce the high wind speed bias near the surface (**changes effective 03 UTC 10 April 2014**).

- Change in forecast quality -- **medium** for near--surface wind forecasts, **small** overall

(8) Added seasonally varying MODIS vegetation fraction and fractional leaf area index for improved surface roughness, and sensible and latent heat fluxes (**change effective 03 UTC 10 April 2014**).

- Change in forecast quality -- **medium** especially for 2 m temperature/dewpoint and 10--m wind forecasts

(9) Added model terrain blending near the lateral boundaries of the HRRR domain for a smoother transition to the coarser--resolution RAP grid and included improved vertical interpolation of RAP data for lateral boundary conditions. These changes enhance the numerical stability in the lateral boundary regions (**change effective 03 UTC 10 April 2014**).

- Change in forecast quality -- **small**

(10) Introduced 6th--order diffusion in regions with very shallow model surface slopes (relatively flat terrain) to reduce grid--scale noise in the mass and momentum fields by at least 25%, particularly in weather regimes with weak flow. This reduction in noise can prevent occurrence of very small scale, generally very weak, model reflectivity structures (**change effective 03 UTC 10 April 2014**).

- Change in forecast quality -- **small**

The ESRL RAP and HRRR data assimilation and model configurations will now remain “frozen”, with the exception of software bug fixes, through the remainder of the spring, summer and early fall with the next changes being applied on or after 01 November 2014.

A summary of all changes effective 03 UTC 10 April 2014 is presented in the following table:

Table A1. Summary of changes to GSD RAP and HRRR for 2014 warm season evaluation

|  ESRL RAPv3/HRRR-2014 Changes | | |
|---|---|---|
| | Model | Data Assimilation |
| RAP-ESRL (13 km) | WRFv3.5.1+ incl. physics changes <u>Physics changes:</u> Grell-Freitas convective scheme MYNN PBL update - Olson version RUC LSM update Thompson microphysics – v3.5.1 RRTMG radiation scheme Shallow cumulus parm w/ rad feed MODIS veg fraction/leaf area index | Merge with GSI trunk Increase ensemble weight in hybrid DA 8m → 2m bkg for sfc Td assim Radiance bias correction New sat assimilation (NOAA-19, METOP-B, GOES, direct readout – RARS) |
| HRRR (3 km) | WRFv3.5.1+ incl. physics changes <u>Physics changes:</u> MYNN PBL update - Olson version RUC LSM update Thompson microphysics – v3.5.1 RRTMG radiation scheme MODIS veg fraction/leaf area index <u>Numerics changes:</u> 6 th order diffusion in flat terrain Smooth terrain @lat BC | 3-km hybrid ens/var assimilation (was var-only in 2013) 8m → 2m bkg for sfc Td assim Radar LH – 4x less intense than 2013 (2x less intense than RAP but more local) <div style="border: 1px solid black; background-color: yellow; padding: 5px; width: fit-content; margin: 10px auto;"> Changes with high/medium importance for <i>overall</i> forecast skill </div> |