P. 90

24th Conference on Weather and Forecasting , AMS
91th Annual Meeting, 23-27 Jan 2011, Seattle, WA

Verifications of Simulated Radar Reflectivity and Echo-Top Forecasts at NCEP

Binbin Zhou[1,*], Jun Du[2], Shun Liu[1] and , G. DiMego[2]

1 IMSG @ EMC/NCEP/NOAA, Camp Springs, Maryland
2 EMC/NCEP/NOAA, Camp Springs, Maryland

1. Introduction

Simulated radar reflectivity and echo top are good indicators for detailed structure of convection and precipitation forecast by mesoscale models, particularly as their resolutions have been increasing in recent years. Although these two products have been generated from NCEP's operational mesoscale models or systems for several years, their forecast performance has not been systematically evaluated nor compared. Recently we have dedicated much effort to do this work over CONUS. This paper summarizes the recent verifications of these two products using hourly 1 km National 88-D Radar Mosaic data as 'truth'. The evaluated mesoscale models include NAM, RUC and 4km-WRF (NMM, ARW of Hi-Res window), 32km-SREF (Short Range Ensemble Forecast System before and after its upgrade in 2009), and SREF in a downscaled 5 km grids. The primary results based on short and long time period data are presented. The objectives of this paper are (1) evaluate the current prediction performance for reflectivity (composite), including NAM, RUC, hi-resolution (4km) WRF-NMM and ARW, SREF's 32 km low resolution NMM and ARW, (2) examine any improvement in reflectivity (composite) and echo-top (height) in both SREF's base WRF models and ensemble predictions after new SREF implementation in November of 2009 (SREF base NMM and ARW WRF upgrades: V2.0 → 2.2; same schemes but raised resolution in both WRF models; number of WRF members were increased) and (3) compare verification of coarse 32 km WRF models over hi-resolution 4 km grid and over coarse grid to see if downscaling SREF from 32km to 4km grid can improve skill.

*Corresponding author address: Binbin Zhou, NOAA/NWS/NCEP/EMC, 5200 Auth Rd. Camp Springs, MD 20746. Binbin.Zhou@noaa.gov

2. Methodology

2.1 Simulated radar reflectivity/echo-top

In NCEP operational models, simulated radar reflectivity products are generated from a unified model post processor. At a given forecast time, three-dimensional radar reflectivity values are calculated at the native model resolution (vertical and horizontal) from the algorithm described below using model output of the three-dimensional mixing ratios of rain and precipitation ice (including variable ice densities to distinguish between snow, graupel, and sleet), and the two-dimensional convective surface precipitation rates. The following two-dimensional radar reflectivity products are derived from the three-dimensional forecast reflectivity fields: (1) lowest level reflectivity - calculated at

the lowest model level above the ground; (2) composite reflectivity - maximum anywhere within the atmospheric column; and (3) 1 km and 4 km AGL reflectivity - interpolated to heights of 1 km and 4 km, respectively, above the ground. In this work only composite reflectivity product is verified. The detailed calculation steps are as following:

$$dBZ = \log_{10} Z_e \qquad (1)$$

where Z_e is the equivalent radar reflectivity factor. It is derived from the 6th-moment of the size distribution for precipitation or snow/ice particles, and assumes that all of the particles are drops. The equivalent radar reflectivity factor is the sum of the radar backscatter from grid-scale rain Z_{grid} *(rain)*, precipitation-sized ice particles Z_{grid} *(ice)*, and parameterized convection *Z(conv)*, in mm^6m^{-3}

$$Z_e = Z_{grid}(rain) + Z_{grid}(ice) + Z_{conv} \qquad (2)$$

$$Z_{grid}(rain) = \sum N(D_r) D_r^{\,6} \qquad (3)$$

where D_r is droplet diameter and $N(D_r)$ is the droplet number ~ size exponential distribution for either rain within Rayleigh scattering.

$$Z_{grid}(ice) = \sum N(D_i) D_i^{\,6} \qquad (4)$$

where D_i is droplet diameter and $N(D_i)$ is the droplet number ~ size exponential distribution for snow or ice within Rayleigh scattering. Although Z_{grid} *(ice)* is also calculated as the 6th- moment of the particle size distributions for ice, but with several correction factors. The first is accounting for the reduced backscatter from ice particles compared to liquid drops. Because equivalent radar reflectivity assumes that the precipitation is in the form of rain, this correction factor is 0.189 (the ratio of the dielectric factor of ice divided by that of liquid water). The second correction factor accounts for deviations in ice particle densities from that of solid ice, in which the radar backscatter from a large, low-density irregular ice particle (e.g., a fluffy aggregate) is the same as that from a solid ice sphere of the same mass (e.g., a small sleet pellet).

$$Z_{conv} \sim Z_{sfc} = 300 \cdot R_{conv}^{1.4} \qquad (5)$$

which is the radar backscatter from parameterized subgrid-scale cumulus convection and $R_{conv}^{1.4}$ is surface rain rate derived from the cumulus parameterization. This so-called Z-R relationship is based on the original WSR-88D algorithm. The radar reflectivity is assumed to remain constant with height from the surface up to the lowest freezing level, and then it is assumed to decrease by 20 dBZ from the freezing level to the top of the parameterized convective cloud. This top is defined as echo-top. The schematic framework for this scheme is shown in Fig. 1

2.2 Verification Method

Since both model and true data 88-D Radar Mosaic) are gridded data over CONUS, NCEP's grid-to-grid verification system (g2g) is a proper tool to fulfill our objectives. The g2g verification tool was developed at NCEP at an attempt to unify all grid-to-grid verifications for all NCEP operational models, either single models or ensemble systems. The output data of the g2g are so-called verification statistic database records (VSDB) that accumulate all comparison results between model data and truth over all grid points in the verification domain or sub-regions. These VSDB records are also called verification partial sum. Based on the VSDB record files, the final statistic scores and plots are then further computed from another NCEP forecast verification system, so-called NCEP-FVS.

The g2g system, the core procedure is editing a user-defined control file that tells the system the verified model, truth data, file paths, grid space, validation regions, forecast and validation times, statistic types, variable names, levels, etc. The comparison between model data and truth can be performed for both single model and ensemble/probabilistic prediction. For single model or deterministic verification, the g2g first configures the data files paths, for both model and truth files, according to the user-defined control file, and then validates the model and truth data at all grids in a region to generate VSDB records based on number of forecast, hit and observed

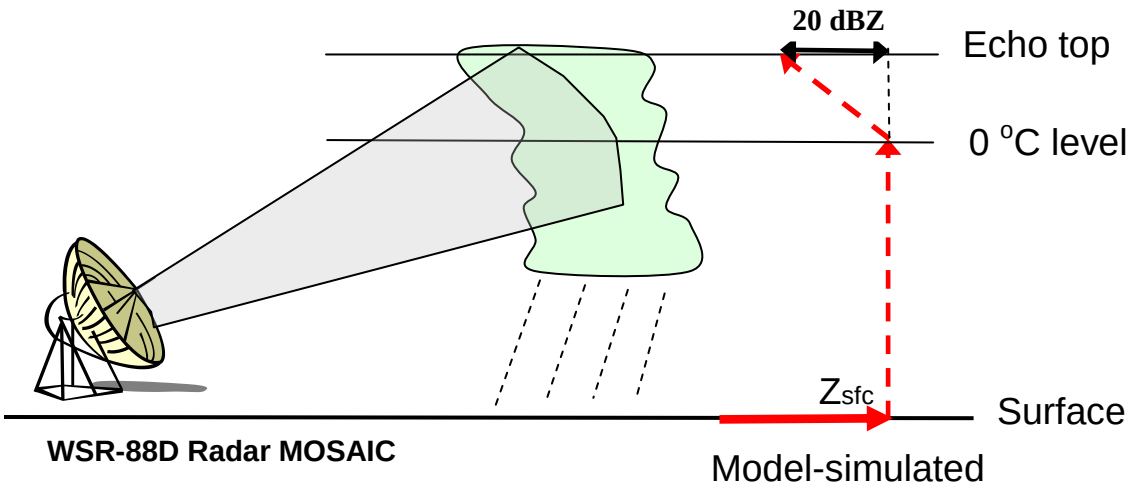

Figure 1, Reflectivity diagnostic scheme

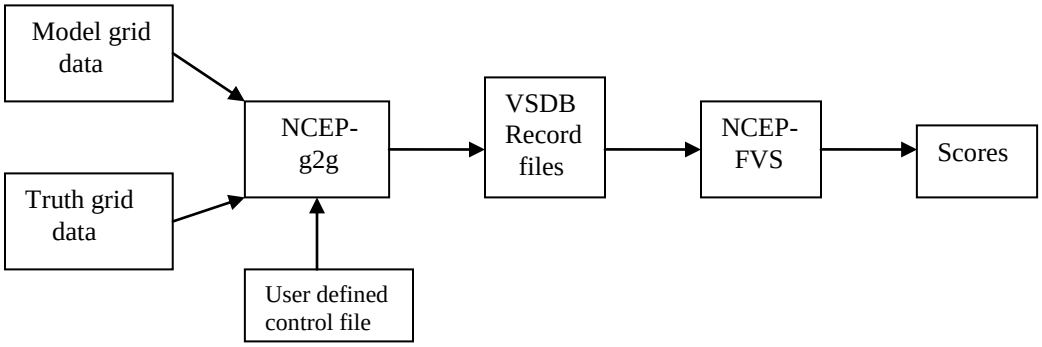

Figure 2, NCEP-grid-to-grid (g2g) procedures

grid points (F.H.O. counts). This g2g procedure can be executed daily in real time or retrospectively run using historical data. After enough VSDB data are accumulated over a period of time, NCEP-FVS system is used to generate traditional verification scores. In this work, bias, POD, FAR, ETS are used as evaluation scores. Since the VSDB has become a standard verification output data format for all operational forecast verifications at NCEP, more and more other verification systems, such as RTVS and MET, have also been enhanced to include the VSDB format files as input to generate verification scores.

In order to compare an ensemble prediction with single model, the ensemble probabilistic verification can be conducted in a deterministic way. The method is setting a specific percentage threshold (such as 50%). A probabilistic forecast can be viewed as a deterministic forecast in the way that an event is expected to occur when the forecast probability is greater than or equal to the selected threshold. Thus, the F.H.O counts can be accumulated in the same way as a verification of a single model. To evaluate the performance on all probability range, multiple ensemble probability thresholds can be selected. In this work, 10, 20, …90% are selected as the probability thresholds.

3. Verification Data

3.1 WSR-88D Radar MOSAIC data

The 3-D WSR-88D Radar Mosaic (refer as MOSAIC) data product was developed by NSSL of NOAA (Zhang et al. 2005) and is currently processed at NCEP to generate 1-km high resolution gridded MOSAIC data over CONUS (Liu:http://www.dtcenter.org/comGSI/users/docs/presentations/2010_tutorial/L11-0630/RadarData_ShunLiu_resaved.ppt.pdf). Current MOSAIC data at NCEP are hourly produced, including radian wind, leveled and composite reflectivity, and echo-top. To save space, it is upscaled to 4-km resolution in GRIB2 format. In this work, totally 8 month MOSAIC data, from Aug. of 2009 to April of 2010, are used. A example of 4-km resolution MOSAIC at a specific time can be seen in Fig. 3a, in which some very detailed structures in a convection in northeast coast can be identified. In spite of some limitations of MOSAIC data, such as its only on land, missing detection at low angle of scanning, and quality control problems, its advantages in evaluation of high resolution model are still obvious in comparison to other types of data in cloud/convection particularly in severe weather verifications. Due to no MOSAIC data over ocean and missing data in those grids caused by low scanning radar beams and earth's curvature, these missing data grids are masked from the F.H.O counting. This may lead to missing some important structure of a convection system in certain places. However, over such a large domain as CONUS over relative long period, the verification scores should have statistical meaning.

3.2 Model data

The models involved in this reflectivity/echo-top forecast evaluations include 12km resolution North American Mesoscale model (NAM: http://www. emc.ncep.noaa.gov), 13km resolution Rapid Update Cycle model (RUC: http://ruc.noaa.gov), 4km High-resolution WRF (both NMM: Janjić *et al.* 2001 and ARW: Skamarock *et al.* 2005) and 32km Short Range Ensemble Forecast System (SREF: Du et al. 2006). NAM runs 4 times (00, 06, 12 and 18Z) per day to provide guidance to regular weather for all of local forecasters in the United States. RUC runs hourly, specifically for aviation weather forecast. To compare improvements in reflectivity forecast performance before and after new SREF implementation in November of 2009, the base models, 32km low resolution WRF NMM and ARW, were also evaluated. Since the echo-top products currently are only produced in the SREF's WRF members, its evaluation was only conducted for the SREF and its base models. Different models for different verifications over different times are summarized in Table 1. Examples of one time MOSAIC and model reflectivity and echo plots are presented in Fig. 3.

4. Result

To evaluate any skill improvement in reflectivity prediction before and after new SREF implementation in Nov 2009, one month data before and after Nov 2009 were selected as shown in Table 1. To compare NAM, RUC, hi-res NMM and ARW, all data as well as MOSAIC data were first converted to 4km hi-resolution WRF's east and west grids by NCEP's copygb utility with nearest neighbor interpolation. To compare the performances of reflectivity forecasts verified at different grids, 32km 212 grid and 4 km high resolution eat and west grids were selected, respectively. When verified over 32km grid 212, copygb, with nearest neighbor interpolation, was used to convert MOSAIC from 4km fine resolution grid to coarse resolution 32km grid, while verified over 4km fine resolution grid, coarse resolution model grids were converted the fine resolution grid with copygb, still with nearest neighbor interpolation.

4.1 Single models simulated reflectivity verified over hi-res (4km) grids

Single model reflectivity performance, over 4 km east and west high resolution grids, are compared, respectively in Fig. 4 and 5. The east high resolution grid covers 70% of CONUS in east and west high resolution covers 70% of CONUS in west. There are more than 50% overlap in central regions between east and west high resolution grids.

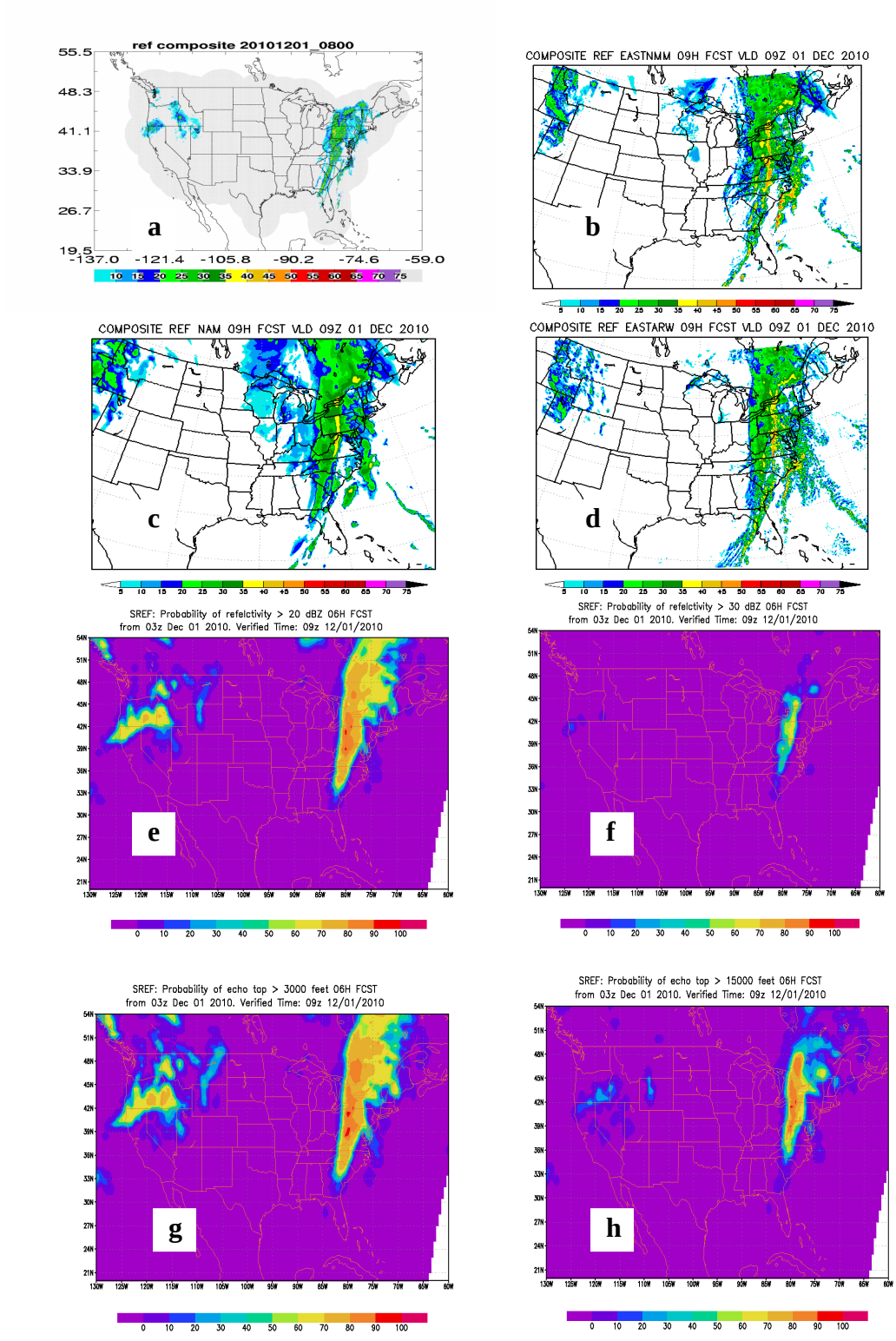

Figure 3. Example of 09Z run, Dec.1, 2010 MOAIC and model reflectivity and echo top Prediction over CONUS: (a) MOSAIC reflectivity, (b) NAM reflectivity 9 hr forecast (c) NMM(4km) reflectivity 9hr fcst, (d) ARW(4km) reflectivity 9hr fcst, (e, f) SREF reflectivity>20,30dBZ probability, and (g, h) SREF eco-top>3000,15000Ft probability

Table 1. Model data correspondent objectives, verification times and grids

Model data used	Objectives	Verification period	Verification grid on
NAM-12km	Reflectivity performance	Aug 12, 2010 , all cycles	4 km hi-res east grid and west grid
RUC-13km	Reflectivity performance	Aug 12, 2010, all cycles	4 km hi-res east grid and west grid
NMM-4km	1.Reflectivity performance	Aug 12, 2010, one cycle (Only one cycle is available)	4 km hi-res east grid and west grid
ARW-4km	1.Reflectivity performance	Aug 12, 2010, one cycle (Only one cycle is available)	4 km hi-res east grid and west grid
SERF-NMM (32km)	1. Reflectivity performance 2. Improvement of new SREF in reflectivity and echo-top 3. Fine~coarse grid comparison	Apr. 3 ~ Apr. 30, 2009 before SREF upgrade Apr. 6 ~ Apr. 27, 2010 After SREF upgrade	32 km 212 grid 4 km hi-res east and west grids
SREF-ARW(32km)	1. Reflectivity performance 2. Improvement of new SREF in reflectivity and echo-top 3. Fine~coarse grid comparison	Apr. 3 ~ Apr. 30, 2009 before SREF upgrade Apr. 6 ~ Apr. 27, 2010 After SREF upgrade	32 km 212 grid 4 km east and west grids
SREF-32km	1. Reflectivity performance 2. Improvement of new SREF in reflectivity and echo-top 3. Fine~coarse grid comparison	Apr. 3 ~ Apr. 30, 2009 before SREF upgrade Apr. 6 ~ Apr. 27, 2010 After SREF upgrade	32 km 212 grid 4 km east and west grids

Fig. 4 displays reflectivity verifications of all single models verified over the east high resolution 4km grid region. MOSAIC data, NAM, RUC, and WRF 32km NMM and ARW are all converted, with copygb, from their coarse resolution to 4km high resolution grid before grid to grid verification. Bias is presented in Fig. 4a where the best value is 1 (no bias), above which indicates over-prediction while below which indicates under-prediction. We can see that all higher resolution models have positive bias, i.e. over-prediction of reflectivity. The NMM (4km) has biggest positive bias which become even larger as the reflectivity threshold is increased. ARW(4km) also shows positive bias and maximum in 30 dBZ. But over 30dBZ its bias decreases. The bias of ARW(4km) is overall smaller than NMM(4km). NAM and RUC have smaller positive bias in comparison to ARW(4k) and NMM(4km) with maximum in 20 dBZ, and decrease gradually, and then become negative in 40 dBZ. RUC also shows lower bias than NAM. Low resolution WRF, NMM (32km) and ARW (32km) show negative bias, that is, under-prediction of reflectivity in all reflectivity thresholds. The negative bias for both coarse WRF models also increase with reflectivity thresholds. Fig. 4b shows probability of detection (POD) or hit rate, from which one can observe that NMM (4km) has highest POD. This is due to that NMM(4km) takes advantage of its high over-prediction. NAM also shows higher POD close to that of NMM(4km) but its bias is much lower that that of NMM(4k). It is not surprising that NAM's ETS, the overall performance score, is the best in Fig. 4c. ARW(4km) has smaller POD that leads to smaller ETS as shown in Fig. 4c. RUC also has smaller POD although its bias is smaller. Its smaller POD cause its smaller ETS than NAM. Comparing ETS scores between high and low resolution WRFs, for both NMM and ARW, in Fig. 4c, we can see that in lower reflectivity threshold range, low resolution WRF models have higher performance than their corresponding high resolution WRFs while in high reflectivity threshold range, high resolution WRF models become better than their low resolution versions. This is interesting in that high resolution models have high skill to capture strong convection centers but are prone to missing weak storm or miss-matching convection edges. Another feature in Fig. 4c is that beyond 40 dBZ threshold, for all models, except for NAM, almost have no skill to predict reflectivity. The ETS values of low resolution WRF models reach zero as the dBZ threshold is over 30 dBZ.

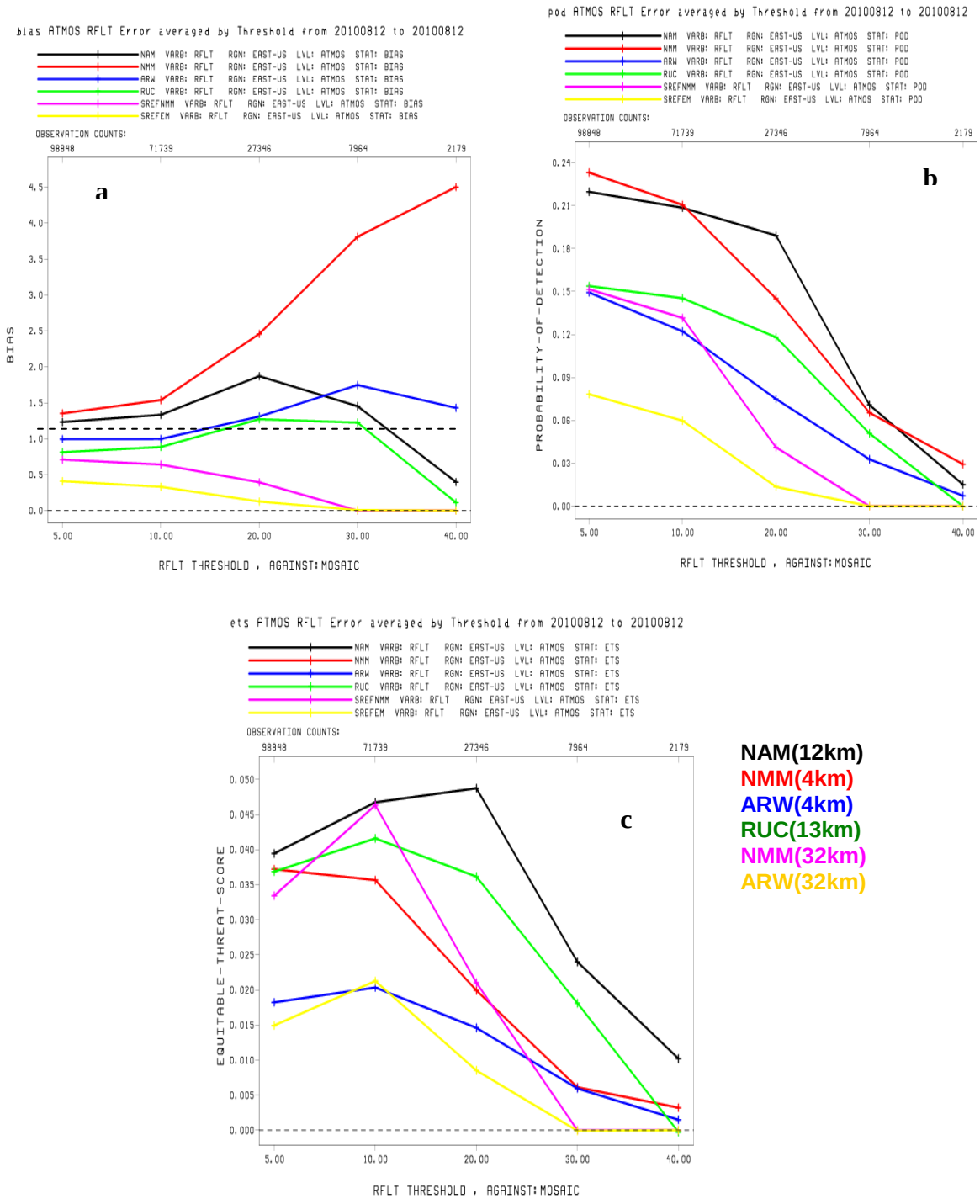

Figure 4. Single model reflectivity forecast scores, (a) Bias, (b) POD and (c) ETS, verified over east 4 km high resolution grid region

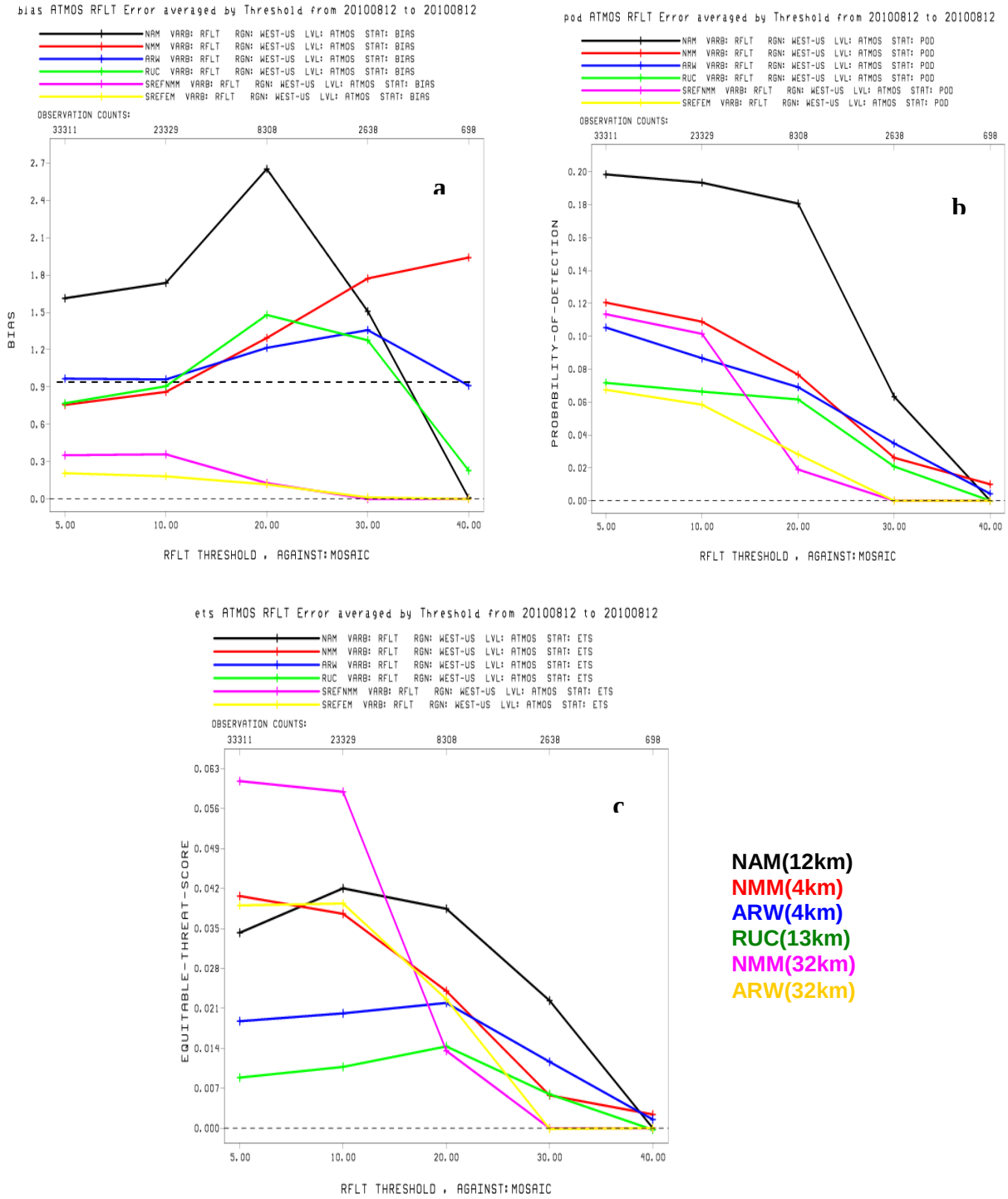

Figure 5. Single model reflectivity forecast scores, (a) Bias, (b) POD and (c) ETS, verified over west 4 km high resolution grid region

Fig. 5 displays reflectivity verifications of various single models verified over the west high resolution 4km grid. As same as the way over the east grid, MOSAIC data, NAM, RUC, and WRF 32km NMM and ARW are all converted from their original grids to the west 4km high resolution grid, with copygb with nearest neighbor interpolation, before grid to grid verification. In Fig. 5a, NAM shows highest positive bias near 20 dBZ threshold, which is quite different from its performance verified over the east grid, then decreases until reaching negative bias (under-prediction). The bias of 4km NMM is still larger than that of 4km ARW, particularly in higher dBZ thresholds, similar to the feature shown in Fig. 4a. Lower resolution WRF NMM (32km) and ARW (32km) show negative bias in all dBZ thresholds as same as in the east grid. Again, both NAM and RUC have smallest bias, same as that shown in the east grid. In Fig. 5b, NAM shows highest POD, and high resolution WRFs have higher POD than low resolution WRFs, same as in the east grid. In Fig. 5c, lower resolution models show better ETS for lower dBZ thresholds and lower ETS in higher dBZ thresholds, which is also similar to the east grid.

4.2 SERF reflectivity/echo-top prediction improvement after new implementation

Fig. 6 presents ETS scores for echo-top and reflectivity prediction from WRF-NMM(32km) and ARW (32km), both the base models in SREF, and from SREF before and after it was upgraded in November of 2009. All verifications were conducted on their original 32km 212 grid over CONUS with nearly one month data as shown in Table 1. The 4km grid MOSAIC data were first upscaled to the same 212 grid using copygb with nearest option before grid to grid verification. In the new SREF, many upgrades were made, including (1) both NMM and ARW were upgraded from V2.0 → 2.2, (2) native resolution of ARW from 45 to 35km while native resolution of NMM from 40 to 32km (On CONUS, both are delivered to WFOs in format of 32 km 212 grid), and (3) the number of WRF members was increased from 3 NMM and 3 ARW members in the old SREF to 5 NMM and 5 ARW members in the new SREF.

Fig. 6a and 6b are one month ETS score accumulations for echo-top and reflectivity forecasts, respectively, from SREF base models (NMM and ARW) under different reflectivity thresholds (in dBZ) and echo-top height thresholds (in m) before and after the SREF upgrade. We can see that the performances of both reflectivity and echo-top forecasts from 2 WRF models are increased for all thresholds except for very high ends. For echo-top, the biggest ETE gain is round threshold 6000 m while for reflectivity the biggest gain is at its smaller threshold range, showing that WRF version upgrade and resolution increase have positive effect on reflectivity and echo-top prediction performance, although at higher threshold range, the ETS scores decrease in both reflectivity and echo-top. This means, for very strong convection, the new SREF's NMM and ARW have no significant improvements in reflectivity and echo-top forecasts. In comparison between NMM and ARW base models, it is shown that ARW has a little better skill for echo-top while NMM a little better skill for reflectivity prediction. This is true for both before and after the SREF upgrade. Another feature for Fig. 6a and 6b is that ETS values for echo-top prediction is much lower than that of reflectivity prediction. So it requires much more effort for us to improve the echo-top prediction.

Fig. 6c and 6d are one month ETS score accumulations for echo-top and reflectivity ensemble forecasts, respectively, from SREF system. Because only WRF members have these two products, the ensemble reflectivity and echo-top computation only included WRF members, no Eta and RSM members were involved. ETS scores of echo-top and reflectivity probability forecasts (converted to deterministic prediction with a probability threshold) for new SREF were significantly improved in all probability thresholds. It is noticeable that such improvements are not only in lower reflectivity and echo-top thresholds but also in high reflectivity and echo-top thresholds although this is not true for single models as shown in Fig. 6a and 6b.

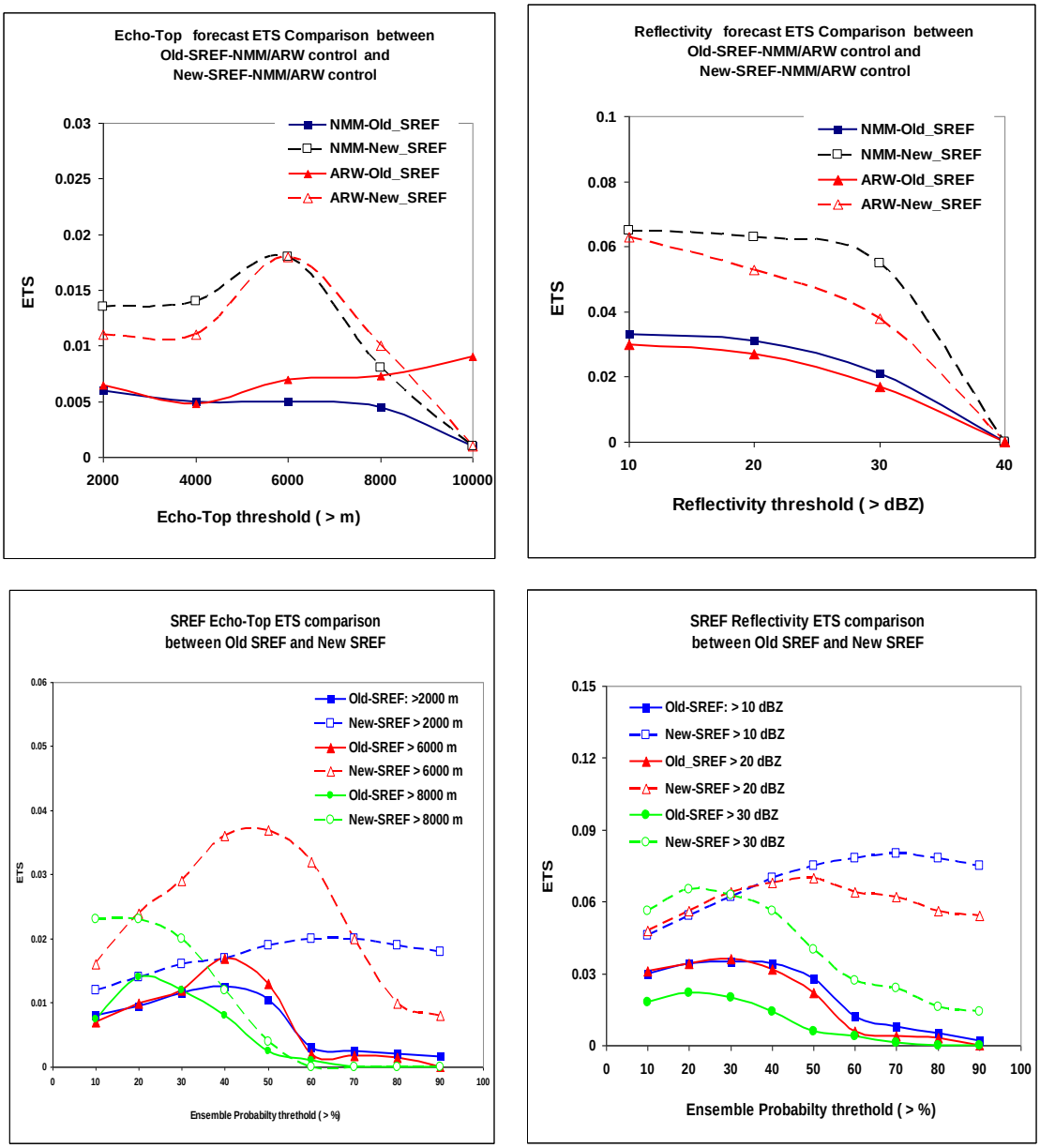

Figure 6. Monthly accumulated ETS scores for Echo-top (a) and reflectivity (b) of SREF-NMM (32km) and ARW(32km) before (Apr. 3~30, 2009) and after (Apr. 6~ 27, 2010) new SREF implementation (Nov. 2009) for different echo-top and reflectivity thresholds. (c) and (d) same as (a) and (b) but for SREF ensemble probability prediction over various ensemble probability thresholds for echo-top and reflectivity.

4.3 Comparison between reflectivity verifications over coarse grid and fine grid

Section 4.1 compares verifications of different resolution models over same 4 km high resolution grid. This section compares same resolution models over different verification grids. In this work, SREF's 32 km base WRF models, were used. To do this comparison, two sets of verifications were made. The first verification is over 32 km 212 grid, the same grid as NMM (32km) and ARW (32km). Only 4 km MOSAIC data were converted to the 212 grid with copygb. The second verification is over high resolution grids, 4 km east and west grids, respectively. In this case, all data including NMM(32km), ARW(32km) and MOSAIC were copygbed to the east and west grids, respectively, before grid to grid verification. The results are presented Fig. 7.

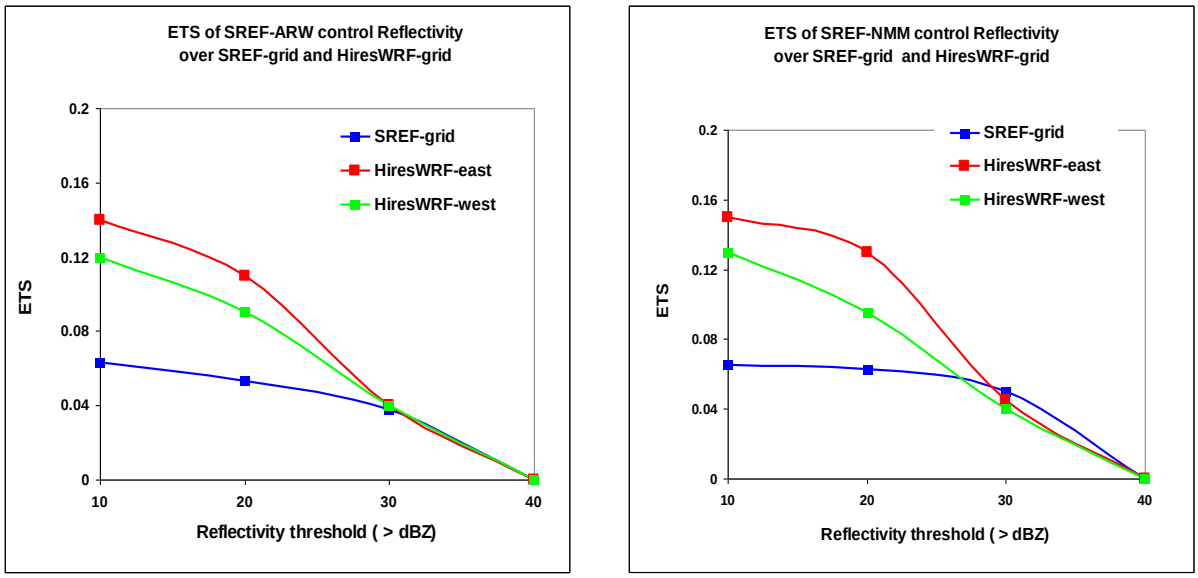

Figure 7. Monthly accumulated ETS scores (Apr. 6~27, 2010) for reflectivity from, (a) SREF's 32 km NMM, and (b) SREF's 32 km ARW, over SREF's 212 coarse resolution grid and 4 km east and west high resolution grids.

From Fig. 7a and 7b, it is interesting to notice that reflectivity forecasts from coarse models verified over high resolution grid have better results than verified over coarse grid in lower dBZ thresholds. But this feature is not true for higher dBZ thresholds. Moreover, for higher dBZ thresholds, verifications over fine and coarse grids have no difference. In other words, verification over high resolution grid can not improve the results for strong reflectivity forecast. These interesting features require further investigations. One possible explanation could be in MOSAIC data or usage of copygb with nearest neighbor option.

To explain the two cases, one verification over coarse grid the other is over fine grid, let's assume the fine 4 km grid is same as the MOSAIC 4 km gird. The grid to grid verification over coarse grid can be expressed in Fig. 8, where left grid is supposed to be coarse grid reflectivity and right is fine grid. The assumed numbers represent the reflectivity values by 10. Since coarse gird values are smoothed, they are usually smaller than fine grid values in many strong convection locations, as assumed in Fig. 8. For the verification over left

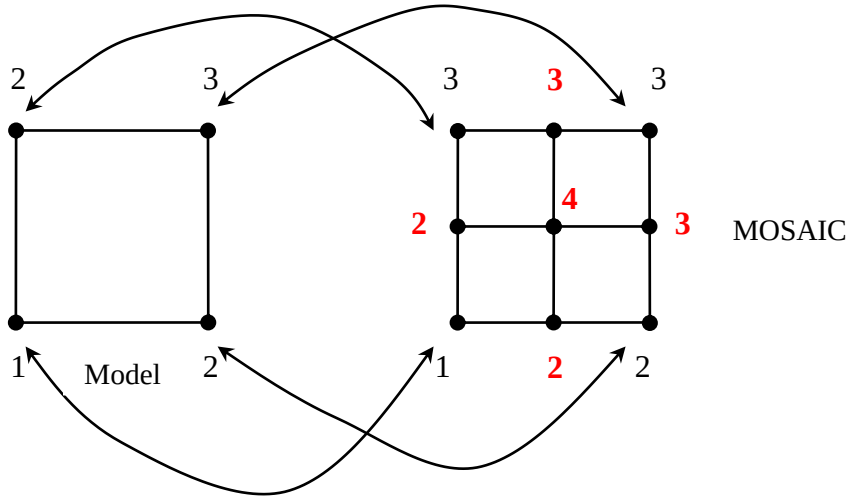

Figure. 8, Conceptual plot for grid to grid verification over coarse grid (left). The values in the grid corner represent dBZ(x10) value.

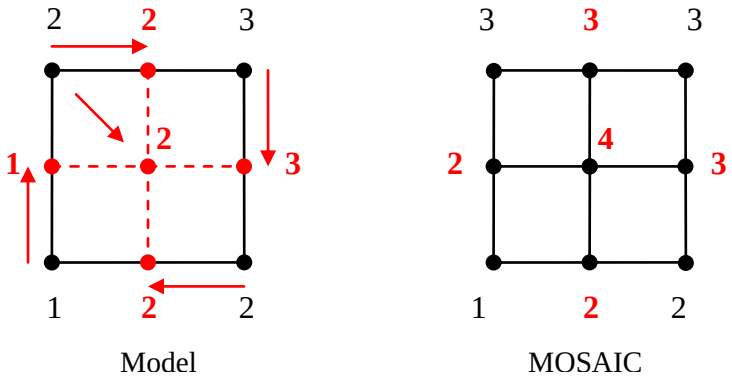

Figure. 9, Conceptual plot for grid to grid verification over fine grid (right). The values in the grid corner represent dBZ(x10) value. The values on the dash line obtained from assumed nearest neighbors

coarse grid, right side 4 points on the 4 corners match the left 4 corner points' values. In this case the values on the other points in the right side grid are not involved. Suppose there are some larger value in side, or there is a strong convection inside, this strong convection grid will be missed in the coarse grid verification.

The second case is verification over fine MOSAIC grid shown in Fig. 9, where the left

coarse grid is divided into several fine grids, and the nearest point values are taken for the new divided grid points (assumed). And then the left girds match the corresponding grids in the right grid. It can be seen that more lower reflectivity values can be captured in Fig. 9 than in Fig. 8. That is why verification over fine grid has better performance than that over coarse grid. However, because the "nearest neighbor" interpolation from coarse grid to fine grid can not increase the smaller values from the coarse grid, the verification over fine grid still miss the strong reflectivity grids (such as the center grid value 4 is still missed). So verification over fine grid can not improve performance for strong reflectivity prediction, but for only weak convection reflection forecast.

5. Conclusion

Single model and ensemble simulated reflectivity and echo top forecasts from NCEP's 12 km NAM, 13 km RUC , 32 km SREF's base NMM and ARW, 4 km hi-resolution WRF NMM and ARW, as well as before and after the new SREF implementation in Nov. 2009 were evaluated with grid to grid verification against the national 88-D Radar MOSAIC data over COUNS. The grid to grid verification tool and FVS of EMC/NCEP were employed in this work. If model data and MOSAIC data are in different grid, NCEP's cobygb utility is used to convert one grid to another or convert both model and MOSAIC grids to a common grid. With this verification, reflectivity/echo-top forecast performances from different models, same model in different resolutions, and same model verified over different verification grids, are evaluated. The results show that (1) hi-resolution models have generally better detection and prediction skill than low resolution models, particularly for higher dBZ range, but this is obtained at a higher positive bias as a cost; (2) for all models, lower dBZ and lower echo-top threshold ranges have higher prediction performance than higher reflectivity/echo-top threshold ranges; (3) for all models, prediction of simulated reflectivity is better than that of simulated echo-top; (4) coarse model reflectivity verified on fine grid has better score than on coarse grid for lower dBZ threshold range, but this does not true for high dBZ threshold; The primary reason is likely due to the way of copygb with nearest neighbor interpolation; (5) the performances of both reflectivity and echo-top from new implemented SREF are improved; and (6) for reflectivity > 40 dBZ, all models and ensemble system have no prediction skills

Reference

Du, J, J. McQueen, G. Dimego, Z. Toth, D. Jovic, B. Zhou and H. Chuang, 2006: New dimension of NCEP SREF system: Inclusion of WRF members. *Report to WMO Export Team Meeting on Ensemble Prediction System, Exeter, UK, Feb. 6-10, 2006*

Janjić, Z. I, J. P. Gerrity Jr., and S. Nickovic, 2001: An alternative approach to nonhydrostatic modeling. *Mon. Wea. Rev.*, **129**, 1164-1178.

Koch S. E., B. Ferrier, M. T. Stoelinga, E. Szoke, S. J. Weiss and S. Kain, 2005: The use of simulated radar reflectivity fields in the diagnosis of mesoscale phenomena from high-resolution WRF model forecasts, *11th Conf. on Mesoscale Processes, AMS, 22-29 Oct, Albuquerque, NM*.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers, 2005: A description of the Advanced Research WRF, Version 2, *NCAR Technical Note*.

Zhang J., K. Howard, and J. J. Gourley, 2005: Constructing Three-Dimensional Multiple-Radar Reflectivity Mosaics: Examples of Convective Storms and Stratiform Rain Echoes, *J. Atmos. Oceanic Technol*., **22**, 30-42.

Zhou B., 2008: NCEP grid to grid verification system, available online: http://www.emc.ncep.noaa.gov/mmb/papers/zhou/NCEPGrid2GridVerificationSystem-V2.doc