J6.1   SUPPLEMENTAL AUTOMATED QUALITY CONTROL OF MRMS REFLECTIVITY PRODUCTS FOR LAMP CONVECTION AND LIGHTNING FORECAST GUIDANCE APPLICATIONS

Jerome P. Charba*
Frederick G. Samplatsky

Meteorological Development Laboratory
Office of Science and Technology Integration
National Weather Service, NOAA
Silver Spring, Maryland

Andrew J. Kochenash
KBRwyle Science, Technology & Engineering Group
Silver Spring, Maryland

1. INTRODUCTION

The quality of meteorological observations applied as predictors and predictands in automated prediction models is ultimately reflected in the quality of the forecasts produced by the models. Observational data used in the experimental Localized Aviation MOS Program (LAMP) convection and total lightning forecast guidance products (Charba et al. 2016; 2017) developed by the National Weather Service Meteorological Development Laboratory (MDL) consist of Multi-Radar Multi-Sensor (MRMS; Smith et al. 2016; Zhang et al. 2016) reflectivity products developed by the NOAA National Severe Storms Laboratory (NSSL) and total lightning measurements provided by the Earth Networks Total Lightning Network (https://www.earthnetworks.com/networks/lightning/). The MRMS reflectivity data in the LAMP convection model have a critical role, as these data are used not only for specifying key convection predictors but also to specify the convection predictand. MRMS data are also used in the "sister" LAMP total lightning prediction model, though their predictive role there is secondary to total lightning observations.

Unfortunately, radar reflectivity data are notoriously contaminated by non-meteorological targets (birds, insects, etc.), anomalous (beam) propagation echoes, partial beam filling, beam occultation due to rugged terrain, diverse electromagnetic scattering by various precipitation types, etc. The lead author encountered such radar data quality concerns long ago in a study of the quality of manually-digitized radar data (Charba and Liang 2005), and more recently involving radar-based quantitative precipitation estimates (Charba and Samplatsky 2011); many other studies on the subject appear in meteorological literature (e.g., see references in OFCM 2005; 2006).

Fortunately, strong advances in radar data quality control (QC) have been developed in recent years to screen most non-precipitation echoes. These procedures are based primarily on dual-polarimetric radar measurements for use at radar sites (see Krause 2016 and references therein) and "nationally" as an integral component of MRMS (Tang et al. 2014). On the other hand, the historical sample used for development of the experimental LAMP convection and total lightning products spans a relatively long period (January 2012 to September 2016), and thus the quality of this sample may be lacking, especially for the earlier years. Also, poor radar coverage over parts of the network, particularly over the mountainous western contiguous United States (CONUS), poses an ongoing radar data quality problem. Thus, at the outset of the LAMP convection and lightning model development effort, it was believed prudent to investigate the quality MRMS data sample for use in these applications.

In this study, the authors investigated the quality of the MRMS data sample, and ensuing findings of data quality issues prompted development of a "MDL" automated supplemental QC process. The article discusses the MRMS data quality

---
* *Corresponding author address:*
Dr. Jerome P. Charba, National Weather Service
1325 East West Highway, Room 10204
Silver Spring, MD 20910-3283
email: jerome.charba@noaa.gov

problems, the MDL supplemental QC process, and its beneficial contribution to LAMP convection forecast performance. The study also examined the quality of quite recent MRMS data, which is also discussed.

2. INVESTIGATION OF MRMS DATA QUALITY

2.1 MRMS Data Gridding

MRMS grids used in the data quality investigation consisted of radar reflectivity product maximum values, whereby the maximum pixel[1] value (for a given variable) in a 5-km grid box[2] is assigned to a centered grid point. Such time-instantaneous grids for three MRMS reflectivity products [composite reflectivity (CREF), vertically-integrated liquid (VIL), and "storm top height 30" (STP)] were specified at 15–min intervals for four "hh:mm" clock times, where hh = 00, 01, …, 23 and mm = 00, 15 (14), 30, and 45 (44) minutes past the top of the hour (the parenthetical values apply to dates beginning with 31 July 2013, where the time resolution of MRMS data changed from every 5 minutes to every 2 minutes).

The focus of the data quality investigation involved the max CREF grids, as they are used most extensively in the LAMP convection model. The study consisted of map inspection of individual max CREF fields and max CREF relative frequency (RF, i.e., echo climatology) fields for several CREF thresholds based on the historical sample.

2.2 Random and Systematic Error in Max CREF Grids

Inspection of individual max CREF maps revealed the sporadic presence of strong non-precipitation echoes, especially early in the archive period, an example of which is shown in Fig. 1 (for the famous eastern United States (US) derecho of 30 June 2012). The map sequence shows anomalous propagation (AP) and microwave interference echoes, both of which are temporally transient across the 1-h map span. Such temporal transience was found typical of such non-precipitation echoes.

---

[1] An MRMS pixel is 0.01 degree x 0.01 degree (approximately 1 km on a side over the CONUS).
[2] 2.5 km and 10 km grid boxes were also considered; the 5-km grid box choice was found to provide an optimal combination of needed spatial resolution and computational economy.

Examination of the RF maps revealed the presence of systematic error in the max CREF grids. Figure 2 shows an example RF map for CREF ≥ 40 dBZ derived from all grids for hh:00 valid times (hh = 00, 01,…, 23) for the warm season months (April – September) of 2012-2016. Note the (unrealistic) flower-shaped spatial pattern in the RF field over much of the radar coverage area, which reflects range dependency of reflectivity measurements for individual radars. This artificial echo climatology pattern is evident everywhere in Fig. 2 except for the southeastern US, where the close spacing of network radars avoids the problem.

Highlighted in Fig. 2 (with text boxes and pointers) are additional echo climatology artifacts that warrant attention. These consist of unrealistic, near-zero RFs across the mountainous western US and southwestern Canada, scattered extensive areas of spuriously high RFs in western Washington (state) and across much of southern Canada, RF "spikes" scattered across parts of the central US, and a RF "streak" near the South Carolina coastline. The near-zero RFs over the western US and Canada are likely due to under-detection of precipitation due to wide spacing of radar sites and radar beam obstruction due to rugged mountainous terrain, the spuriously high RFs over Washington (state) and Canada are likely due to AP echoes, and RF streak in South Carolina Coast likely reflects microwave interference echos, as seen in Fig. 1.

Regarding the RF spikes over the central US, a magnified map view is shown in Fig. 3 where selected spikes are enclosed with diamond-shaped boxes. For the same map area, Fig. 4 depicts 5-km grid boxes with wind turbines and/or other tall surface structures (yellow spots), where the diamond boxes in Fig. 3 are superimposed. Note that the highlighted RF spikes and turbine/tall structure locations coincide in these maps, which suggests the spikes reflect AP echoes due to these ground targets.

2.3 Improved Quality of Recent MRMS Data

In the Introduction, it was suggested the quality of the MRMS data may have improved over the course of the 1 January 2012 – 30 September 2016 sample due to locally-applied and MRMS-applied QC upgrades. To investigate this premise the historical time series of CREF grids was divided into two equal parts and separate RF maps were derived for the first half and second half

samples. A comparison of these RF maps revealed that spurious RF spikes, as seen in Fig. 2, were somewhat weaker in the second half sample than in the first half sample (not shown).

Pursuing this premise further, note that the most recent MRMS upgrade occurred in October 2016, which is after the endpoint (30 September 2016) of the sample studied here. Thus, to investigate the cumulative benefit of upgrades from January 2012 to the present, the historical sample was extended to 31 March 2017, and echo RF maps based on the earliest and latest warm and cool seasons were compared (Fig. 5). For the warm season, we see striking mitigation of spurious RF maxima over Canada and the US from 2012 to 2016. For the cool season the corresponding mitigation is even more evident (note the absence of RF spikes over the central US in Fig. 5d), which may reflect an additional benefit of the most recent MRMS QC upgrade.

3. MDL SUPPLEMENTAL QUALITY CONTROL

3.1 QC Framework

Because of data quality concerns with the MRMS 2012–2016 base sample (discussed in the previous section), development of an automated supplemental QC procedure to support development of LAMP experimental convection and lightning forecast guidance products (Charba et al. 2016; 2017) was undertaken. This "MDL supplemental QC" is comprised of a two-step process. The first step is called dynamic QC, as it consists of a "decision tree" of heuristically-formulated meteorological consistency checks designed to identify non-precipitation echoes in the max CREF grids spaced at 15-min intervals across the historical sample. The consistency checks apply to CREF ≥ 35 dBZ; thus, lower CREF values (valid or not) are not affected by the dynamic QC process. The second step is called "static QC", as it consists of the application of warm season and cool season grid masks, as developed from the dynamically-QC'd grids. Essential details regarding the decision tree and seasonal grid masks are discussed below.

3.2 Dynamic QC Decision Tree

A schematic flowchart of the consistency checks comprising the decision tree is shown in Fig. 6, and details pertaining to each decision module are summarized in Fig. 7. The first module, the total lightning (TL) flash consistency check,

determines whether a max CREF value for a given grid box is associated with at least one TL flash in the immediate neighborhood, whereas the subsequent spike echo check determines whether a CREF value ≥ 45 dBZ forms the apex of a sharp CREF peak characteristic of a non-precipitation "spike echo." Note from Figs. 6 and 7 that both of these decision modules are "global," as the associated decision criteria apply to any 5-km grid box over the full radar coverage domain and to any 15-min grid over the historical sample. Note also that the TL consistency check can result in a "keep" decision, and the spike echo check can produce a "reject" decision; otherwise each module results in no decision, which leads to application of two additional decision modules.

The additional decision modules, the MRMS consistency check and the HRRR[3] consistency check (Fig 6), ultimately determine whether the previously-retained CREF value is kept or rejected. As indicated in Fig. 7 the MRMS consistency check determines whether the max CREF value is consistent with prescribed MRMS VIL and STP threshold values, where unique VIL and STP consistency criteria apply to each of five CREF intervals, to the warm or cool season, and to each of three geographical domains (Canada, CONUS, and the CONUS tropical sub-domain) as shown in Fig. 8).

The HRRR consistency check determines whether a max CREF value is consistent with short-range HRRR CREF forecast values in the neighborhood of the central 5-km grid box (Fig. 6). As for the MRMS consistency module the HRRR consistency criteria are stratified both seasonally and geographically. Lastly, the tropical domain decision module (Figs. 6 and 7) determines whether the MRMS and HRRR consistency checks are repeated with unique tropical consistency criteria. While discussion of the many individual consistency criteria comprising the MRMS and HRRR consistency modules are beyond the scope of this article, it is noted these criteria were carefully "tuned" to yield optimal CREF "keep" or "reject" decisions (based on subjective judgements of the authors) for a comprehensive set of convective storm cases (including two hurricanes).

---

[3] HRRR refers to CREF forecasts from the High Resolution Rapid Refresh model (Benjamin et al. 2016).

3.3 Static QC Grid Mask

The static grid-masking process rejects max CREF values for a fixed set of pre-determined grid points in all 15-min grids. The mask points were specified subjectively from visual examination of seasonal RF (echo climatology) maps, which were derived from CREF grids previously subjected to the dynamical QC. The mask specification procedure involved ingesting a grid of seasonal RFs (such as that shown in Fig. 2) into a Geographical Information System (GIS) program and applying associated map editing tools to select the mask points. A given grid point is assigned to the mask if the associated RF value is either inconsistent (meteorologically) with neighboring values or the RF value is unrealistically small (indicating precipitation under-detection). Such decisions were based on the authors' subjective judgements.

Separate grid masks were developed for the warm season (April – September) and the cool season (October – March) to reflect seasonal uniqueness in the echo climatology patterns. This uniqueness is linked to predominantly lower radar reflectivities and lower precipitating cloud heights during the cool season versus the warm season. Also, it was advantageous to use separate CREF thresholds for each season; for the warm season the ≥ 40 dBZ threshold was most effective, whereas for the cool season ≥ 30 dBZ was most used.

The warm and cool static grid masks are shown in Fig. 9. A comparison of the warm season mask with the corresponding RF map in Fig. 2 reveals that almost all of the masked area corresponds to anomalously low RFs in the western US and along the entire outer perimeter of the radar coverage area. Note that the mask also includes a few small spots mostly over the central US, where the dynamic QC did not result in full removal the spurious RF spikes (section 2.2).

The corresponding cool season mask (Fig. 9) is generally similar to that for the warm season, except it has increased geographical coverage, especially over northern and western (mountainous) parts of the radar coverage area and the outer fringe of the entire radar coverage area. The increased coverage reflects lower precipitating cloud heights and thus poorer long-range radar detection of cool season precipitation.

3.4 Use of QC'd Grids in LAMP Convection and Lightning Models

As discussed in Charba et al. 2017, the QC'd max CREF (or VIL) grids are used for specifying LAMP convection and lightning (observational) predictors and, in the case of the convection model, the predictand as well. Thus, follow-on procedures to the dynamic and static QC applications address how missing data values (due to the QC) are treated in these predictor and predictand specifications.

In the case of the LAMP predictor specification, a missing MRMS CREF (or VIL) grid point value is replaced with a 75-min HRRR CREF (or VIL) forecast from the latest available cycle. Thus, these "observational" predictor variables are actually comprised a mix of MRMS measurements and HRRR forecasts, though the fraction of the latter is very small across the eastern US due to the small coverage of the static mask there. These "QC'd radar predictors" are henceforth called MRMS-HRRR hybrid predictors.

In the case of the LAMP convection predictand, the missing MRMS CREF values are retained in the convection specification. This can result in missing predictand values and thus their removal from the LAMP regression equation development and convection probability verification[4]. This convection specification strategy has two benefits: (1) it ensures the integrity of the LAMP convection predictand definition, and (2) it avoids a procedure-induced statistical correlation between the MRMS-HRRR hybrid predictors and the convection predictand. Still the approach introduces the drawback of a small bias toward total lightning occurrences in the samples used for the convection probability (regression equation) development and the ensuing probability verification.

4. IMPACT OF MDL QC

4.1 Improved Echo Climatology

---

[4] Since the convection predictand is also specified from total lightning measurements (Charba et al. 2017), an occurrence is assigned (rather than missing) in the case of one or more total lightning flashes during the predictand valid period. Thus, in the case of a missing MRMS CREF value, the convection predictand is assigned missing only where lightning did not occur.

The warm season RF of ≥ 40 dBZ reflectivity (echo climatology) based on the non-QC'd MRMS 5-km max CREF grids was discussed in section 2.2 (Fig. 2 shows the associated RF map). The corresponding RF field based on the QC'd grids (includes both the dynamic grid screening and static masking) is shown in Fig. 10. Comparing these RF maps, we find that the spurious RF maxima seen in Fig. 2 across southern Canada and Washington (state), as well as the spurious RF spikes over the central US are removed in Fig. 10. Also the RF streak along coastal South Carolina is clearly mitigated, as are questionable, localized RF maxima in the immediately vicinity of network radar sites. This indicates the dynamic QC correctly removed most systematically-recurring non-precipitation echoes.

Outside the locations of the spurious RF maxima in Fig. 2, we see that the pre-QC and post-QC RFs are rather similar, with only a slight general reduction in magnitudes of the latter RFs[5]. This implies the dynamic QC correctly retained the vast majority of valid precipitation echoes.

4.1.1   QC Impact for Recent MRMS Data

In section 2.3, it was shown that MRMS-based echo climatology fields for the most recent warm season (2016) and cool season (2016-17) are relatively devoid of indications of false precipitation echoes, which contrasts with findings based on the historical sample used in this study. This finding was attributed to improved QC procedures recently implemented at radar sites and within MRMS, which now appear to be removing most non-precipitation echoes. To examine this premise further, the MDL supplemental QC was applied for these recent warm and cool seasons and the corresponding echo climatology maps were compared with the non-QC'd versions. This map comparison (not shown) revealed that the QC'd warm and cool season climatologies very closely matched non-QC'd versions, which provides further evidence of the effectiveness of currently-operational false echo screening procedures.

----

[5] A notable exception applies to the RF maximum offshore coastal North Carolina, where a clear reduction in the RF peak is shown in the post-QC map. Note that since this RF peak is located well offshore (and thus at long from the nearest radars), it is reasonable to expect that the MRMS consistency check might not perform well there.

4.2 Effect of QC on LAMP Convection Probability Performance

4.2.1   Subjective Assessment

The subjective assessment was performed by visually comparing QC and no-QC LAMP convection probability maps for selected convection cases. The QC probabilities are from the experimental LAMP convection probability regression equations (Charba et al. 2016; 2017), whereby MRMS predictor variables are specified from the MRMS-HRRR CREF (or VIL) hybrid grids (section 3.4). The corresponding no-QC probabilities are produced by same regression equations, but where the MRMS predictors are instead based on the corresponding non-QC'd MRMS grids. With this approach the QC and no-QC probabilities are identical where MRMS-HRRR hybrid grid values and non-QC'd grid values are the same; the former and latter probabilities will be different where rejected MRMS grid values (due to the QC) are replaced with HRRR-forecasted grid values.

Figure 11 shows the effect of the QC for the strong eastern US derecho of 30 June 2012 (discussed in section 2.2). The figure shows the no-QC MRMS 5-km max CREF field, where intense AP echoes over southeast Virginia and northeast North Carolina are highlighted, as well as the corresponding QC map, which shows masking of the AP (by the dynamic QC process). Also shown are the QC and no-QC LAMP convection 0-1 h forecast probability maps, which respectively incorporate the no-QC and QC CREF (and VIL) fields. Note that erroneous no-QC convection probabilities near 100% coinciding with the AP echoes are reduced to less than 30% in the QC probability map due to the AP echo rejection combined with the follow-on replacement of missing CREF (and VIL) values with HRRR-forecasted values.

A more recent eastern US example (for 05 June 2014) is shown in Fig. 12. In this case small AP echoes are scattered from northern Illinois to just north of Lake Superior. These resulted in small erroneous probability maxima, with peak values over 50%. Note that replacement of the masked AP echoes with HRRR forecasts resulted in near-complete removal of these erroneous probability peaks.

Fig. 13 is for a western US (cool season) convection example for 23 December 2015. Here, the MRMS max CREF maps are not shown (for brevity); instead convection probability maps are shown

for two forecast projections (3-4 and 9-10 hours) to illustrate, for the no-QC case, localized, stationary probability maxima along the West Coast. These meteorologically-unrealistic probability maxima are due to highly non-uniform radar coverage in the area, which results in repeated examples such as this during the cool season. This problem provided strong motivation for developing the static QC masks, which not-surprisingly have extensive coverage over the western US, particularly for the cool season (see Fig. 9). In Fig. 13, the benefit of the static QC masking together with follow-on HRRR-forecast filling of masked areas is reflected as more plausible spatial patterns in the QC convection probability maps. The forecast pattern improvement seen here is typical of such cases.

4.2.2 Objective Assessment

The approach used for the objective assessment is the similar to that used in the subjective assessment, except that a key departure is that separate no-QC convection probability regression equations were developed and tested to evaluate the QC impact. The development and application of these no-QC test equations is identical to that for the baseline (QC) convection probability regression equations, except that the MRMS predictor variables are based on the no-QC MRMS grids rather than the QC grids. Thus, with this approach, the no-QC and QC convection probabilities may be different throughout the forecast domain. Still, it is important to note that the convection predictand data used in both the equation development and the comparative forecast performance scoring is not ideal, since the QC results in missing convection events (see section 3.4) and thus removal of these cases from the both the equation development as well as the verification. On the other hand, since this drawback applies equally to the no-QC and QC probabilities it should not favor (or disfavor) the performance scores for either set of probabilities.

The verification test is for the 0600 and 1800 UTC LAMP cycles, and the (independent) verification sample consists of 216 days, where 72 days are taken from each of three calendar periods: 01–18 September of 2012-2015, and 01-18 February 2013-2016, and 14-31 May of 2013-2016. This date selection strategy ensures the verification sample is (roughly) uniformly distributed across the MRMS historical archive used in the study.

The Brier Skill Score (bss; Brier 1950; Wilks 2006) is used to measure the comparative performance of the QC and no-QC convection probabilities for the test sample. In Fig. 14 no-QC and QC bss versus forecast projection curves are shown separately for the geographical area east and west of the US Continental Divide because of the large contrast in the static mask coverage between these areas (Fig. 9). Also, while the maximum LAMP forecast projection is 25 hours, the scores are shown only for the first 10 hours because the MRMS predictors are not useful at longer projections.

For the eastern US, Fig. 14 shows a very small skill improvement for the QC probabilities over the no-QC probabilities at the 1-h projection, which gradually erodes to no improvement at 5 hours and beyond. Contrastingly, for the western US the QC probabilities exhibit a substantial skill improvement for the early projections, and a slight skill advantage extends out to 9 hours. This result implies the combination of extensive static masking together with HRRR filling is an effective QC strategy for the western US.

It was noted at the beginning of this sub-section that the QC introduces missing values in the convection predictand, which results in undesirable removal of cases from the regression equation development and verification. Since it cannot be determined whether this may have impacted the comparative scoring, as an alternative the scoring was repeated where total lightning (predictand) validation data (which does not contain missing values) is substituted for the convection validation data. [Note that convection is defined from a combination of MRMS reflectivity and total lightning observations (Charba et al. 2017)]. While the validity of this substitution is open to question, the verification results obtained (not shown) were remarkably similar to the those obtained with the convection validation data. This result strengthens the findings based on the convection verification data[6].

4.3 Discussion

The weak statistical forecast skill benefit of the QC for the eastern US, discussed in the previous sub-section, is consistent with related findings

---

[6] While the objective scoring is admittedly weakened by the lack of sound validation data, the tests performed are believed to be the best that can done with data available.

discussed earlier in this article. In particular, the aggregate number of eastern US cases with substantial QC benefit was found to be relatively small across over the full historical period studied. Since the verification sample consisted of a small sub-set of this sample, the number of high-QC-impact cases within it was likely quite small. Also, the aggregate areal coverage of non-precipitation spikes in echo climatology maps removed by the QC (Fig. 10) over the eastern US is also relatively small. These findings (which apply to the historical sample used in LAMP model development) along with findings of strongly improved removal of false precipitation echoes in recent MRMS samples (section 2.3) suggest the forward benefit of the MDL supplemental QC for the eastern US will be quite small.

For the western US, on the other hand, the evidence indicating a substantial positive impact of the supplemental QC on LAMP convection probability performance is more compelling. Recall from section 2.2 that a major deficiency in MRMS reflectivity data in the west evidently stems from poor coverage by network radars there. The substantial convection forecast skill improvement achieved there implies the quality of MRMS reflectivity data was (in effect) improved through the extensive static masking together with the HRRR CREF (and VIL) forecast replacement of masked data.

5. SUMMARY AND FINDINGS

Since MRMS reflectivity products based on a historical sample spanning January 2012 to September 2016 have a critical role in the development and application of the LAMP convection model, the data quality for the sample was investigated. The investigation revealed data quality deficiencies, which include AP contamination particularly over southern Canada, wind farm AP over the central US, microwave interference in South Carolina, and extensive under-detection of precipitation across the mountainous western US and southwestern Canada. This led to the development and application of an MDL supplemental QC process, which contains a decision tree comprised of diverse meteorological consistency checks to dynamically detect and screen non-precipitation echoes and grid masking for areas with poor radar coverage. Application of the dynamic QC process revealed that it captured most non-precipitation echoes while having little effect on valid echoes. The static grid masking had important QC role across western portion of the MRMS radar coverage area.

For predictor usage in the LAMP convection model, missing MRMS reflectivity product data resulting from application of MDL QC are replaced with 75-min HRRR model forecasts from the most recent hourly cycle. For the convection predictand specification, the missing values are retained to avoid an internal correlation with the MRMS-HRRR hybrid predictors.

Testing of this QC strategy for selected cases showed improvement in the convection probability forecast patterns due to the QC. Objective forecast performance scoring for a historical sample indicated a slight skill improvement for the eastern US and substantial improvement for the western US. The stronger benefit of the QC for the western US is attributed to the extensive static masking combined with HRRR forecast replacement there.

Very recent MRMS samples were also examined for evidence of data quality improvement due to recently-implemented QC upgrades at radar sites and within MRMS. This investigation revealed that most non-precipitation echoes are removed in recent and current MRMS data. Thus, going forward it is expected the MDL QC will have little benefit in the LAMP model over the eastern US. For the western US, the substantial QC benefit achieved here may continue since the problem of poor radar coverage is an ongoing problem.

7. REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1-3.

Benjamin, S. L., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669-1694.

Charba, J. P, and F. Liang, 2005: Qaulity control of gridded national radar reflectivity data. Preprints, *21st Conference on Weather Analysis*

and Forecasting, Washington, D. C, Amer. Meteor. Soc., **6A5**.

___, and F. G. Samplatsky, 2011: High resolution GFS-based MOS quantitative precipitation forecasts in a 4-km grid. *Mon. Wea. Rev*, **139**, 24-38.

___, ___, P. E. Shafer, J. E. Ghirardelli, and A. J. Kochenash, 2016: LAMP convection probability and "potential" guidance: An experimental hires upgrade. Preprints, 23rd *Conference on Probability and Statistics in the Atmospheric Sciences*, New Orleans, LA, Amer. Meteor. Soc., **1.3**.

___, ___, ___, ___, ___, 2017: Experimental upgraded LAMP convection and total lightning probability and "potential" guidance for CONUS. Preprints, 18th *Conference on Aviation, Range, and Aerospace Meteorology*, Seattle, WA, Amer. Meteor. Soc., **2.4**.

Krause, J. M., 2016: A simple method to discriminate between meteorological and non-meteorological radar echoes. *J. Atmos. Oceanic Technol.*, **33**, 1875-1884.

OFCM, 2005: Doppler radar meteorological observations: Doppler radar theory and meteorology. Federal Meteorological Handbook No. 11 Part B, Office of the Federal Coordinator for Meteorological Services and Supporting Research, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, (Available online at http://www.ofcm.gov/publications/fmh/FMH11/fmh-11B-2005.pdf).

___, 2006: Doppler radar meteorological observations: WSR-88D unit description and operational applications. Federal Meteorological Handbook No. 11 Part D, Office of the Federal Coordinator for Meteorological Services and Supporting Research, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, (Available online at http://www.ofcm.gov/publications/fmh/FMH11/FMH11D-2006.pdf).

Smith, T. M, , V. Lakshamanan, G. J. Stumpf, K. L. Ortega, K. Hondl, K. Cooper, K. M. Calhoun, D. M. Kingsfield, K. L. Manross, R. Toomey, and J. Brogden, 2016: Multi-Radar Multi-Sensor (MRMS) Severe Weather and Aviation Products. *Bull. Amer. Meteor. Soc., 97, 1617-1630.*

Tang, L., J. Zhang, C. Langston, J. Krause, K. Howard, and V. Lakshmanan, 2014: A physically based precipitation-nonprecipitation radar echo classifier using polarimetric and environmental data in a real-time national system. *Wea. Forecasting*, **29**, 1106-1119.

Wilks, D. S., 2006: *Statistical Methods in the Atmospherics Sciences*. 2nd ed. Academic Press, 627 pp.

Zhang, J., and Coauthors, 2016: Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.,* **97**, 621-637.
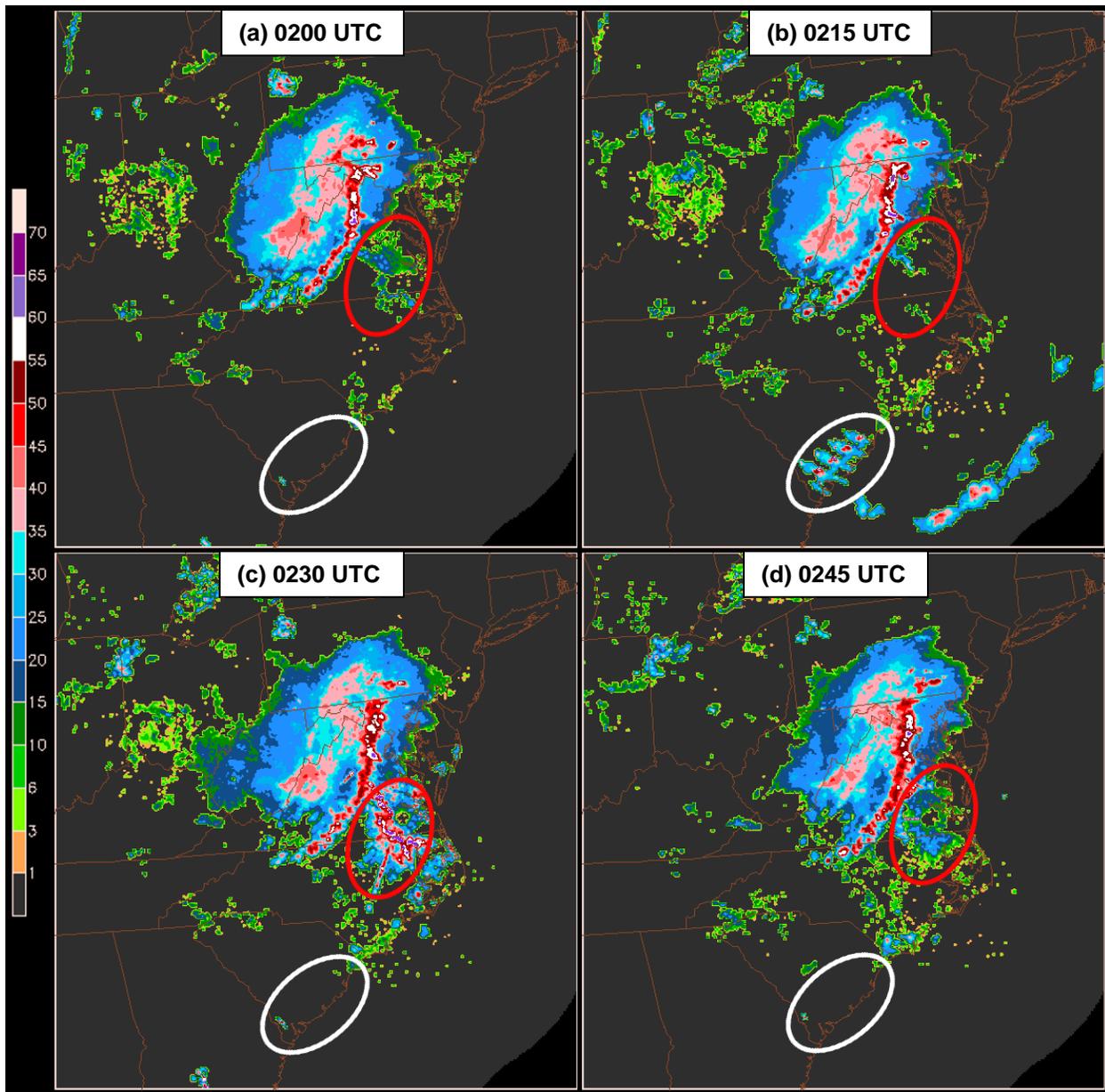
Figure 1. Example of temporally-transient false precipitation echoes in MRMS 5-km max CREF grids at 15-min intervals on 30 June 2012 (famous eastern US derecho). Anomalous propagation echoes just ahead of derecho (within the red ellipse) are quite weak at (a) 0200 UTC, almost non-existant at (b) 0215 UTC, very intense at (c) 0230 UTC, and moderate at (d) 0245 UTC. Microwave interference echoes within the white ellipse are present only at (b) 0215 UTC.
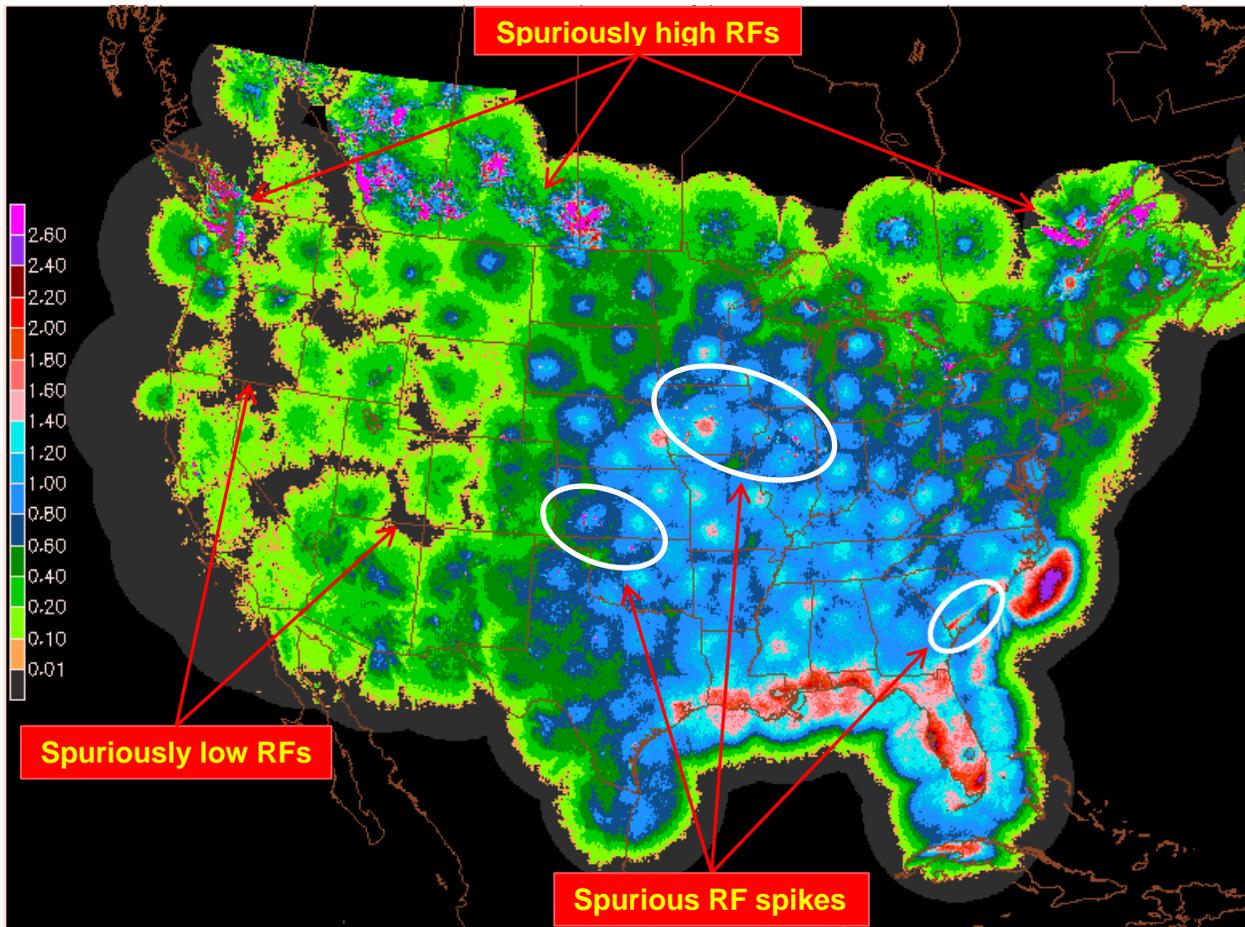
Figure 2. Relative frequency (RF; %) of max CREF ≥ 40 dBZ in 5-km grid boxes for the the warm season (April – September) of 2012-2016. Annotated RF features are discussed in the text.
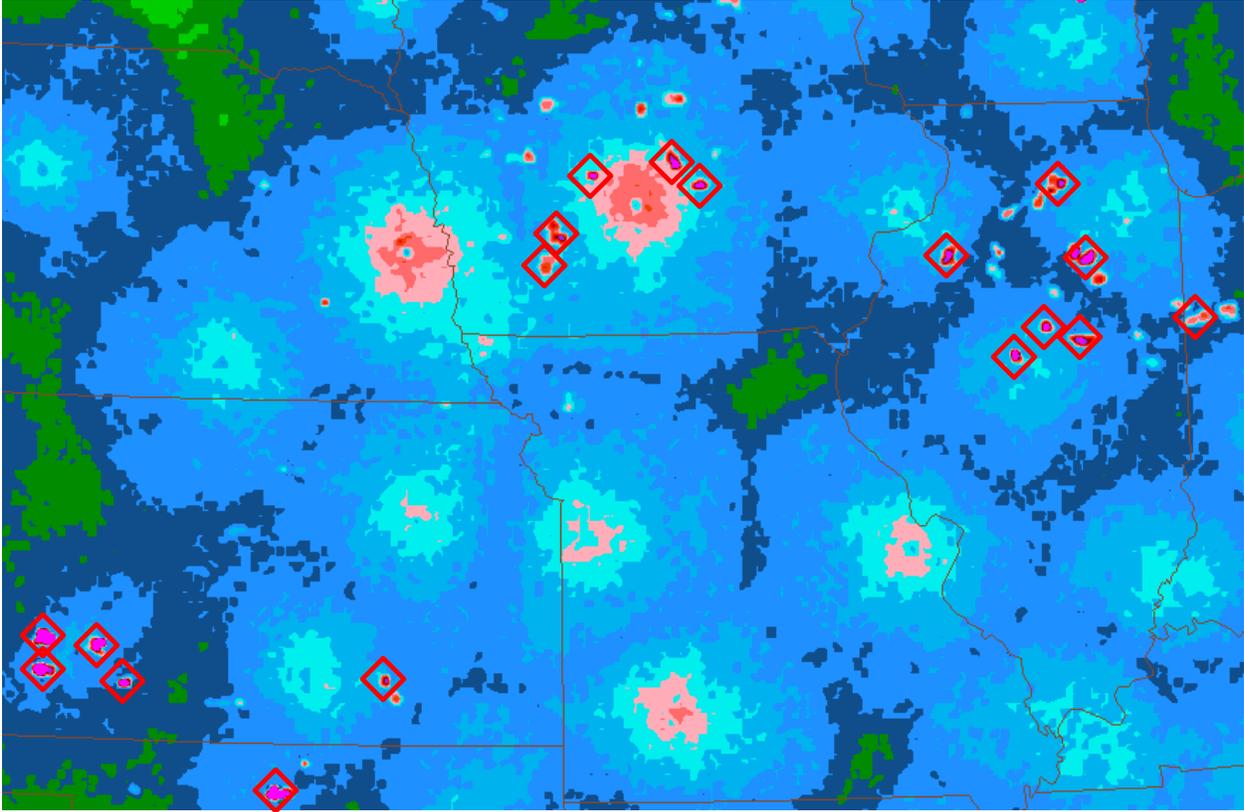
Figure 3. Magnification of warm season RF map (Figure 2) for the central US. Selected RF spikes are enclosed with red diamond symbols.
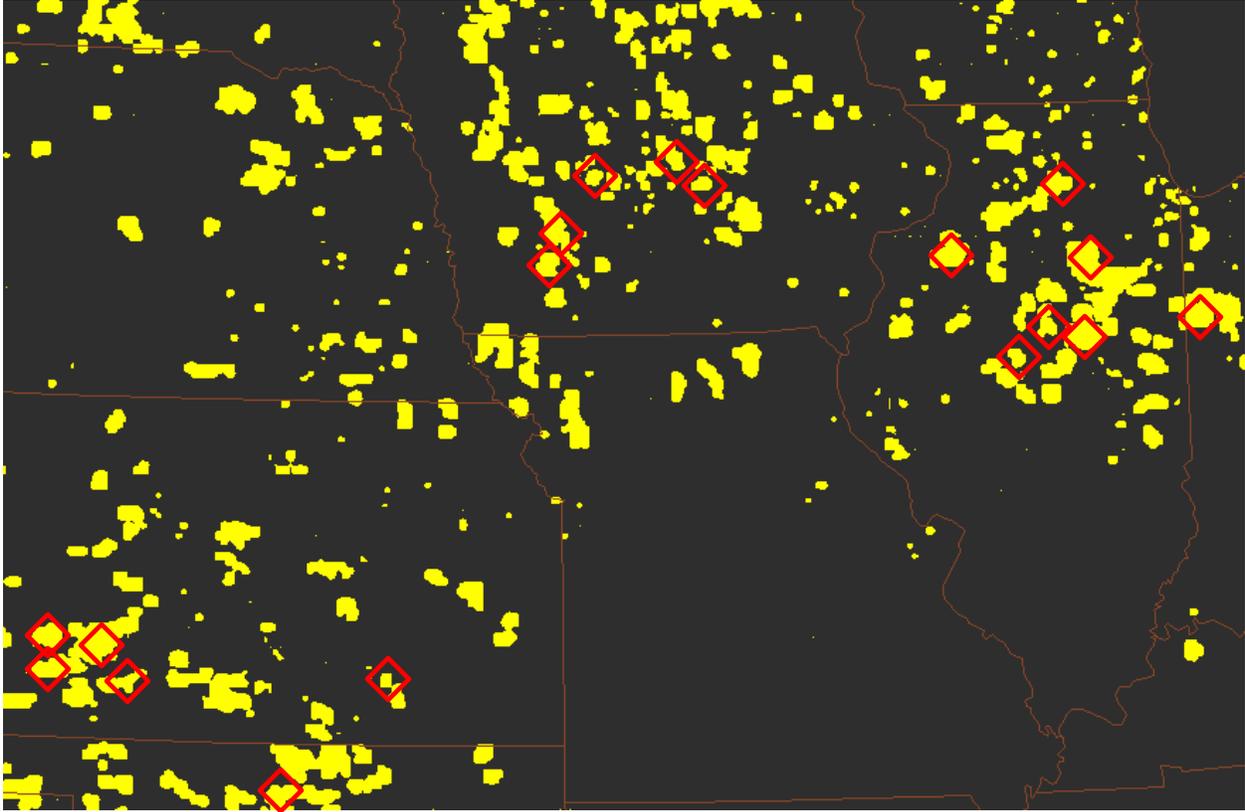
Figure 4. Yellow spots are 5-km grid boxes with either wind turbines and/or other tall surface structures, based on the database https://www.fws.gov/southwest/es/Energy_Wind_FAA.html. The superimposed red diamond symbols are the same as those in Figure 3.
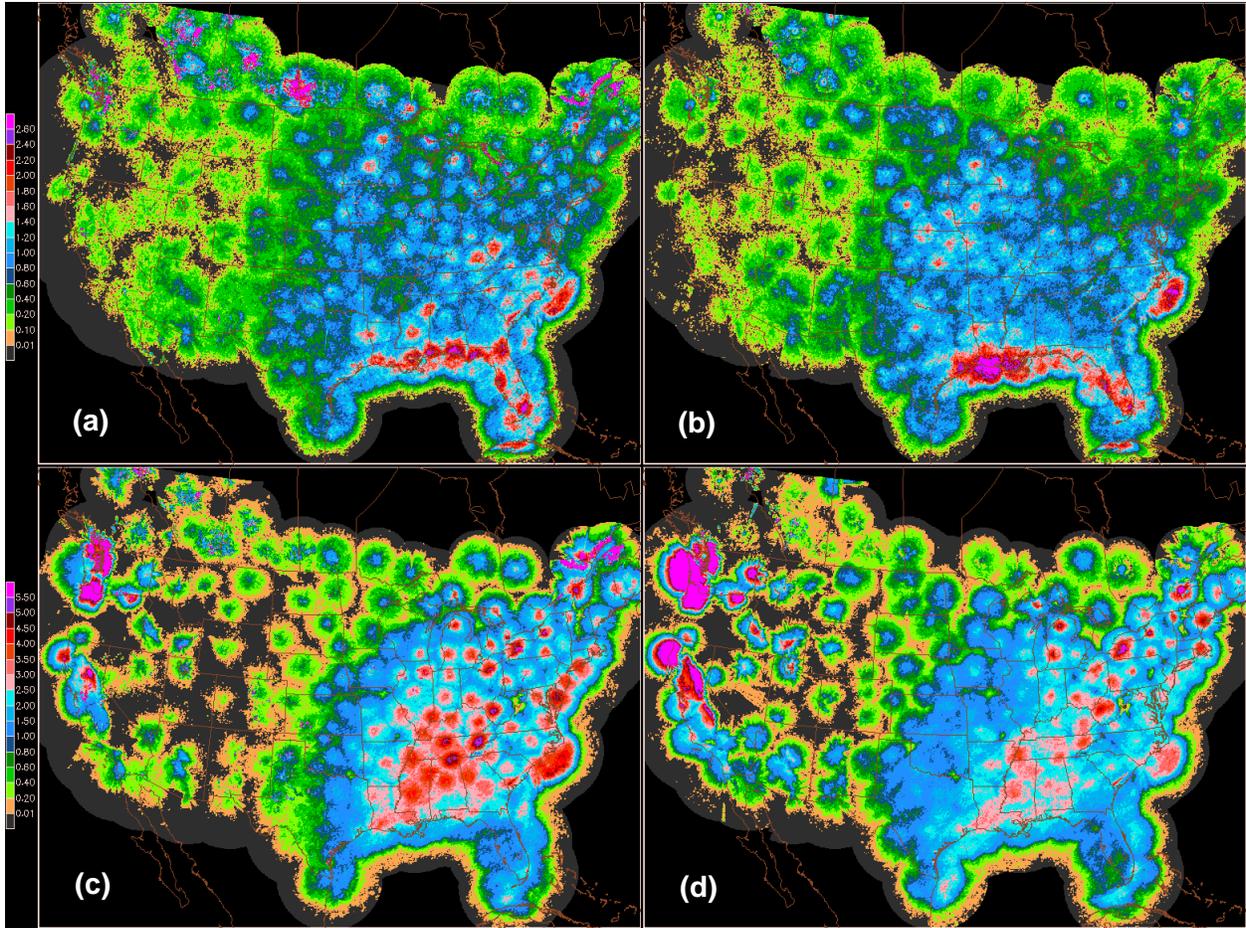
Figure 5. Relative frequency (RF; %) of max CREF ≥ 40 dBZ in 5-km grid boxes for the (a) 2012 warm season (April – September) and (b) 2016 warm season. (c) As in (a) except for ≥ 30 dBZ and the 2012-2013 cool season (October – March) and (d) as in (c) for 2016-2017.
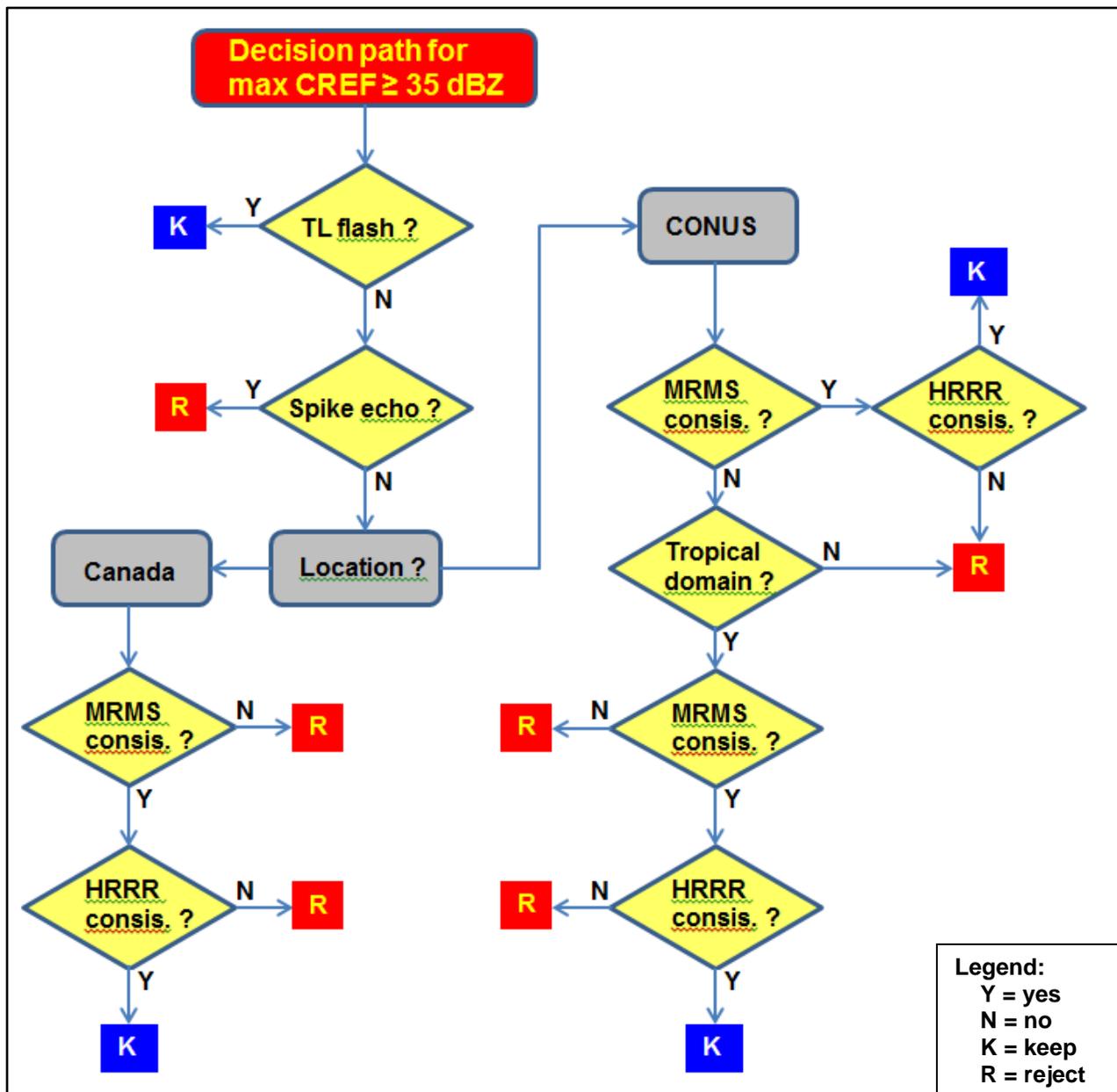
Figure 6. Dynamic QC decision tree, which is applied to maximum composite reflectivity (max CREF) ≥ 35 dBZ in a 5-km grid box. Diamond shapes with yellow fill denote individual tree decision modules, where "consis." is an abbreviation for consistency. See Figure 7 and text for details pertaining to each decision module.

TL flash ?   At least one total lightning (TL) flash in 15-km square centered on grid point ?

Spike echo ?   Do 7 of 8 adjacent gridboxes have max CREF < 30% of central value (for central max CREF ≥ 45 dBZ) ?

MRMS consis. ?   Is max CREF (≥ 35 dBZ) consistent with max VIL and STP criteria?
- Separate VIL and STP criteria apply to 5 max CREF intervals
- Separate VIL and STP criteria apply to warm and cool seasons, CONUS, and Canada geographical areas
- Separate VIL and STP criteria are used for CONUS domain and CONUS-tropical sub-domain

HRRR consis. ?   Is MRMS max CREF consistent with HRRR CREF values within 200 km of grid point?
- Consistency met where ≥ 8 grid points satisfy HRRR minimum CREF threshold
  - Both MRMS and HRRR CREF values must satisfy minimum thresholds, which have geographical and seasonal variations

Tropical domain ?   Are conditions for tropical sub-domain met?
- Date falls within May 15 – November 30
- Grid point lies within "tropical" region of CONUS
- GFS precipitable water ≥ 45 mm and lifted index ≥ 0 deg C within 300 km of grid point

Figure 7. Details pertaining to each decision module in Figure 6. Note that a "grid point" is centered on a 5-km gridbox; gridboxes are used in the MRMS data gridding (see text). The geographical domains for which MRMS and HRRR consistency criteria are stratified are shown in Figure 8.
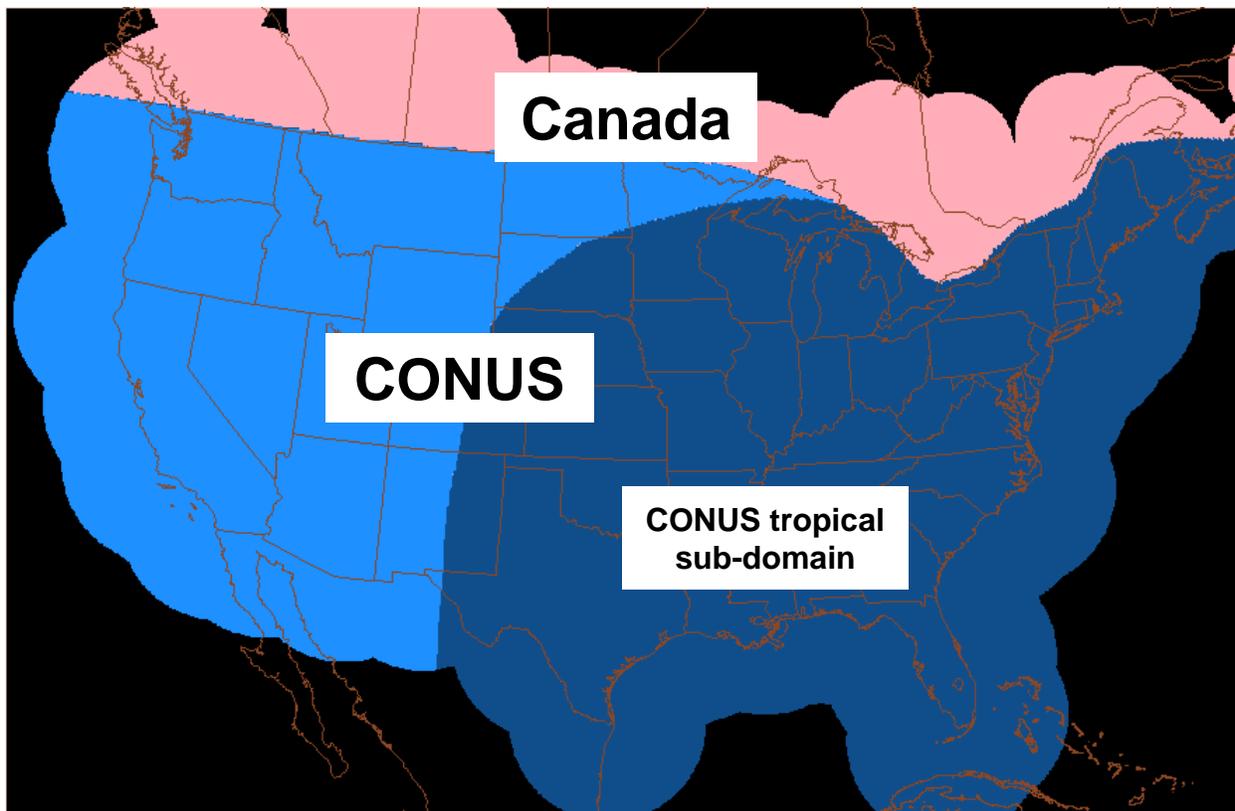
Figure 8. Geographical domains by which MRMS and HRRR consistency criteria in Figure 7 are stratified.
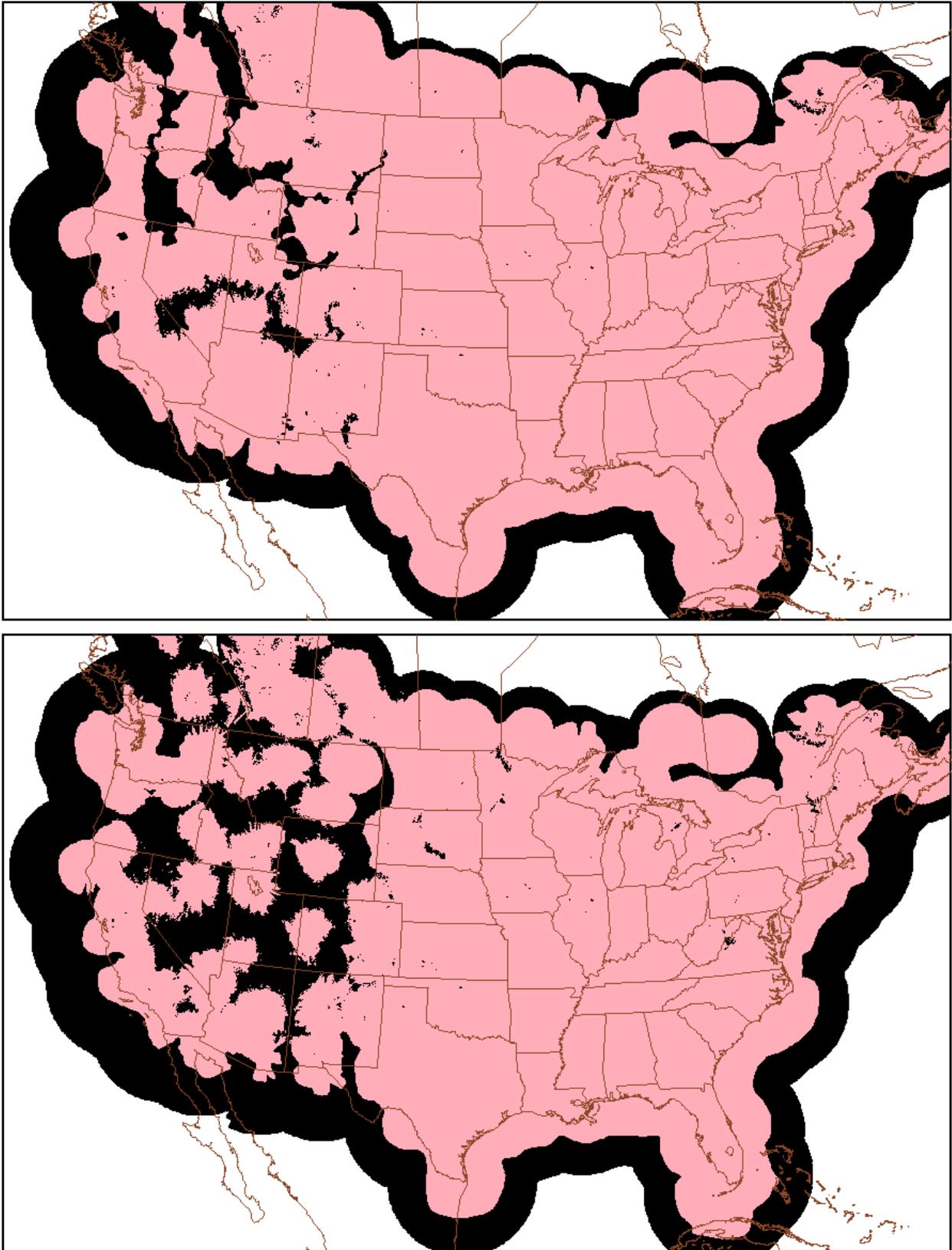
Figure 9. Warm season (top) and cool season (bottom) static QC grid mask (black shading).
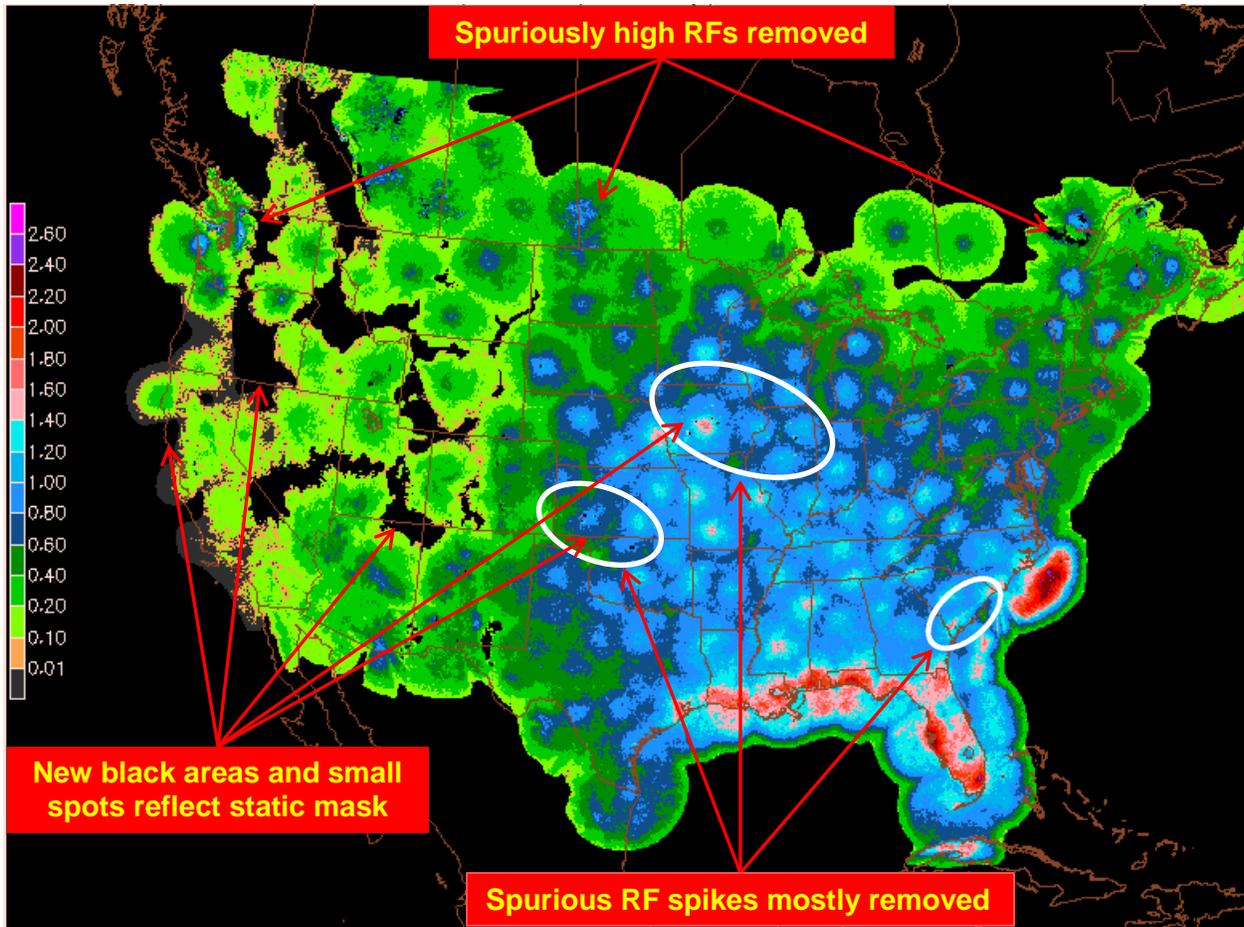
Figure. 10. Relative frequency (%) of CREF ≥ 40 dBZ for warm season after dynamic QC and static QC processes are applied to the base 5-km max CREF grids. Annotations are discussed in the text.
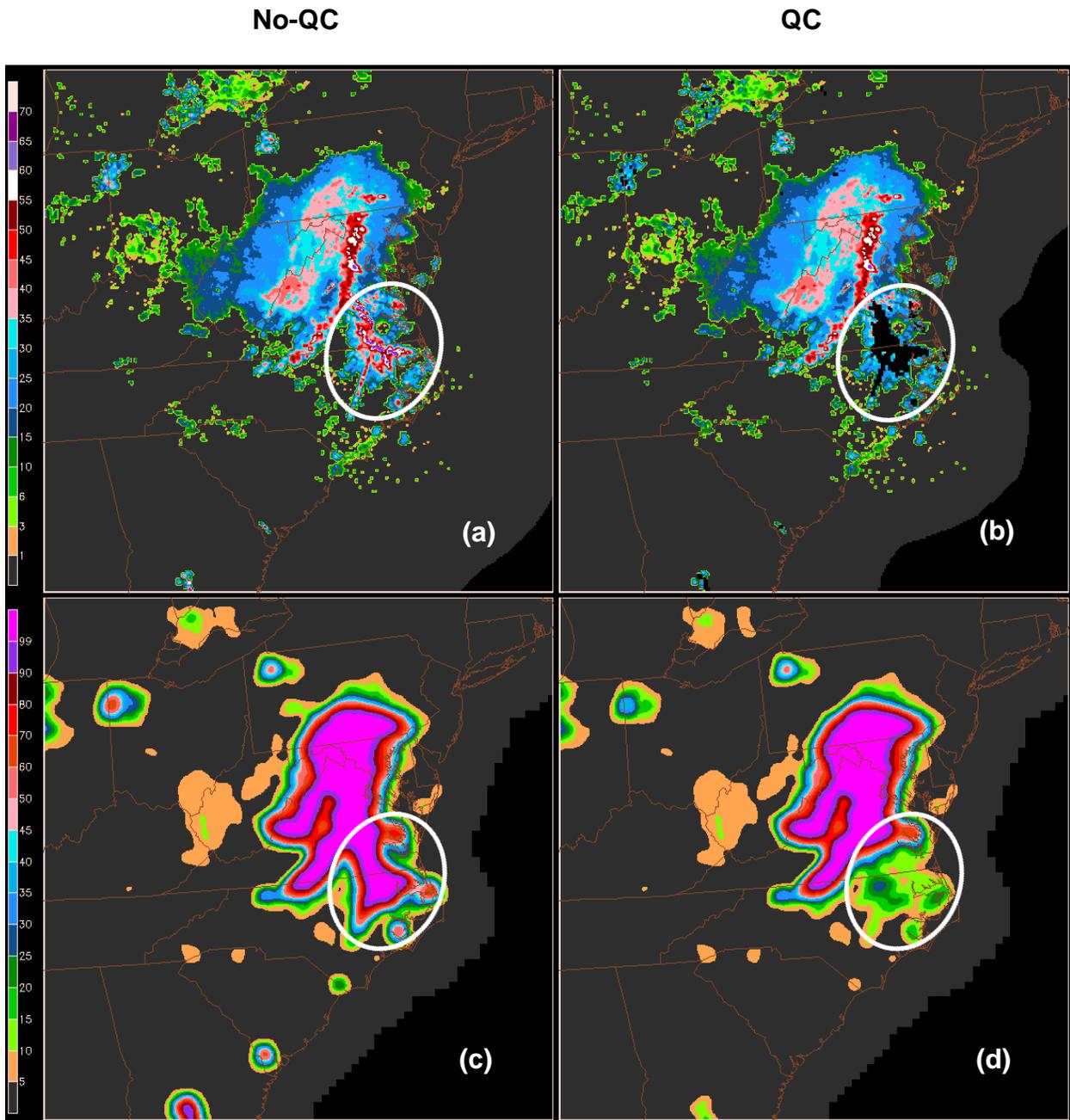
No-QC　　　　　　　　　　　　　　　　**QC**

Figure. 11. (a) No-QC MRMS 5-km max CREF (dBZ) valid for 0215 UTC on 30 June 2012, (b) same as (a) except QC CREF, where dynamic QC masking of AP echoes ≥ 35 dBZ over southeast Virginia and northeast North Carolina is depicted as black shading, (c) 0-1 h LAMP convection probability valid 1200-1300 UTC from the 1200 UTC LAMP cycle [uses no-QC CREF input from (a)], and (d) as in (c) except QC CREF input from (b) following HRRR CREF replacement of missing (rejected) MRMS CREF.values (HRRR replacement CREF not shown). See text for explanation of "no-QC" and "QC".
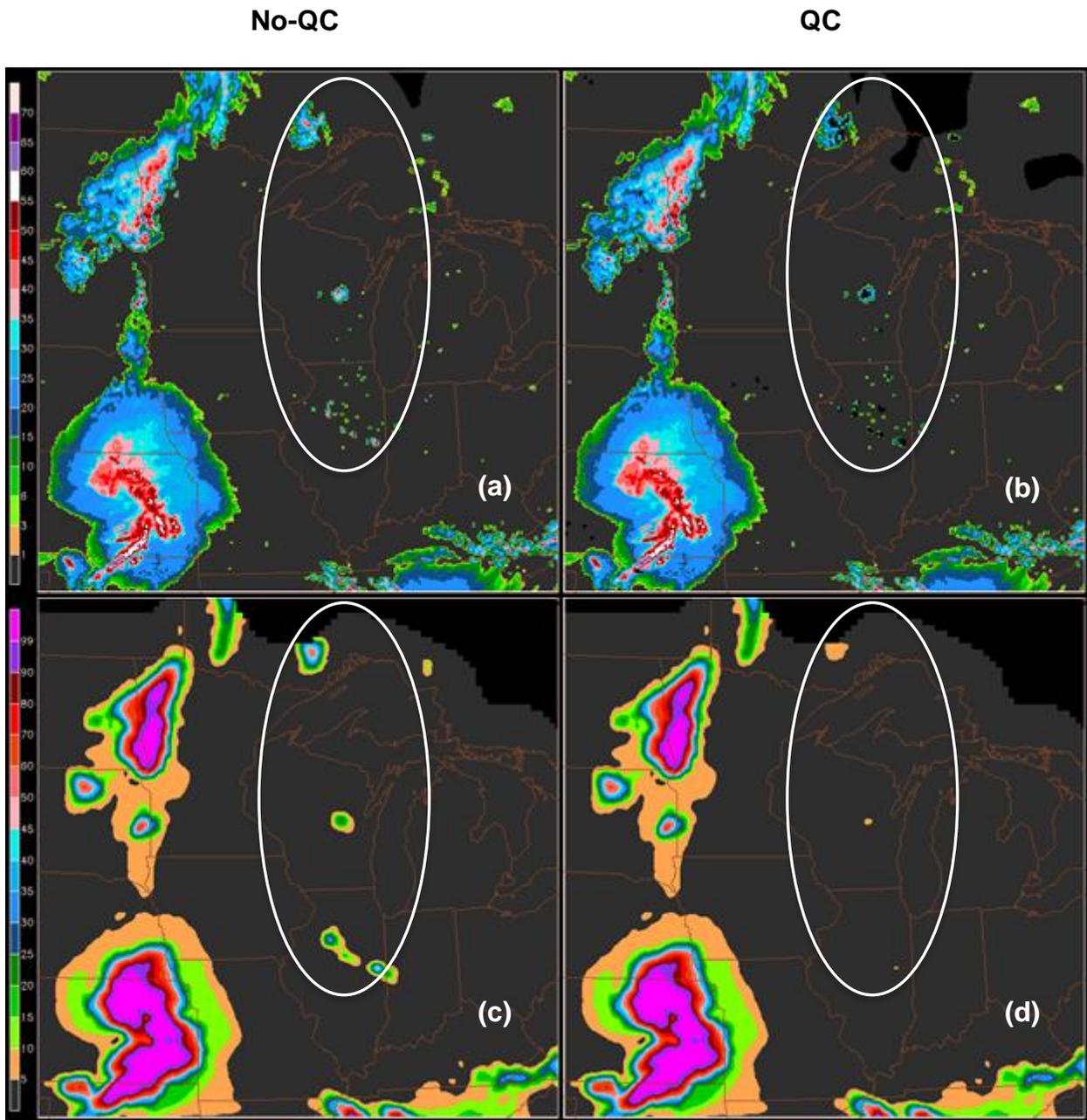
No-QC   **QC**

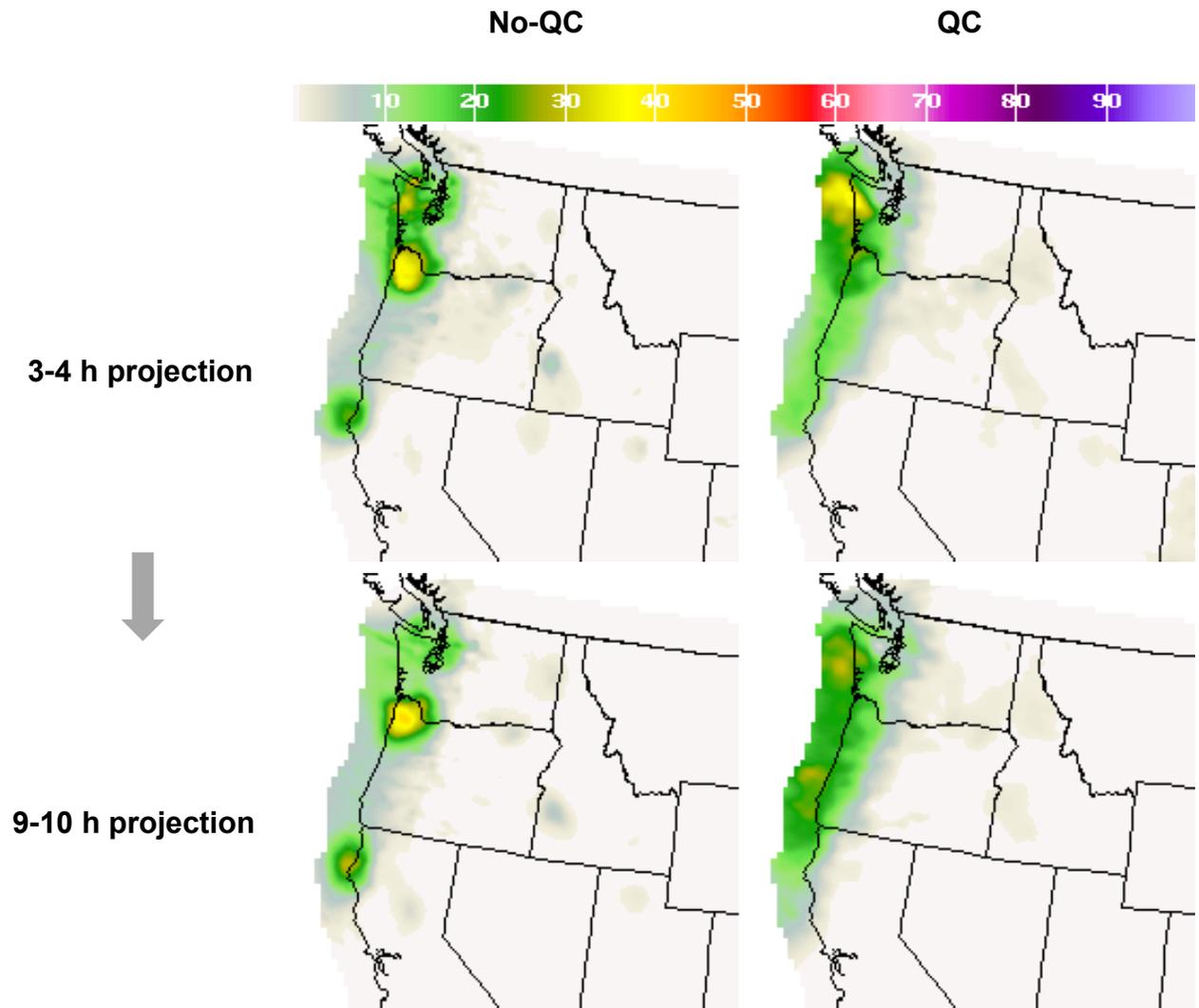Figure 12. Same as in Fig. 11, except for 05 June 2014.

Figure 13. No-QC (left) and QC (right) LAMP 1-h convection probability (%) for the 3-4 h forecast projection (top) and 9-10 h forecast projection (bottom) from 1800 UTC cycle on 23 December 2015. See text for explanation of "no-QC" and "QC" LAMP convection probability.
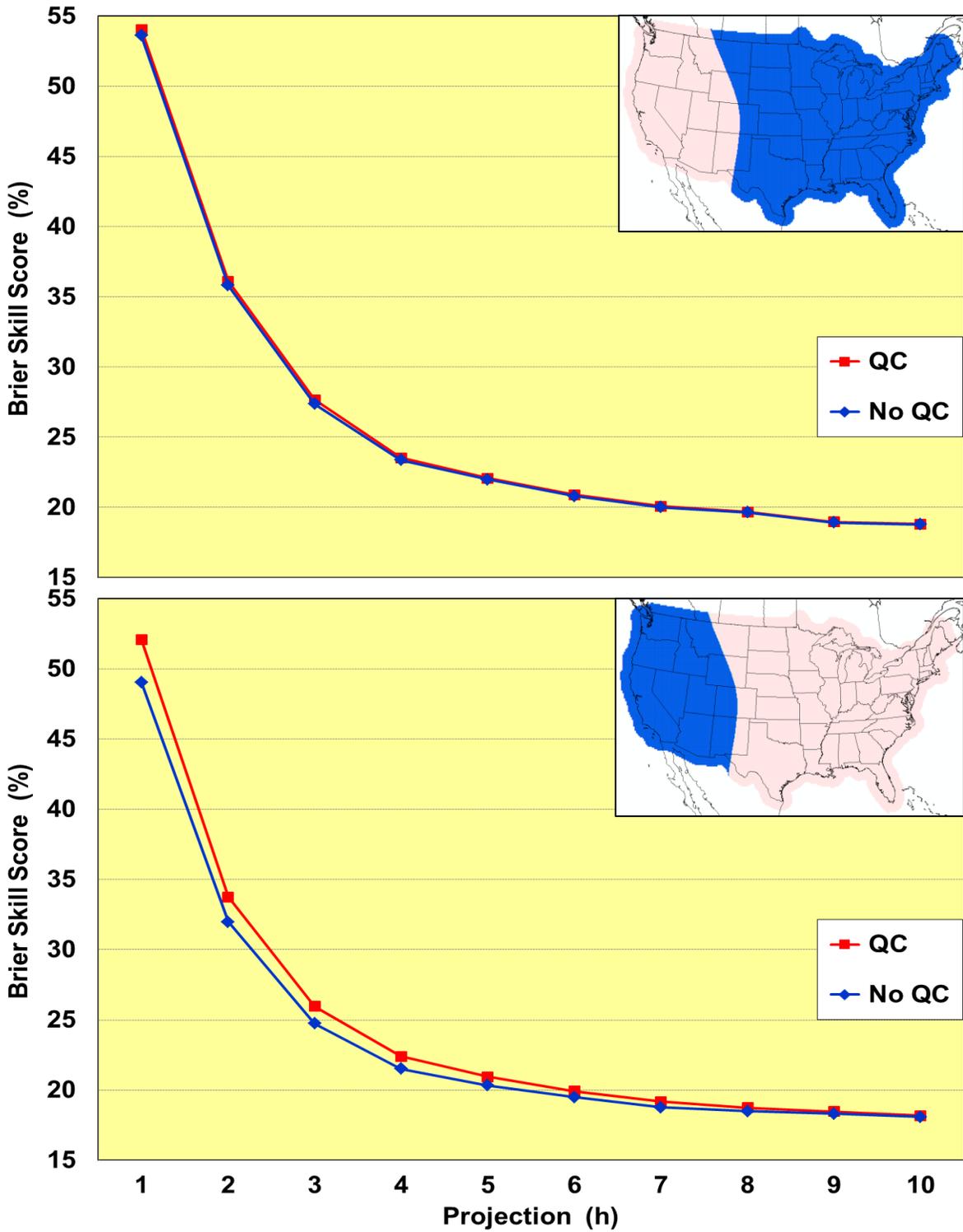
Figure. 14. Brier Skill Score versus forecast projection for QC and no-QC LAMP convection probability forecasts for the 0600 and 1800 UTC cycles (combined), where the scores are shown separately for eastern (top) and western (bottom) sub-divisons of the CONUS forecast area (shown with blue shading in the map insets). See text for explanation of "QC" and "no-QC" LAMP convection probability.