1.2 LAMP CONVECTION PROBABILITY AND "POTENTIAL" GUIDANCE: AN EXPERIMENTAL HI-RES UPGRADE

Jerome P. Charba*
Frederick G. Samplatsky
Phillip E. Shafer
Judy E. Ghirardelli
National Weather Service, NOAA
Office of Science and Technology Integration
Meteorological Development Laboratory
Silver Spring, Maryland

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

1. INTRODUCTION

Localized Aviation MOS Program (LAMP) convection probability and "potential" guidance forecasts have been operational in the National Weather Service (NWS) since April 2014 (http://www.nws.noaa.gov/mdl/gfslamp/convection.php; Charba et al. 2011, hereafter cited as CSS). Therein LAMP convection is defined as the occurrence of ≥ 40 dBZ radar reflectivity and/or one or more cloud-to-ground (CG) lightning flashes over a 2-h period in a 20-km grid box. Forecasts are issued every hour for projections in the 1- 25 hour range for the contiguous United States (CONUS).

LAMP convection guidance was developed for various applications in the weather enterprise, though targeted for aviation operations interests. Feedback from various users has indicated significant field usage, but some aviation-oriented users have indicated a need for higher spatial and temporal resolution. This is not surprising since fine spatial and temporal detail along with acceptable forecast skill is important, especially for inflight operations and various aviation operations at airports. So, a primary goal of this study is to increase the spatial and temporal resolution of the LAMP convection guidance. Note that as resolution in forecasts increases, skill generally decreases since the targeted event becomes "smaller" and thus more difficult to accurately forecast. Thus, an equally important goal is that forecast skill for the new "hi-res" product should be at least comparable to that for the current product. This article focuses on how new datasets and new techniques are applied for achieving increased resolution and skill in the new, experimental LAMP convection product

* *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

2. UPGRADED CONVECTION PREDICTAND

2.1 Attributes of the Upgraded Predictand

The convection predictand was upgraded in several ways. (1) The temporal resolution was doubled, as the valid period was reduced from two hours to one hour. (2) The spatial resolution was also doubled, as the base grid mesh was reduced from 20 km to 10 km. At the same time, the valid area (a 20-km square) of the convection event remained unchanged. While this results in 10-km overlap of neighboring predictand grid boxes, the overlap does not adversely impact the convection definition[1]. (3) The lightning dataset, which comprises one of two databases used jointly to specify the convection predictand, was upgraded from CG flashes to total lightning (TL) flashes [consisting of both CG and in-cloud (IC) flashes]. The new TL data are from the Earth Networks Total Lightning Network, which was recently developed and implemented by Earth Networks, Inc. (4) The radar reflectivity database, which is used together with the TL database in the predictand specification,

[1]The rationale for maintaining the 20-km valid area is to avoid an undesirable reduction in convection occurrence relative frequency that inherently accompanies the 50% valid period reduction.

was upgraded from obsolete, coarse-resolution RCM (Radar Coded Messages) reflectivity data (OFCM 1991; Kitzmiller et al. 2002) to high resolution radar reflectivity data from the Multi-Radar Multi-Sensor System (MRMS; http://nmq.ou.edu/, Zhang, et al. 2011; 2015), which became operational in the NWS in October 2014.

Note that radar reflectivity data have a greater role in specifying convection occurrences than lightning flashes, as reflectivity of ≥ 40 dBZ (a proxy for convection occurrence) generally occurs more often than lightning flashes. Unfortunately, reflectivity data are notorious for being contaminated with various types of random and systematic error. In automated applications of radar data, these errors may be ultimately transferred to the product of the application. Thus, the quality of the MRMS data was carefully examined prior to its application here.

2.2 Quality Control (QC) of MRMS Reflectivity Grids: Needs/Benefits

Underlying our decision to examine the quality of the MRMS data is the foundational role these data have in this work. Also, in the lead author's previous experience with automated radar data applications, the data quality was found to have deficiencies which warranted the development and application of supplemental quality control procedures (Charba and Liang 2005). While the quality assurance procedures built into the MRMS dataset may be more comprehensive than for radar data used previously, perusal of fine grid composite reflectivity maps based on the MRMS archive used in this study revealed the presence of quality flaws similar to those encountered in the past.

The analysis of MRMS data quality began by first gridding the data, which consists of tabulating the maximum instantaneous composite reflectivity (1-km MRMS grid resolution) within 5-km square grid boxes on a Lambert Conformal Conic grid[2]. These 5-km grids were produced at 15-min intervals for the period spanning January 2012 to March 2015.

With this grid archive, maps of seasonal relative frequency for various thresholds of reflectivity were developed. Figure 1a shows such a radar echo climatology map for the ≥ 40 dBZ composite reflectivity threshold based on all hh:00 times (hh = 00, 01,…, 23 UTC) for warm season months (April - September) of 2012 - 2014. The map shows expected relative frequency (RF) "flower patterns" around individual radar sites, which results from the inherent radar-range dependency of precipitation echo detection. The map also shows areas of unrealistically low RFs (in gray color) mostly across the western U.S., which result from radar beam blockage and "precipitation overshooting" in this mountainous region. Such anomalously low RFs also appear in locations of extended radar range not only along the perimeter of the network but even in interior locations of widely-spaced radar sites. From past experience with similar echo climatology maps (Charba and Liang 2005), these low RF anomalies were expected.

What we did not expect are large areas of spuriously high RFs (magenta areas in Fig. 1a), which appear throughout southern Canada and into Washington. Examination of individual grids revealed these high RF anomalies arise from the frequent presence of anomalous propagation (AP) echoes. We also did not expect to find small "spikes" of spurious RF maxima, which can be seen in the magnified rectangular area of the central U.S. In individual 5-km maximum reflectivity grids these RF spikes coincide with stationary "spike echoes," which likely result from radar detection of wind farms.

These types of systematic reflectivity errors together with random, occasionally-occurring false echoes (noted earlier) motivated us to pursue the unenviable, arduous task of developing automated processes to detect and remove such error from the 5-km maximum composite reflectivity grids. This effort is referred to as the Meteorological Development Laboratory (MDL) automated supplemental quality control (QC) of MRMS data.

The MDL QC processes consist of two basic components. The first process, which consists of detection and screening of suspected false echoes in individual grids, is called dynamic QC. It is comprised of an extensive series of computer-code consistency checks between the 5-km maximum composite reflectivity grids and corresponding grids based on other MRMS reflectivity product parameters, TL flashes, and fine scale and large scale numerical model output. A subsequent process, which consists of automated "grid masking" for a pre-determined set of grid points, is called static QC. While even a cursory description of the dynamic and static QC processes is beyond the scope of this article (a separate article devoted to

[2]5-km grid boxes yielded better data error diagnoses than smaller or larger boxes

this topic is pending), the benefits resulting from its application (briefly discussed below) seem to justify its development.

In particular, comparing the after-QC echo climatology map (Fig. 1b) with the before-QC counterpart (Fig. 1a) we find complete removal of the spuriously high RFs across the northern portion of the map and strong mitigation of spike RFs in the central U.S. These QC benefits resulted from throw-out of suspected AP echoes by the dynamic QC process. Also, close comparison of the before- and after-QC RFs (Figs. 1a and 1b) reveals a very slight reduction in RF magnitude over the entire radar coverage domain, which results from throw-out of randomly-occurring suspected false echoes. Careful inspection of individual grids with echo throw-outs reveals (a) most tossed echoes are obviously false, (b) a small fraction of truly-false echoes were not thrown out, and (c) and a much smaller fraction of tossed echoes were actually valid echoes (not shown). Note that the reduction in magnitude of the RFs all across the map in Fig. 1b is quite small, which suggests the fraction of tossed valid echoes is probably negligible.

Of course, the value of the MDL QC ultimately resides in whether it yields improved LAMP convection forecasts, i.e., the MRMS data application at hand. This question is briefly addressed in section 5.4 of this article.

2.3 Specification of Upgraded Predictand

Following application of the MDL supplemental QC to the 5-km maximum composite reflectivity grids, these grids (along with corresponding TL flash count grids) were used to specify the upgraded convection predictand. The first step consists of tabulating the maximum reflectivity for each 5-km grid box across the four 15-min times falling within the 1-h valid period. The corresponding TL component is obtained by tabulating the TL flash count within the same 5-km grid boxes for the 1-h period. Then, the maximum reflectivity or TL flash count is obtained for 20-km grid boxes by scanning over 4x4 5-km reflectivity or flash count sub-grids. Note that these tabulation processes are performed at alternate 5-km grid points to obtain 20-km grid box values centered on grid points spaced 10 km apart. Thus, 20-km grid boxes for adjacent 10-km grid points overlap by 10 km.

The second step of the convection predictand specification involves combining the reflectivity and TL components. Thus, the predictand for a 20-km grid box is specified as a convection occurrence (1 value) if either (or both) the reflectivity is ≥ 40 dBZ or the TL flash count is greater than zero; otherwise the convection predictand is specified as a non-occurrence (0 value). Note that this predictand specification is not impacted by the 10-km overlap of the neighboring predictand grid boxes.

3. UPGRADED CONVECTION PREDICTORS

3.1 Upgraded Predictor Databases

The upgraded convection predictor databases consist of three types. (1) Fine scale observational data (OBS), which consist of 5-km grids of the most recent QC'd MRMS reflectivity and TL flash count parameters (the same gridded databases used for the convection predictand specification discussed in section 2). (2) Fine scale model output from the HRRR (High Resolution Rapid Refresh model; Smith et al. 2008; Benjamin et al. 2016), which became operational in the NWS in September 2014 (http://rapidrefresh.noaa.gov/hrrr/; http://ruc.noaa.gov/pdf/RAPHRRR_WCOSS_2016Q1_Final-sb-12oct2015.pdf). (3) Large scale model forecasts comprised of GFS- (Global Forecast System; Kanamitsu et al., 1991) and NAM- (North American Mesoscale Model; Rodgers, et al., 2005) based Model Output Statistics (MOS) 1-h convection probabilities. Note that the MOS convection predictand and associated 10-km forecast grid are the same as for LAMP.

3.2 Specifying OBS and HRRR Predictor Grids

The specification of OBS and HRRR predictor variables on the LAMP 10-km forecast grid is comprised of a two-step process. The first step involves use of 5-km base (working) grids of the type noted above. In the case of the TL data, these working grids contain either 30- or 60-min TL flash counts for the 5-km grid boxes. Then, to specify a given TL flash count predictor variable on the 10-km grid (section 3.4), these flash counts are summed over 2x2 sub-grids of 5-km grid boxes to obtain counts for 10-km grid boxes.

The 5-km working grids used for deriving MRMS predictors on the 10-km grid (section 3.4) are the output grids from the MDL supplemental QC of the MRMS data. Recall from section 2.2, the maximum pixel value of an MRMS reflectivity product variable was tabulated for the 5-km grid box to fully capture indications of convection occurrences by CONUS network radars. The corresponding 5-km grid box tabulation procedure for

HRRR variables is identical, except the mean (rather than the maximum) over all native 3-km HRRR grid point values is tabulated. [The averaging (i.e., smoothing) is done to mitigate possible random error in the HRRR forecasts.] Finally, for each of the MRMS and HRRR predictor variables, the value tabulated for the 10-km grid box is the maximum value over the 2x2 sub-grid of 5-km grid box values. Thus, the MRMS value is the actual maximum native grid point value within the 10-km grid box, whereas the corresponding HRRR value is an area-averaged maximum.

Note that the use of the 5-km working grids for specification of the TL predictor variables was motivated by programming convenience (these 10-km predictor grids could have been specified without use of the 5-km working grids), whereas for MRMS variables the MDL supplemental QC procedure dictated use of the working grids. For the HRRR variables, on the other hand, the use of the 5-km grids provided a convenient mechanism for (slight) areal smoothing of the fine scale HRRR grids. Finally, note that use of all TL observations and all native MRMS and HRRR grid point values in the LAMP predictor specification may optimize predictor usage of these fine scale datasets.

3.3 Advection of OBS Variables

The OBS candidate predictor variables consist of various radar reflectivity product and TL flash count parameters (see following sub-section) based on the most recent observations. In addition to these "persisted-observations" predictors, simple advection is applied to the OBS grids using GFS lower tropospheric wind forecasts as the advecting wind. The advection program used herein is identical to that described in (Glahn and Unger 1986, pp. 1319-20), with one exception. Since the program operates on a 10-km Polar Stereographic grid, it is necessary to interpolate the OBS variables from the LAMP 10-km Lambert grid to the Polar Stereographic grid prior to the advection, and then perform the reverse interpolation afterwards. Visual comparison of non-interpolated and interpolated advected grid maps suggested the very slight impact of the interpolation should have a negligible effect on the quality of the advected predictors.

3.4 Candidate OBS, HRRR, and MOS Convection Predictors

A complete list of candidate OBS, HRRR, and MOS convection predictor variables is shown in Table 1. Several comments are noteworthy. (1) The OBS variables consist of "initial" and "advected" types, where "initial" denotes persistence of the most recent observation and "advected" refers to the application of the advection process noted above to "synchronize the observation" with the ending time of the 1-h predictand valid period. (2) All of the HRRR variables in Table 1 are obtained directly from an "experimental archive"[3] provided by NOAA/Earth System Research Laboratory/Global Systems Division, except moisture divergence which is computed from HRRR 10-m wind and 2-m specific humidity. (3) While the HRRR model (as for LAMP) runs on an hourly cycle, the HRRR ingest into a given LAMP cycle is from the 1-h old HRRR cycle, which is necessary (for real time application) to account for an 80-85 min latency of HRRR model output. Also, because the longest forecast projection from the operational HRRR model is just 15 hours, we use "persisted" 15-h HRRR forecasts for LAMP projections in the 15-25 h range. While the predictive value of these persisted HRRR forecasts is obviously limited, especially for the upper end of this LAMP projection range, their use was judged to be a better option than the alternative of no HRRR predictor input. (4) Light grid smoothing was applied for all variables in the table, the degree of which is specific to each variable.

4. DEVELOPMENT OF UPGRADED CONVECTION REGRESSION EQUATIONS

4.1 New Aspects of Regression Equation Development

As in CSS, the 1-h LAMP convection probabilities are produced by multiple linear regression equations based on the new 1-h convection predictand discussed in section 2 and candidate predictors in Table 1. The historical sample, most of which was used for the equation development and the remainder for testing, spans 01 January 2012 to 30 September 2015. Note that separate regression equations, each with a unique set of predictors, are developed for each hourly LAMP cycle, each of three LAMP seasons, and each of seven geographical regions over the CONUS. The LAMP seasons are "spring" which applies to March 16 to June 30, "summer" to July 01 to October 15, and "cool" to October 16 to March 15. The seven regions are shown in Fig. 2, which is a reduction

[3] The "experimental archive" is based on the most current experimental version of the HRRR model for each date-time contained in the archive (http://rapidrefresh.noaa.gov/hrrr/).

from 13 used in CSS. The fewer (larger) regions used here offsets the sample size reduction resulting from the shorter development sample. Note that the regions have slight overlap, which is necessary to avoid regional discontinuities in the convection probabilities (Charba and Samplatsky 2011).

4.2 Ranking of Convection Predictors

The regression equations are developed such that a fixed set of predictors is used across the entire 1-25 h LAMP forecast range, which enhances consistency in the probabilities therein. Preliminary (forecast skill) testing of equations with varying numbers of predictors revealed that 20 is an optimum number.

It is useful to examine the ranking (i.e., the order of selection in the screening regression process) of predictors in the regression equations, as a predictor's rank is directly related to its nominal contribution to the forecast probability. The objective ranking scheme involves assigning a rank number to a predictor to reflect its order of selection. Since the equations contain 20 predictors, the assigned rank number is 20 (1) where a predictor was selected first (last) and accordingly for predictors selected between these bounds. Recall that while a given LAMP cycle and season involves seven regional equations (sets of predictors), for brevity the predictor ranking is discussed where the individual predictor rankings are summed over the seven equations for a combined-region ranking. Also, the discussion is mostly limited to relative ranking trends of the three basic predictor types (OBS, HRRR, and MOS in Table 1) across the three LAMP seasons and four LAMP cycles (00, 06, 12, and 18 UTC) developed to date.

A major predictor trend is that the OBS variables (both initial and advected) generally have top ranking across the four LAMP cycles. Note the contribution of the OBS variables to the probabilities is mainly in the 1-to-4 hour LAMP forecast projection range (supporting evidence is provided in the following section). Also, within the OBS group the MRMS variables generally have higher ranking than the TL variables, where the main exception is for the 00 UTC cycle during spring and summer where a TL predictor has top ranking. The latter finding is consistent with the relatively high lightning frequency at this time of the day during the warm season of the year. The typically-higher ranking of MRMS predictors is consistent with generally higher frequency of ≥ 40 dBZ radar echoes than lightning flashes. Note that for the LAMP cool season, when the nation-wide seasonal lightning frequency is at a minimum, we find the CONUS ranking of the TL predictors is especially low.

Another major trend is that the HRRR and MOS predictors have roughly equal ranking across all three seasons and four cycles. Further, we find a clear trend that both types of model forecasts have a higher ranking during summer than during the cool or spring seasons. Conversely, the weak relative ranking of OBS predictors during summer is consistent with the predominance of relatively short space and time scales of convective systems during that time of the year, i.e., short space-time scales likely weaken the predictive value of persisted and advected observations.

Note also that while the (fine-scale) HRRR and (large-scale) MOS predictors have generally similar overall ranking they each have opposite effects on spatial detail in individual probability maps. Specifically, the HRRR predictors infuse fine detail in the probability fields, while the MOS variables control the large scale variability. Still, fine detail contributed by the HRRR predictors is mostly limited to 4-16 h LAMP forecast projections. For shorter LAMP projections, detail and skill in the LAMP probabilities is provided mainly by OBS predictors. For LAMP projections of 16 hours and higher, the weak contribution of the HRRR predictors is undoubtedly due to the limited predictive value of the persisted 15-h HRRR forecasts in this range (see section 3.4). Thus, for long LAMP projections, where the predictive contribution of the OBS variables is likely negligible and the HRRR contribution is small, the convection probabilities are almost entirely controlled by the large scale MOS variables. (Each of these predictor impact assertions is supported by findings from forecast performance assessments, which are discussed in the following section.)

5. FORECAST PERFORMANCE OF THE UPGRADED CONVECTION PROBABILITIES

In this section, the performance of the new 1-h convection probabilities is examined from several perspectives. In section 5.1 regional variations in the performance are explored to provide insights into the quality of the convection predictand and predictors for different geographical locations. In section 5.2 the CONUS-wide quantitative skill for the new LAMP 1-h probabilities is compared against the corresponding skill for 1-h MOS and the LAMP operational 2-h probabilities. Lastly, in

section 5.3, the forecast performances of the 2-h and 1-h probabilities are compared subjectively via a case study to address forecast performance from a field user perspective. The quantitative skill measure used is the Brier Skill Score (BSS), which is defined as the percentage improvement in ½ the Brier Score (Brier 1950) for the convection probabilities over the corresponding score for convection relative frequency (Wilks 2006, pp 284-285).

5.1 Geographical Variations in Probability Performance

Figure 3 shows the seasonal BSS versus LAMP projection for geographical regions shown in Fig. 2. The developmental sample is used here, as this sample is sufficiently long to provide robust regional scores. Note that these BSS's are not true skill measures as they are based on dependent samples.

In Fig. 3 we see substantial regional diversity in BSS over the three seasons: the diversity is quite strong for the cool season, moderate for spring, and weak for summer. Note that the BSS's for the two western regions (Pacific Coast and Rocky Mountain regions) are rather low for both the cool and spring seasons. To a large degree this may be a reflection of the poor quality of the MRMS data in this geographical area. Note that the quality of the convection predictand and OBS predictors in this area are largely dependent on quality of the radar reflectivity data, as lightning flashes (where the quality of the underlying database is higher) are relatively rare in this area during the LAMP cool and spring seasons.

Contrastingly, the BSS's are relatively high for the Central Plains region throughout the year and for the Southeast region (Fig. 3) during the cool season, which suggests a BSS linkage to the predominant convection spatial scales. That is, convection tends to occur on relatively large scales in the central U.S. throughout the year, whereas in the southeastern U.S. that is true mainly during cool season. Lastly, the relatively low BSS's for most regions during the summer are consistent with the predominance of smaller convection spatial scales during this season compared to other seasons. Note that an awareness of these regional and seasonal convection probability performance trends should be especially beneficial to field users of these guidance probabilities.

5.2 Conus-Wide Probability Skill

The skill of the convection probabilities was examined for the 18 UTC LAMP cycle, where the most recent 30 days of each seasonal sample was withheld from the regression equation development. The BSS for these seasonal samples is plotted as a function of LAMP forecast projection in Figs. 4a - 4c. The skill curves for the 1-h LAMP convection probabilities feature quite high skill for the 1-h forecast projection for all three seasons[4]. Following this initial skill maximum, the BSS's exhibit a steep fall to about 4 hours, they level off from that point out to roughly 16 hours, and thereafter they show a gradual fall-off. Note that BSS curves for the corresponding developmental samples are quite similar to these, as are corresponding curves for the 00, 06, and 12 UTC cycles (not shown).

The characteristic shape of the LAMP BSS curves in Fig. 4 is controlled by the diverse relative importance of the OBS, HRRR, and MOS predictors across the LAMP forecast projections. Specifically, the very high skill at the 1-h projection reflects the high predictive effectiveness of the initial and advected OBS variables. Thereafter, the sharp skill drop to 4 hours results from the rapid loss in predictive value of these predictors combined with a minimal contribution from HRRR predictors. Contrastingly, the contribution of HRRR predictors is evidently substantial in the 4-16 h forecast range, as indicated by the strong improvement in BSS's of the 1-h probabilities over BSS's for the 2-h probabilities in this forecast range. (Note that HRRR predictors are not used for the 2-h probabilities; also see a qualification for this skill comparison below.) Lastly, the slow skill fall-off with projection beyond 16 hours likely reflects the weak predictive value of persisted 15-h HRRR forecasts and thus a near-total skill reliance on MOS predictors.

Included in Figs. 4a – 4c are corresponding BSS curves for the GFS-based and NAM-based MOS 1-h convection probabilities as well as the corresponding BSS curves for LAMP operational 2-h convection probabilities. However, a quantita-

[4]Note that while the MRMS and TL data cutoff time for this LAMP cycle is 1800 UTC, the LAMP convection probabilities are not available until after 1830 UTC, which makes these probabilities largely a nowcast. The delay in availability of the probabilities is due to latency of HRRR predictors from the 1700 UTC HRRR cycle (see section 3.4).

tive skill comparison of the 1-h LAMP probabilities against the 1-h MOS probabilities is not valid because the verification sample at hand was included in the sample used for development of the MOS regression equations and thus the MOS samples are not "independent." A quantitative skill comparison with the 2-h LAMP probabilities is also not valid because the 1-h and 2-h convection predictands are substantially different from one another (see section 2.3). Nevertheless, a qualitative comparison of scores of the 1-h and 2-h convection probabilities is both valid and useful.

In Figs. 4a – 4c we see that BSS's for the 1-h probabilities are higher than those for the 2-h probabilities for all projections to at least 18 hours. Afterward the relative scores for the two probability products are mixed across the three seasons. Note that the improvement in the 1-h probability scores is substantial in the 4-16 hour range, reflecting the strong contribution of HRRR predictors there.

The 1-h LAMP versus 1-h MOS BSS comparison (Figs. 4a – 4c) shows a clear improvement of LAMP over MOS for all projections. Note that the improvement is especially strong for the shortest LAMP projections, which is expected because of the strong contribution of the OBS predictors, as noted above. What might not be expected is the small though clear improvement of LAMP on MOS for the longest forecast projections, where the contributions of the OBS and HRRR predictors is evidently negligible. This LAMP improvement is likely due to the summation of several individual benefits that result from predictor usage of the GFS-based and NAM-based MOS convection probabilities in the corresponding LAMP regression equations. One benefit is that the predictive skill of these two MOS predictors is combined in the LAMP regression equations, whereas the MOS BSS's in Fig. 4 apply to each MOS predictor individually. Further, these MOS predictors are used to derive supplementary MOS product predictors (see Table 1), which may result in additional LAMP convection probability skill. Each of these asserted LAMP skill benefits was demonstrated in conjunction with the 2-h operational LAMP convection probabilities in CSS.

Another skill benefit that may result from the incorporation of the MOS predictors in the LAMP equations is due to he "LAMP regional calibration of non-regionalized MOS convection probabilities." That is, the MOS regression equations were developed with the generalized operator approach (Glahn and Lowry 1972), wherein a single equation applies to all points in the CONUS domain. Then, as the MOS probabilities are incorporated as predictors in the regionalized LAMP equations, the MOS probabilities may effectively undergo regional calibration. The skill benefit that may result from such regional calibration was demonstrated for the 2-h operational probabilities in CSS.

Finally, note the true skill advantage of LAMP over MOS is expected to be larger than indicated in Figs. 4a – 4c, as the MOS BSS's are based on dependent samples while the LAMP samples are independent. Thus, a better indication of the true LAMP skill advantage is where both MOS and LAMP scores are based on the same dependent sample, as provided in Fig. 4d. Note that the LAMP BSS advantage here is larger than that shown in Figs. 4a – 4c. Also, similar results were obtained for other seasons for this LAMP cycle and for the three other LAMP cycles developed to date. Thus, the LAMP skill improvement on MOS appears substantial at all projections.

5.3 Case Study

From a field user perspective, the forecast performance improvement of the 1-h convection probabilities over the operational 2-h probabilities may be best judged by examining individual probability maps side-by-side. An example of such a comparison is shown in Fig. 5 for the selected case of 18 UTC 23 December 2015, where the 2-h probabilities are shown in the left panel and the 1-h probabilities to the right. The verifying map used here for both forecasts is the MRMS instantaneous composite reflectivity (from http://nmq.ou.edu/applications/qvs_2d_maps_main.html) valid near the temporal center of each predictand. Note that while the true verification map for each probability product (not shown) is substantially different than this surrogate, composite reflectivity is a principal constituent of each predictand and also this surrogate does not favor either forecast.

In Fig. 5 the 2-h and 1-h probability maps are compared for three LAMP forecast projections (the projection applies to the end of the valid period for each predictand). Note that the 1-h probability map for the 3-h (short range) projection (Fig. 5a, right panel) shows finer spatial detail than for the corresponding 2-h probability map (left panel) over the entire eastern U.S. At the same time peak probabilities are similar between the two maps despite the shorter valid period for the 1-h probabilities. Also, probabilities in the 50-100% range for

the 1-h map show better spatial focus and positioning for the intense convection line that extends from eastern Arkansas to southern Illinois. These 1-h probability improvements reflect upgrades to the radar and lightning predictors discussed in section 3.

For the corresponding 12-h (medium range) projection (Fig. 5b) the improved spatial focus for the 1-h probabilities is quite strong, especially for two intense echo lines labeled A and B in the verifying map. Also, the 1-h forecast map shows probabilities peaking near 40% over southern Lake Superior which coincides (in the verifying map) with reflectivites above the 40 dBZ convection occurrence threshold; contrastingly the corresponding 2-h map shows quite low (< 5%) probabilities there. Note further the 1-h probability pattern generally resembles the verifying MRMS map to a far better degree than for the 2-h probabilities. These findings evidently reflect the power of HRRR predictors in the 1-h probability regression equations.

For the 24-h (long range) projection (Fig. 5c), the 1-h and 2-h probability maps show a much closer resemblance to one another. This similarity arises as the 1-h probability map exhibits a loss in spatial detail, which is present for the 3- and 12-h projections. This loss in spatial resolution along with the substantial loss in probability skill evident in Figs. 4a – 4c reflects the marginal value of the persisted HRRR 15-h forecasts for this extended LAMP projection. *This finding points a critical need for a forecast range extension of operational HRRR forecasts beyond the present cutoff of 15 hours.*

5.4 Contribution of MDL MRMS QC to Convection Probability

In section 2.2 we broached the question whether the MDL supplemental QC of MRMS reflectivity data results in improved LAMP convection forecasts. Here, we very briefly address this question. This was done whereby we derived separate "before-QC" 1-h convection probability regression equations, where the only difference from the "after-QC" regression equations used in this study was that the MDL supplemental QC was not applied to the 5-km MRMS reflectivity product grids prior to specification of the convection predictand and MRMS candidate predictors.

In Fig. 6 before-QC and after-QC 1-h convection probabilities for a 1-h LAMP forecast projection are shown for a selected case in which (suspected) AP echoes were present in the base 5-km MRMS maximum reflectivity grid at the LAMP model initial (cycle) time (lower-left panel). These AP echoes, which appear within the oval area shown, are reflected in the "before-QC" 1-h convection forecast as small, widely scattered "probability spots," with peak probabilities near 50%. In this case the dynamic QC process (section 2.2) detected the AP echoes with reflectivities ≥ 35 dBZ and masked them out. This false echo removal is reflected in the after-QC forecast map as a near-complete elimination of the contaminated convection probabilities, while probabilities in neighboring areas are undisturbed. Thus, the QC resulted in small, meaningful improvement in the convection probabilities in this case.

Additionally, seasonal CONUS BSS's for the before-QC and after-QC 1-h convection probabilities were computed over the full developmental sample (01 January 2012 – 30 September 2015) for each of the three LAMP seasons and four LAMP cycles (00, 06, 12, and 18 UTC) addressed in this study. The upgraded 1-h convection predictand, which uses the QC'd MRMS maximum composite reflectivity grids discussed in section 2.2 was used as verifying data for both sets of forecasts. Preliminary BSS results obtained to date are presently being assessed and will be reported in a future paper.

6. CONVECTION POTENTIAL

In the previous section, we showed that skill of the convection probabilities varies strongly with forecast projection and geographical region. What we did not show (for brevity) is that the sharpness (i.e., range) of the probabilities exhibits similarly strong variations. The strong sharpness variability together with a common lack of understanding of forecasts expressed in probability form makes using the convection probabilities guidance rather challenging.

The conventional "remedy" to the problem is to convert the probabilities to categorical yes/no occurrence forecasts by applying a pre-determined threshold probability. At MDL, the threshold is specified objectively with an iterative scheme where the threat score (TS; same as Critical Success Index; Schaefer 1990) is maximized with bias restricted to a narrow range. For the operational LAMP 2-h convection probabilities, CSS extended the technique to produce four convection risk/threat categories (no, low, medium, and high)

instead of just the usual two "yes/no" categories. The four category product is called convection potential. The following sub-section expands on how convection potential is derived, and the subsequent sub-section discusses its performance.

6.1 Deriving and Interpreting Convection Potential

Principal features of convection potential are summarized in Table 2. Note that the three convection potential categories (low, medium, and high) are based on three pre-determined threshold probabilities, where the threshold probability value is relatively small for low potential, it is higher for medium potential, and it is highest for high potential. Also the prescribed high bias for low and above potential ("low potential threshold"; bias in the 2.70 – 2.83 range, where unbiased forecasts have a 1.0 value) implies strong convection over-forecasting. Just slight over-forecasting (1.03 – 1.13 bias range) is associated with the medium potential threshold, and strong under-forecasting (0.38 – 0.43 bias range) is associated with high potential.

A key feature of convection potential is that each of the three threshold probabilities (which are derived from the same sample as used for development of the regression equations) is derived separately for each LAMP cycle, forecast projection, season, and geographical region. Thus, the threshold probability associated with a given potential category varies with the LAMP model cycle, projection, etc. Also, since the bias associated with a given potential threshold is restricted to a narrow range, this threshold potential will exhibit about the same bias regardless of the LAMP cycle, projection, etc. For example, the bias associated with the medium potential threshold will be in the 1.03 – 1.13 range regardless of the LAMP model cycle, forecast projection, season, and geographical location. The recognition of this attribute of convection potential should benefit users of both the potential and the probabilities.

6.2 Example Convection Potential Forecast Maps

Example maps of 1-h (and 2-h) convection potential are shown in Fig. 7. The most obvious property seen here is the progressively decreasing areal coverage from the low potential threshold, to the medium potential threshold, and to high potential. This property is a reflection of the decreasing bias across the three potential thresholds discussed above.

Comparing the 1-h and 2-h convection potential maps (Fig. 7), we see general similarity between them, which is expected since potential is specified identically across the two predictands. On the other hand, greater spatial detail is present in the 1-h potential, which is a reflection of the increased spatial detail in the underlying convection probabilities (Fig. 5). In fact, this increased detail results in instances of a small map feature in the 1-h convection potential which is entirely absent in the corresponding 2-h potential map. Examples of this appear over Lake Superior at the 12-h projection and around Cape Cod at the 24-h projection.

Close cross-comparison of the 1-h probability and 1-h potential maps in Figs. 5 and 7, respectively, shows how the probabilities associated with a potential category vary across different map locations and forecast projections. For example, medium potential over Iowa at the 3-h projection is associated with probabilities in the 35-45% range, whereas medium potential over the southern Appalachian Mountains at the 24-h projection is associated with probabilities in the 20-30% range. Similarly, high potential over Iowa at the 3-h projection is associated with probabilities near 50%, whereas high potential around the Alabama Coast at the 24-h projection is associated with probabilities as low as about 40%.

6.3 Convection Potential Scores

TS and bias associated with the three convection potential thresholds are shown in Fig. 8, which (for brevity) are shown only for the spring season. Several features are noteworthy. The peak TS value near 0.5 for the medium potential threshold at the 1-h projection indicates high forecast accuracy, as it means about 50% of the forecast and observed convection "envelope" area is correctly forecast. Recall, however, the 1-h LAMP forecast is largely a "nowcast," as noted in section 5.2. Another feature in Fig. 8 is that the TS's for high potential are much lower than for the low and medium potential thresholds. This result reflects a commonly-known property of the TS, which is that peak values result from slight over-forecasting. Recall that high convection potential strongly underforecasts convection, which is reflected in the bias scores in Fig. 8. Note also that for the (short) independent sample at hand, the bias for the medium threshold is around 1.0 and that for the low threshold is around 2.5; each of these bias levels

is quite close to the prescribed values used for deriving the associated probability thresholds (see previous subsection).

7. SUMMARY, FINDINGS, AND COMMENTS

The object of this study was to develop LAMP convection probabilities and potential with improved temporal-spatial resolution and skill at least as high as that exhibited by the currently operational 2-h LAMP convection product. The temporal and spatial resolution of the new 1-h convection predictand was doubled, and the first time use of fine-scale predictors based on MRMS and total lightning observations and HRRR model output resulted in a clear increase in spatial detail in the 1-h product over that exhibited by the 2-h operational product. Also, objective and subjective forecast performance evaluations revealed that skill of the new product exceeds that for the current operational product out to about 18 hours after LAMP cycle time; afterward skill for the two probability products appears similar. The forecast performance improvement is most evident in the 1-16 h projection range, which is due to strong predictor contributions from the observational and HRRR predictors.

The skill of the 1-h probabilities is quite high in the first few hours after LAMP model cycle time, which is a reflection of the dominant predictor contribution of the fine scale MRMS and total lightning observational predictors. Still, the skill drop-off with projection in this range is steep, indicating a rapid loss in predictive value with time of these observational predictors. For longer forecast projections, a strong contribution from HRRR predictors results in a levelling-off of LAMP skill with projection over about the next 12 hours. Afterwards the skill exhibits a slow fall, which indicates weak predictive value of "persisted" 15-h HRRR forecasts and thus a near total reliance on large scale MOS convection probability predictors.

To improve the quality of the 1-h convection predictand and MRMS predictors, a supplemental quality control process was developed and applied to fine-scale MRMS reflectivity grids. Evidence was presented to show that the QC effectively removed random and systematic non-precipitation echoes. The QC process also includes selective grid masking to remove MRMS data where the radar coverage is poor. An example case was shown to illustrate the positive impact of the QC on the convection probabilities.

Similar to that for the currently-operational 2-h LAMP convection product, the 1-h LAMP convection probabilities incorporate newly developed GFS-based and NAM-based MOS 1-h convection probabilities. Skill comparisons between the LAMP and MOS probabilities reveal that LAMP clearly improved on MOS at all projections. For long LAMP projections, where the predictive contributions from the fine-scale observational and HRRR predictors are small or non-existent, the LAMP improvement on MOS is evidently due to several individual skill benefits that result from the inclusion of non-regionalized MOS probabilities (and associated supplementary predictors) in the regionalized LAMP convection probability regression equations.

Lastly, a convection potential product was derived from the 1-h probabilities, as done previously for the 2-h probabilities. Attributes of convection potential were discussed to show how the product can aid interpretation of the probabilities, thus enhancing their guidance value.

Presently, the LAMP 1-h convection probability and potential guidance is produced experimentally in real time for the 00, 06, 12, and 18 UTC cycles, and forecast maps are posted at http://www.nws.noaa.gov/mdl/lamp/cnv1h.php for field evaluation. In the near future the product will be expanded to include additional LAMP cycles and eventually to cover all 24 hourly cycles. Ultimately, we anticipate the new 1-h product will operationally replace the present 2-h product by early 2017.

Finally, two comments deserve mention. (1) The non-availability of HRRR model forecasts beyond 15 hours results a significant loss in the quality of the LAMP convection forecasts in the 15–25 h LAMP projection range. A forecast range extension of the operational HRRR model is needed to remedy this problem. (2) The especially high skill of LAMP convection forecasts in the first several hours after MRMS and total lightning data cutoff time suggests substantial forecast guidance gains from additional LAMP model runs (with short projections) between the standard hourly runs. Thus we are considering running the model experimentally three additional times per hour at intervals of 15 minutes with a maximum forecast projection of four hours.

8. ACKLOWLEDGEMENTS

Archived MRMS data and HRRR model output were provided by NOAA's National Severe Storms

Laboratory (NSSL) and the NOAA/Earth Systems Research Laboratory/Global Systems Division, respectively. Archives of GFS and NAM model output were provided by MDL's Statistical Modelling Branch and archived total lightning data were furnished by Earth Networks, Inc. This paper is the responsibility of the authors and does not necessarily represent the views of the NWS or any other governmental agency.

9. REFERENCES

Benjamin, S. L., and Coauthors, 2015: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.* doi:10.1175/MWRD-15-0242.1, in press.

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

Charba, J. P., and F. Liang, 2005: Quality 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: Regionalization in fine grid GFS MOS 6-h quantitative precipitation forecasts. *Mon. Wea. Rev.*, **139**, 24-38.

___, F .G. Samplatsky, and P.E. Shafer, 2011: Experimental LAMP 2-h convection guidance on a 20-km grid. Preprints, *24th Conference on Weather Analysis and Forecasting/20th Conference on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., **J19.3**.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203-1211.

___, and D. A. Unger, 1986: A Local AFOS MOS Program (LAMP) and its application to wind prediction. *Mon. Wea. Rev.*, **114**, 1313-1329.

Kanamitsu, M., and Coauthors 1991: Recent changes implemented into the Global Forecast System at NMC. *Wea. Forecasting*, **6**, 425-435.

Kitzmiller, D. H., F. G. Samplatsky, and D. L. Keller, 2002: Production of a national radar reflectivity mosaic and automated radar observations from WSR-88D radar coded messages. NOAA Tech. Memo. NWS MDL 83, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 23 pp.

OFCM, 1991: Doppler radar meteorological observations: Part C, WSR-88D products and algorithms. Federal Meteorological Handbook 11, FMC-H11C-1991, Office of the Federal Coordinator for Meteorological Services and Supporting Research, Silver Spring, MD, 210 pp.

Rodgers, E., Y. Lin, K. Mitchell, W. S. Wu, B. Ferrier, G. Gayno, M. Pondeca, M. Pyle, V. Wong, and M. Ek, 2005: The NCEP North American Mesoscale Modelling System: Final Eta model/analysis changes and preliminary experiments using the WRF-NMM. Preprints, *21st Conf. on Wea. Analysis and Forecasting,* Washington, D.C., Amer. Meteor. Soc., **4B.5**.

Schaefer, T. J., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570-575.

Smith, T. L., S. G. Benjamin, J. M. Brown, S. Weygandt, T. Smirnova, B. Schwartz, 2008: Convection forecasts from the hourly updated, 3-km High Resolution Rapid Refresh (HRRR) Model. Preprints, *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., **11.1**.

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

Zhang, J., and Coauthors., 2011: National mosaic and multi-sensor QPE (NMQ) system: Description, results and future plans: *Bull. Amer. Meteor. Soc.*, **92**, 1321-1338.

___, 2015: Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.* doi:10.1175/BAMS-D-14- 00174.1, in press.

Table 1. Candidate OBS, HRRR, and MOS predictors. Abbreviations: OBS = observational; HRRR = High Resolution Rapid Refresh Model; MOS = Model Output Statistics; hh = clock hour, which is the same as LAMP model cycle time; max = maximum; cref = composite reflectivity; cape = convective available potential energy; vil = vertically integrated liquid; TL = total lightning; GFS = Global Forecast System; NAM = North American Mesoscale model; topography = gridded terrain elevation; RF = relative frequency. Note that (1) the valid time is specified for each OBS predictor. For example, the notation hh:00 indicates the valid time is the top of the hour hh; hh-1:30 means the observation is valid 30 minutes past the previous hour. Also, each OBS predictor is specified as "initial" and "advected" (see text); (2) each of the OBS and HRRR predictors is used in continuous value and grid binary forms; (3) MOS convection probabilities are valid for the same 1-h valid period as for LAMP.

	OBS	HRRR
1.	MRMS max cref at hh:00	cref
2.	MRMS max cref at hh-1:30	vil
3.	MRMS max cref at hh:00 – MRMS max cref at hh-1:30	1-h total precipitation amount
4.	MRMS max vil at hh:00	surface moisture divergence
5.	60 min TL count ending hh:00	cape
6.	30 min TL count ending hh:00	precipitable water
7.	30 min TL count ending hh:00 – 30 min TL count ending hh-1:30	lifted index
8.		lightning threat

	MOS
1.	GFS-based convection probability
2.	NAM-based convection probability
3.	GFS-based convection probability x NAM-based convection probability
4.	GFS-based convection probability x convection monthly RF
5.	NAM-based convection probability x convection monthly RF
6.	GFS-based convection probability x topography
7.	NAM-based convection probability x topography

Table 2. Basic aspects of convection potential

Threshold probability	Convection threat category	Bias range
low	low	2.70 – 2.83
medium	medium	1.03 – 1.13
high	high	0.38 – 0.43

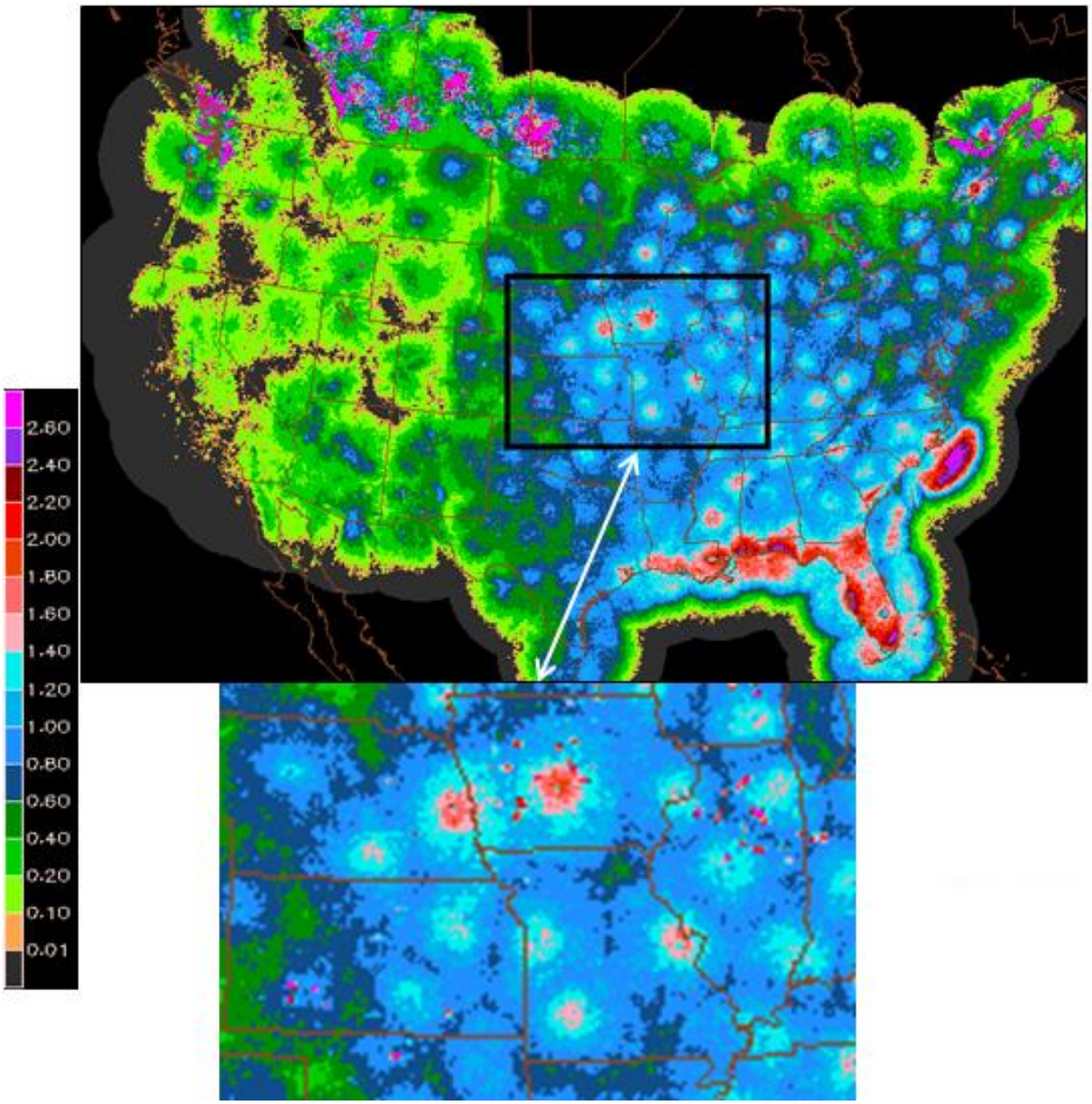


Figure 1a. Relative frequency (%) of "raw MRMS" ≥ 40 dBZ maximum composite reflectivity in 5-km grid boxes at hh:00 UTC (hh = 00, 01, …, 23) during April – September of 2011 - 2014. "Raw MRMS" denotes MRMS data as obtained from NOAA's National Severe Storms Laboratory (NSSL). The rectangular area bounded by bold black lines is magnified below the CONUS map. Black areas are outside the MRMS coverage area.

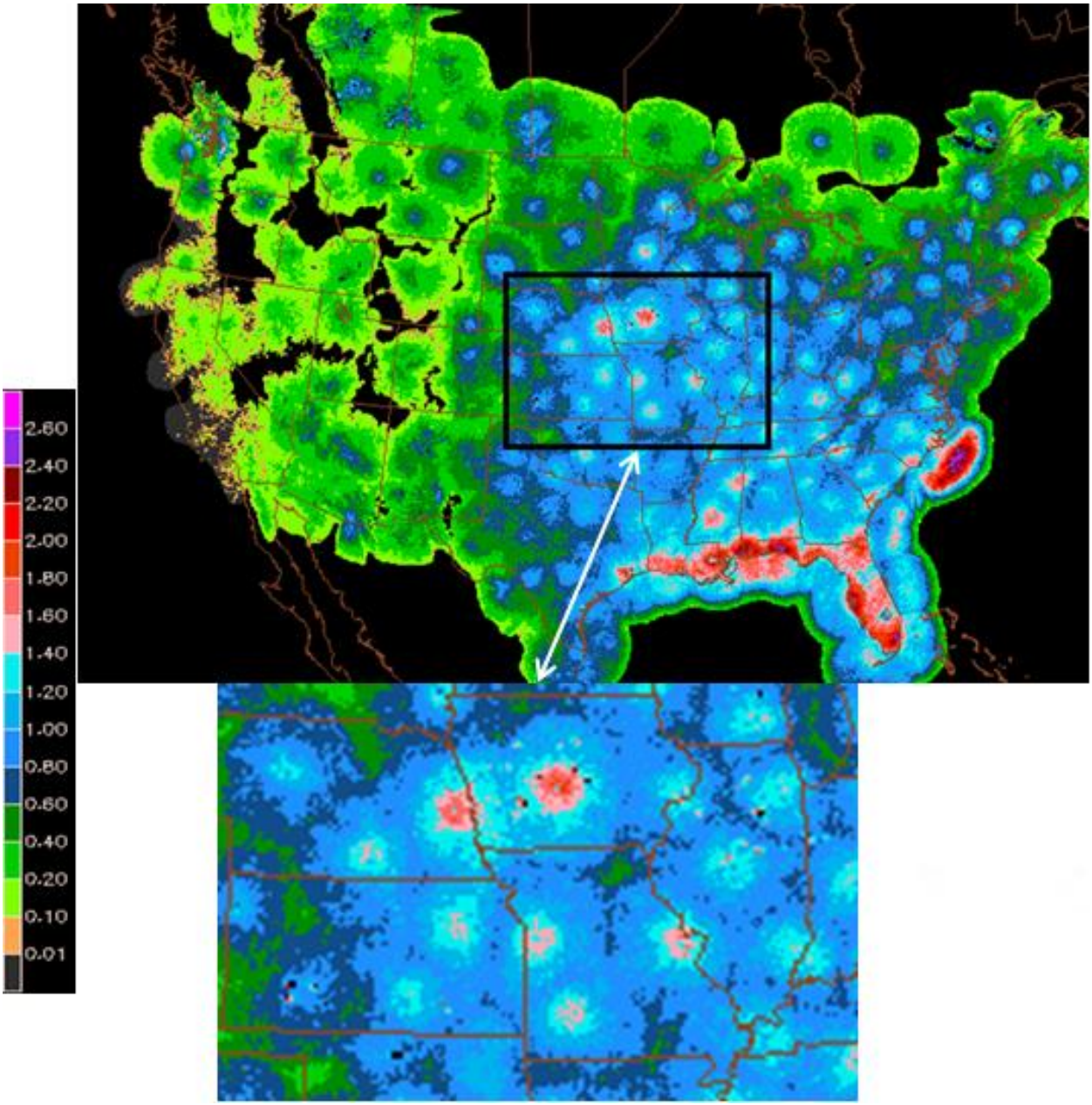
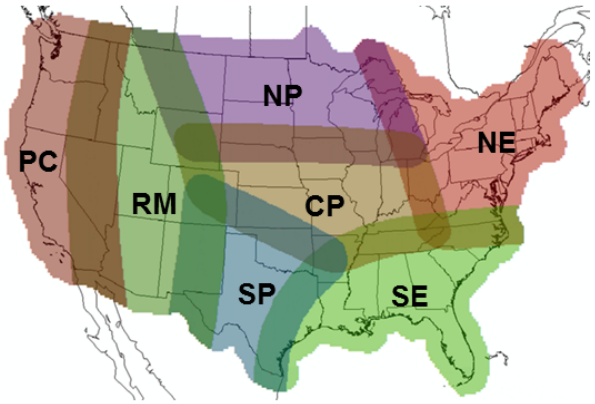


Figure1b. As in Fig. 1a, except after MDL supplemental quality control (QC) processes were applied. Black (missing data) areas inside the MRMS coverage area result from the application of a "static mask" -- a component of the QC processes (see text).



PC = Pacific Coast

RM = Rocky Mountains

SP = Southern Plains

CP = Central Plains

NP = Northern Plains

SE = Southeast

NE = Northeast

Large PC and RM regions provide adequate samples there

Note: Region overlap avoids inter-region discontinuities in convection probabilities

Figure 2. LAMP convection overlapping geographical regions and CONUS forecast domain.

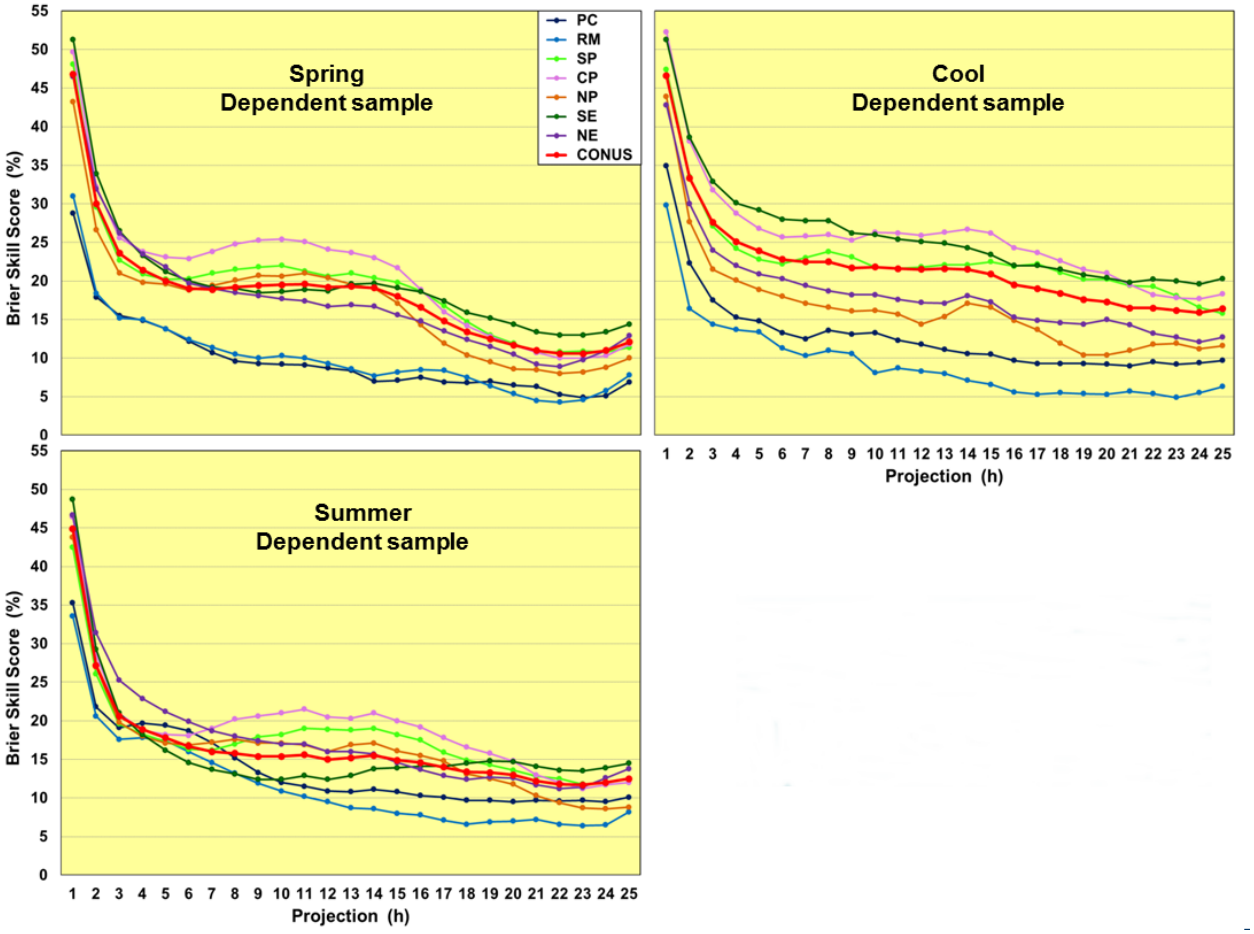


Figure 3. Seasonal Brier Skill Score for individual regions and for the CONUS (regions listed in legend are as in Fig. 2). The dependent sample for "spring" (upper left panel) consists of 16 March – 30 June, 2012 – 2015 , except for 2015 when the period ends 31 May; for "summer" (lower-left pane) the sample consists of 01 July – 15 October, 2012 – 2015, except for 2015 when the period ends 31 August; for "cool" (upper right panel) the sample consists of 16 October – 15 March, 2012 – 2015, except for 2015 when the period ends 14 February.

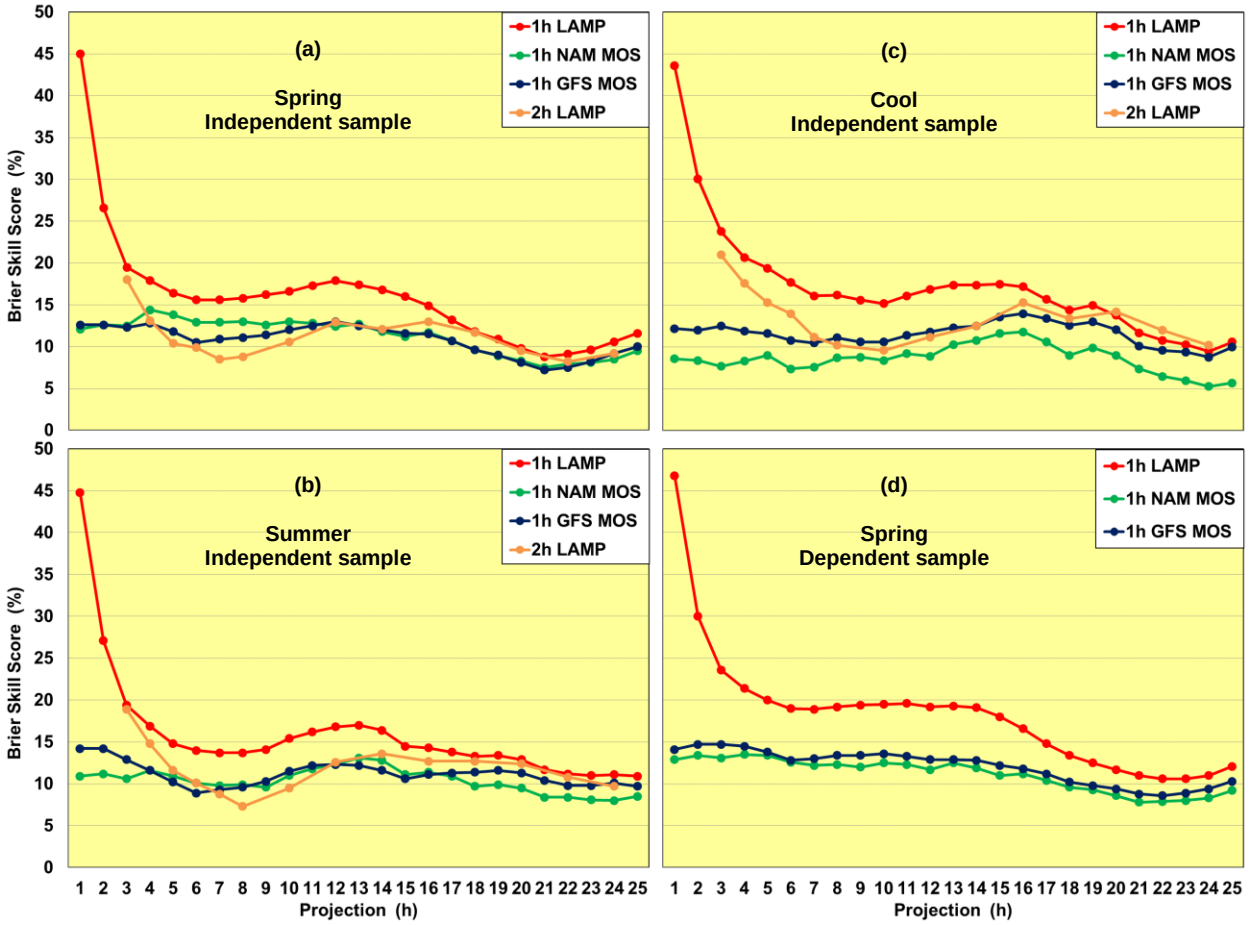


Figure 4. Seasonal mean Brier Skill Score for CONUS for convection probabilities for the LAMP 18 UTC cycle and the associated 12 UTC MOS cycle. The legend notation is: "1h LAMP" denotes 1-h LAMP probability; "1h NAM (GFS) MOS" denotes 1-h NAM- (GFS-) based MOS probability; "2h LAMP" denotes 2-h LAMP probability. In (a) the 30-day independent "spring" sample is for 01 – 30 June 2015; the corresponding "summer" sample in (b) is for 01 – 30 September 2015, and the "cool" sample in (c) is for 15 February – 15 March, 2015. The dependent spring season sample used in (d) is for 16 March – 30 June of 2012 – 2015. [Note that for "1h LAMP" in (d) the last 30 days of the "dependent" sample is actually the same independent sample used in (a).]

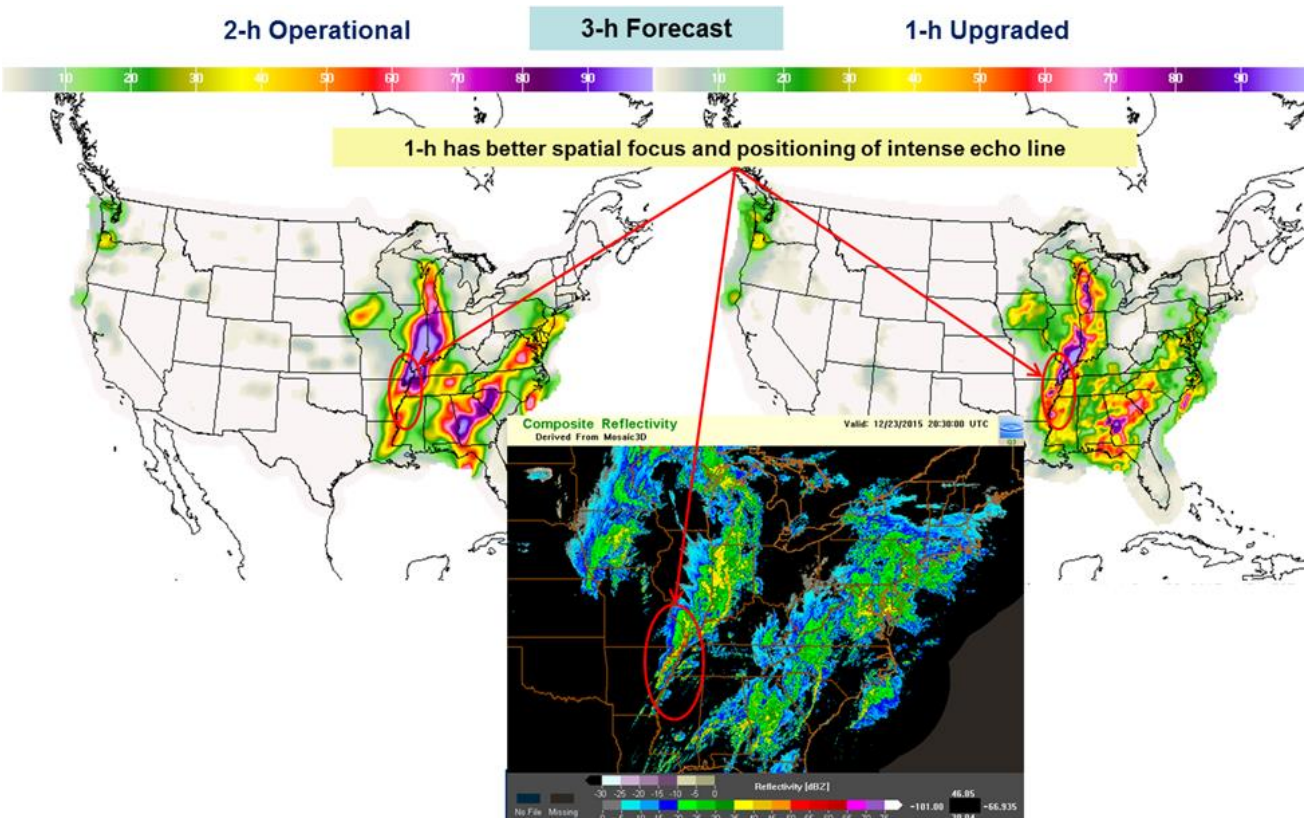


Figure 5a. LAMP 2-h operational convection probability (left) and 1-h upgraded (experimental) convection probability (right) for the 3-h forecast projection from 1800 UTC, 23 December 2015. Note that the 2-h and 1-h valid periods each end at the forecast projection time. The "verifying map," which is obtained from http://nmq.ou.edu/applications/qvs_2d_maps_main.html, is the MRMS composite reflectivity at 2030 UTC (30 minutes beyond the center of the 2-h valid period and at the center of the 1-h valid period). The red ellipse superimposed on all three maps points to an intense convection line in the MRMS map, for which the 1-h probability pattern "has better spatial focus and positioning."

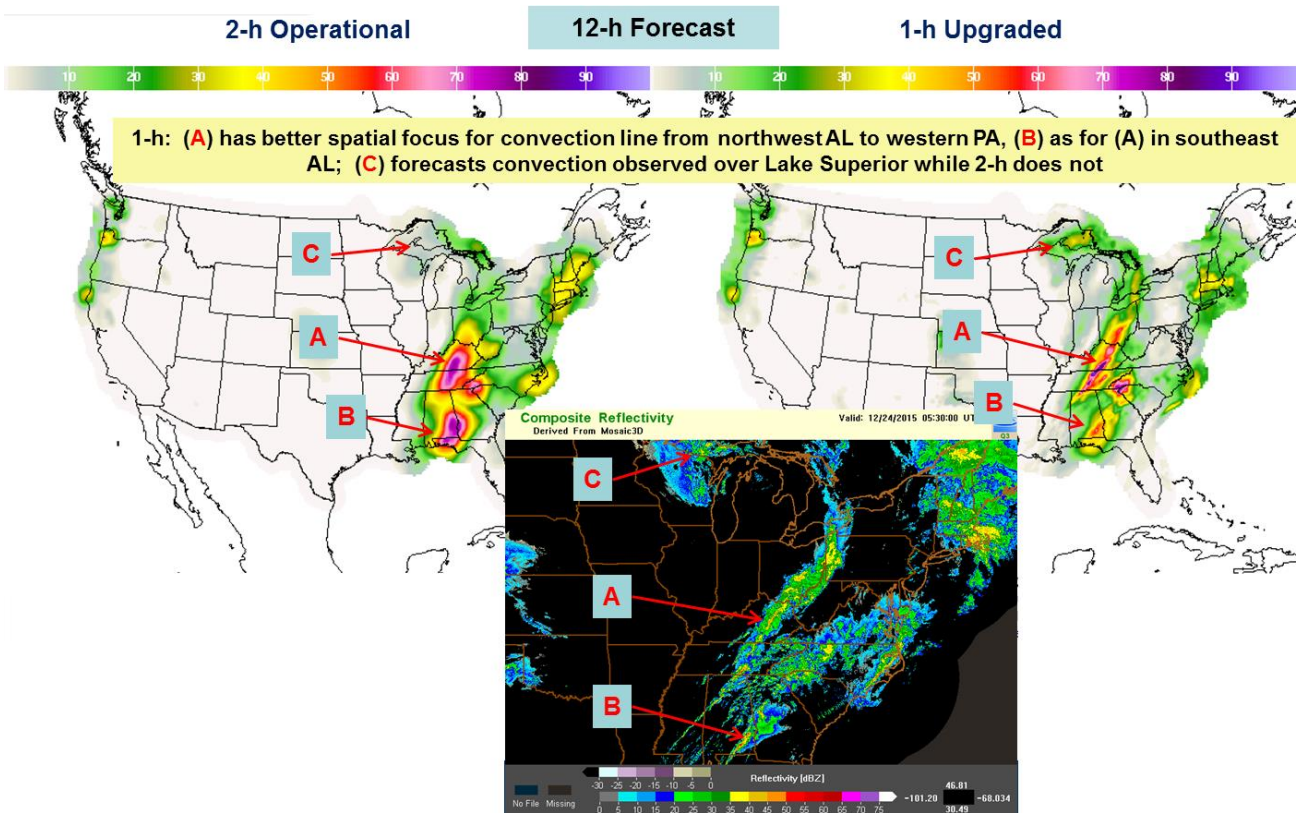


Figure 5b. As for Fig. 5a, except for the 12-h forecast projection. The labels “A”, “B”, and “C” point to specific map features noted in the highlighted text above the probability maps and also discussed in the text body.

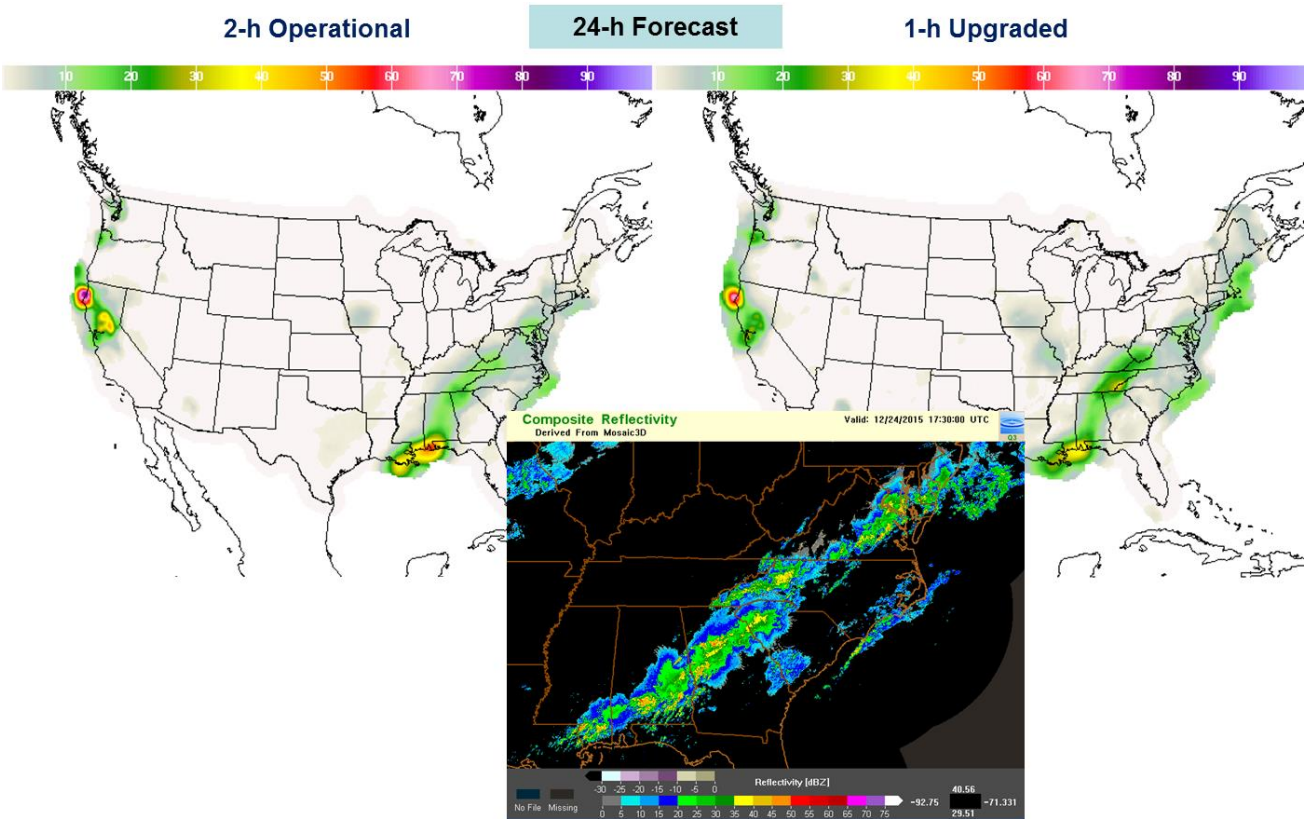


Figure 5c. As in Fig. 5a except for the 24-h forecast projection.

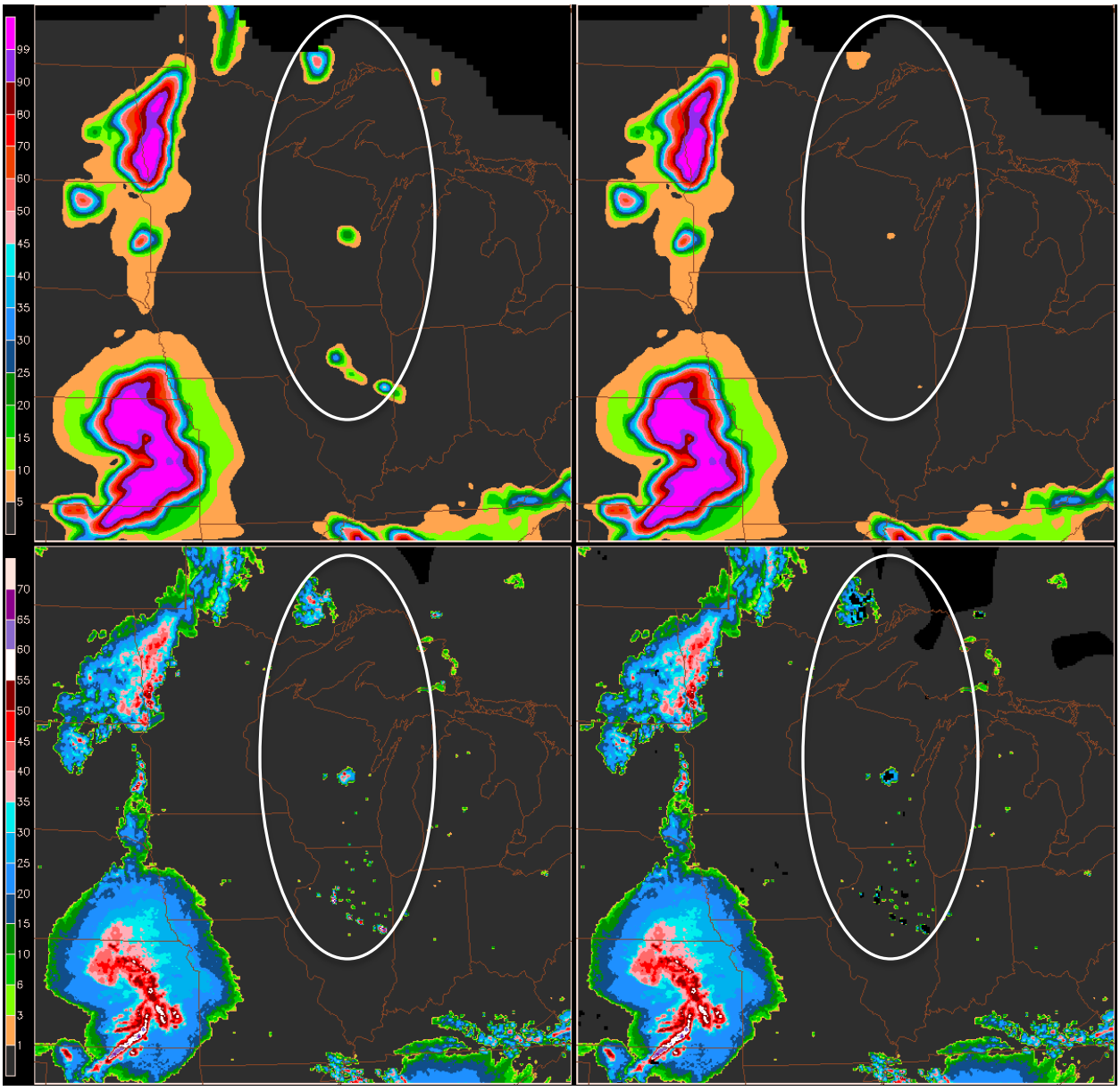


Figure 6. 1-h LAMP convection probability (%) for 1200 – 1300 UTC, 05 June 2014 (top) before (left) and after (right) the MDL supplemental quality control (QC) program was applied to the base 5-km MRMS maximum composite reflectivity grid valid at the 1200 UTC LAMP model cycle time (bottom left). The corresponding reflectivity map following the QC application is shown in the lower right panel. Within the oval shown (bold white outline), the dynamic QC process (see text) masked out anomalous propagation (AP) echoes with reflectivities ≥ 35 dBZ (small black spots within oval in lower right map).

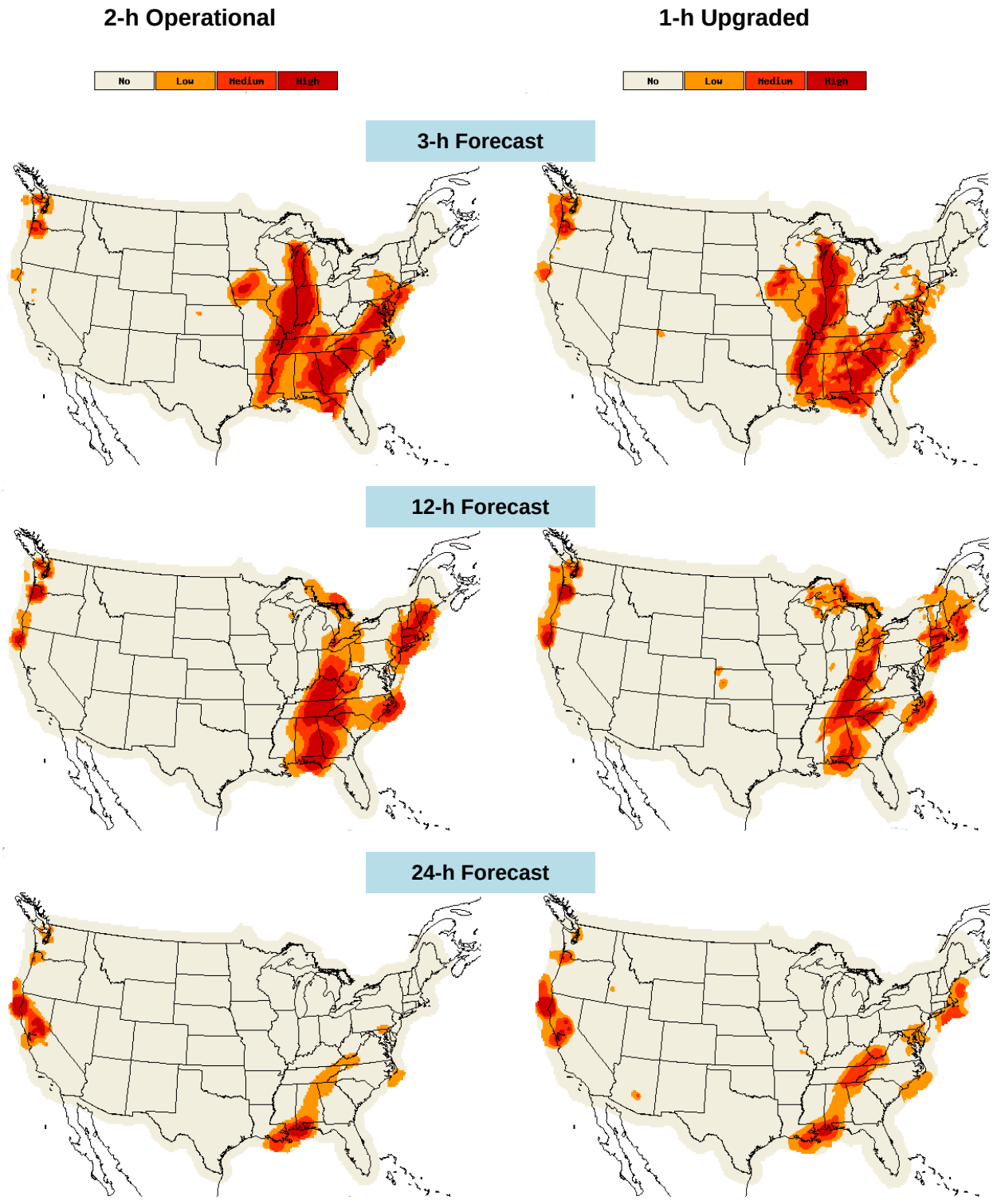


Figure 7. As for Fig. 5, except convection potential and verifying map not shown.

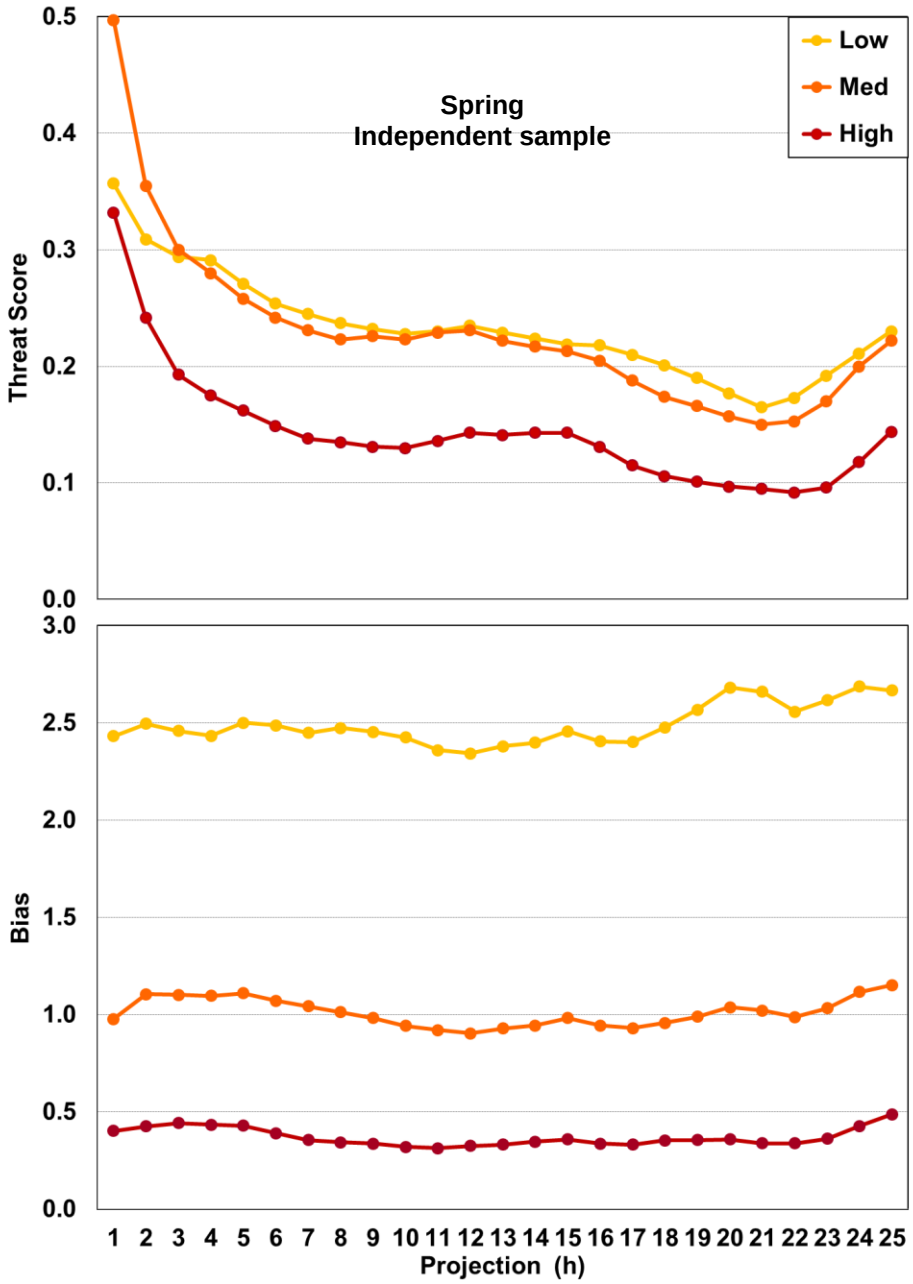


Figure. 8. CONUS threat score and bias for three LAMP convection potential thresholds indicated in the legend: “Low” denotes low potential threshold, “Med” denotes medium potential threshold, and “High” denotes high potential. The scores apply to the 18 UTC LAMP cycle and for the same 01 June – 30 June 2015 independent spring season sample used for the Brier Skill Score in Fig. 4a.