1.1 IMPROVED AVIATION GUIDANCE WITH A LAMP/HRRR MELD

Bob Glahn,* Adam D. Schnapp, Judy E. Ghirardelli, and Jung-Sun Im
Meteorological Development Laboratory
Office of Science Technology Integration
National Weather Service
Silver Spring, Maryland

1. INTRODUCTION

The National Weather Service (NWS) has been disseminating a suite of weather forecast guidance products from the current version of the Localized Aviation MOS Program (LAMP) since 2006 (Ghirardelli and Glahn 2010). The primary purpose of LAMP is to support aviation interests, and included in that suite are forecasts of ceiling height and visibility at specific sites that report those variables, predominantly METAR (OFCM 1995) sites. LAMP provides forecasts each hour, available about 40 minutes after the hour, at projections each hour out to 25 h. More recently since 2010, LAMP gridded forecasts over the conterminous United States (CONUS) have been put into the National Digital Guidance Database (NDGD), the guidance companion to the National Digital Forecast Database (NDFD) (Glahn and Ruth 2003). A number of numerical models also produce forecasts of ceiling and visibility, including some that are run operationally at the National Centers for Environmental Prediction (NCEP). We at the Meteorological Developmental Laboratory (MDL) have developed a method to meld the LAMP forecasts with those produced by the High Resolution Rapid Refresh (HRRR) model to produce probabilistic and categorical values of ceiling height and visibility. The categorical forecasts improve over LAMP alone by a substantial amount for projections 4 through 17 hours, as measured by the threat score (TS) (Palmer and Allen 1949; Wilks 2011)[1]for both warm (April-September) and cool seasons (October-March).

2. THE LAMP MODEL

LAMP is described in Ghirardelli and Glahn (2010). Basically, it follows the MOS (Glahn and Lowry 1972) paradigm, whereby a predictand, usually composed of observations (obs) of a weather variable, is related to a variety of predictors. The predictors used in LAMP for ceiling and visibility prediction come from three sources: (1) the current observation of the variable being forecast, (2) the output from simple advective models, and (3) the MOS forecasts based on NCEP's Global Forecast System (GFS) (GFS MOS) (Dallavalle et al. 2004). Very short range forecasts (i.e., on the order of an hour or two) must be heavily based on the current observation for the forecasts to compete favorably with the observation itself as a forecast (persistence). Essentially, the LAMP model furnishes a blending mechanism from the obs at initial time to MOS at the longest projections.

When dealing with violently non-normal distributions such as ceiling and visibility, MDL has found the Regression Estimation of Event Probabilities (REEP) (Miller 1958; Wilks 2011) method of development works better than dealing with a continuous predictand (e.g., Bocchieri and Glahn 1972, p. 877; Unger and Glahn[2]). The predictand is divided into several categories, say M, and REEP estimates the probability of occurrence of each category. A predictand category that occurs is given the value of 1, and 0 if it doesn't; this defines the binary predictand necessary for REEP. The categories can be either discrete or cumulative (from above or below). For development purposes, it is better to use cumulative binaries (Glahn 1965, p. 125, 126), but for provision to users, discrete categories are many times preferred. It is also customary for many or all of the predictors in this regression to be binary, and generally, cumulative binary.

The M REEP equations are used to estimate the probability of the M predictand categories. However, usually a specific, single value forecast of ceiling and of visibility is preferred, even required, by users of aviation forecasts. In order to produce such categorical forecasts, a probability threshold for each category is computed in such a manner that the bias[3] of the category falls within prescribed limits, and within those limits, the TS is maximized. These thresholds are then used to make the cumulative forecasts from which the discrete forecasts of the M categories can be derived. The categories used by LAMP are indicated in Table 1; the cumulative categories are used for development. The lowest category of ceiling and of visibility were the lowest for which sufficient observations were available to develop stable equations.

The LAMP forecasts are made from REEP equations developed on a regional basis. That is, stations within a geographic region for which it was thought the

[1] Palmer and Allen suggested the name because the event being forecasted and evaluated was thought to be a threat. The TS is the same as the critical success index proposed by Donaldson et al. 1975 and discussed by Shaffer 1990.

* Corresponding author address: Bob Glahn, Meteorological Development Laboratory, National Weather Service Scientist Emeritus, 1325 East West Highway, Silver Spring, MD 20910; email: harry.glahn@noaa.gov.

[2] Unpublished. The developers did much work in the early part of the LAMP project using various transformations of the quasi-continuous visibility and ceiling height observations as predictands. This work was largely unsuccessful; reliable and skillful forecasts could not be made, especially of the lowest values.

[3] Bias for a categorical variable (event) is defined as the number of forecast events divided by the number of observed events.

predictand/predictor relationships were similar were grouped, and all such stations share the same equations. The predictand data for producing the LAMP equations were the METAR obs. Equations for ceiling and also for visibility were developed for all projections 1 through 25 h simultaneously so that the predictors for all projections were as consistent as could be achieved (see Glahn and Wiedenfeld 2006 and Ghirardelli and Glahn 2010 for details). The predictors were selected by specialized software (Glahn and Dallavalle 2000, Chapter U602 with attachments and updates). Forecasts are made each hour for hourly projections out to 25 h for about 1562 stations, the stations that had METAR obs when the equations were developed.

3. THE HRRR MODEL

The HRRR numerical (dynamic) model is described in: http://ruc.noaa.gov/pdf/NCEP_PSR_2013_RAP_FINAL_v5.pdf. It produces ceiling and visibility forecasts according to internal algorithms for projections 1 through 15 hours. The forecasts are for specific values in meters, and ceiling height is above sea level.

4. DATA AVAILABILITY AND PREPARATION

LAMP probability and categorical forecasts are made for specific locations (stations) and are archived. Gridded specific value forecasts are also available on the NDGD grid, but not gridded probability forecasts; however, gridded probability forecasts could be produced for the sample if needed. HRRR forecasts are available on a 3-km grid and could be interpolated either to stations or to the NDGD grid. The obs are available at stations, but could be put (analyzed) onto a grid. Therefore, the matching of predictands and predictors for the statistical analysis could be done either at stations or at gridpoints. Because the predictand is at stations, there is no reason to grid the obs and do the statistical analysis at gridpoints, because all the predictand information is in the station values; an analysis of them adds no information, and the information at gridpoints, not being obs but being interpolated values, would be less accurate than the station values themselves. Therefore, the regression analysis was done at stations.

4.1 LAMP Forecasts

Operational LAMP ceiling and visibility forecasts have been archived in both the probabilistic and categorical forms for the first seven cumulative categories of ceiling and the first six categories of visibility shown in Table 1.

After development of the regression equations based on the then available data, some other stations were later added within the CONUS regions and also over southern Canada as extensions of the adjacent regions. For those added stations that do not have obs, LAMP "backup" equations are used that do not include obs as predictors. No LAMP equations could be developed for locations over water because of lack of obs, but some forecasts over water have been added by using nearby land backup equations. These point forecasts are gridded with the BCDG method (Glahn et al. 2009; Im and Glahn 2012; Glahn and Im 2015) for guidance for forecasters in preparing grids for the NDFD; an example of these gridded forecasts is shown in Fig. 1. The example shown in Fig. 1 was chosen without reference to forecasts, but rather on the basis of a well-defined frontal system in the central part of the U.S., as shown in Fig. 2. However, neither these gridded LAMP forecasts nor the forecasts produced by backup equations for the added stations were used in the regression meld of LAMP and HRRR data.

4.2 HRRR Forecasts

Two years of HRRR forecasts were available starting in March 2013. We divided the sample into warm seasons (April –September) and cool seasons (October-March), characteristic of MOS and LAMP development in MDL. Of the 12 months in each seasonal sample, we used eight for development and four for testing, as indicated in Tables 2a and 2b.

The HRRR ceiling and visibility forecasts are available each hour at hourly increments on a 3-km Lambert conformal grid covering the CONUS for projections 1 through 15 h. To furnish the regression dataset, interpolation was done into the HRRR grid to the LAMP points. The meld of HRRR and LAMP forecasts should be distributed very shortly after LAMP is currently available, about 40 minutes after the top of the hour. The HRRR run is not completed for nearly an hour later, so for any given LAMP start time (cycle), the HRRR must be used from the hour previous. For instance, for the 1200 UTC LAMP cycle, the HRRR 1100 UTC cycle is used. The HRRR ceiling forecasts are in reference to sea level, so the HRRR terrain was used to adjust the forecasts to above ground level, the way ceiling heights are expressed for aviation uses. In addition, visibility was converted from m to mi and ceiling was converted from m to hundreds of ft, the conventional units used in aviation.

The HRRR forecasts have much detail, detail that looks synoptically realistic, but much of it is beyond the realm of predictability at the present time. For instance, visibilities that vary from 8.0 mi to 0.5 mi within the space of 10 km or so are possible, but are not generally observed or forecastable on this scale. Therefore, a preprocessor (to the melding) was run on the HRRR forecasts after converting them to the predictor categories (discussed later) that essentially eliminated "spots" of $\leq$ 7.5 km. This has the effect of coalescing the smaller spots into larger ones, which are still of marginal predictability, but more plausible.[4] Figures 3 (before) and 4 (after) show the effect of this "spot removal." The HRRR

[4] While the spot removal has some characteristics of smoothing, it is not smoothing in the usual sense where averages are computed. The integrity of "unusual" values is maintained when the area covered is of sufficient size or a number of unusual values are close together, even though not contiguous. No change of value is made unless the elevation difference among the points involved is < 100 m, so that variations that may be due to terrain are maintained.

Lambert grid on the files available remapped to the LAMP/NDGD grid does not fully fill the NDFD rectangle.

4.3 Observations

METAR and other obs have been archived by MDL for many years in standard aviation units. They were accessed to extract the needed data.

5. REGRESSION ANALYSIS

REEP was used to develop equations with predictors from the LAMP and HRRR models and obs to produce Meld forecasts for projections 1 through 25 h. The predictors are the same in the Meld equations for each projection, except that the model predictor projections "march" with the predictands. For instance, for the 1200 UTC cycle, and for the 6-h projection, the observation at 1800 UTC (the predictand) is matched with the LAMP 6-h forecast made with 1200 UTC data and the HRRR 7-h forecast made from 1100 UTC data. As noted earlier, a 1-h old HRRR run has to be used to meet timeliness requirements. The predictors in the Meld regression equations were chosen by forward selection. At each selection step, the next predictor was chosen based on the highest added reduction of variance (RV) afforded by any potential predictor for any projection and any predictand category. The selection stopped when no potential predictor reduced any predictand variance by $\geq$ 0.5%.

In order to keep the process reasonably simple, and especially because of the limited data sample, a generalized approach was used, where all stations were grouped together. In an initial study, Glahn et al. (2014) determined that the LAMP probability forecasts are much better predictors than the categorical ones, so only the probabilities were used for the Meld equations.

The LAMP forecasts have only a few categories, sufficient for providing forecasts to users in matrix form. However, for a gridded product, more definition is desirable, so we used an expanded set of categories shown in Table 3. Two categories of visibility and one category of ceiling were added below those for which LAMP forecasts are available. For visibility, there is a category for each reportable value below 10 mi, except the very lowest ones. For ceiling, every reportable value has a category below 1,000 ft, and at meaningful thresholds above that. The Meld produces a probability of each category. Using the same procedure as was used in LAMP, we developed thresholds to produce categorical forecasts with biases in the range 1.0 to 1.2. This process is explained fully in Ghirardelli and Glahn (2010). Because some of the categories cover more than one reportable value, the values put on the grid are sometimes averages; the values for the grid are shown in the 3rd and 5th columns of Table 3. Also shown in Table 3 are the upper limits for Instrument Flight Rules (IFR), Low IFR (LIFR), Very Low IFR (VLIFR), and Marginal Visual Flight Rules (MVFR); any value above MVFR indicates Visual Flight Rules.

It is of considerable importance that the forecasts are not only consistent from projection to projection, but also from the analysis (0-h projection) to the 1-h projection. Much care was taken in developing the LAMP regression equations in this regard. To enhance continuity of the Meld, the initial obs were used in developing the Meld equations for all projections, as they had been in developing the LAMP equations.

We were also concerned about the possible lack of continuity between the 14-h projection, the longest projection for which the HRRR is available, and the 15-h and following projections. Therefore, we used the HRRR 14-h projection, not only for the 14-h Meld projection, but for all projections 15 h through 25 h.

5.1 Ceiling Height

Grouping all stations together gave a large number of predictand-predictor pairs (sample size) varying for the warm season from about 335,000 for the 1-h projection to 297,000 for projections 14 to 25 h. The decrease of sample size with projection was due to missing HRRR data. The low relative frequencies of low ceilings restricted the development method to generalized operator (Bocchieri and Glahn 1972, p. 970) as mentioned earlier. For instance, there were < 100 occurrences of ceiling < 100 ft and < 300 occurrences of ceiling < 200 ft in the 8-month sample for all stations combined, so further spatial stratification would not be feasible unless the two lower categories were eliminated.

We were concerned that if all potential predictors—LAMP, HRRR, and obs—were offered together for selection, the HRRR might be overwhelmed by the obs, which are well-known for their importance in the early projections. Therefore, we made an initial screening of only the seven LAMP predictors and the 12 binary HRRR predictors shown in Table 4 for projections 1 through 14 h. For the warm season, all seven LAMP predictors and five of the 12 potential HRRR predictors were selected with the 0.5% RV cutoff criterion. We then forced these 12 predictors and added the 15 potential obs predictors. The six observation categories indicated in Table 4 were selected. Another regression run was made for projections 14 through 25. All 18 of those previously selected were "forced," but were included only if the additional RV was > .01%. One of the obs, < 8 mi, was not included in the equations for these projections. These are the equations used for the independent verification. The development for the cool season followed the same general process (see Table 4).

One could speculate why these specific predictors were chosen. It is clear that the obs were furnishing information for the very low categories, for which LAMP and HRRR did not do an adequate job. Also, they were chosen for the very short-range projections. The RVs for the categories below which LAMP is available were higher than for the other categories indicating the equations were likely somewhat unstable because of the low number of cases.

5.2 Visibility

The developmental process was the same for visibility as for ceiling.

Besides the six LAMP predictors, the HRRR and obs used as predictors are shown in Table 5. Previous work for the cool season (see Glahn et al. 2014) showed that higher HRRR thresholds were not useful. For the warm season, a trial regression run was made where all LAMP and HRRR predictors were screened together; all six LAMP predictors were selected and only three HRRR predictors. The final regression run was made by forcing the six LAMP and the three HRRR predictors selected in the trial run. Five observation predictors were selected from the set shown in Table 5. The HRRR and observations chosen as predictors are shown in Table 5 in red and marked with an asterisk.

As with ceiling, the lower categories of observations were chosen for the low categories. In addition, three others were chosen, indicating the importance of persistence in visibility prediction. Also, similarly to ceiling, the lower two categories had unexpectedly high RVs showing them to likely be fitting the data too closely.

6. EVALUATION ON INDEPENDENT DATA

As described earlier, the development was done at stations—discrete points where the predictand data applied. For implementation and evaluation, three options were considered:

(1) Interpolate the HRRR forecasts to the LAMP stations, apply the equations and thresholds at the LAMP stations, and analyze the probabilities (if they are desired) and categorical forecasts to the LAMP grid,

(2) analyze the LAMP station probabilities and observations to the LAMP grid, interpolate the HRRR forecasts to the same grid, and apply the equations and thresholds on the grid, and

(3) interpolate the HRRR forecasts to the LAMP stations, evaluate the equations at the LAMP stations, analyze the Meld probabilities and apply the thresholds at the gridpoints.

Any one of the three processes will work and it is not known which is best; we chose (2) for the implementation process, but for the test sample verification, we applied the equations and thresholds at stations.

We applied the implementation process to the April 11, 2013, 7-h forecast from 1200 UTC data. The results looked reasonable. Features of both LAMP and the HRRR could be seen, the LAMP being more apparent because LAMP furnished better predictors than did HRRR. However, in concert with the suspected instability of the lowest category equations, some “blobs” of category 1 forecasts were made in unexpected places. Such features detract from the overall usefulness of the Meld. Rather than not use the suspect equations, we chose to mitigate the effect by developing thresholds with biases between 0.4 and 0.6 for the two lower categories.

The developmental equations were evaluated on the 4 months of test data indicated in Table 2 for both warm and cool seasons. The primary scores were bias and TS for several categories, although the probability of detection, false alarm ratio, and Gerrity score (Gerrity 1992) were also computed. In all the verification graphs shown here, LAMP means the original LAMP forecasts; the equations on which the forecasts are based were developed several years before the test sample. The HRRR forecasts were interpolated from the HRRR grid to LAMP stations and for verification did not include the preprocessing that was done for the regression analysis. All comparisons were on matched samples, differing only by projection. As discussed above, the predictand categories were defined as cumulative from below. Verification scores were also computed for cumulative categories. The primary verification used the categories for which LAMP forecasts were available, and comparative verification could be done.

6.1 Cool Season, Ceiling Height

Figures 5 through 8 show the biases and TSs for events < 1,000 ft and for events < 500 ft for the cool season. The LAMP and Meld biases are generally good, the HRRR less so. For this cycle, 1200 GMT, persistence is high biased except for short projections and 24 h later, peaking around projection 12 h. LAMP, Meld, and persistence are nearly equal at 1 h; they all decline rapidly, persistence more rapidly than LAMP, and the Meld less rapidly than LAMP. The uncalibrated HRRR is not competitive for several hours from start time. The Meld is better than LAMP at all projections except for their near equality at 1 h. Even though there are HRRR forecasts only up until 14 h, their influence lingers and gradually diminishes.

6.2 Cool Season, Visibility

Figures 9 through 12 show the biases and TSs for events < 3 mi and for the events < 1 mi for the cool seasons. Although the TSs are lower for visibility than for ceiling height, the comments above for the ceiling height apply for visibility as well.

6.3 Warm Season, Ceiling Height

Figures 13 through 16 are similar to Figs. 5 through 8 except for the warm season. Although the TSs are somewhat lower for the warm season than for the cool season, the comments above still hold except that the effect of the 14-h HRRR fades more quickly and gives little or no improvement past about 18 hours. It is surprising the LAMP biases are above 1.5 for the later projections; this may be because the GFS MOS has changed since the development of the LAMP equations.

6.4 Warm Season, Visibility

Figures 17 through 20 are similar to Figs. 9 through 12 except for the warm season. The TSs are lower than

for the ceiling height, and for the cool season. The TSs for the HRRR are more similar to persistence than to the LAMP and Meld except for the first 2 hours where the HRRR is not as good as persistence; however, by 14 h, the HRRR is between persistence and the Meld. The LAMP biases show remarkable diurnal variation, being quite low then quite high with projection.

7. EQUATIONS FOR DAILY USE

7.1 Ceiling Height

The equations for daily use were developed on all 12 months of data. For projections 1-14 h, the full set of potential predictors was offered for selection. Nearly same sets of predictors were selected for the 12-month equations as for the 8-month equations shown in Table 4. These final predictors are also in the equations for projections 15-25 h.

A Meld forecast, depicted in Fig. 21, was made with the 12-month equations for the same case as shown in Figs. 1 through 4; features of both LAMP and HRRR can be seen. The Meld forecast contains some very small-scale features that are not forecastable, so spot removal software[5] was applied to produce the slightly less "choppy" one shown in Fig. 22. The frontal detail shown by HRRR in Fig. 3 is generally present in Fig. 22. The small blue spot in northeastern Texas is caused by one LAMP station having a low ceiling forecast, and the spot is larger than what the software will remove; being a valid LAMP forecast, it is not obvious that it should be removed, even though it does not agree with its neighboring stations. Projection 7, depicted in these maps, is one where HRRR is expected to contribute strongly. Both verification and maps (not shown) indicate that the HRRR is much less influential at very short projections, and also past about projection 18.

7.2 Visibility

As with ceiling, the equations for daily use were developed on all 12 months of data. For projections 1-14 h, the six LAMP and 11 HRRR predictors were offered for selection. All six LAMP predictors and only three HRRR predictors were selected. These nine predictors were then forced when developing for all 25 projections. Five obs predictors were chosen, making a total of 14 predictors in the equations

A Meld forecast, shown in Fig. 23, was made with the 12-month equations for the same case shown in Figs. 1 through 4. As with ceiling, a few small spots can be the result of the binary process we use for making the forecasts. The probability forecasts made directly from the equations are thresholded to make specific value forecasts. When the probability is near the threshold for that category, it may get "tripped" for one gridpoint, but not for a neighboring one. The spot remover postprocessing routine was run on the grid depicted in Fig. 23 to give the one shown in Fig. 24.

8. SUMMARY AND CONCLUSIONS

A system for making objective ceiling height and visibility forecasts at gridpoints based on a meld of the LAMP and HRRR predictions of those weather elements has been developed, tested, and readied for daily use. Observations at initial time were also included in the regression equations, primarily for continuity from the analysis of observations at initial time to the 1-h forecast. The results shown here are for the 1200 UTC cycle and are consistent with earlier work for the cool, 0000 UTC cycle (Glahn et al. 2014)

Overall, the Meld approach seems to be viable, the Meld biases and TSs being generally markedly better than HRRR or persistence alone, except for the 1-h forecast where the Meld could not improve upon persistence. The Meld is also better than LAMP alone except for the first hour or two and after about 18 h when the 14-h HRRR forecast is no longer very useful. The Meld forecasts show characteristics of both LAMP and HRRR. The HRRR has much very small-scale detail, some of which needs to be disregarded for specific point forecasts. While such detail might be reasonable at a 1-h projection, HRRR is not good at that range. At projections of several hours, where HRRR is closer to competitive with LAMP, pinpointing variations in ceiling and visibility on the order of 10-km is beyond forecasting ability, and the smaller spots of this size are removed. However, larger-scale detail, such as the low ceilings and visibilities associated with the frontal structure east of the lower Mississippi River is kept as shown in Figs. 3, 4, 21, and 22.

The HRRR does not show up particularly well in the Figs. 5 through 20. This is partly because the forecasts are not calibrated, and the individual category verifications do not indicate the overall usefulness of the HRRR, which shows up in the calibrated Meld.

Persisting the HRRR past 14 h can cause a feature to remain stationary for a time until the equation coefficients render the HRRR non-effective for larger projections. The alternative would have been to not use the HRRR past 14 h, then a feature due to the HRRR would disappear immediately. This problem will be corrected when a HRRR archive of longer projections becomes available.

The Meld is a combination of models. LAMP itself incorporates three advective models and GFS MOS. In some sense, persistence can be called a "model," as it furnishes extremely useful information. Heretofore, we have not incorporated a dynamic, mesoscale model because of lack of an adequate developmental sample and our belief that the model and its output would change significantly before operational implementation could be achieved. However, the HRRR has been developed to the point we believe it (or some similar model) should be used. This is in concert with the National Blend of Models (Gilbert et al. 2016), although the characteristics

[5] This postprocessing removes spots as large as 12.5-km across, while the preprocessing removes 7.5-km spots.

of the Blend and the Meld are necessarily and strikingly different. It is unusual for an "improvement" to an existing operational product (LAMP) to be as large as the HRRR affords, and the improvement justifies the HRRR's use in the LAMP suite of products.

Acknowledgements. We are indebted to Gordana Sindic-Rancic and Chenjie Huang for assistance in data preparation. We also thank David Rudack for a careful proofreading. A portion of the work was funded by NOAA's Nextgen Weather Program. This paper is the responsibility of the authors and does not necessarily represent the views of the NWS or any other governmental agency.

REFERENCES

Bocchieri, J. R., and H. R. Glahn, 1972: Use of Model Output Statistics for predicting ceiling height. *Mon. Wea. Rev.,* **100**, 869-879.

Dallavalle, J. P., M. C. Erickson, and J. C. Maloney III, 2004: Model output statistics (MOS) guidance for short- range projections. *Preprints, 20th Conf. on Weather Analysis and Forecasting/ 16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., **6.1.**

Donaldson, R., R. Dyer, and M. Krauss, 1975: An objective evaluator of techniques for predicting severe weather events. *Preprints, Ninth Conf. on Severe Local Storms*, Amer. Meteor. Soc., Norman, OK, 321-326.

Gerrity, J. P., 1992: A note on Gandin and Murphy's Equitable Skill Score. *Mon. Wea. Rev*., **120**, 2709-2712.

Ghirardelli, J. E., and B. Glahn, 2010: The Meteorological Development Laboratory's aviation weather prediction system. *Wea Forecasting*, **25,** 1027-1051.

Gilbert, K. K., J. P. Craven, T. M. Hamill, D. R. Novak, D. P. Ruth, J. Settelmaier, J. E. Sieveking, and B. Veenhuis, Jr., 2016. The national blend of models, version one. *Preprints, 23th Conf. on Probability and Statistics in the Atmospheric Sciences*, New Orleans, LA, **1.3**.

Gilbert, K. K., J. P. Craven, T. M. Hamill, D. R. Novak, D. P. Ruth, J. Settelmaier, J. E. Sieveking, and B. Veenhuis, Jr., 2016. The national blend of models, version one. *Preprints, 23th Conf. on Probability and Statistics in the Atmospheric Sciences*, New Orleans, LA, **1.3**.

Glahn, H. R, 1965: Objective weather forecasting by statistical methods. *The Statistician*, **15,** 111-142.

_____, and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor*., 11, 1313-1329.

_____, and J. P. Dallavalle, eds., 2000: Computer programs for MOS-2000. *TDL Office Note* **00-2**. Techniques Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce.

Glahn, B, and D. P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.*, **84**, 195-201.

_____, and J. Wiedenfeld, 2006: Insuring temporal consistency in short range statistical weather forecasts. *Preprints, 18th Conf. on Probability and Statistics in the Atmospheric Sciences*, Atlanta, GA, **6.3**.

_____, K Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, 2009: The gridding of MOS. *Wea. Forecasting,* **24**, 520-529.

____, R. Yang, and J. Ghirardelli, 2014: Combining LAMP and HRRR visibility forecasts. *MDL Office Note* 14-2. Meteorological Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce, 20 pp.

____, and J.-S. Im, 2015: Objective analysis of visibility and ceiling height observations and forecasts. *MDL Office Note* **15-2**. Meteorological Development Laboratory, National Weather Service, NOAA, U.S. Department of Commerce, 17 pp.

Im, J.-S., and B. Glahn, 2012: Objective analysis of hourly 2-m temperature and dewpoint observations at the Meteorological Development Laboratory. *Natl. Wea. Dig*., **36(2),** 103-114.

Miller, R. G., 1958: Regression estimation of event probabilities, U.S. Weather Bureau, Contract Cwb-10704, *Tech. Rep.* No. **1**, The Travelers Research Center, Inc., Hartford, CN.

OFCM, 1995: Surface weather observations and reports. *Federal Meteorological Handbook* **1**, NOAA/Office of the Federal Coordinator for Meteorological services and Supporting Research, 104 pp.

Palmer, W. C., and R. A. Allen, 1949: Note on the accuracy of forecasts concerning the rain problem. Weather Bureau Manuscript, Washington, D. C., 2 pp.

Shaffer, J. T., 1990: The Critical Success Index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570-575.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 676 pp.

Table 1. Category definitions of LAMP ceiling height and visibility. Ceilings are observed (reported) in hundreds (hd) of feet (ft). Visibilities are observed to fractions of a mile (mi) when the visibility is low.

Category Number	Discrete Categories		Cumulative Categories	
	Ceiling (hd ft)	Visibility (mi)	Ceiling (hd ft)	Visibility (mi)
1	< 2	< 0.5	< 2	< 0.5
2	2-4	≥ 0.5 and < 1.0	< 5	< 1.0
3	5-9	≥ 1.0 and < 2.0	< 10	< 2.0
4	10-19	≥ 2.0 and < 3.0	< 20	< 3.0
5	20-30	≥ 3.0 and ≤ 5.0	≤ 30	≤ 5.0
6	31-65	> 5.0 and ≤ 6.0	≤ 65	≤ 6.0
7	66-120	> 6.0	≤ 120	
8	>120			

Table 2a. The warm season months with an X are those used for independent testing.

Year	April	May	June	July	August	September
2013	X				X	
2014			X			X

Table 2b. The cool season months with an X are those used for independent testing.

Year	October	November	December	January	February	March
2013-2014		X			X	
2014-2015			X			X

Table 3. The 16 predictand cumulative from below category upper limits for visibility and 24 for ceiling, and the associated values for the grid used in the Meld. There is a category above the last one in the table of ≥ 10 mi for visibility and > 12,000 ft for ceiling, the last including unlimited ceiling. The categories for which LAMP forecasts exist are in red and marked with an asterisk.

Category No.	Visibility (mi)		Ceiling (hd ft)	
	Nominal (mi)	Value on Grid	Nominal (ft)	Value on Grid
1	= 0	0	< 100	0
2	< 1/4	.125	< 200 (VLIFR)*	1
3	< ½ (VLIFR)*	.25	<300	2
4	< 3/4	.5	< 400	3
5	< 1 (LIFR)*	.75	<500 (LIFR)*	4
6	< 1 ½	1.12	< 600	5
7	< 2*	1.62	< 700	6
8	< 2 ½	2.0	< 800	7
9	<3 (IFR)*	2.5	< 900	8
10	≤ 3	3	< 1,000 (IFR)*	9
11	≤ 4	4	<1,200	11
12	≤ 5 (MVFR)*	5	<1,500	13
13	≤ 6*	6	<1,700	15
14	≤ 7	7	<2,000*	18
15	≤ 8	8	< 2,500	22
16	< 10	9	≤ 3000 (MVFR)*	27
17			≤ 4,000	35
18			< 5000	45
19			≤ 6,500*	58
20			≤ 8,000	73
21			≤ 9000	85
22			≤ 10,000	95
23			≤ 11,000	110
24			≤ 12,000*	120

Table 4. The 12 HRRR ceiling height forecasts and 15 ceiling height observations offered as predictors for predicting ceiling. The HRRR and obs predictors selected by screening on the 8-month developmental sample are shown in red and marked with an asterisk.

Predictor No.	HRRR Predictor		Observation Predictor	
	Warm Season	Cool Season	Warm Season	Cool Season
1	< 2	< 2*	< 2*	< 2*
2	< 3	< 3	< 3*	< 3*
3	< 5*	< 5*	< 5	< 5
4	< 6	< 6	< 6*	< 6*
5	<8*	< 8*	< 8*	< 8*
6	< 10	< 10	< 9	< 9
7	<15	< 15*	< 10	< 10
8	< 20*	< 20	< 15*	< 15*
9	≤ 30*	≤ 30*	< 20	< 20
10	≤ 65	≤ 65*	≤ 30	≤ 30
11	≤ 100	≤ 100	≤ 50*	< 50*
12	≤ 120*	≤ 120	≤ 65	≤ 65
13			≤ 80	≤ 80*
14			≤ 100	≤ 100
15			≤ 120	≤ 120

Table 5. The 11 HRRR visibility forecasts and 15 visibility observations offered as predictors for predicting visibility. The HRRR and obs selected by screening on the 8-month developmental sample are shown in red and marked with an asterisk.

Predictor No.	HRRR Predictor		Observation Predictor	
	Warm Season	Cool Season	Warm Season	Cool Season
1	< 0.25	< 0.25	< 0.25*	< 0.25*
2	< 0.5	< 0.5*	< 0.5*	< 0.5*
3	< 1.0	< 1.0	< 0.75	< 0.75
4	< 2.0*	< 2.0*	< 1.0	< 1.0
5	< 3.0	< 3.0*	< 1.5	< 1.5
6	≤ 4.0	≤ 4.0	< 2.0	< 2.0
7	≤ 5.0	≤ 5.0	< 2.5	< 2.5
8	≤ 6.0*	≤ 6.0*	<3.0	<3.0
9	≤ 7.0	≤ 7.0*	≤ 3.0	≤ 3.0*
10	≤ 8.0*	≤ 8.0	≤ 4.0*	≤ 4.0
11	≤ 10.0	≤ 10.0*	≤ 5.0	≤ 5.0
12			≤ 6.0	≤ 6.0
13			≤ 7.0*	≤ 7.0*
14			≤ 8.0	≤ 8.0
15			≤ 10.0*	≤ 10.0*

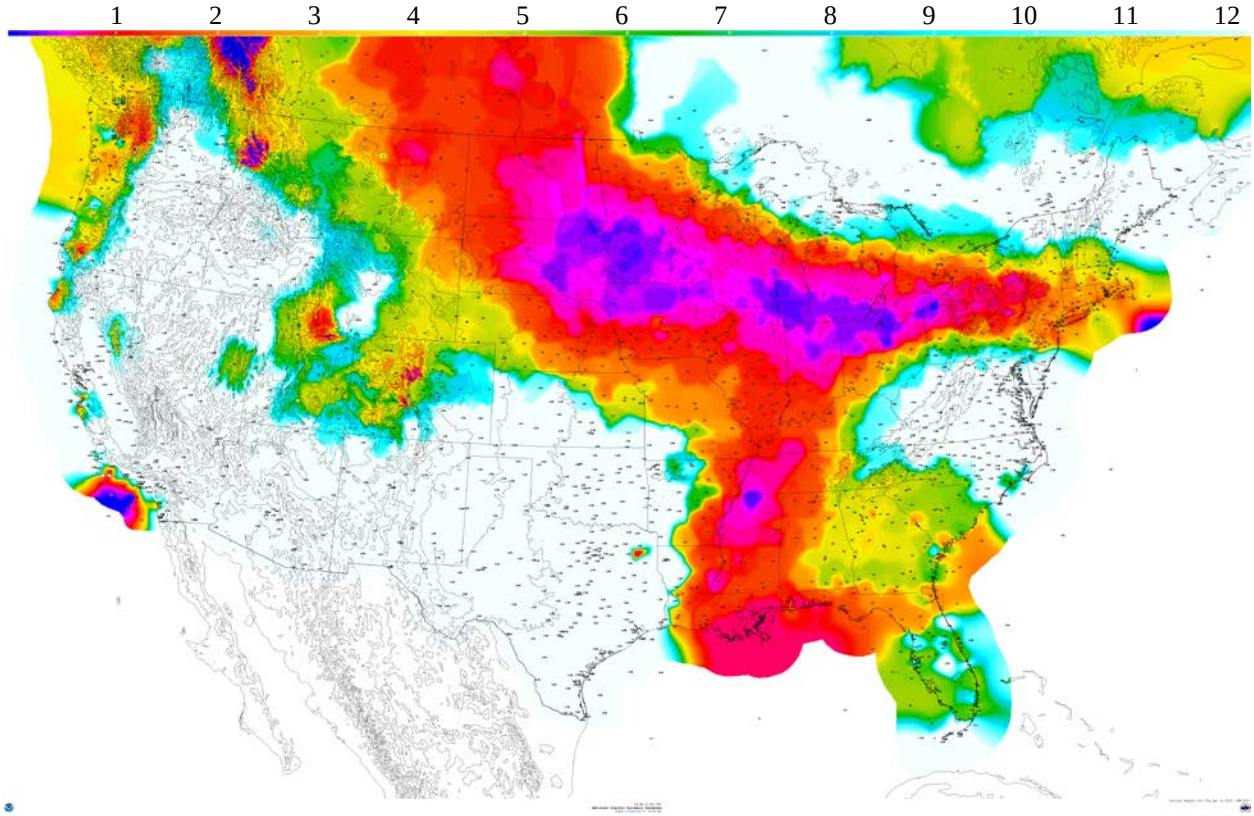

Figure 1. The LAMP categorical ceiling height forecast, 7-h projection from April 11, 2013, 1200 UTC. Color bar is in thousands of ft.

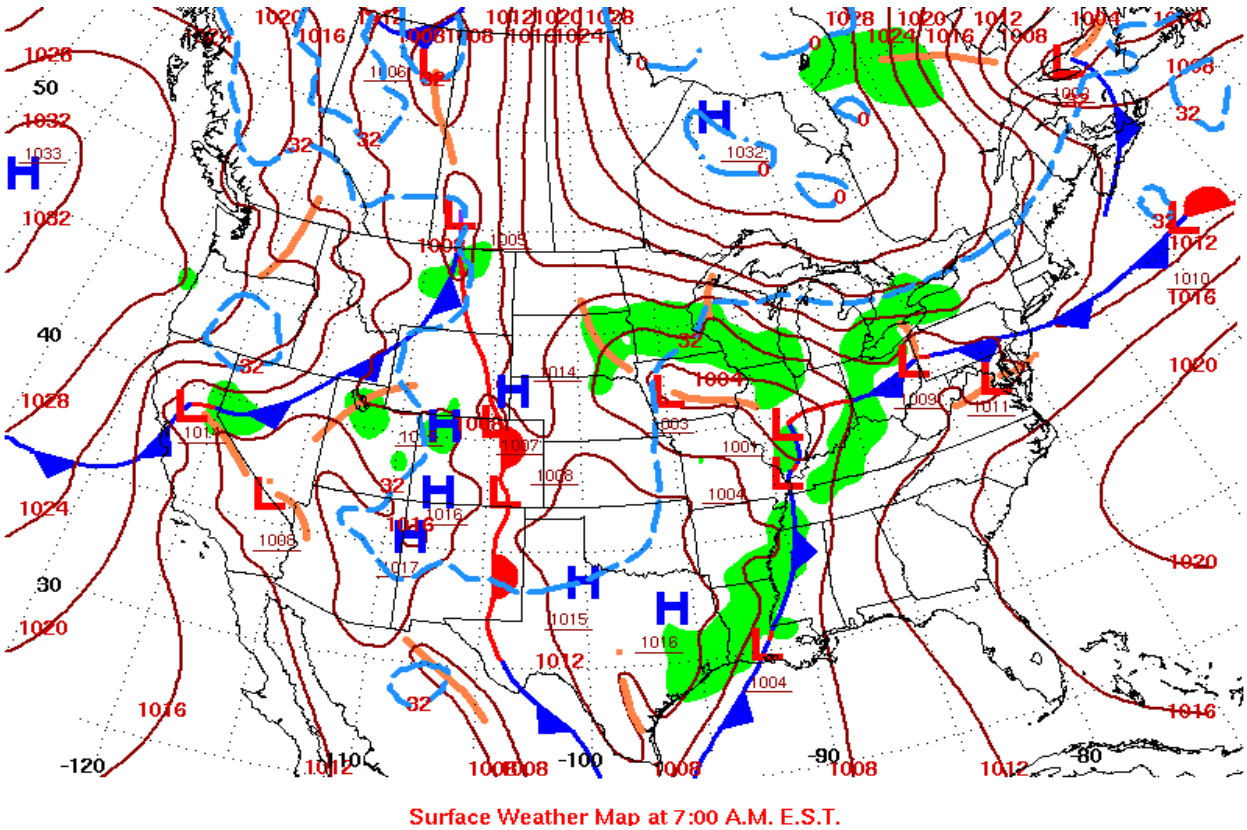

Figure 2. Sea level pressures and fronts for April 11, 2013, 1200 UTC.

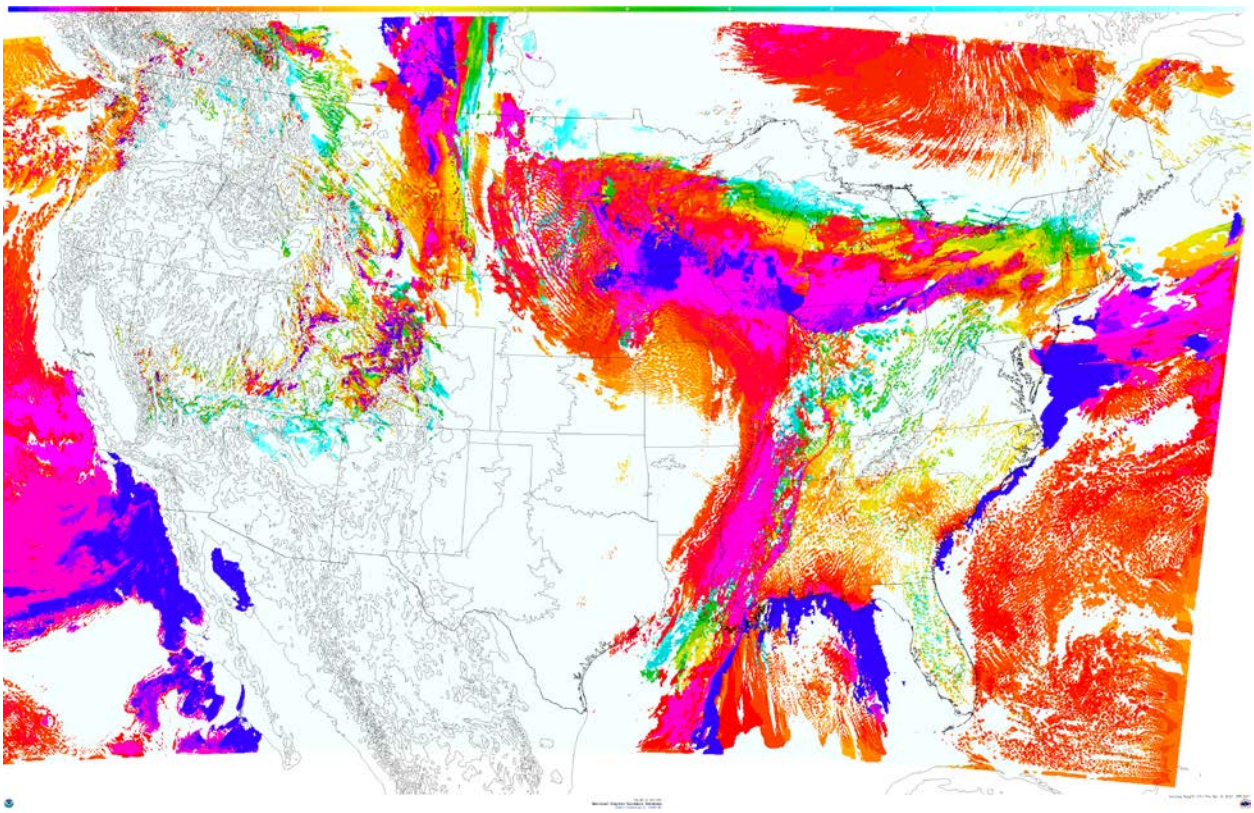

Figure 3. The HRRR ceiling height forecast for April 11, 2013, 8-h projection from 1100 UTC. Color bar is in thousands of ft (see below).

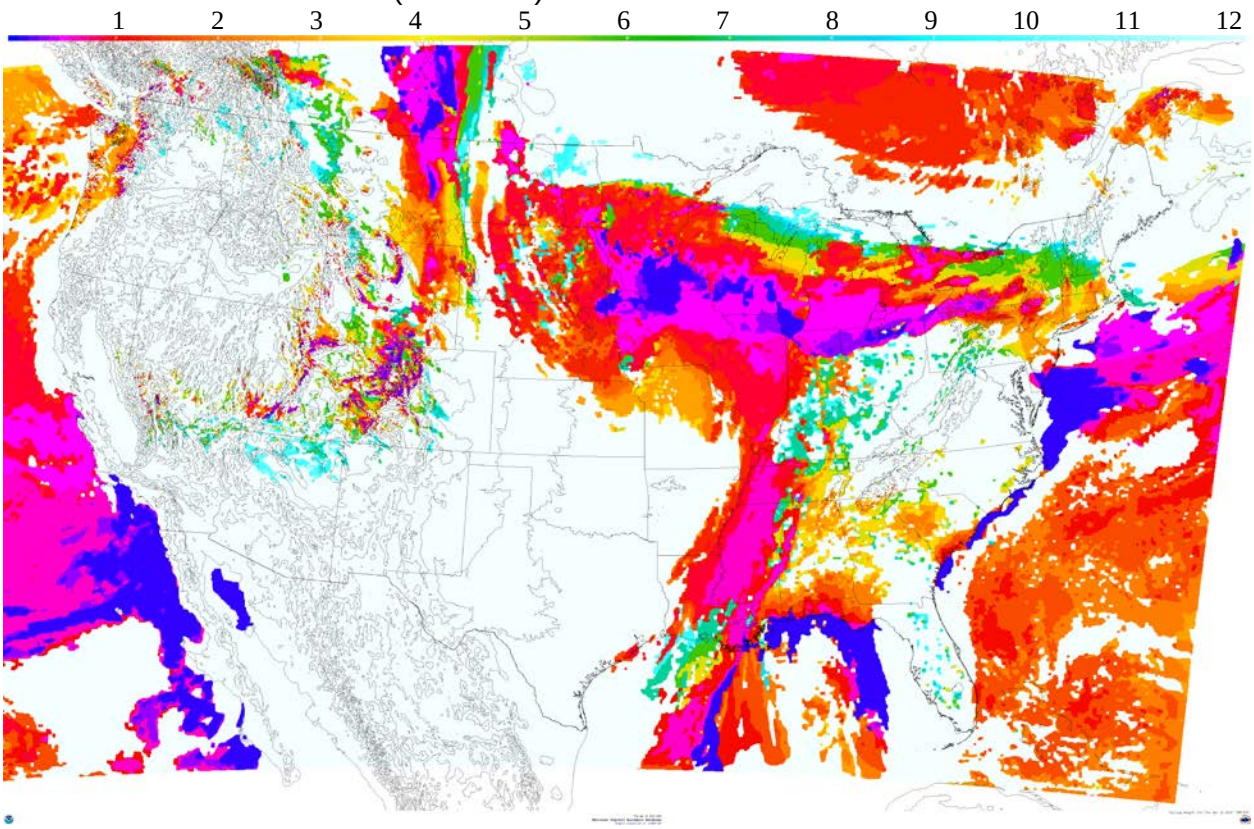

Figure 4. The HRRR ceiling height forecast as shown above but after removal or coalescing of small spots. Color bar is in thousands of ft.

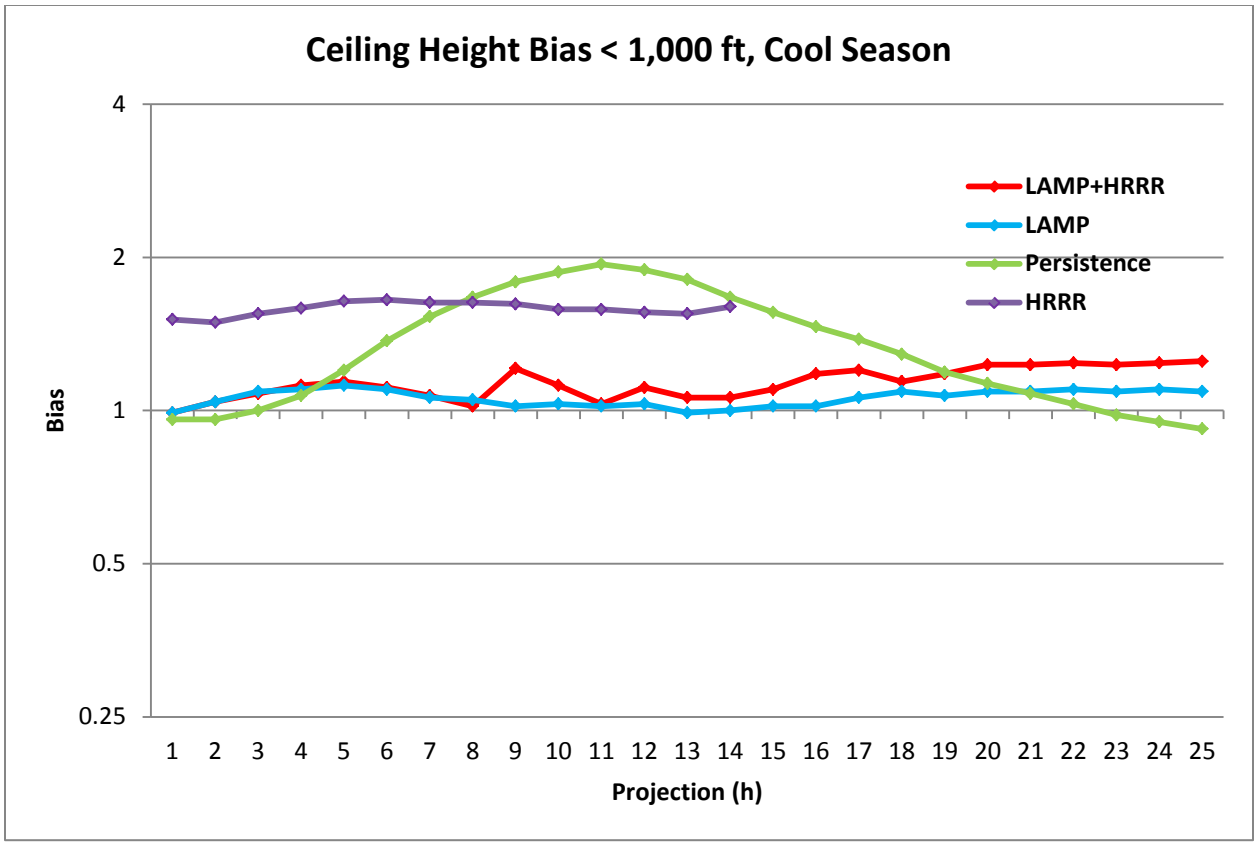

Figure 5. Ceiling height bias for events < 1,000 ft, cool season, 4 months independent data.

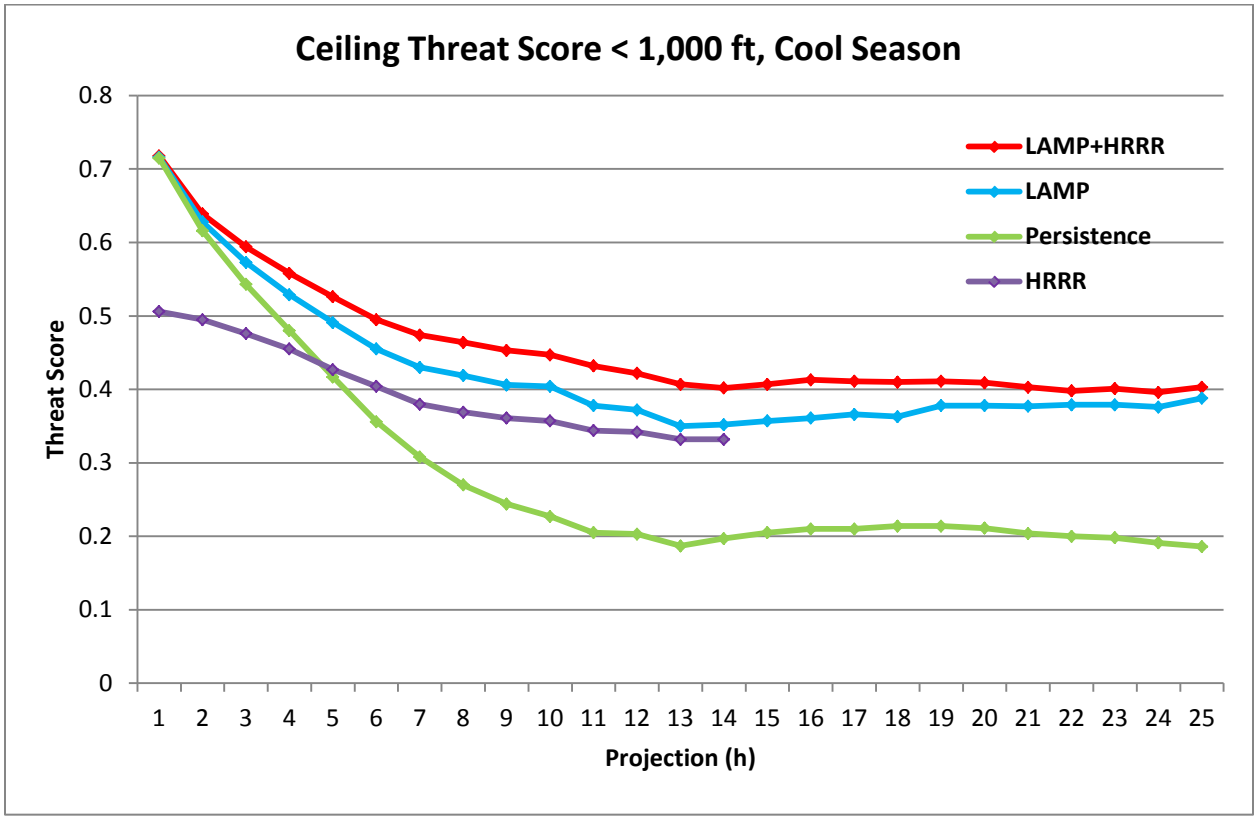

Figure 6. Ceiling height TS for events < 1,000 ft, cool season, 4 months independent data.

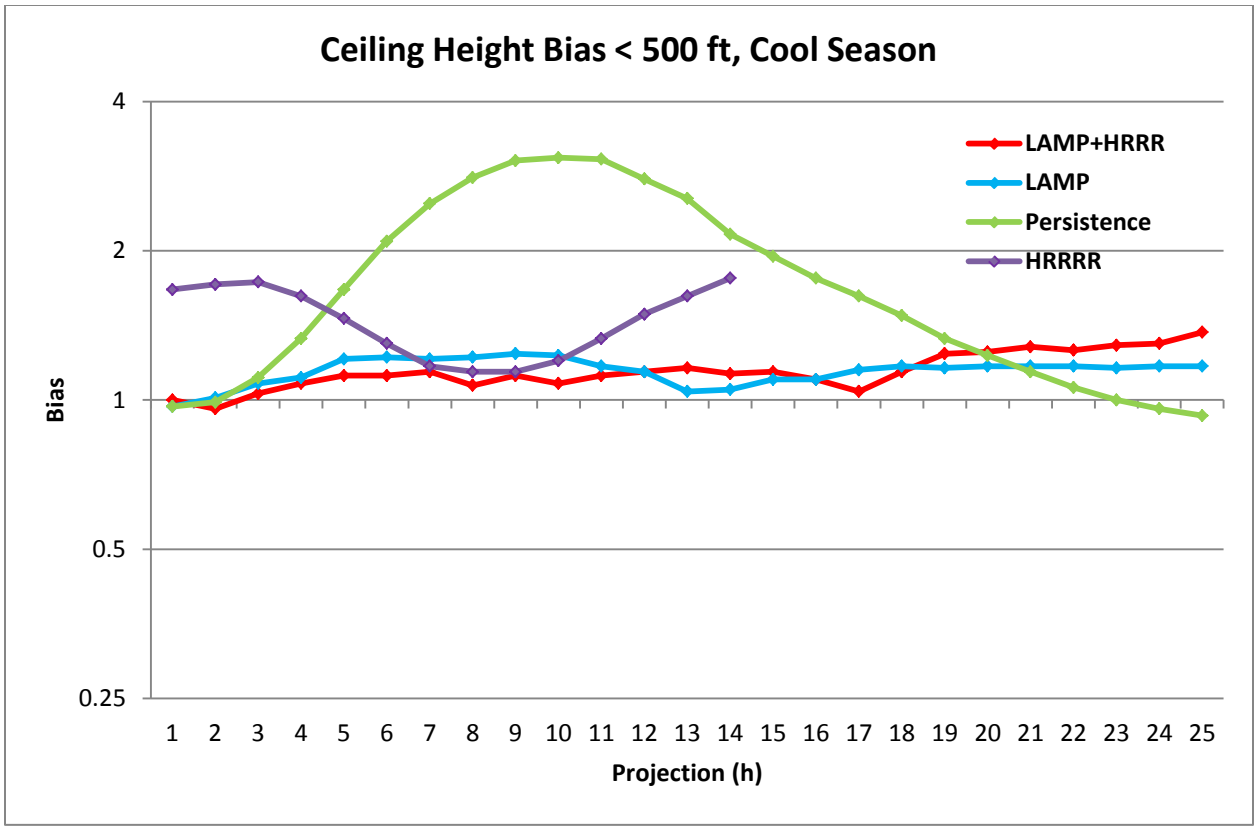

Figure7. Ceiling height bias for events < 500 ft, cool season, 4 months independent data.

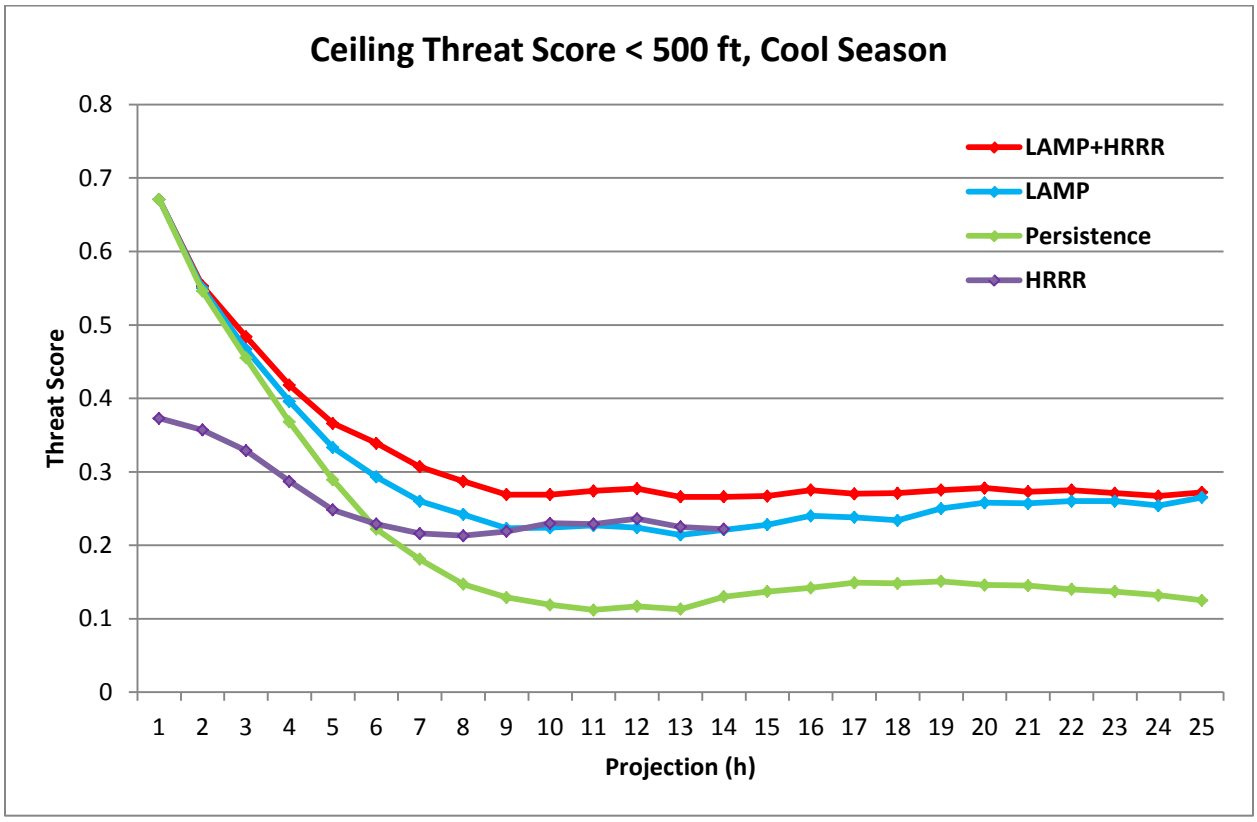

Figure 8. Ceiling height TS for events < 500 ft, cool season, 4 months independent data.

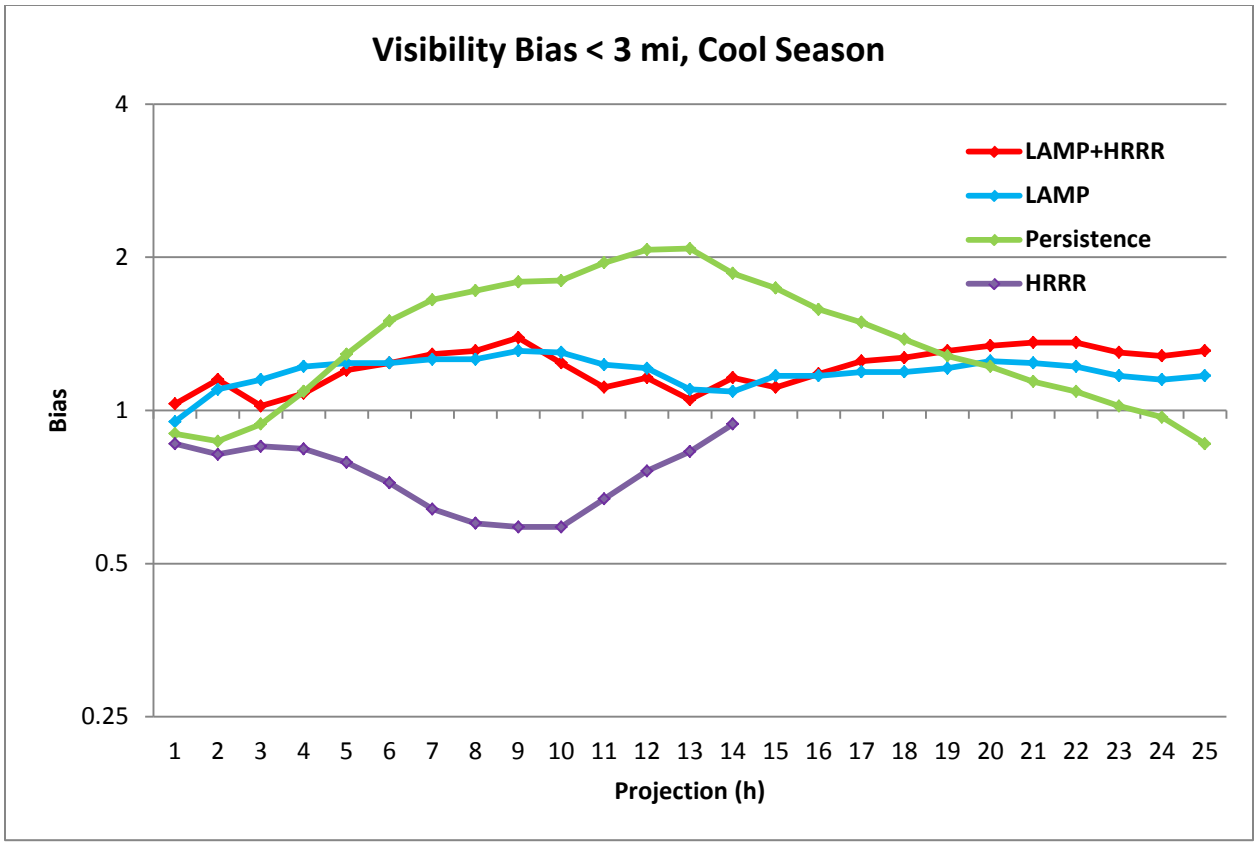

Figure 9. Visibility bias for events < 3 mi, cool season, 4 months independent data.

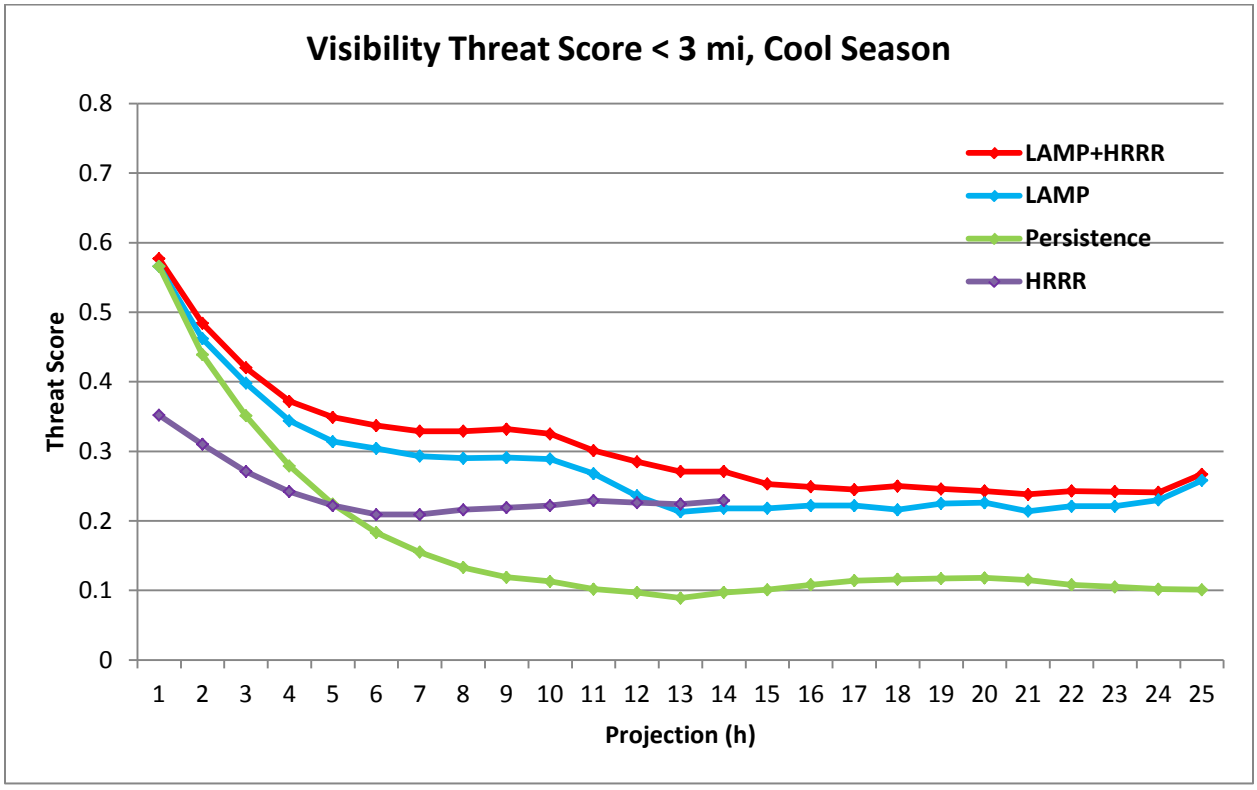

Figure 10. Visibility TS for events < 3 mi, cool season, 4 months independent data.

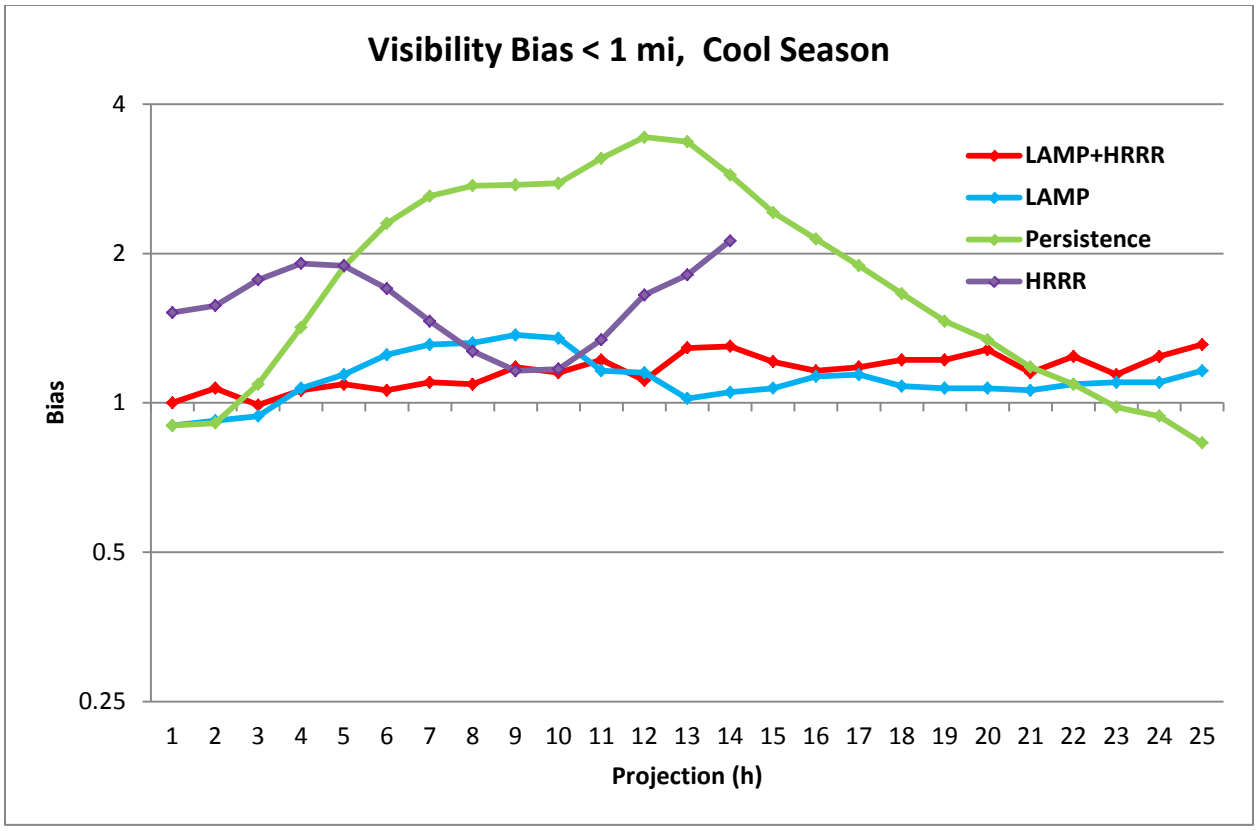

Figure11. Visibility bias for events < 1 mi, cool season, 4 months independent data.

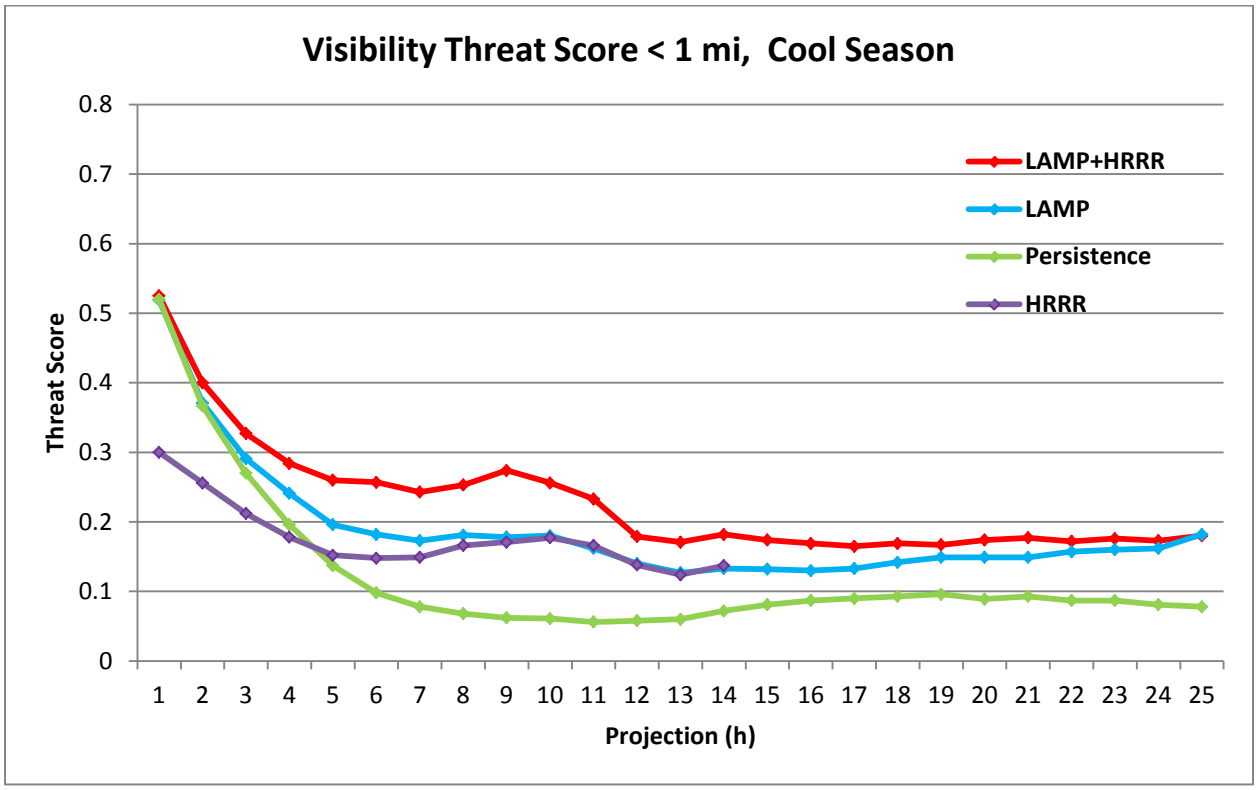

Figure 12. Visibility TS for events < 1 mi, cool season, 4 months independent data.

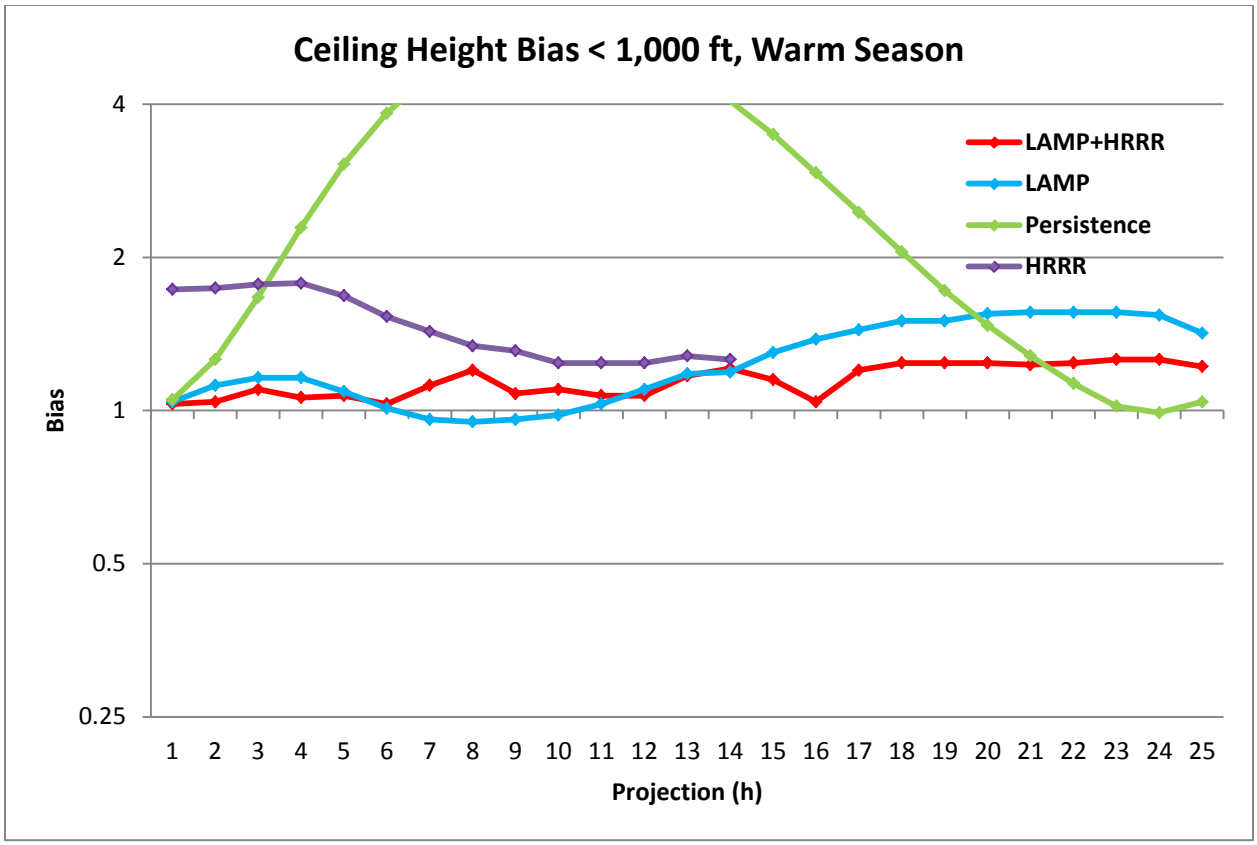

Figure 13. Ceiling height bias for events < 1,000 ft, warm season, 4 months independent data.

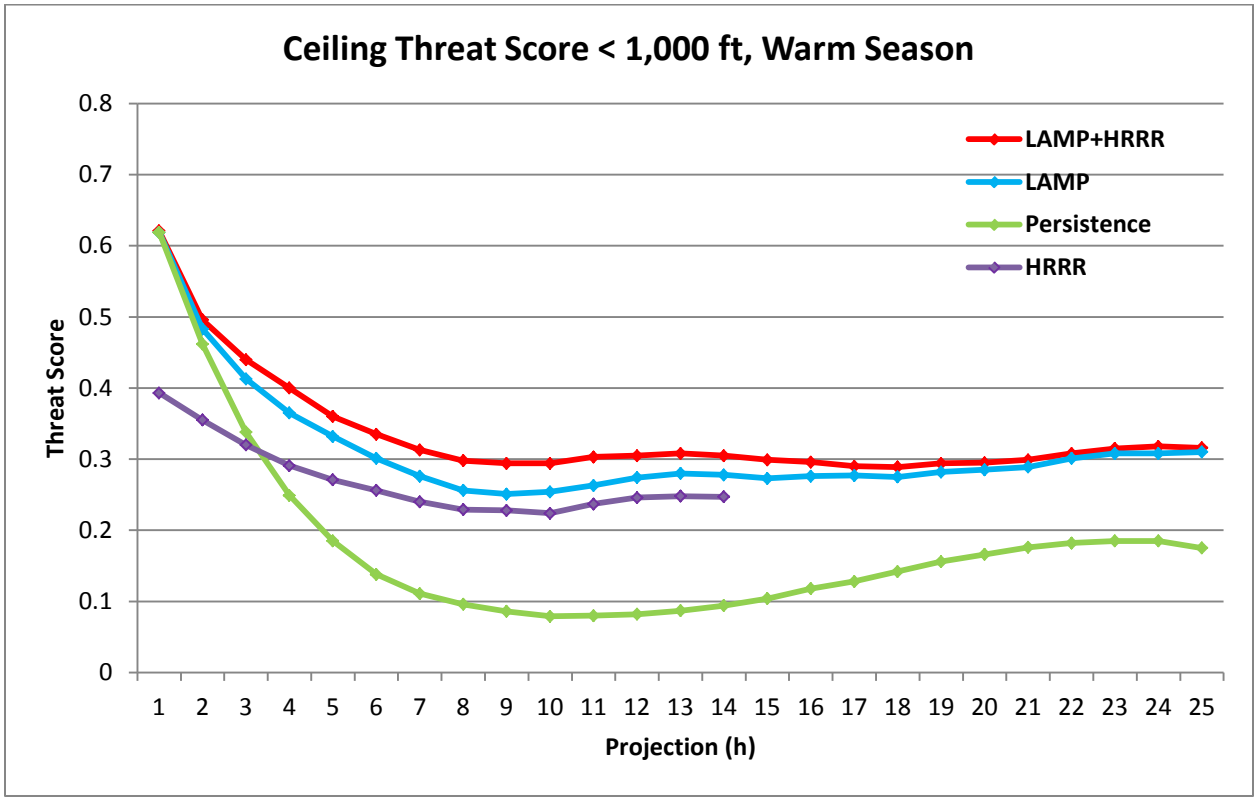

Figure 14. Ceiling height TS for events < 1,000 ft, warm season, 4 months independent data.

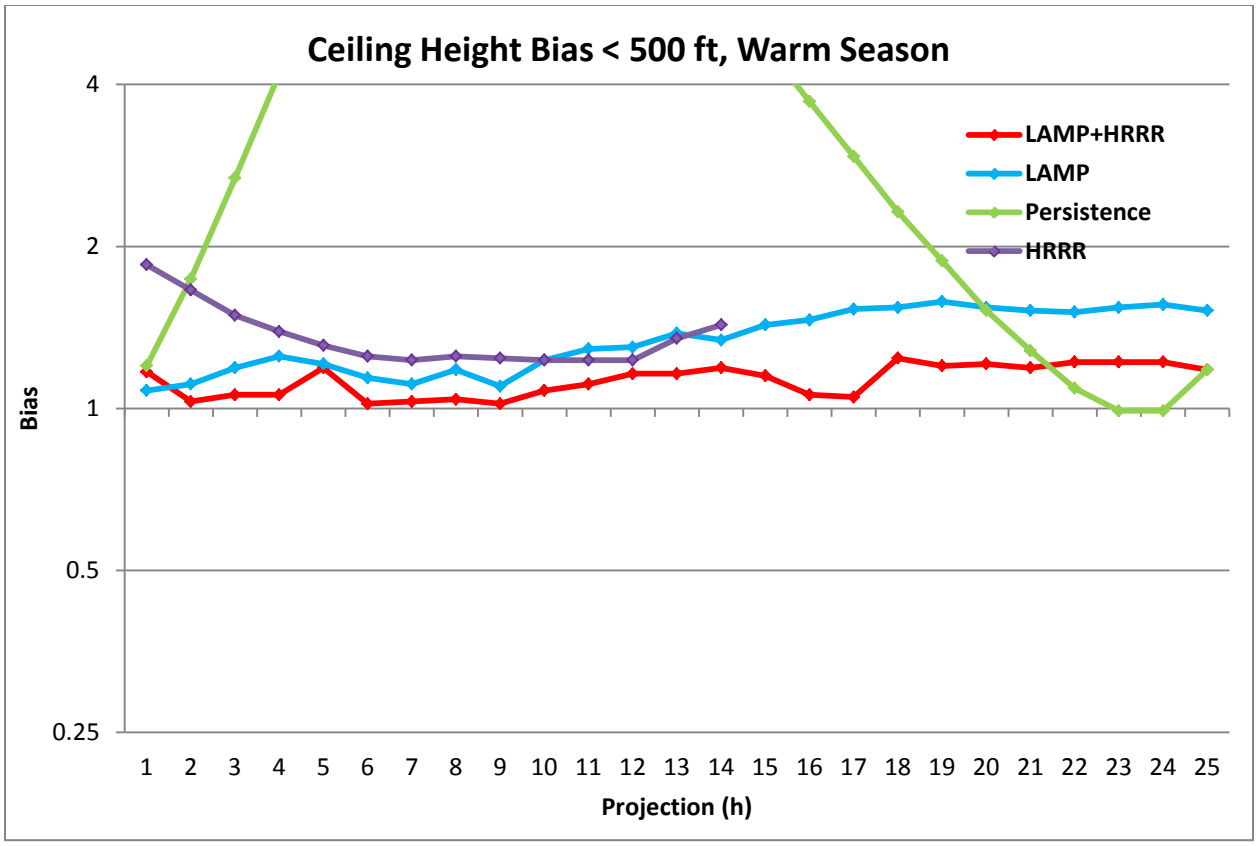

Figure 15. Ceiling height bias for events < 500 ft, warm season, 4 months independent data.

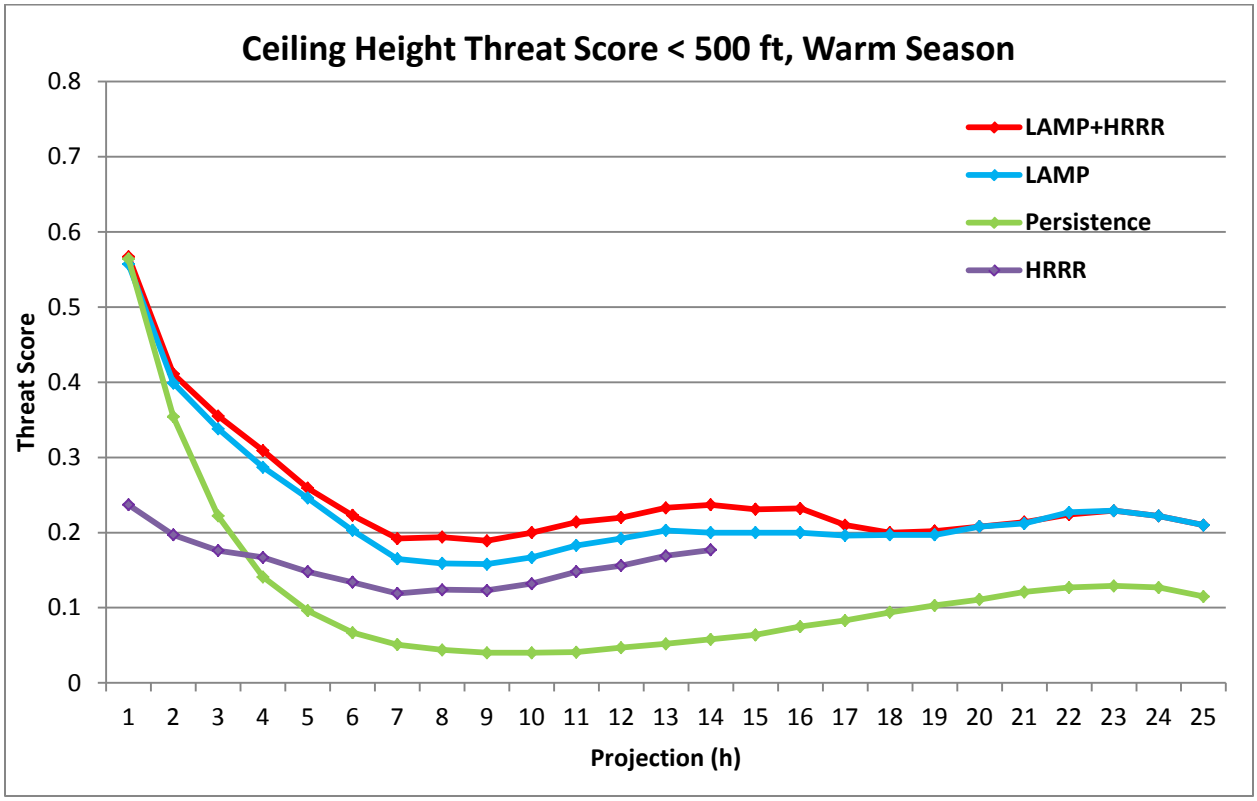

Figure 16. Ceiling height TS for events < 500 ft, warm season, 4 months independent data.

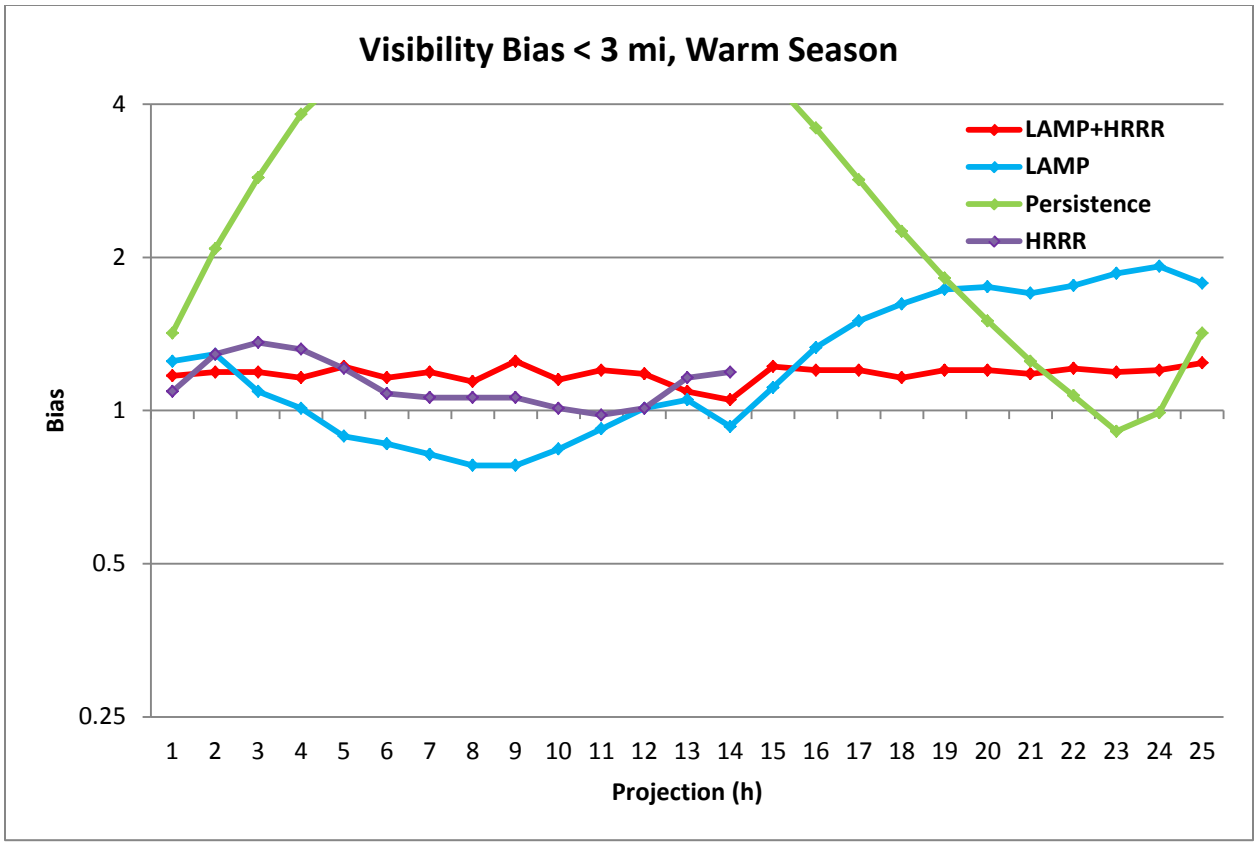

Figure 17. Visibility for events < 3 mi, warm season, 4 months independent data.

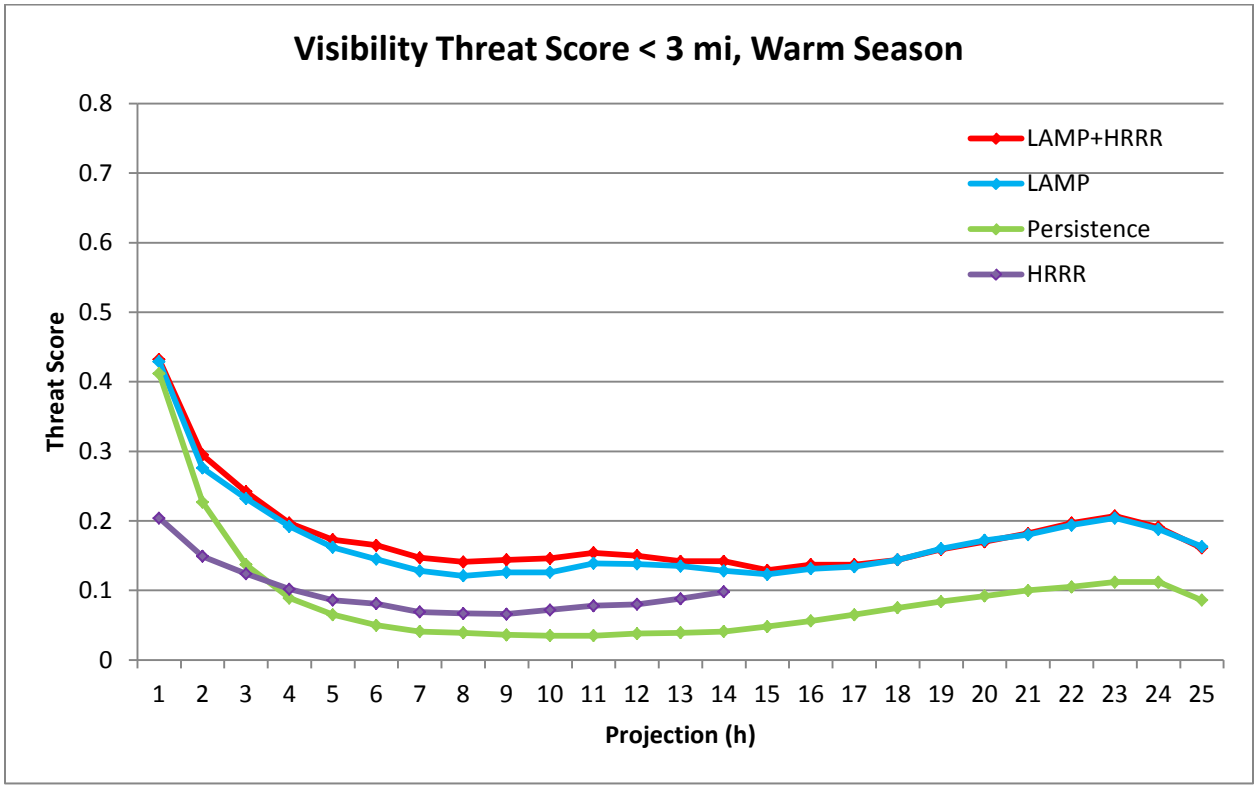

Figure 18. Visibility TS for events < 3 mi, warm season, 4 months independent data.

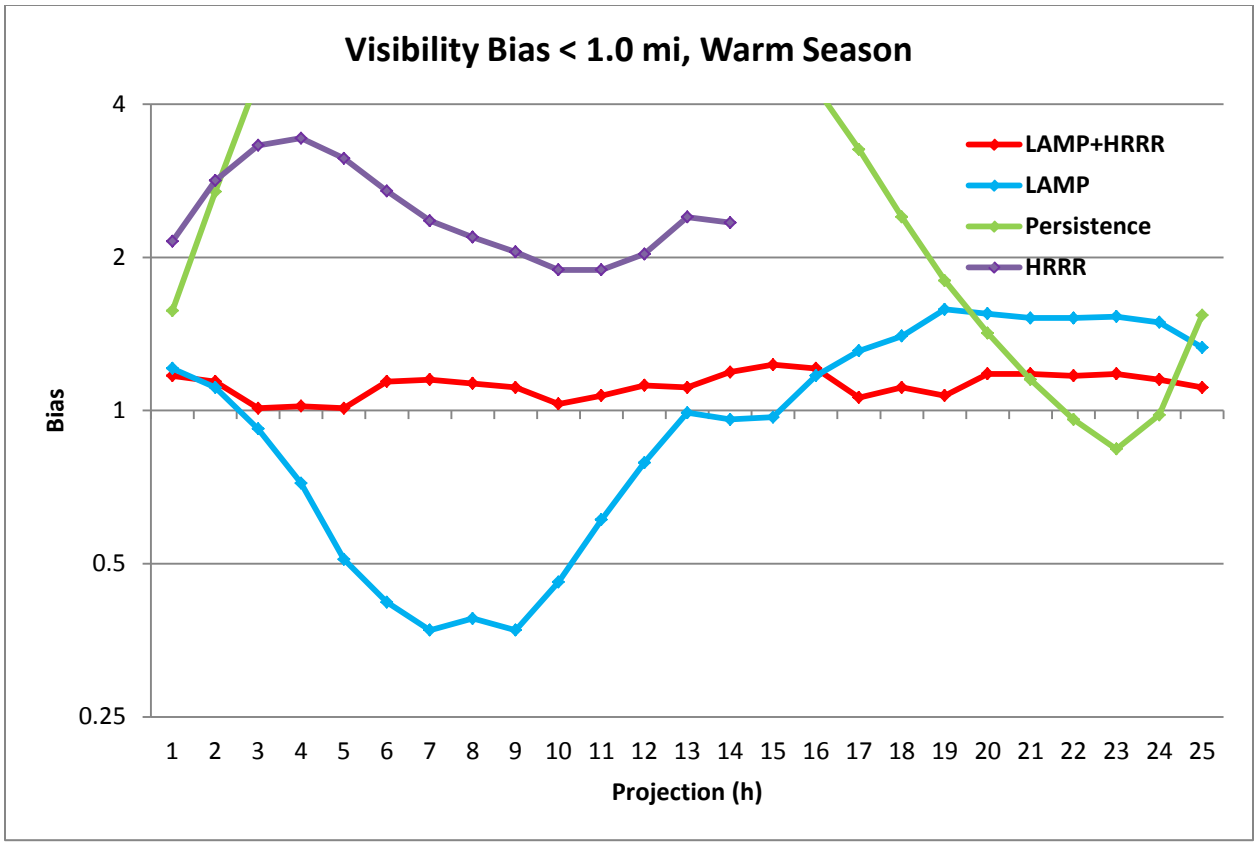

Figure 19. Visibility bias for events < 1 mi, warm season, 4 months independent data.

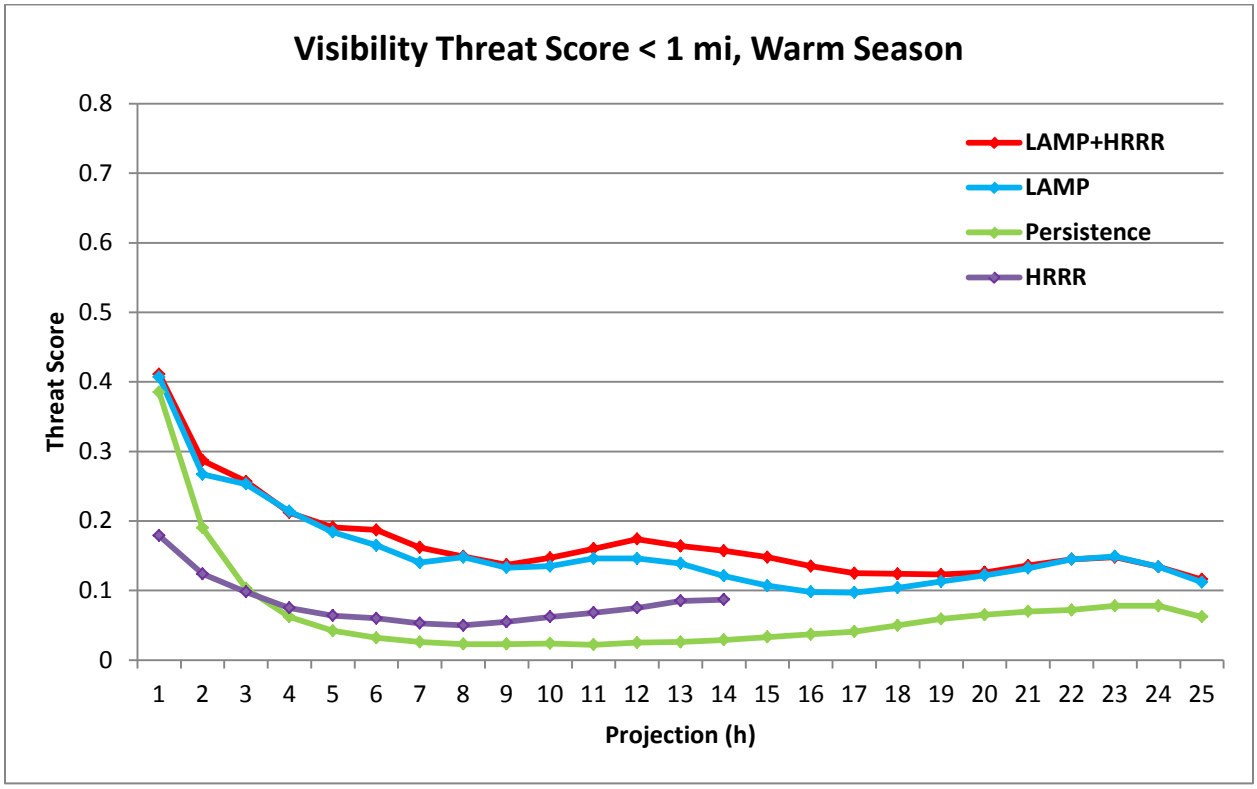

Figure 20. Visibility TS for events < 1 mi, warm season, 4 months independent data.

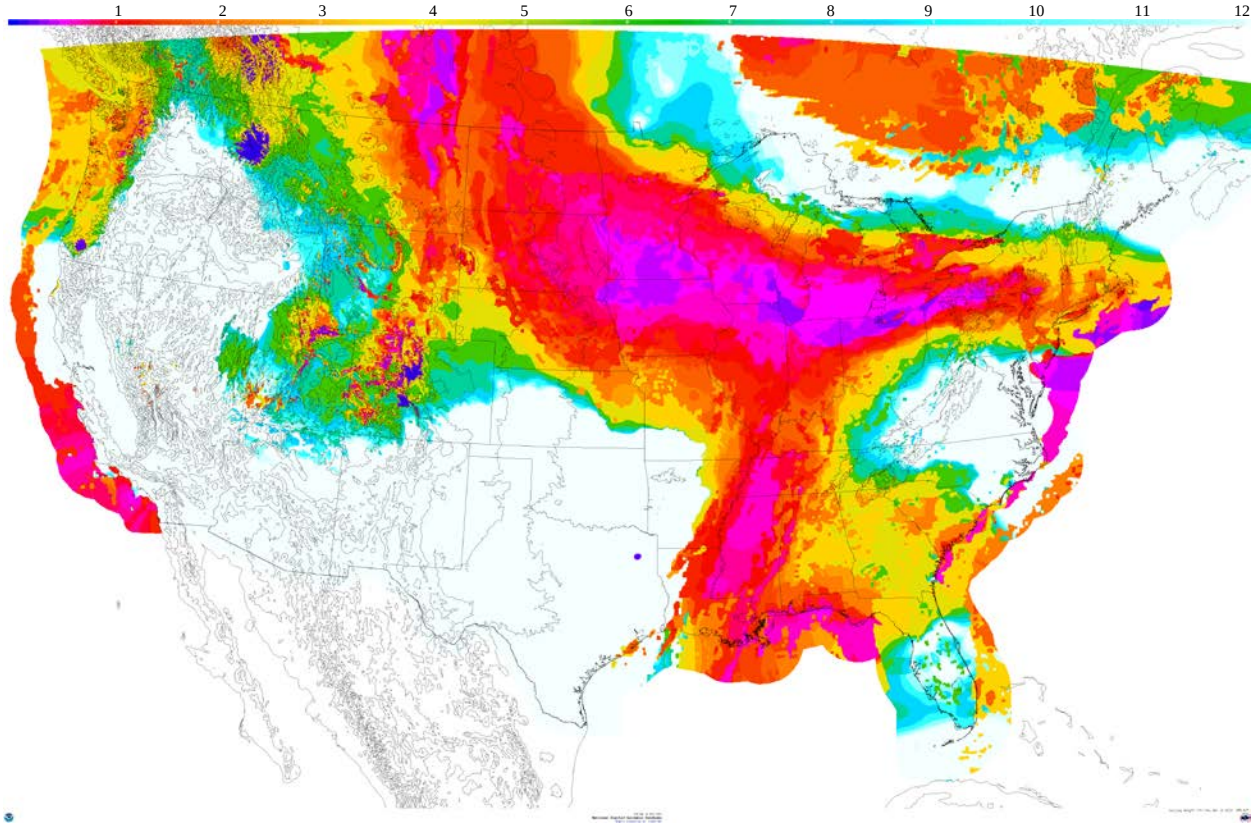

Figure 21. The ceiling 7-h Meld forecast from April 11, 2013, 1200 UTC. Color bar in thousands of ft.

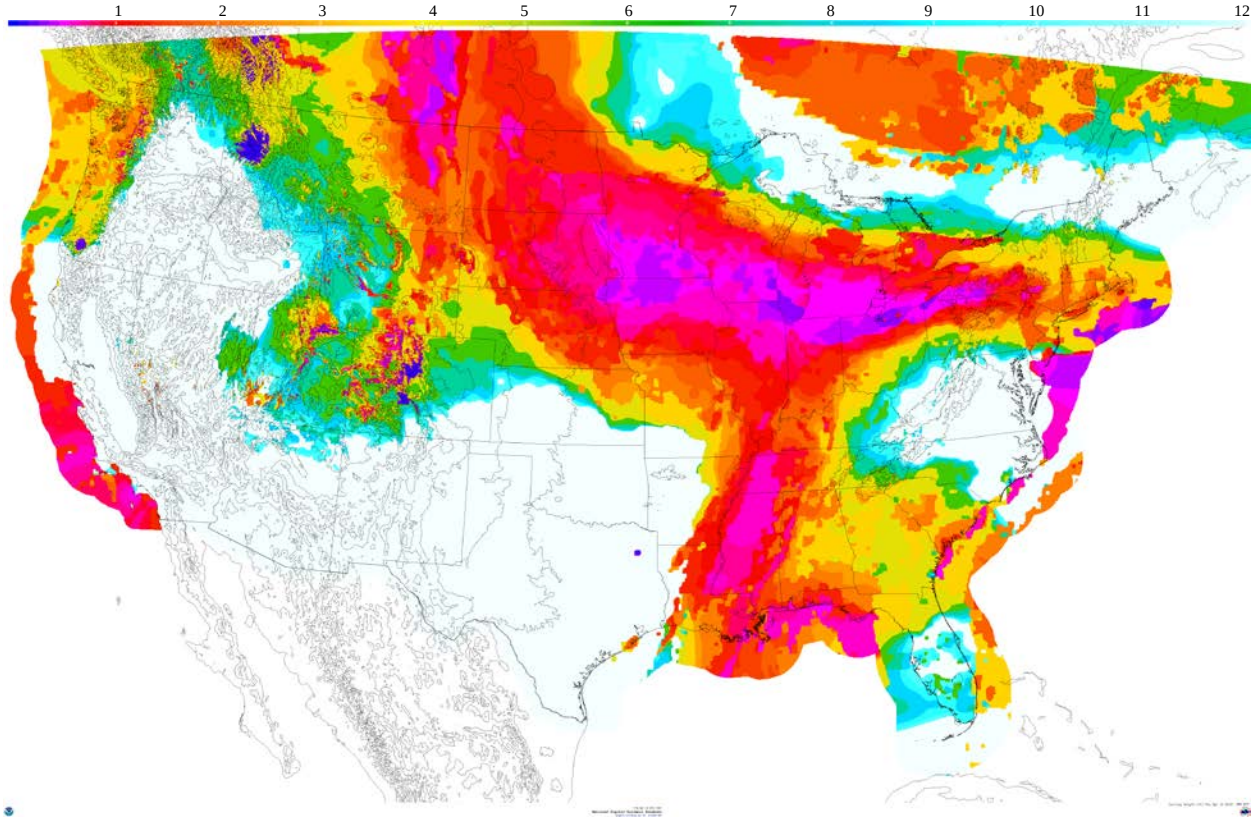

Figure 22. The same as Fig. 23, except after removal or coalescing of spots. Color bar in thousands of ft.

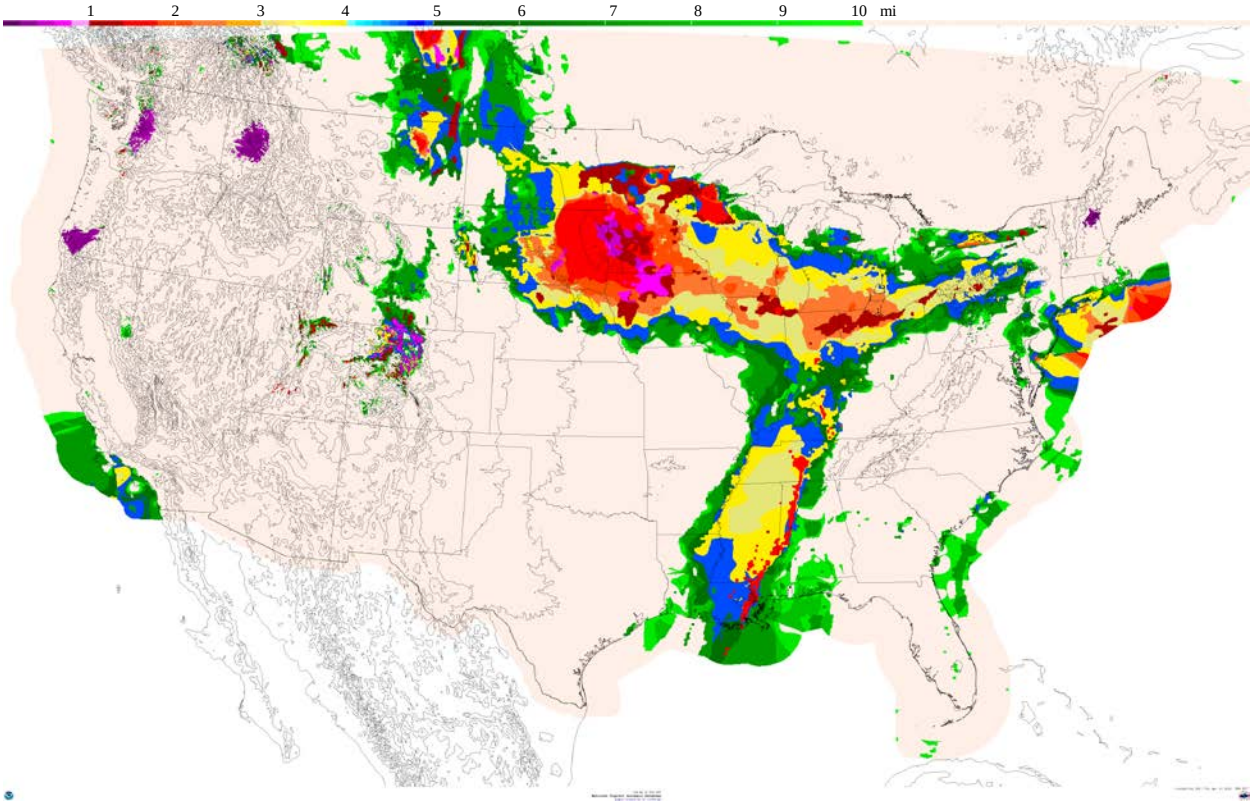

Figure 23. The visibility 7-h Meld forecast from April 11, 2013, 1200 UTC. Color bar in miles.

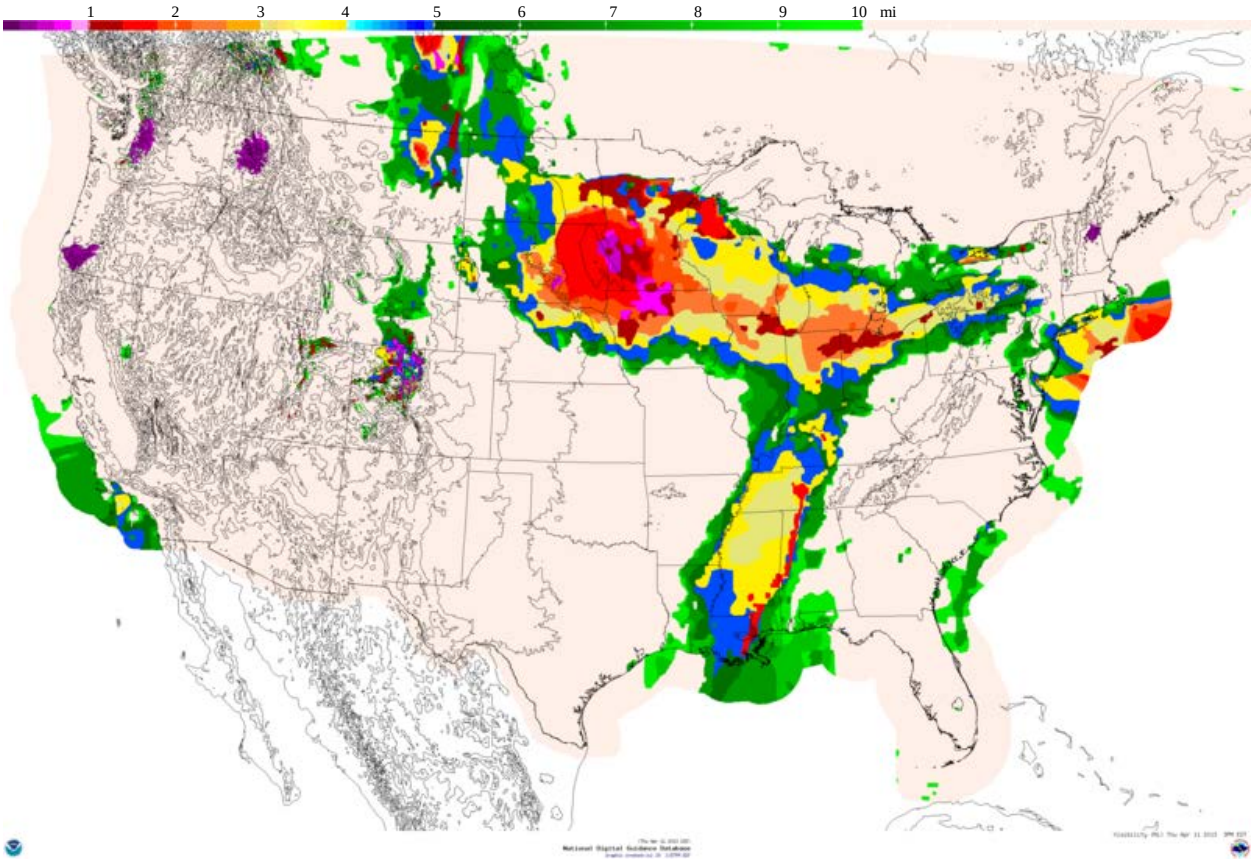

Figure 24. The same as Fig. 25, except after removal or coalescing of spots. Color bar in miles.