Quality Assessment Report:

National Ceiling and Visibility Analysis Product

Prepared By:

Quality Assessment Product Development Team
NOAA/ESRL/GSD

Andrew F. Loughe[2], Brian P. Pettegrew[2], Judy K. Henderson[1], Joan E. Hart[2],
Sean Madine[3], and Jennifer L. Mahoney[1]

30 November 2010
(enhanced analysis of sub-regions)

Affiliations:

1 – National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Global Systems Division (NOAA/ESRL/GSD)

2- Cooperative Institute for Research in Environmental Sciences (CIRES) and NOAA/ESRL/GSD

3 – Cooperative Institute for Research in the Atmosphere (CIRA) and NOAA/ESRL/GSD

Corresponding Author:

J.L. Mahoney (NOAA/ESRL/GSD, 325 Broadway, Boulder, CO 80303; Jennifer.Mahoney@noaa.gov)

Table of Contents

List of Tables

List of Figures

Executive Summary

This report summarizes a formal quality assessment of the National Ceiling and Visibility Analysis product (NCVA), a gridded analysis that evaluates reported ceiling and visibility information for the purpose of improving the flight planning process. On behalf of the Federal Aviation Administration's Aviation Weather Research Program, and in support of an Aviation Weather Technology Transfer (AWTT) D4 (operational) decision point, this study was carried out to examine the following:

- The quality of the NCVA product with respect to a baseline analysis.
- The effect on the NCVA of utilizing a satellite-based cloud mask.
- The potential value of NCVA's frequent update-cycle to the flight planning process.
- NCVA's performance compared to the operational Weather Depiction Analysis, a product specifically referenced in the NCVA Concept of Use.

Constrained to the use of METARs for verification, this assessment employs a cross-validation technique to create independence between the observational set and the input set that is utilized by the NCVA algorithm. Cross-validation statistics produced from METAR data that are withheld from the NCVA should not be interpreted as measures of the algorithm's skill *at* METAR sites, but rather as a measure of performance of the NCVA grid at points *away from* METAR sites. The NCVA algorithm retains information from all operationally available METARs used as input, never altering data at grid points associated with available METAR reports.

In the absence of a secondary gridded product that represents current operational planning information, the analysis team created a proxy baseline product for comparison. This product, referred to in this report as the Nearest Neighbor Analysis (NN-A), simply uses the METAR closest to each evaluation point in determining ceiling, visibility, and flight category values. The study period consists of the summer months of 2008 and the winter months of 2008-09, chosen to provide significant data for stratification of results by season and time of day.

Quality of the NCVA product with respect to a baseline analysis

Overall, the NCVA could add significant value to the planning process compared with the baseline analysis (NN-A):

- By more effectively detecting IFR events and reducing risk throughout the airspace (NCVA Probability of Detection of 0.71 vs. 0.60 for the NN-A).
- By more effectively reducing false alarms of IFR events, resulting in more efficient use of the airspace (NCVA False Alarm Ratio of 0.25 vs. 0.39 for the NN-A), with a lower False Alarm Ratio being more favorable.
- Continuous measures of the error in ceiling and visibility attributes concur with the observations stated above.

The quality of the NCVA and the NN-A differ during the wintertime and summertime:

- NCVA has a significantly higher detection rate of IFR events in the wintertime than in the summertime (0.77 vs. 0.60).
- The NCVA has a slightly lower (better) False Alarm Ratio in the wintertime than in the summertime (0.22 vs. 0.31).
- Continuous measures of the error in ceiling and visibility attributes concur with the observations stated above.

Quality measures differ between normal and low confidences:

- While the IFR detection rates for *normal* and *low* flight category confidences are similar (0.69 vs. 0.71), the False Alarm Ratio for the *normal* confidences is lower (0.17 vs. 0.25).
- For ceiling and visibility values, *normal* confidence analysis points were more accurate than *low* confidence analysis points: Root Mean Square Errors of 1527 ft. and 1.88 SM for *normal* confidences vs. 1821 ft. and 2.41 SM for *low* confidences.

NCVA performance was found to vary greatly by region, sub-region, and weather regime:

- Performance of the NCVA was found to be most favorable along the East Coast, the Southwest Coast, and in the Midwest. It performed much less favorably in the Intermountain West, High Plains region, and during the summertime along the Gulf Coast. Consistent among the sub-regions is demonstrated superiority of the NCVA over the NN-A.

- Skill was found to be greater in sub-regions possessing high METAR station density and where long-lived, large-scale IFR events occur frequently. This is evident along the West Coast in the summertime (sub-regional detection rate of 0.89) and in the Northeast in the wintertime (sub-regional detection rate of 0.91).

Effect of the NCVA algorithm's use of a satellite-based cloud mask

The study indicates that when the cloud mask is utilized there is**:**

- An overall decrease in false alarms (5.9%) that outweighs an increase in misses (0.5%), resulting in more efficient use of the airspace while only slightly increasing the risk; the NCVA risk is still significantly lower than that of the baseline.
- A measurable difference between the daytime and nighttime, as a large number of analysis grid points possessing *low* flight category confidence are actively re-assigned to *normal* flight category confidence during the daytime (24.3%), with only a negligible change during the nighttime.

Potential value to the flight planning process of the NCVA frequent update rate

Incremental changes in the updates of NCVA every five minutes appear to contain information useful to planners:

- The linear correlation between an NCVA issuance of flight category to its successor one hour later is ~0.8, representing a significant change in flight conditions over the CONUS.
- The five-minute updates that occur between hourly issuances appear to effectively capture the incremental change in flight category over CONUS, as indicated by linear changes in correlation between the initial and intermediate issuances.

Comparison of the NCVA to the Weather Depiction Analysis

This study shows that the NCVA, in accordance with its Concept of Use, performs as well as and is consistent with the Weather Depiction Analysis:

- NCVA and the Weather Depiction Analysis have an overall correlation greater than 0.6, indicating that the two products may perform in a consistent manner.
- With no suitable way to directly measure the quality of the Weather Depiction Analysis, the study found indications that NCVA performs at least as well as the Weather Depiction Analysis.

1. Introduction

In the context of pre-flight planning, the goal of this evaluation is to determine how well the National Ceiling and Visibility Analysis product (NCVA) characterizes categorical flight conditions, ceiling, and visibility across the Continental United States (CONUS). In performing this assessment, we evaluate individual attributes of the NCVA, including categorical flight conditions, confidence fields, ceiling, visibility, and issuance frequency. The assessment is performed primarily using a cross-validation technique (Neter 1996) in which portions of the observational dataset are selectively removed to arrive at a measure of interpolation error *in between* individual METAR reporting sites. This will be explained in detail in Section 3.1. During the course of this evaluation, four main questions will be addressed.

The first question is to assess the overall performance of the NCVA and to determine if its performance is a function of confidence level, season, or analysis time. To achieve this aspect of the assessment, we utilize a baseline measure of skill against which the analysis product itself is compared. By measuring relative skill of the product, we are able to answer questions regarding its overall usefulness. Secondly, we assess the effect of the use of satellite cloud-mask data by the NCVA product. This part of the analysis was motivated by previous work (QA PDT member M. Kay, personal communication) in which questions were raised regarding the skill of an earlier version of NCVA when the cloud mask was being used to identify regions of clear skies or no ceiling. The third question addresses whether the 5-minute update frequency of the NCVA provides an added benefit over less-frequently issued analyses. We will measure correlation of the analysis using successive 5-minute issuances over the course of an hour. Finally, the Concept of Use (CONUSE) document for the NCVA stipulates that the automated product must perform as well as manually generated products and must be consistent with those products (FAA/ATO 2007). We will directly compare the NCVA with the Weather Depiction Analysis product that is provided hourly to forecasters via the Advanced Weather Interactive Processing System (AWIPS).

In addressing these questions, the evaluation will take on the following form: Sections 2 and 3 will highlight the data and methodology used in the study, Section 4 provides a description of performance measures, Section 5 presents results and discussion, and Section 6 provides concluding remarks.

2. Data

2.1 NCVA Product Version 1.1

In the context of pre-flight planning and go/no-go decision making, the NCVA product utilizes METAR observations from across the CONUS and portions of Canada and Mexico to create an estimate of ceiling and visibility values within the 5-km National Digital Forecast Database (NDFD) domain. The NDFD domain covers the CONUS and portions of Canada and Mexico. According to the NWS, the NDFD provides a seamless mosaic for forecast and analysis of sensible weather elements (Glahn et al. 2003). A categorical flight condition is based upon ceiling and visibility values reported at METAR observation points and at interpolated grid points which lie in between the METAR sites. The interpolation does not extend into oceanic regions. The NCVA product is updated every 5-minutes to ensure that significant weather changes are captured when surface observing stations issue off-hour special reports.

According to the NCVA description document (Herzegh et al. 2009), approximately 2,100 METAR reporting stations from Canada, Mexico, and the United States are used to create the NCVA product. Stations with valid reports that are no more than 80-minutes old are eligible to be included for a given analysis time. The NCVA uses a nearest neighbor interpolation scheme (Skiena 1997) that produces a somewhat patchy and discontinuous field of ceiling, visibility, and flight category every 5 minutes. Visibility observations are not altered for the interpolation; however, the ceiling height, reported in feet above ground level (AGL), is converted to mean sea level (MSL) to account for natural atmospheric processes. It is then interpolated and converted back to feet AGL using terrain data from the grid. A cloud mask developed by the NASA Global Hydrology and Climate Center (Jedlovec et al. 2003) utilizes infrared and visible channels of the GOES-11 and GOES-12 satellites to augment portions of the ceiling and visibility grid for "clear" conditions. A flight category is derived based on the worst conditions of ceiling and visibility at each grid point, following thresholds defined in **Table 2.1**. For example, if a grid point has a ceiling of 4,000 ft. AGL and a visibility of 4 SM, it is assigned a flight category of MVFR, based primarily on the reported visibility. Finally, two user confidence levels for ceiling, visibility, and flight category are assigned to each grid point, either "normal" confidence or "low" confidence. These levels are based upon different analysis attributes, including grid point distance from the METAR observation, dew point depression, elevation differences, and sky conditions. The use of the cloud mask also affects the confidence field, based on those portions of the grid that are in daylight and those that are not. See **Figure 2.1** for an example of NCVA output.

Table 2.1: Determination of flight category from separate values of ceiling and visibility (FAA, 2008).

Flight Category	Ceiling Height (ft AGL)	Visibility (SM)
Visual Flight Rules (VFR)	ceiling > 3000	vis > 5
Marginal Visual Flight Rules (MVFR)	1000 ≤ ceiling ≤ 3000	3 ≤ vis ≤ 5
Instrument Flight Rules (IFR)	500 ≤ ceiling < 1000	1 ≤ vis < 3
Low Instrument Flight Rules (LIFR)	ceiling < 500	vis < 1

2.2 METARs

The Aviation Routine Weather Report (METAR) and Aviation Selected Special Report (SPECI) are disseminated from automated surface observing stations, including the Automated Surface Observing System (ASOS) and the Automated Weather Observation System (AWOS). Conditions reported by these automated stations include sky conditions, visibility, precipitation, precipitation type (ASOS only), temperature, dew point, wind, and atmospheric pressure (NWS 2000). Visibility is measured using a forward scatter visibility meter that measures the clarity of the air. METAR visibility reports are obtained from a 10-minute average of 1-minute visibility measurements. The measurement is deemed representative of visibility within 2-3 miles surrounding the site. A physical limitation of 10 statute miles (SM) exists based on the limits of human visibility. Sky conditions, which include cloud base height and ceiling measurements are determined by a laser beam ceilometer, which calculates a 30-minute average from 30-second cloud base height measurements. Sky conditions are representative of a 3-5 mile radius surrounding the site. In order to account for abrupt sky condition changes, the last 10 minutes of

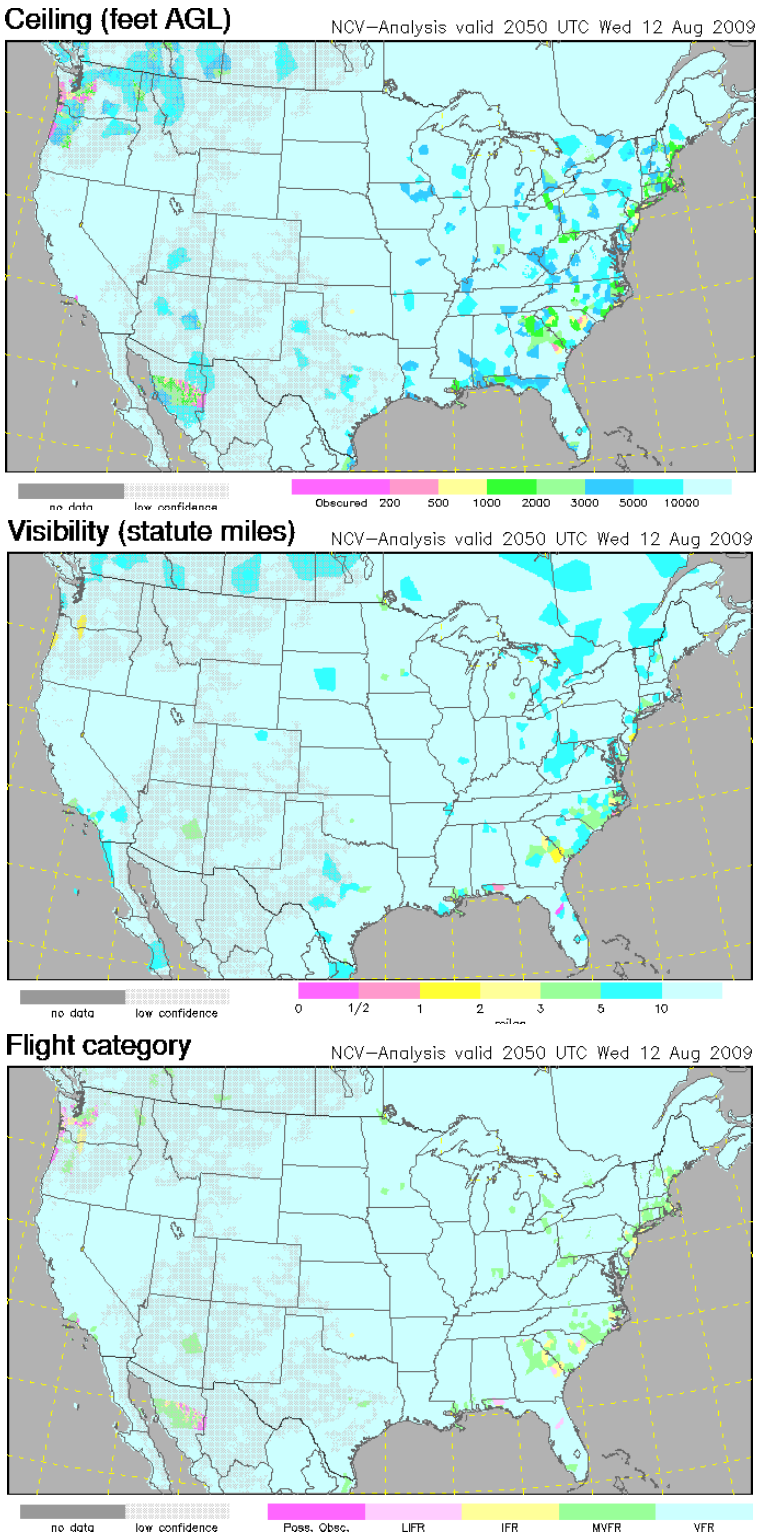

Figure 2.1: Sample NCV Analysis product from 2050 UTC on 12 August 2009 from the Experimental Aviation Digital Data Service (*http://weather.aero*). Displayed are Ceiling (ft. AGL; top), Visibility (SM; middle) and Flight Category (bottom). Note that hatched areas represent areas of low confidence. Ceiling, visibility, and flight category are color coded based on severity.

ceilometer measurements are double weighted in the algorithm. The ceilometer has a physical measuring limitation of 12,000 ft. AGL.

2.3 Weather Depiction Analysis

The Weather Depiction Analysis is an hourly interpolated product generated by the National Centers for Environmental Prediction (NCEP) for the purpose of providing flight category information. The analysis displays regions of IFR or worse conditions within shaded contours, and regions of MVFR conditions within un-shaded contours (**Figure 2.2**). All other regions with a high cloud base or high visibility are otherwise deemed to represent VFR conditions and are not contoured (Senior Duty Meteorologist, SDM@noaa.gov personal communication). These hourly analyses are augmented every 3 hours by forecasters from NCEP's Hydrometeorological Prediction Center (HPC) by adjusting contours and adding fronts using the NCEP Advanced Weather Interactive Processing System (NAWIPS) software. This augmented product, the Weather Depiction Chart (*http://aviationweather.gov/data/iffdp/2020.gif*; **Figure 2.3**), is distributed as an online graphic file which is available to forecasters and is specifically identified in the NCVA CONUSE document as an official manual analysis product for ceiling and visibility. The hourly analysis files are distributed through NOAAPORT, a broadcast system disseminating real-time environmental data to NOAA and other external users, and are accessible to forecasters and decision makers through AWIPS as a graphics file.

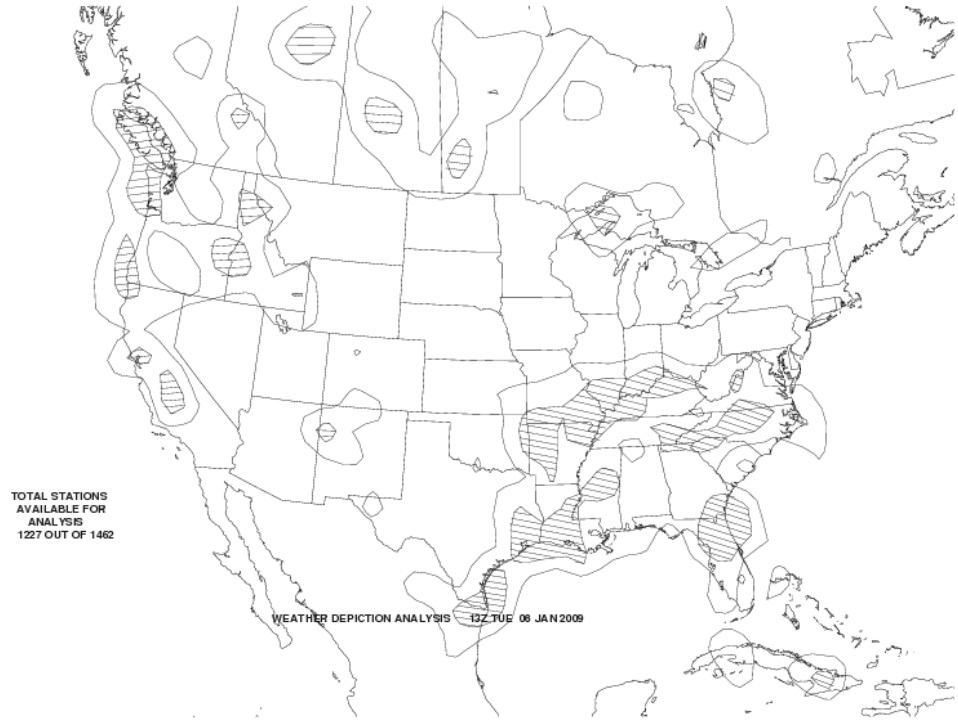

Figure 2.2: Hourly Weather Depiction Analysis as obtained from NOAAPORT valid at 1200 UTC for 06 Jan. 2009. Displayed are IFR (shaded with contour), MVFR (contoured) and VFR.

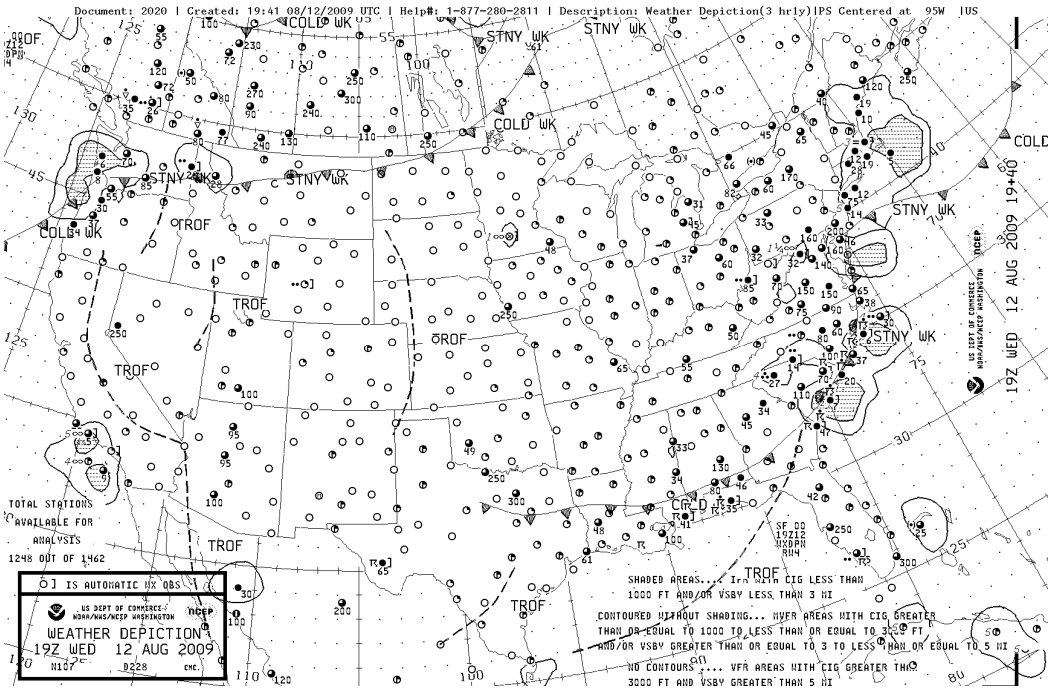

Figure 2.3: Weather Depiction Chart valid at 1900 UTC on 12 AUG 2009. Displayed are IFR (shaded with contours; ceiling and visibility observations labeled), MVFR (contoured; ceiling and visibility observations labeled) and VFR (ceiling and visibility observations labeled where observed, but not contoured).

The hourly graphics files for the months of July 2008 and January 2009 were obtained from an archive of NOAAPORT data. The graphics files were then converted into a vector graphics format, from which valid latitude and longitude values of contoured information were obtained. The contour information was then projected onto the NDFD grid for direct comparison with the NCVA product. By transforming the hourly analysis graphic onto the NDFD grid, we are able to perform grid-to-grid correlations between the Weather Depiction Analysis and the NCVA product.

3. Methodology

3.1 Cross-Validation Technique

A cross-validation technique is used to create independence between METARs used to verify the NCVA product and the input set that is utilized by the NCVA algorithm itself. By design, NCVA skill is perfect at individual METAR reporting sites (with respect to operationally available data) because the analysis does not deviate from values reported *at* those sites. The cross-validation technique, utilizing METAR data that are selectively withheld, provides a measure of how well the analysis performs *in between* METAR reporting sites. In utilizing cross-validation, we assume that there is a minimal effect on the quality of the NCVA by withholding a portion of the METAR input, and that the entire grid domain is being effectively sampled.

The specific method utilized for this study is referred to as random repeat cross validation. We perform 10 iterations, wherein 20% of METAR stations are withheld during each run. For

this study, we utilize a pool of 1700 stations over the CONUS and Canada that report at least 90% of the time. From these frequently-reporting stations, we create 10 separate lists of randomly selected withholding sites, comprising 340 stations per iteration. For each valid time in the study, 10 separate analyses are computed using the remaining 1360 stations (**Figure 3.1**). Once the selective data holes have been created within the analysis grids, we gather analysis output from the 10 iterations and compare results with reported values from the METARs that were held out. In this report, we assess the error in estimating ceiling, visibility, and flight category at each of the withheld METAR locations. This study is carried out over June-August 2008 (summer) and December 2008-February 2009 (winter). **Figure 3.2** shows a climatological distribution of IFR events by month, with a noticeable peak during the wintertime.

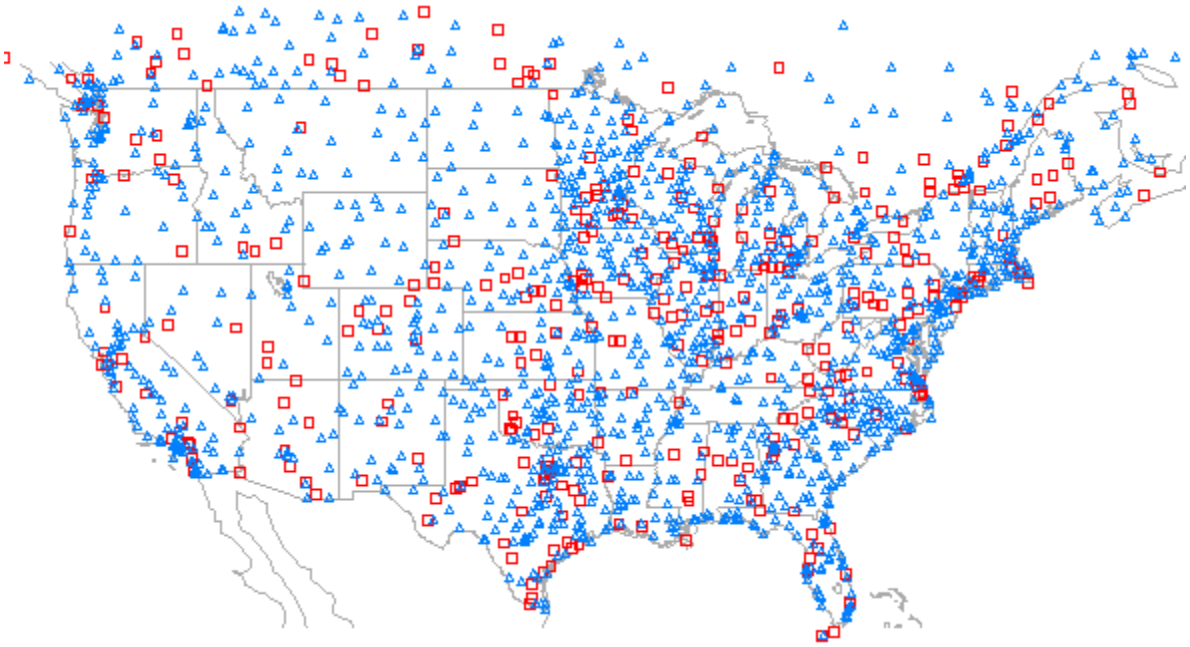

Figure 3.1: Distribution of METAR sites throughout the CONUS. These METARs have been determined to report at least 90% of the time. Blue triangles represent those METARs that were included in the set #10 cross-validation runs (~1360 stations), while red squares are those METARs that were withheld from that particular run (~340 stations).

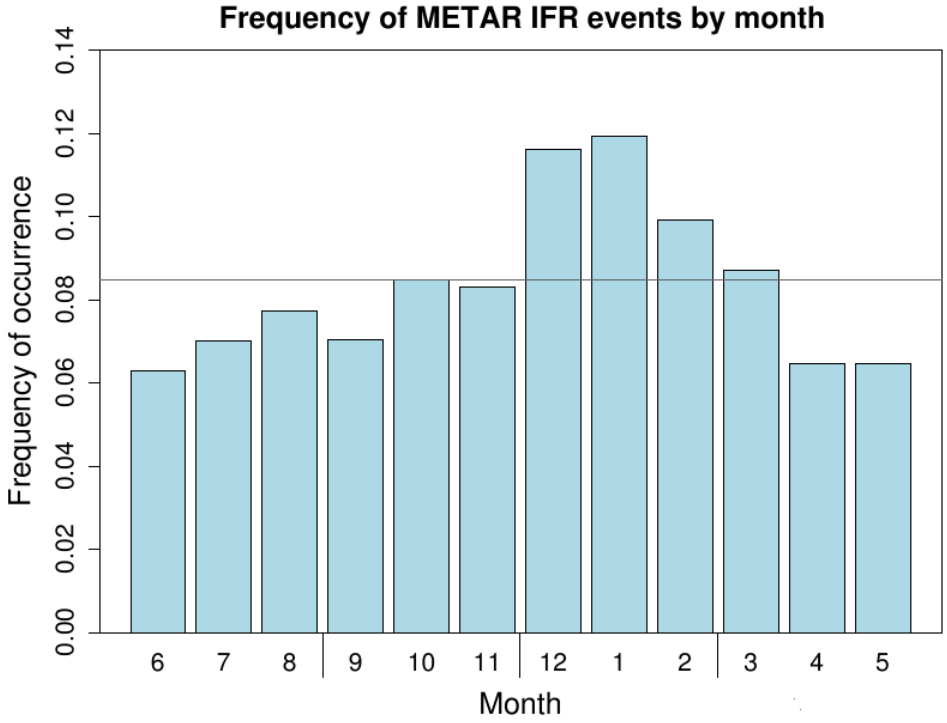

Figure 3.2: METAR IFR event climatology by month, computed over 2004-2008. Note the peak frequency during the winter months which is above the expected frequency of 1/12 (0.08333), as noted by the horizontal line within the figure.

3.2 Matching

When comparing the NCVA gridded product to METARs, we match grid values to specific station values in several different ways, including use of the 4 grid points surrounding a METAR site, and using all analysis points that are within a given radius of a METAR (R=10-, 20-, 30-km). Additionally, we tested a variety of matching criteria such as "best flight category" match, which is a value comparison, and the "worst weather condition" match, which is the lowest independent flight category, ceiling, and visibility value.

Testing of skill measures using each of these matching methods and matching criteria provided sufficient evidence to elucidate methods and criteria that could best be utilized for this study. For the balance of this study, we focus on radius 10 (R=10-km) "best flight category" matching based on physical attributes of METAR reporting. The radius 20 and 30 (R=20-, 30-km) results were best overall, but these radii are larger than desirable, and not as representative of ceiling and visibility forecast products, like the Terminal Aerodrome Forecast (NWS 2005).

3.3 Baseline Skill

In order to assess usefulness of the NCVA product, a level of skill (positive or negative) must be explicitly defined. In this evaluation, a baseline measure is computed by creating a surrogate analysis that is compared with actual values from the withheld METARs, just as the NCVA is compared with values from the withheld METARs. The approach utilizes a comparison of the

withheld METAR to its nearest neighbor METAR in arriving at a measure of baseline skill. In this report, we refer to the baseline analysis as the nearest-neighbor analysis (NN-A). The actual skill of the NCVA against withheld METARs is compared with skill of the NN-A against withheld METARs. The goal is to ascertain performance of the NCVA compared with the NN-A.

4. Skill Assessment

4.1 Dichotomous Assessment of Flight Category

Over the time period of June-August 2008 and December 2008-February 2009, a 4x4 contingency table of results is constructed for the analysis of categorical flight conditions using withheld METAR/NCVA pairs obtained from the cross-validation technique. However, statistics are not computed directly from the 4x4 contingency table. This multi-dimensional table is reduced to a dichotomous table representing conditions for IFR and non-IFR (**Figure 4.1**). Traditional statistics (i.e. PODy, PODn, Bias, False Alarm Rate, False Alarm Ratio, etc; **Table 4.1**) are derived from this dichotomous field.

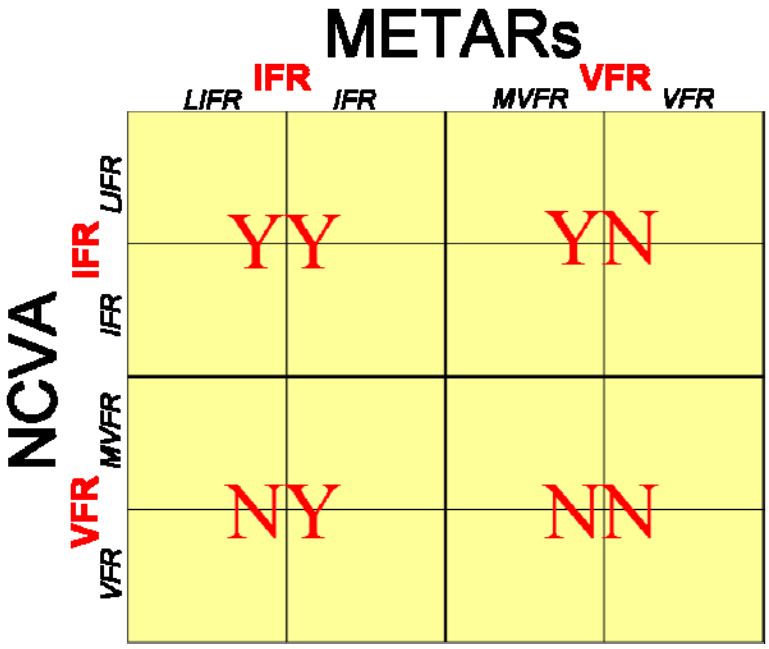

Figure 4.1: 4 X 4 contingency table for categorical flight conditions for the NCVA product vs. METAR observations (in italics). A reduced 2 X 2 (dichotomous) contingency table is labeled in red within the context of the larger 4 X 4 contingency table.

If the NCVA were perfectly able to detect all instances of observed IFR conditions over the NDFD grid, it would possess a hit rate (PODy), threat score (CSI), and bias of 1.0. For the hit rate, this means that all observed IFR conditions were properly analyzed, and that there were no misses. For CSI, this would indicate that there were no misses and no false alarms (*i.e.*, situations where the analysis indicated that there would be IFR conditions but such conditions did not exist). For the bias, a perfect score means that the frequency of analyzed IFR conditions is equal

to the frequency of observed events, or that there is no over/under-analysis of the event. The false alarm ratio (FARatio), or fraction of analyzed events that did not occur, would be zero (0.0) for a perfect analysis.

Table 4.1: Verification statistics and formulas produced from the dichotomous statistical analysis used in the NCVA product evaluation (Wilks 2006).

Statistics	Formula
PODy (Probability of Detection for Y, Hit Rate, also called Detection Rate)	YY / (YY + NY)
PODn (Probability of Detection for N)	NN / (NN + YN)
Bias	(YY + YN) / (YY + NY)
FARatio (False Alarm Ratio)	YN / (YN +YY)
FAR (False Alarm Rate)	1 - PODn
TSS (True Skill Statistic)	PODy + PODn -1
CSI (Critical Success Index, Threat Score)	YY / (YY + YN + NY)

This contingency table methodology is also utilized for the aforementioned baseline measurement using withheld METAR/NN-A pairings. Similar observations and statistics may also be calculated. The purpose is to provide a baseline measure against which to compare performance statistics and trends of the withheld METAR/NCVA. By providing a baseline measure of statistics, the performance of the NCVA at sites *in between* observations can be defined as having either positive or negative skill.

Multiple stratifications of the data will be investigated. Primarily, dichotomous data will be stratified by confidence level, season, and region. As defined in the NCVA product description, two confidence levels exist in the analysis based on differing factors such as distance from nearest METAR observing site. Further data stratifications include time-of-day and cloud-mask augmentation. By stratifying on time-of-day, one can observe any diurnal trends in performance of the NCVA product. This is further enhanced by stratifying the time-of-day by use of the cloud mask. As will be mentioned later in this section, separate runs of the NCVA without the cloud-mask augmentation will be analyzed for measurement of the performance degradation.

4.2 Continuous Assessment of Ceiling and Visibility

Ceiling and visibility are analyzed as separate attributes from the categorical flight conditions in the NCVA product. While being treated independently, a continuous performance measure is applied such as measuring bias and agreement/disagreement between the NCVA and METAR pairs. A full analysis of these error statistics will be carried out in Section 5.2 of the report. These statistics are calculated based on the equations outlined in **Table 4.2** (Wilks 2006).

Table 4.2: Grid-to-grid correlation and verification statistics (computational form) used for studying correlation of the NCVA product to the weather depiction analysis. These same measurements are used for the grid-to-grid temporal calculation analysis, and in the calculation of continuous measurement statistics (Wilks 2006).

Statistic	**Formula**
Sample Size	N
NCVA Value	N
Observed Value	O
Mean value of N and O	$\overline{\mathrm{N}} = 1/n\sum_{i=1}^{n} N_i \quad ; \quad \overline{\mathrm{O}} = 1/n\sum_{i=1}^{n} O_i$
Second moment quantities	$\overline{\mathrm{NN}} = 1/n\sum_{i=1}^{n}(N_i * N_i) \quad ; \quad \overline{\mathrm{OO}} = 1/n\sum_{i=1}^{n}(O_i * O_i)$ $\overline{\mathrm{NO}} = 1/n\sum_{i=1}^{n}(N_i * O_i)$
Mean Error (Bias)	$\overline{\mathrm{N}} - \overline{\mathrm{O}}$
Root Mean Square Error	$\sqrt{\overline{\mathrm{NN}} - 2 * \overline{\mathrm{NO}} + \overline{\mathrm{OO}}}$
Standard Deviation $_{\mathrm{N,\,O}}$	$\sqrt{\frac{n*(\overline{\mathrm{NN}} - \overline{\mathrm{N}}^2)}{(n-1)}} \quad ; \quad \sqrt{\frac{n*(\overline{\mathrm{OO}} - \overline{\mathrm{O}}^2)}{(n-1)}}$
Covariance	$\frac{n*(\overline{\mathrm{NO}} - \overline{\mathrm{N}} * \overline{\mathrm{O}})}{n-1}$
Correlation	$\frac{\text{Covariance}}{(StdDev_N * StdDev_O)}$

4.3 Cloud Mask

In this portion of the overall analysis of the NCVA product, we will determine the influence of satellite-derived cloud mask data for analysis points that lie *in between* METAR reporting sites. By directly comparing results from the NCVA cross-validation study with results from cross-validation runs that do not use cloud mask data, we are able to determine whether use of the cloud mask degrades performance of the NCVA or appreciably alters confidence levels in the NCVA. This analysis is performed for July 2008 and January 2009 at valid times 0600, 0900, 1200, 1500, and 1800 UTC. These hours represent both complete sunlight and darkness over the CONUS, and additionally represent times when the solar termination region lies across the

CONUS. When comparing daytime use of the cloud mask with nighttime use, the 1200 UTC analyses are not included in the comparison.

4.4 Temporal Correlation

An analysis is performed on the NCVA product to measure the reduction in correlation over time. The change in grid correlation is measured from the top-of-the-hour analysis versus increasing 5-minute off-hour analysis increments (00-05, 00-10, 00-15, ..., 00-65 minute comparisons) in order to determine the critical time period over which significant weather changes are captured by the NCVA product. This temporal correlation study is performed for a summer month (18 August - 18 September 2008) and a winter month (January 2009), using 5-minute issuances of NCVA from 0600 through 0900 UTC and from 1800 through 2100 UTC. The correlation analysis is first performed over the entire NCVA grid, and then also over sub-portions of the domain, utilizing only those analysis grid points that are within a specified radius of a METAR site. The purpose of conducting the correlation study on smaller portions of the grid is to highlight the local change in analyses around METAR sites, using radii of 10-, 20-, and 30-km. The statistics are calculated from the equations shown in **Table 4.2**.

4.5 Weather Depiction Correlation

A direct measure of skill cannot be derived from the hourly Weather Depiction Analysis because we are unable to carry out a series of cross-validation withholding experiments using NCEP's operational code. Therefore, a correlation is computed between the NCVA product and the hourly Weather Depiction Analysis using direct grid-to-grid comparisons to obtain a measure of consistency between the two products. This grid-to-grid correlation is compared to the accuracy of the NCVA product in order to obtain a measure of agreement/disagreement between NCVA skill and NCVA consistency with the Weather Depiction Analysis. This analysis is performed over July 2008 and January 2009. The grid-to-grid statistics and correlations are calculated via the equations in **Table 4.2**.

5. Results

5.1 Dichotomous Analysis of IFR Flight Category Conditions

In seeking to assess NCVA's ability to produce accurate ceiling and visibility values at analysis grid points over the CONUS, we utilize a cross-validation technique to selectively withhold METARs from the analysis procedure. We then estimate how well the analysis identified IFR conditions *at points in between* METAR sites. A dichotomous approach is employed, wherein the full 4x4 contingency table of analyzed vs. observed LIFR, IFR, MVFR, VFR is thresholded (reduced) to a 2x2 contingency table for determining NCVA's ability to correctly analyze IFR conditions. There are numerous performance measures and combinations of measures that can be utilized to typify the success of any analysis or forecast system (Appendix A). The measures referenced within the body of this report will suffice to ascertain value of the NCVA product in the context of pre-flight planning.

5.1.1 Findings from Analysis of IFR Flight Category Conditions

In this section, we see that the overall NCVA performance measures for normal and low flight category confidences combined, closely match those for the low confidence level, since about 80% of the analysis points are assigned low confidence.

In this report, we found overall that the NCVA could add significant value to the planning process compared with the NN-A:

- By more effectively detecting IFR events and reducing risk throughout the airspace (with a higher PODy) and more effectively reducing false alarms of IFR events, resulting in more efficient use of the airspace.

We also determined that the quality of the NCVA and the NN-A differ during the wintertime and summertime:

- NCVA has a significantly higher PODy of IFR events in the wintertime than in the summertime and a slightly lower (better) FARatio in the wintertime than in the summertime.

Regarding differences between normal and low flight category confidence values:

- We found that while the IFR detection rates for normal and low flight category confidences are similar, the FARatio for the normal flight category confidences is significantly lower.

5.1.2 Dichotomous Analysis Results

Figure 5.1 shows a graphical representation of computed performance measures by analysis type (NN-A or NCVA), by season (all, summer, winter), and by region (CONUS or northeast). The NCVA values are further stratified by confidence level: All (A), Low (L), and Normal (N). These results show that the NCVA can add significant value to the flight planning process compared with the NN-A by more effectively detecting IFR events and reducing risk throughout the airspace. The PODy of the NCVA is 0.71 compared to 0.60 for the NN-A. When considering hits relative to events that were analyzed and/or observed, we see that the NCVA again outperforms the NN-A for all seasons, with a CSI of 0.58 for the NCVA and only 0.43 for the NN-A (see **Table 5.1**).

Results from **Figure 5.1** show that the NCVA is also effective at reducing false alarms, which provides for more efficient use of the airspace. The FARatio for NCVA is 0.25, compared to 0.39 for the NN-A, with the lower value being more favorable. When considering the frequency of analyzed IFR conditions to observed conditions (bias), we note that the NCVA bias is just slightly lower than that of the NN-A. One important attribute of both the NCVA and the NN-A is that they maintain nearly ideal biases of 1.0 no matter how we stratify the results.

5.1.3 Results by Season, Region, and Confidence Level

Regarding seasonal differences, we observe in **Figure 5.1** and **Table 5.1** that the value added by the NCVA over the NN-A differs during the wintertime and the summertime. The NCVA has a higher PODy in the wintertime than in the summertime (0.77 vs. 0.60, respectively), and the difference (NCVA minus NN-A) between NCVA's PODy and that of the NN-A is nearly the same in the wintertime (0.11) and summertime (0.13).

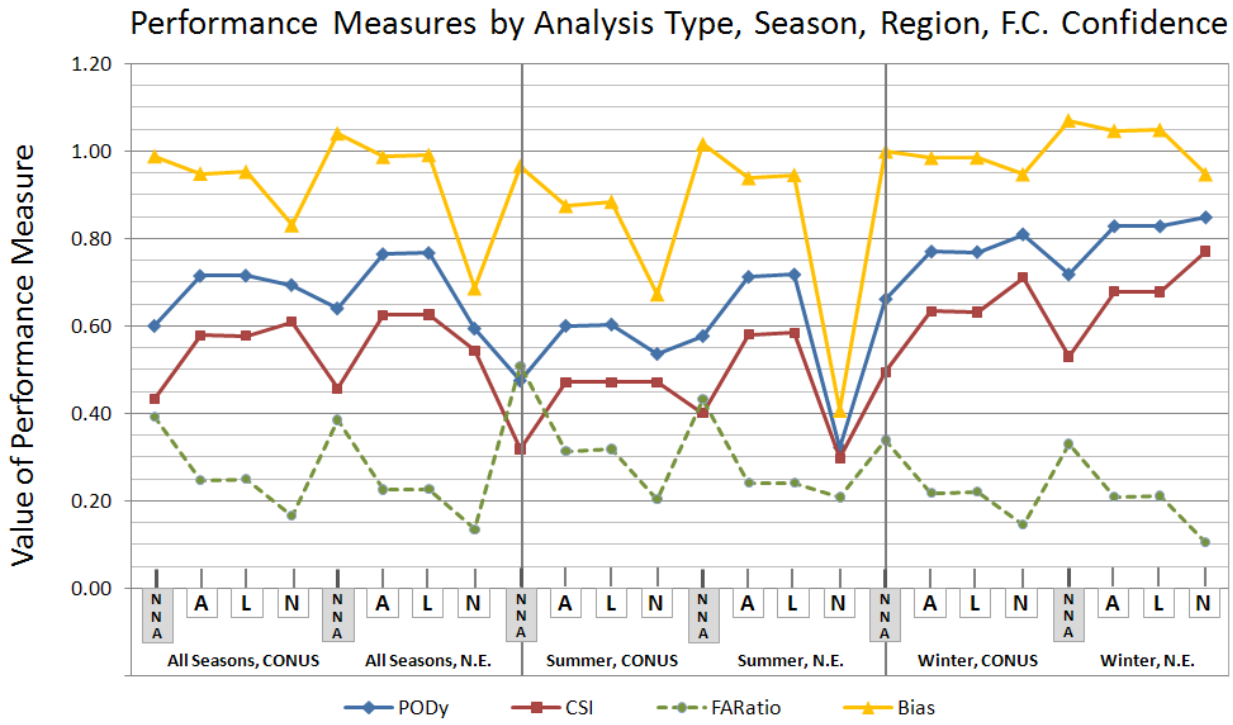

Figure 5.1: Performance measures for the baseline analysis (NN-A), and then for the NCVA at All (A), Low (L), and Normal (N) flight category confidence levels. Results are presented by season and by region (CONUS and northeast only). These results are derived from the radius 10-km "best flight category" matching mechanic.

When considering hits relative to all events that were analyzed and/or observed, we see again that the NCVA has a higher CSI in the winter than in the summer (0.63 vs. 0.47, respectively). For the CSI, the difference between NCVA and NN-A is nearly identical in the winter and summer (0.14 vs. 0.15, respectively). NCVA has a lower FARatio in the wintertime than in the summertime (0.22 vs. 0.31, respectively). The difference (NCVA minus NN-A) between NCVA's FARatio and that of the NN-A is less in the winter (-0.12) than it is in the summer (-0.20), the more favorable situation for NCVA being realized in the summertime. Regarding diurnal effects, we note generally that there is no gross difference in performance measures with time of day.

In examining results across all seasons and regions, the N.E. region's PODy is found to be higher than over the CONUS (0.76 vs. 0.71), while maintaining a FARatio that is nearly identical to the CONUS value (0.23 and 0.25, respectively). Relative to all events that were analyzed and/or observed, the N.E. region's CSI is slightly higher than it is over the CONUS (0.62 vs. 0.58). Consistent with the CONUS results, NCVA in the N.E. region adds significant value to the flight planning process compared with the NN-A, by providing a PODy of 0.76 compared to only 0.64 for the NN-A (0.62 vs. 0.46, respectively, for the CSI), and a FARatio of 0.23 vs. 0.39 for the NN-A.

Table 5.1: Scalar measures summarizing NCVA's ability to detect IFR flight category conditions compared with the baseline (NN-A) analysis. Note that higher detection rates and threat scores are desirable, as are lower false alarm ratios (FARatio). Results from the bias statistic are included in **Figure 5.1** and in the body of the text.

	Detection Rate (PODy)	**Threat Score (CSI)**	**FARatio**
All Seasons	0.71 (NCVA), 0.60 (NN-A)	0.58 (NCVA), 0.43 (NN-A)	0.25 (NCVA), 0.39 (NN-A)
Summer	0.60 (NCVA) NCVA minus NN-A: 0.13	0.47 (NCVA) NCVA minus NN-A: 0.15	0.31 (NCVA) NCVA minus NN-A: -0.20
Winter	0.77 (NCVA) NCVA minus NN-A: 0.11	0.63 (NN-A) NCVA minus NN-A: 0.14	0.22 (NCVA) NCVA minus NN-A: -0.12
CONUS	0.71 (NCVA), 0.60 (NN-A)	0.58 (NCVA), 0.43 (NN-A)	0.25 (NCVA), 0.39 (NN-A)
Northeast	0.76 (NCVA), 0.64 (NN-A)	0.62 (NCVA), 0.46 (NN-A)	0.23 (NCVA), 0.39 (NN-A)
Normal Confidence	0.69 (NCVA)	0.61 (NCVA)	0.17 (NCVA)
Low Confidence	0.71 (NCVA)	0.58 (NCVA)	0.25 (NCVA)

Differences between normal and low flight category confidences reveals that the PODy for normal and low confidence are roughly identical, noting only a small difference in risk, 0.69 and 0.71, respectively. When considering all events that were analyzed and/or observed, the situation is somewhat reversed, as the normal confidence CSI slightly exceeds that of the low confidence (0.61 and 0.58, respectively). The FARatio for the normal flight category confidences are moderately lower than for the low flight category confidences (0.17 vs. 0.25 for low confidence). Wintertime improvement of the NCVA is more pronounced for normal flight category confidences than it is for low flight category confidences. In the summertime, NCVA performance for normal flight category confidences is actually less than for low flight category confidences (see **Figure 5.1**). As noted, this is opposite to the wintertime situation. The high FARatio for NN-A is accentuated in the summertime, compared to the wintertime.

5.1.4 Regional and Sub-regional Results

In this section we present dichotomous performance measures (PODy and FARatio) for 16 climatologically significant regions throughout the CONUS and also for eight specific sub-regions. CONUS averages are computed for the summer and winter months and for the combined months (All). The results show generally better performance in the Midwest, East, and along the Southwest Coast than in the Intermountain West where performance is generally poor (**Figure 5.2 and 5.3**). A full array of regional skill measures is presented in **Tables A.3** and **A.4** of the appendix. Note that in all cases the NCVA PODy is higher than that of the NN-A, and the FARatio of the NCVA is consistently lower than that of the NN-A. For the sub-regional analysis in **Figure 5.4**, note the poor summertime performance from Houston-to-Pensacola, and the favorable performance from New York-to-Boston and from Los Angeles-to-San Francisco.

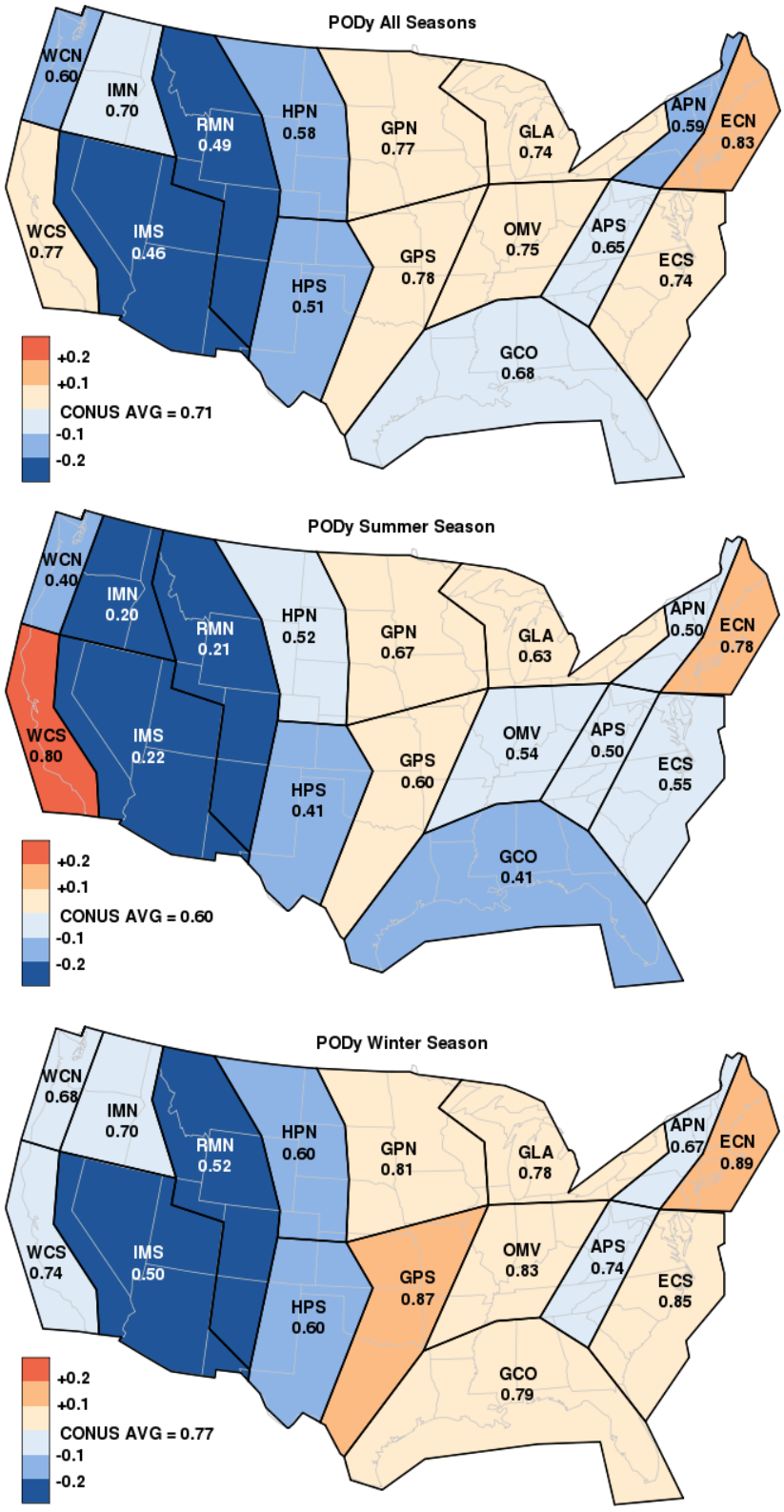

Figure 5.2: Regional performance (PODy) by season (All, Summer, Winter). Colors represent deviation from CONUS averages. Note the relatively high PODy values in the Midwest, East, and Southwest Coast. The Intermountain West reflects a much lower PODy, especially during the summertime.

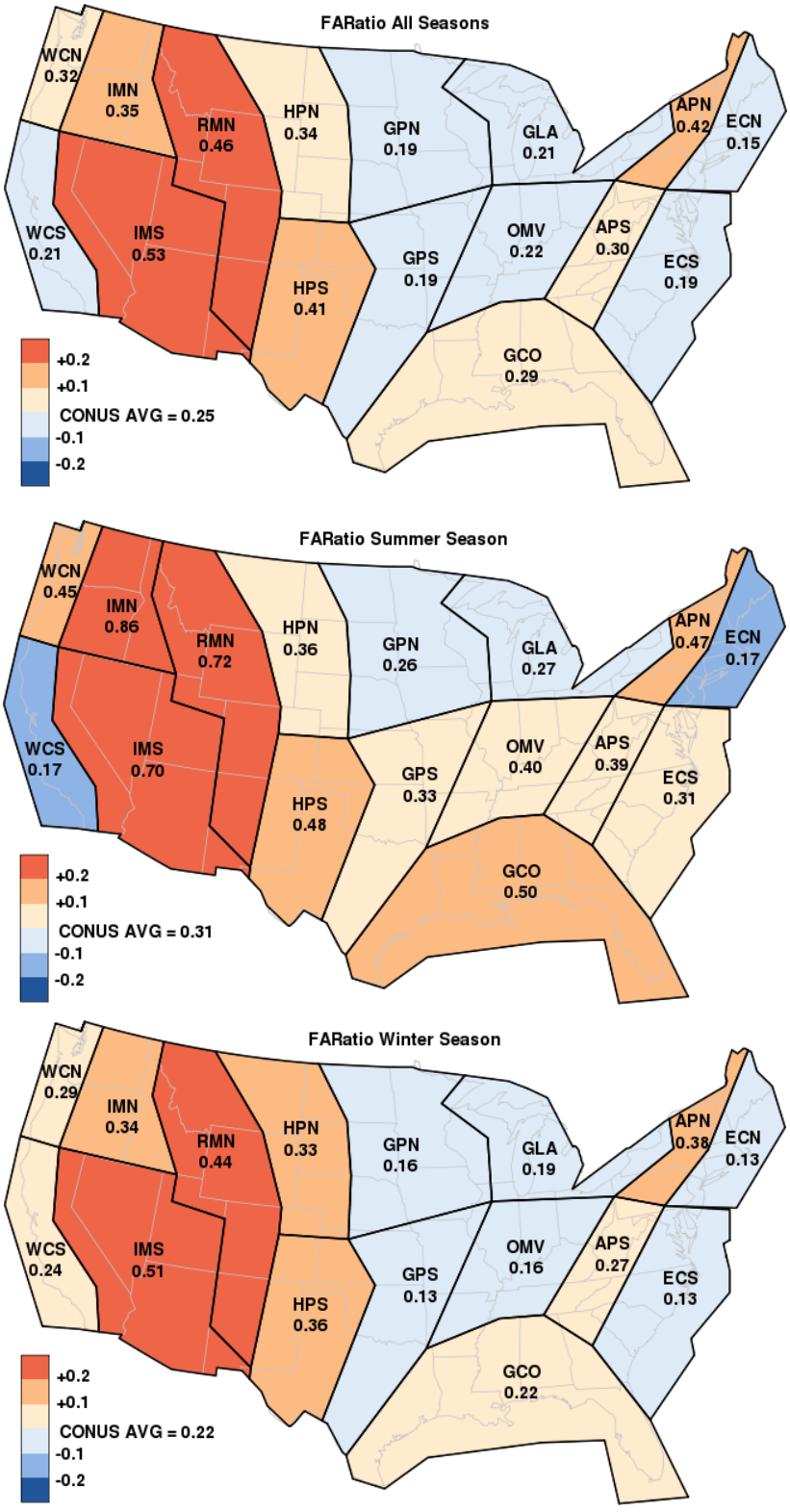

Figure 5.3: Regional performance (FARatio) by season (All, Summer, Winter). Colors represent deviation from CONUS averages. Note the relatively low FARatio values in the Midwest, East, and Southwest Coast. The Intermountain West reflects a much higher FARatio, especially in the summertime.

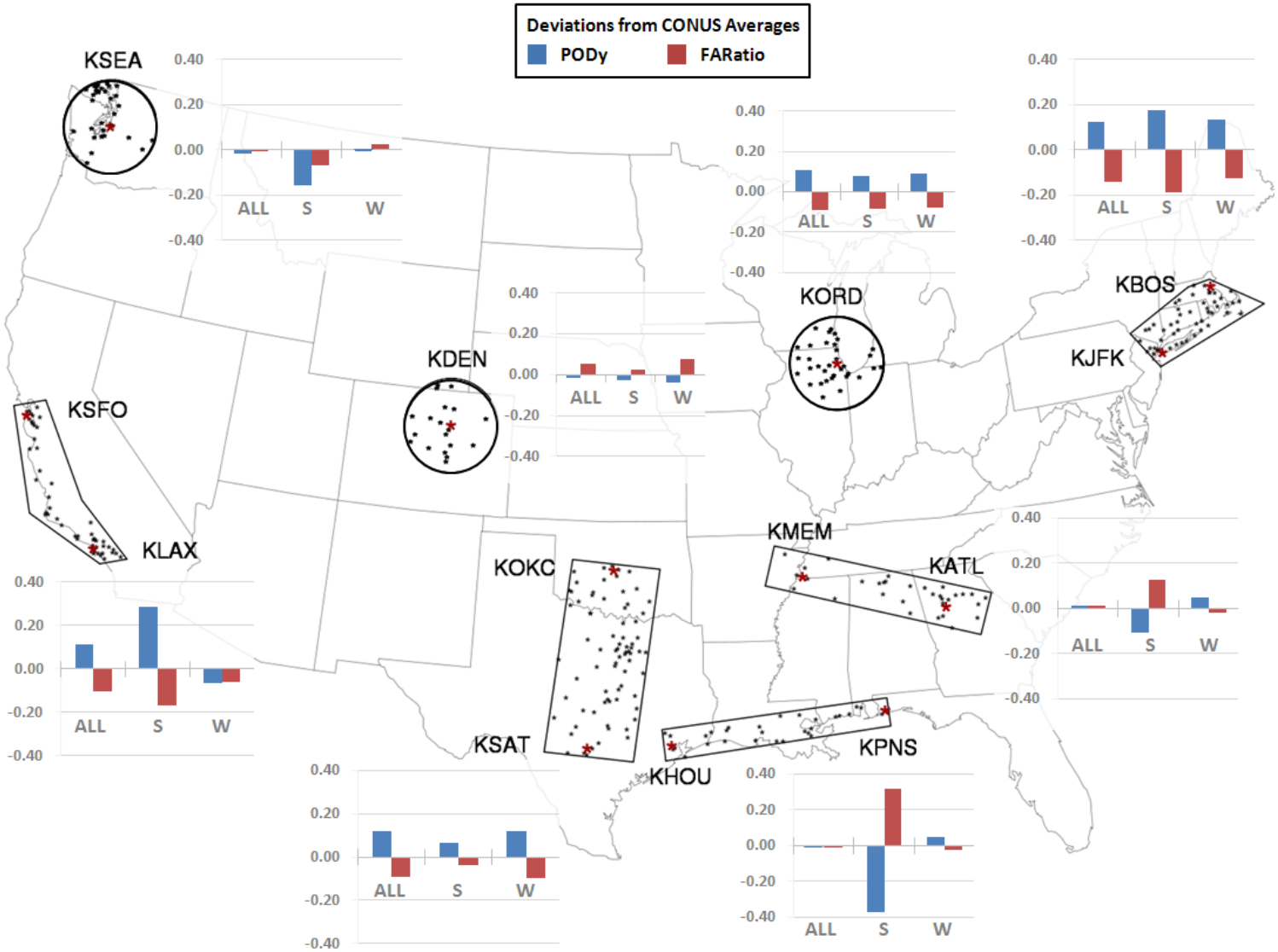

Figure 5.4: Deviations of performance measures from CONUS average for the PODy and FARatio for eight specific sub-regions for summer (S) and winter (W). METAR stations are indicated by dots. Note the favorable performance in sub-regions where the PODy is higher and the FARatio is lower than the CONUS average. This is particularly evident from KJFK–KBOS and from KSFO–KLAX. Sub-regions with less favorable results possess a lower PODy and a higher FARatio than the CONUS average. These less favorable results are noted in the summertime for KHOU–KPNS, KMEM–KATL, and KSEA. Note also that poorer performing sub-regions have a reduced station density.

5.1.5 Sub-regional Results and Weather Regimes

We now investigate the relationship between measured performance of the NCVA and weather that drives characteristics of ceiling and visibility in the various sub-regions. We will show how NCVA performance is a function of the frequency of poor weather, the duration of such conditions, and the time-of-day that unfavorable weather typically develops.

Low ceiling and poor visibility are often the result of fog and stratus which develop in the lowest levels of the atmosphere (FAA 1975). Fog occurs frequently at night, when there is sufficient moisture, clear skies for rapid radiational cooling, and light winds that inhibit turbulent mixing from above. These weather conditions usually clear just a few hours after sunrise. Fog and low-level stratus also develop due to advection of cold air over a warm surface, or the advection of warm air over a cold surface. Such conditions occur frequently in coastal regions. Other weather that leads to low ceiling and poor visibility includes boundary layer saturation from intense or sustained precipitation; upslope flow in a moist environment; and local sea-breeze circulations which are initiated during the daytime (COMET 2004).

For this study, we define an ***IFR event*** to be a period of time during which the ceiling drops below 1,000 feet or the visibility falls below 3 statute miles. Such events are noted by their onset time, cessation time, and overall duration. A climatology of IFR events was assembled by recording raw observations of flight rules conditions from routine and special METAR reports. Transitions to IFR were noted whenever the observed ceiling or visibility dropped below the stated limits at a time when previous conditions had been more favorable than prescribed by these limits. Once the ceiling and visibility rise above these thresholds, we mark the ending time of the IFR event, and compute the duration.

Analysis of results for the onset and duration of IFR across the eight sub-regions reveals many interesting differences. Frequency plots illustrating preferred times of day (local time) for IFR onset are displayed in **Figure 5.5** for selected sub-regions. These climatological results are derived from an extensive database of IFR events that covers the time period from January 2004 – October 2010. Areas on the plot where hexagons are large, represent preferred times of the day for the onset of IFR (x-axis), that last for the duration indicated by the corresponding value along the y-axis. These hexagonal frequency plots show, for example, a propensity for short-duration IFR events to be initiated throughout the night, and around local sunrise. Included with these climatological results are NCVA performance measures from the summer of 2008 and the winter of 2008-09, as discussed previously in this report (see **Table A.3**). Results focus on three different regimes, with distinct diurnal and seasonal signatures: the west coast (KLAX to KSFO), gulf coast (KHOU to KPNS), and northeast sub-region (KJFK to KBOS).

West coast: The summertime initiation of low-level stratus along the west coast (KLAX to KSFO) is a large-scale phenomenon that yields a relatively high PODy (0.89) and CSI (0.77), and a low FARatio (0.14), see **Figure 5.5a**. Coastal waters are known to be quite cool over the eastern Pacific Ocean due to upwelling in the warm season which, along with persistent northwesterly winds, leads to late-day seabreezes and widespread marine stratus development (Leipper 1994). Note from these results, that long-lived IFR events begin in the early evening, persisting for as long as 19 hours. During summer, there is a high frequency of IFR events initiated throughout the nighttime, with new onset ending around 7:00 a.m. local time. The frequency of initiation is actually highest from 2:00 – 5:00 a.m., with these events typically lasting no more than 4 hours in length. There are substantially more IFR events recorded in the

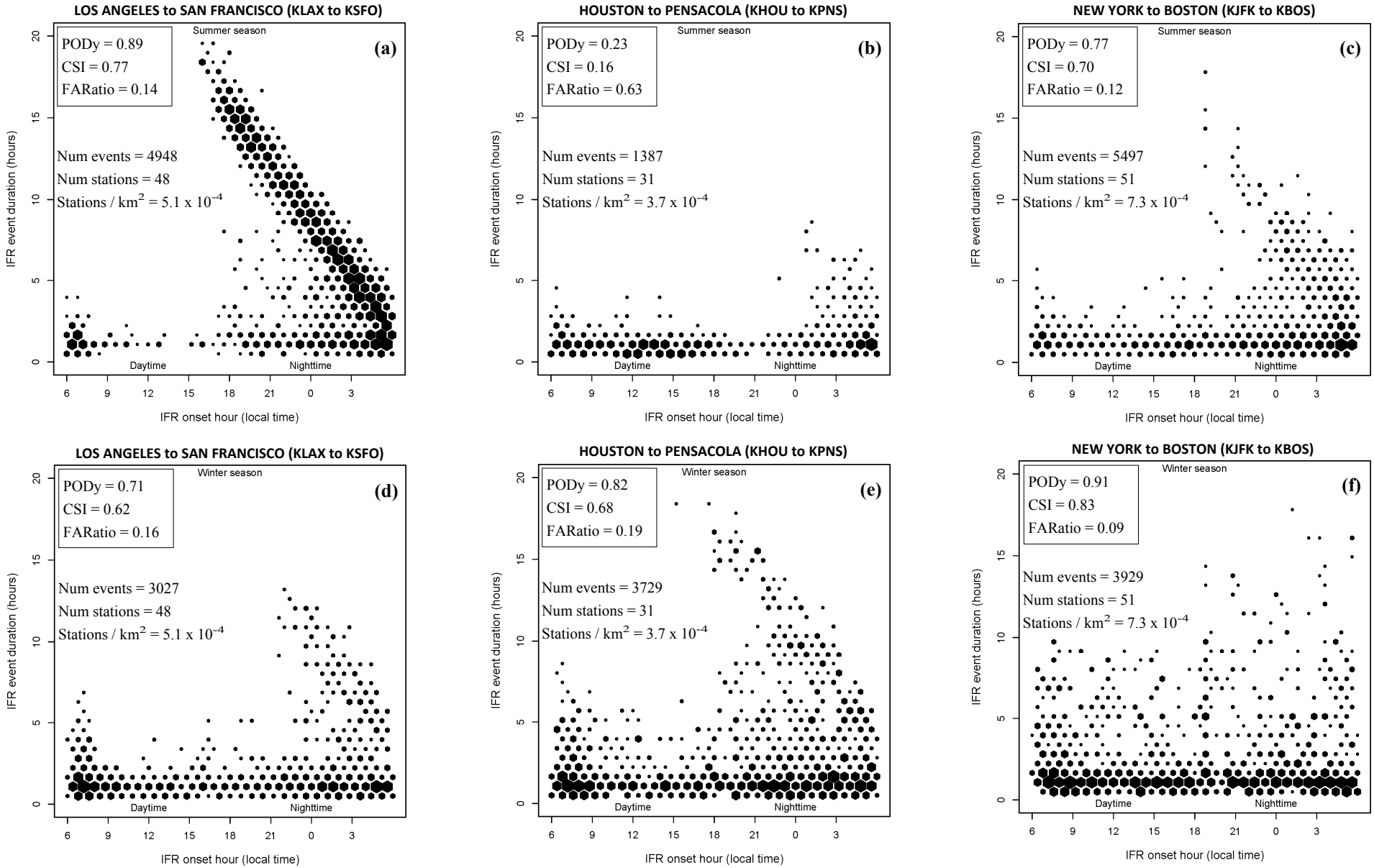

Figure 5.5: Frequency analysis of IFR event duration vs. onset hour, with NCVA performance measures, for the summer (top row) and winter (bottom row). Note from KLAX to KSFO during summer that long-duration IFR events are initiated with great frequency during nighttime. For other regions in both seasons, there are numerous short-duration IFR events initiated throughout the day, with greater frequency usually noted during the nighttime. The diurnal signal is notably weaker during winter, especially from KJFK to KBOS, where IFR is initiated with little regard to time of day. For KHOU to KPNS, events occur frequently during daytime in the summer, and are of longer duration in winter.

summer season (4,948) than in the winter season (3,027). During winter (**Figure 5.5d**) one can see that IFR events occur throughout the day and are of relatively short duration. In this season, the maximum frequency for onset is in the early morning hours, from about 4:00 to 7:00 a.m., with these early-morning events lasting no longer than about 2-3 hours. These short-duration, small-scale wintertime events yield lower performance measures than in the summer, with a PODy of 0.71, CSI of 0.62, and FARatio of 0.16.

Gulf coast: From KHOU to KPNS, IFR is often initiated in the presence of southeasterly winds that develop around a persistent Bermuda High. Warm coastal winds from the southeast collide with relatively cool air over land, which leads to the development of this type of fog in the fall and winter (Vines 2010). In this sub-region, the wintertime PODy of 0.82, CSI of 0.68, and FARatio of 0.19 are significantly better than their summertime counterparts: PODy of 0.23, CSI of 0.16, and FARatio of 0.63 (see **Figure 5.5b,e**). Skill along the Gulf Coast is considerably worse during summer. A contributing factor is the short-lived nature of IFR events that are mostly associated with diurnal processes: fog developing early in the day when the soil is moist and winds (and mixing) are light (Carrin et al. 2008). IFR in this sub-region is a rare occurrence during summer. The climatological analysis reveals nearly three times as many IFR events in the winter (3,729) than in the summer (1,387). This region also has the lowest density of METAR stations (see **Figure 5.5**). The impact this can have on overall performance is discussed in more detail in the next section.

Northeast: Consistent with results we have seen in other sub-regions, performance in the northeast region (KJFK to KBOS) is very favorable in the presence of large synoptically driven precipitation events that impact the area (Tardif et al. 2007). These occur primarily during wintertime, when performance is rather exceptional, possessing a high PODy (0.91) and CSI (0.83), and a very low FARatio (0.09). Results for the summertime are also quite good, with a PODy of 0.77, CSI of 0.70, and FARatio of 0.12 (see **Figure 5.5c,f**). During winter, long-lived IFR events are initiated without regard to time of day. Compared with the gulf coast, IFR events in the northeast summer are of much longer duration, lasting as long as 12-17 hours, which is not observed in the south. Those summertime events in the northeast that are initiated from 7:00 p.m. to 5:00 a.m. are of reasonably long duration. It is also important to recall that NCVA performance measures would be perfect *at* individual METAR sites, and that the northeast does have the highest concentration of stations of these three sub-regions, with nearly twice the station density as in the gulf coast (see **Figure 5.5**).

Overall: Skill is seen to be much higher whenever the frequency of long-lived, large-scale IFR events is greatest, as is the case from KLAX to KSFO in the summer, and from KHOU to KPNS and KJFK to KBOS in the winter. When IFR is the result of more diurnally-driven processes, or these events are of smaller scale, the NCVA performance is shown to be not as favorable. Station density is also an important consideration, with the lowest density measured in the gulf coast region, followed by the western sub-region, and the greatest station density found in the northeast. The summertime extremes along the west coast are impressive from a meteorological standpoint, and a great difficulty from an air traffic management standpoint, with very long-duration IFR events being initiated in the warm season throughout the nighttime, and with those initiated during the early evening hours lasting well into the next day.

5.1.6 Performance vs. Distance for NN-A and Associated NCVA

For the NN-A and associated NCVA (see **Figure 5.6**), notice that as we compute results with increasing distance to the nearest neighbor METAR (out to 100 km), the separation in measured performance between the NCVA and the NN-A remains nearly constant. Note the dramatic separation in performance throughout the distances illustrated. In all measures computed, the full NCVA closely matches the NCVA low confidence performance since over 80% of the overall number of analysis points possess low confidence.

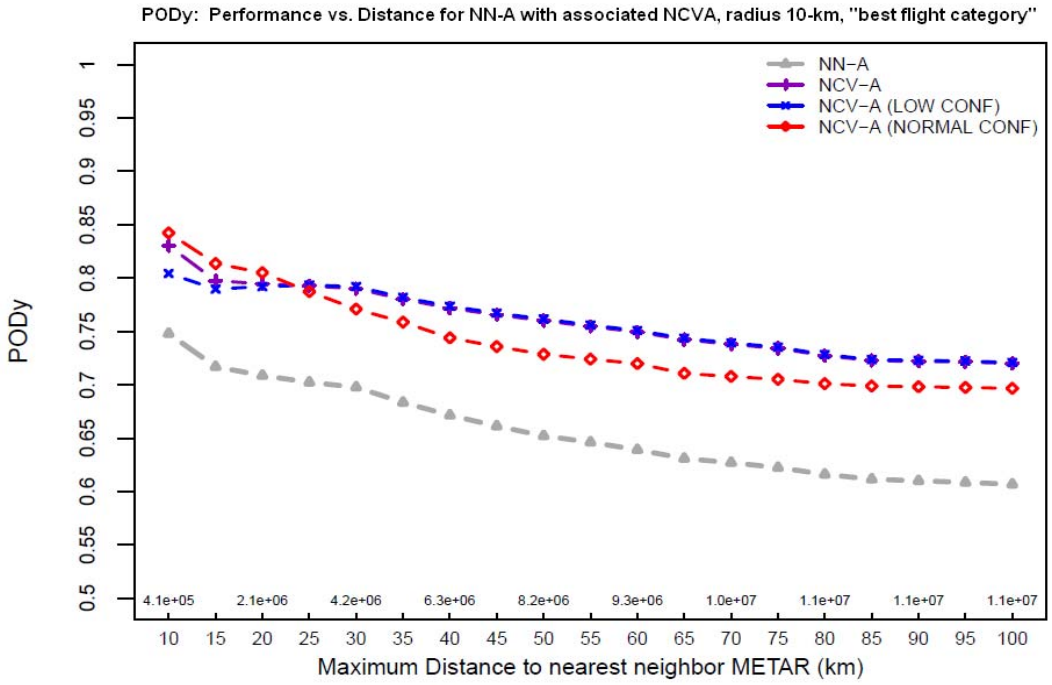

Figure 5.6: Sample plot of performance vs. integrated distance to the nearest neighbor METAR. This result is for the Hit Rate (PODy), with the number of analysis points listed for every other distance measure along the x-axis.

5.1.7 Results between Matching Mechanics

All observations within Section 5.1.1, which are based upon radius 10-km, "best flight category" matching mechanic, are consistent with results derived using the more spatially constrained 4 pt, "best flight category" matching mechanic (see **Table A.2** in Appendix A).

5.2 Ceiling and Visibility

5.2.1 Ceiling Distribution

Continuous measures of skill (see **Table 4.2**) in both the ceiling and visibility attributes of the NCVA product over the CONUS are insensitive to the matching method chosen. As determined in a sensitivity study, the ceiling and visibility analysis will be verified utilizing the 10-km radius and "best flight category" matching method. The results of continuous

measurements are displayed via the root mean square error (RMSE). The mean square error (MSE) is a way of measuring the difference between an estimator and the truth, in this case NCVA and METAR observations.

Because of the physical limitations of automated observation systems, continuous measurements of ceiling are limited to 12,000 ft. and measurements at or above 12,000 ft. are considered to be unlimited. Similarly, visibility measurements are limited to a horizontal distance of 10 SM. Values of error are calculated for measurements below the limitations as these are considered meaningful measurements of ceiling and visibility. This stratification of values is event equalized between NCVA, METARs, and nearest-neighbor METARs, meaning only the existence of all three products at a given valid time were included in the calculation.

Distributions of event equalized ceiling height data, as discussed above, reveal that measurable ceiling data (<12,000 ft.) consists of 23% of available ceiling data. A higher percentage of low ceiling measurements occur during the winter and are attributed with primarily low confidence grids. The percentage of data where either the METAR observations are unlimited and NCVA and NN-A ceilings are not, or the NCVA and NN-A ceilings are unlimited and METAR observations are not, consist of less than 10% of the data. This is less than half the quantity of data within the physical range of measurable ceilings.

5.2.2 Continuous Ceiling Results

Overall, NCVA ceilings have less error (1820 ft.) than baseline NN-A ceilings (1981 ft.). This is a difference in error of 161 ft. The NCVA also had a measurable positive bias (24 ft.) indicating slightly higher analyzed ceilings than are observed by METARs. As indicated by a standard deviation of the distribution of analysis values (2725 ft.), the NCVA has less variation than the NN-A (2729 ft.), indicative of slightly improved consistency and lower risk in ceiling analysis. See **Table 5.2** for overall RMSE values for NCVA and NN-A.

Table 5.2: Root Mean Square Error Values for CONUS continuous cross-validation evaluation for ceiling values in feet for NCVA and NN-A. Results are presented for All Seasons, Summer, and Winter and for All Confidence, low confidence and normal confidence. ΔRMSE is presented as NCVA minus NN-A and is in feet. Values in parentheses indicate the sample size.

	Overall			Summer			Winter		
CONUS	**NCVA/ METAR**	**NN-A/ METAR**	**ΔRMSE**	**NCVA/ METAR**	**NN-A/ METAR**	**ΔRMSE**	**NCVA/ METAR**	**NN-A/ METAR**	**ΔRMSE**
All Confidence	1819.87	1980.72	-160.8 (2,617,587)	2282.21	2439.6	-157.4 (728,596)	1739.69	1902.1	-162.4 (1,888,991)
Low Confidence	1820.62	1981.76	-161.1 (2,482,750)	2284.51	2442.37	-157.9 (681,824)	1740.33	1903	-162.7 (1,800,926)
Normal Confidence	1527.23	1558.83	-31.6 (134,837)	1745.36	1775.33	-30.0 (46,772)	1456.53	1488.83	-32.3 (88,065)

5.2.3 Continuous Ceiling Measurements by Flight Category

NCVA and NN-A ceilings were conditioned on the flight category range of observed METAR ceilings. For instance, the RMSE of all NCVA and NN-A ceiling values were measured where the associated METAR ceiling observations were within the range for a specific flight

category. To see the ceiling flight category specifications, refer to **Table 2.1**. When evaluated based on METAR flight categories, NCVA showed better skill and lower RMSE over the NN-A (see **Table 5.3**). The difference in RMSE (NCVA minus NN-A) for these conditions is greater than 230 ft. in IFR conditions. An increased bias measurement is observed for NCVA when conditioned on LIFR. Measures are better than those of the NN-A when conditions were IFR and LIFR. When conditioned on VFR, the bias measurements are still reduced with lower variance in analyzed ceilings, but a negative bias (-503 ft.) indicates that the NCVA analyzes lower ceilings than observed by METARs.

Low confidence stratifications in this conditioning show similar results to the overall RMSE values. Where grids have normal confidence there is a decrease in RMSE differences between NCVA and NN-A. Additionally, the RMSE for the NCVA is lower at normal confidence for flight category conditioning than low confidence inferring an improvement in ceiling analysis in the NCVA product.

Table 5.3: Root Mean Square Error Values for CONUS continuous cross-validation evaluation for ceiling values in feet for NCVA and NN-A conditioned on observed METAR flight category for all seasons and confidence levels. Δ RMSE is presented as NCVA minus NN-A and is in feet. N is the sample size.

	N	NCVA/METAR	NN-A/METAR	Δ RMSE
LIFR Condition	264,035	1471.60	1725.77	-254.20
IFR Condition	299,389	1373.79	1612.46	-238.70
MVFR Condition	906,351	1486.12	1625.03	-138.90
VFR Condition	1,147,812	2220.01	2367.56	-147.54

5.2.4 Continuous Ceiling Measurements by Confidence, Season, and Region

There is added value in the ceiling confidence fields (**Table 5.2**) as shown by less error (lower RMSE and Bias) in the NCVA normal confidence grids for ceiling analyses (1527 ft.). Error in the NCVA at both confidence levels is less than that of the NN-A. However, the difference in error between NCVA and NN-A is less in normal confidence grids (32 ft.) than low confidence grids (162 ft.).

Seasonally, the NCVA has less error during the winter (1740 ft.) than during the summer (2282 ft.). Both seasons maintain consistent improvement over the NN-A (162 ft. and 157 ft.). The NN-A ceilings also have a reduction in RMSE from summer (2440 ft.) to winter (1902 ft.). Seasonal stratifications change little by confidence with low confidence grids matching the RMSE of the aggregate while the error for normal confidence grids is reduced rather significantly. Also, the difference in RMSE between NCVA and NN-A for the normal confidence grids is drastically reduced.

Regionally, the CONUS region versus the NDFD region has the same amount of error for both the NCVA ceiling and the NN-A ceiling. It was found in this analysis that a greater difference exists between the NCVA over the NN-A ceilings in the N.E. region (~200 ft.). The N.E. region has a greater RMSE than the CONUS region and NDFD region, which can be attributed to less variability in ceiling measurements over this region. Low confidence stratification per selected domain changes little, but the RMSE for normal confidence grids are reduced.

For a more complete table of results reflective of this summary, please see the tables in Appendix B.

5.2.5 Visibility Distribution

The distribution of visibility observations and analyses were measured according to the physical limitations of the ASOS. On average, 13% of all visibility measurements are within the 10 SM limitation while greater than 70% of the time in this study, visibility measurements are all unlimited. Analysis visibility of less than 10 SM and unlimited METAR visibility observations and vice versa each contain on average 7-8% of the study data.

Due to less available measurable data and a smaller scale and variability of observed values, error in visibility appears small in the aggregate. For this same reason, the difference in error between the NCVA and the NN-A are small. In this continuous analysis of visibility, the measured error, however, is enough to influence flight category determinations by at least 1 category.

5.2.6 Continuous Visibility Results

Overall, the NCVA analyzed visibility has lower RMSE (2.41 SM) than the NN-A visibility (2.57 SM). This is a difference in RMSE of 0.16 SM (**Table 5.4**). The NCVA has a positive bias (0.09 SM.) indicating slightly higher analyzed visibilities than reported by METAR observations. The NCVA and the NN-A have the same standard deviation and variance (2.66 SM). The scale of visibility is less variable and orders of magnitude smaller than that of ceiling observations, thus, smaller differences are expected, but can still be significant.

Table 5.4: Root Mean Square Error Values for CONUS continuous cross-validation evaluation for visibility values in SM for NCVA and NN-A. Results are presented for All Seasons, Summer, and Winter and for All Confidence, low confidence and normal confidence. Δ RMSE is presented as NCVA minus NN-A and is in SM. Values in parentheses indicate the sample size.

	Overall			Summer			Winter		
CONUS	**NCVA/ METAR**	**NN-A/ METAR**	**ΔRMSE**	**NCVA/ METAR**	**NN-A/ METAR**	**ΔRMSE**	**NCVA/ METAR**	**NN-A/ METAR**	**ΔRMSE**
All Confidence	2.41	2.57	-0.16 (1,529,612)	2.45	2.63	-0.18 (669,000)	2.38	2.51	-0.13 (860,612)
Low Confidence	2.41	2.57	-0.16 (1,418,163)	2.45	2.64	-0.19 (595,396)	2.38	2.51	-0.13 (822,767)
Normal Confidence	1.88	1.96	-0.08 (111,449)	1.83	1.9	-0.07 (73,604)	2.1	2.18	-0.08 (37,845)

5.2.7 Continuous Visibility Measurements by Flight Category

When evaluated based on METAR flight category conditions, NCVA shows less error over the NN-A when flight category conditions are LIFR and IFR (**Table 5.5**). The difference in RMSE in these conditions is 0.24 and 0.20 respectively. Additionally, the NCVA has a lower bias and variance showing much better consistency and less risk when conditions are IFR. A negative bias in the NCVA when METAR conditions are VFR (-1.17) reflects a lower visibility analysis value in the NCVA than observed METAR visibilities, but the NCVA variance (4.51 SM^2) is significantly lower than that of the NN-A (4.87 SM^2) supporting the lower risk analysis of the NCVA.

Table 5.5: Root Mean Square Error Values for CONUS continuous cross-validation evaluation for visibility values in SM for NCVA and NN-A conditioned on observed METAR flight category for all seasons and confidence levels. ΔRMSE is presented as NCVA minus NN-A and is in SM. N is the sample size.

	N	NCVA/METAR	NN-A/METAR	Δ RMSE
LIFR Condition	158,857	2.54	2.78	-0.24
IFR Condition	299,631	2.20	2.41	-0.20
MVFR Condition	493,625	2.18	2.28	-0.10
VFR Condition	577,499	2.54	2.68	-0.14

5.2.8 Continuous Visibility Measurements by Confidence, Season, and Region

There is added value in the confidence field as shown by lower RMSE in the NCVA normal confidence grids for visibility (1.88 SM; **Table 5.4**). Error in the NCVA at both visibility confidence levels is less than that in the NN-A. The difference in error between the NCVA and the NN-A is less in normal confidence grids (0.08 SM) than low confidence grids (0.16 SM).

Seasonally, the NCVA has less error during the winter (2.38 SM) than during the summer (2.45 SM). Both seasons maintain consistent improvement over the NN-A visibilities (0.13 SM and 0.18 SM respectively; **Table 5.4**). The cloud mask, utilized more effectively during periods of extended daylight, might be responsible for the reduced error in visibility for the summer.

Regionally, the CONUS region and NDFD region have similar RMSE for both the NCVA visibility and the NN-A visibility. Overall error is slightly reduced in the NCVA when masked for the CONUS region. When analyzed on the N.E. region, there is a greater difference between the NCVA and the NN-A visibilities (0.21 SM). When confined to the N.E. region, the RMSE is increased; however, this is due to the lower variability in visibility values. This follows the pattern displayed in the ceiling analysis. A similar pattern is also observed within the NN-A visibility.

For a more complete table of results reflective of this summary, please see the tables in Appendix C.

5.3 Cloud Mask / No-Cloud Mask Analysis Comparison

In this section of the report, we seek to determine the effects of the cloud mask on skill and confidence of the NCVA. Results are presented for a set of cloud mask/no cloud mask analysis runs that were carried out for a summer month and a winter month (July 2008 and January 2009). The night/day results presented in this section include the 0600, 0900 UTC analyses as part of the nighttime, and the 1500, 1800 UTC analyses as part of the daytime. The 1200 UTC analyses are entirely omitted from the daytime/nighttime comparison. The cloud mask can alter ceiling or visibility values at a location where satellite data strongly indicates that the skies are clear while a distant yet influential METAR has reported otherwise. At NCVA grid points, the cloud mask has only a clearing effect on ceiling and visibility. Overall, this evaluation shows that the cloud mask does not harm the NCVA with regard its ability to identify IFR flight conditions.

In the aggregate for all flight category confidence levels and for the situation referencing IFR/non-IFR flight category classifications, it can be shown from the viewpoint of a statistical contingency table (**Figure 5.7**), that the migration of NCVA analysis points from hits to misses and from false alarms to correct negatives are the only changes possible with application of the cloud mask. This results in an ***unfavorable*** decrease in the hit rate (PODy), and a ***favorable*** decrease in the false alarm rate (FAR). Since the TSS measures the difference between the analysis hit rate and false alarm rate (TSS = PODy - FAR), if this score changes little with application of the cloud mask (or actually increases), then we can conclude that use of the cloud mask does not adversely affect the overall performance of the NCVA. Raw numbers of analysis points, in addition to percentages, are shown to highlight the actual migration of analysis points ***to*** the "miss" cell of the contingency table versus those that migrate ***from*** the "false alarm" cell.

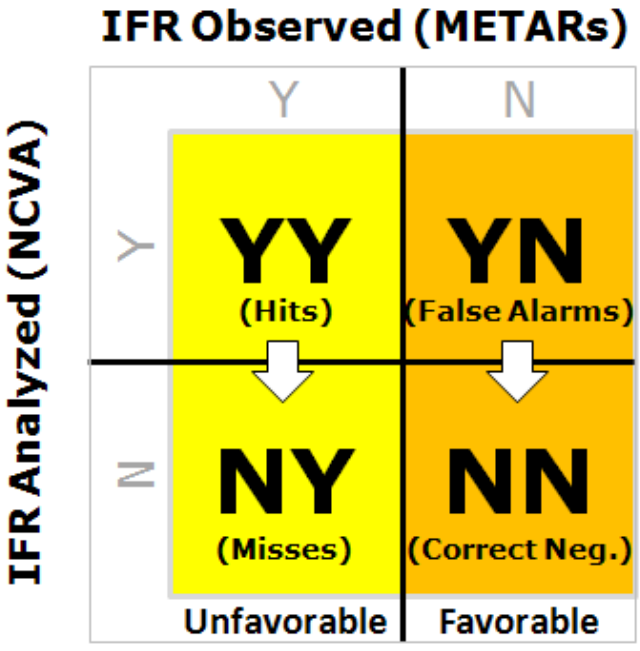

Figure 5.7: The cloud mask has the potential for altering analysis grid point flight category classifications from IFR to non-IFR, but *not* from non-IFR to IFR. Thus in the aggregate for all flight category confidence values, application of the cloud mask can move NCVA values from the top row to the bottom row, which leads to an unfavorable increase in misses, but a favorable decrease in the number of false alarms.

With application of the cloud mask, we measure a small decrease in hits with an associated increase in missed events, along with a more appreciable decrease in false alarms with associated increase in correct negatives, thus yielding a slight increase in analysis risk (decrease in hit rate), and a more significant increase in analysis efficiency (measurable decrease in false alarms) over the air space (**Table 5.6**). Note, however, even with this slight increase in risk, the overall NCVA risk measurement is still significantly lower than that of the baseline (better PODy for the NCVA represents lower risk). Utilization of the cloud mask for both the summer and winter months yields an increase of only 137 (0.52%) misses migrating from the hit cell of the contingency table, and a much larger decrease in the number of false alarms (1356, a 5.87% reduction).

One effect of utilizing the cross-validation technique, and NCVA's response to this technique, is a large re-assignment of NCVA analysis flight category confidence levels from low confidence to normal confidence with application of the cloud mask (see **Table 5.6**). For the summer and winter months combined, 9.26% (74,504) of the low flight category confidence analysis points were re-assigned a normal flight category confidence value with utilization of the cloud mask.

5.3.1 Seasonal Cloud Mask Effects

The percent decrease in hits and associated increase in misses with application of the cloud mask is nearly the same in the summer as it is in the winter. In summer there is a 0.53% increase in misses, while in the winter the increase is 0.51%, manifest in each season by a decreased hit rate (**Table 5.6**). The decrease in false alarms with associated increase in correct negatives, as a percentage of the total with application of the cloud mask, is more pronounced in summer than in winter. In summer there is a 7.35% decrease in the number of false alarms, while in winter the decrease is 5.0%, manifest in each season by a decrease in the false alarm rate (FAR). Regarding risk and efficiency of the airspace, these percentage differences reflect a small increase in missed IFR events (63 for summer and 74 for winter, total=137), but a significantly larger number of reduced false alarms (635 for summer and 721 for winter, total=1356). Thus, during winter, there is an appreciable decrease in the number of false alarms issued compared to the decrease measured during the summer, while the increase in misses is nearly identical.

The net re-assignment of flight category confidence from low confidence to normal confidence is greater as a percentage during the summertime than during the wintertime, with 12.3% of the low confidence analysis points migrating to normal confidence in the summer, and only 6.4% in the winter. This represents 48,440 analysis values in the summer and 26,064 in the winter (**Table 5.6**).

5.3.2 Diurnal Cloud Mask Effects

The percent increase in misses is slightly greater at night than during the day (**Table 5.6**), with a 0.72% (81) increase in misses at night compared to a 0.33% (28) increase in misses during the day. The percent decrease in false alarms is nearly the same during the day and night, with a 5.52% (524) decrease at night and a 6.52% (503) decrease during the day. It is interesting to note the strong similarity in results for this particular daytime/nighttime comparison.

The net re-assignment of flight category confidence from low confidence to normal confidence with application of the cloud mask is reversed at night, with 0.054% (180) of the nighttime low confidence analysis points having been re-assigned ***from*** the normal confidence classification at night, and 24.3% (72,659) of the daytime low confidence analysis points having

been re-assigned ***to*** the normal confidence classification during the day. These results, which concentrate on the net re-assignment of NCVA values between confidence levels, are not as strongly correlated as the results above which concentrated on the increase in misses and decrease in false alarms with application of the cloud mask. At night, the cloud mask is only slightly more active in converting analysis values from IFR classification, as noted above, but the cloud mask is much less active at night in altering confidence values of the analysis field (**Table 5.6**).

Table 5.6: Change in misses and false alarms for the statistical contingency table as drawn in **Figure 5.7**, along with the net decrease in low flight category confidence points with re-assignment to normal flight category confidence through utilization of the cloud mask.

Time Period of the Study	**Increase in Misses with Application of the Cloud Mask**	**Decrease in False Alarms with Application of the Cloud Mask**	**Net Decrease in *Low* Flight Category Confidence Values (re-assignment to *Normal* Flight Category Confidence)**
Summer and Winter	137 (0.52%)	1,356 (5.87%)	74,504 (9.26%)
Summer	63 (0.53%)	635 (7.35%)	48,440 (12.3%)
Winter	74 (0.51%)	721 (5.0%)	26,064 (6.4%)
Daytime (excludes 1200 UTC)	28 (0.33%)	503 (6.52%)	72,659 (24.3%)
Nighttime (excludes 1200 UTC)	81 (0.72%)	524 (5.52%)	-180 (-0.054%)

5.3.3 Summary of Cloud Mask Results

When analyzing cloud mask/no-cloud mask results, we have determined that the decrease in the number of false alarms is generally much larger than the increase in misses; however, even with a slight increase in risk, the overall NCVA risk measurement is still significantly lower than that of the baseline (better PODy for the NCVA represents lower risk). The statistical net effect of these competing changes to the NCVA is a very slight change in the TSS. The large decrease in false alarms is somewhat overwhelmed by the magnitude of correct negatives, thus yielding only a modest decrease in the FAR, even though there is an observably larger decrease in false alarms than increase in misses. The change in TSS for all seasons combined, for individual seasons, and by time of day is on the order of $1x10^{-3}$, and so it is concluded that utilization of the cloud mask does not harm the NCVA with regard its ability to identify IFR flight conditions.

The re-assignment of analysis values from low to normal confidence via the cloud mask obviously does not affect skill in forecasting detrimental fight conditions, but as a diagnostic, it appears to behave as expected: with a much larger net re-assignment of confidence values during the day than at night, and a greater influence noted during the summertime than during

the wintertime. We expect the impact to be greater during the day because the satellite can more readily detect clear skies when the sun angle is higher, as noted in the NCVA description documentation (Herzegh et al. 2009). The summertime increase in confidence re-assignment by the NCVA is most likely due to the increased length of day (yielding high sun-angles and greater satellite utilization throughout the day), a greater horizontal variability of cloud cover during the summertime, and resultant impacts of increased relative humidity on calculations for the radius of influence of a METAR by the NCVA during the summertime.

5.4 Temporal Correlation of the NCVA Product

In this section of the paper, we compare analyses that are issued in successive 5-minute intervals from the top of one hour through to the top of the next hour. Our aim is to measure the correlation of a fixed reference analysis (from the top of the hour), against those that are issued every 5 minutes thereafter (00-05, 00-10, ..., 00-65 minutes). In performing this study for a summertime month in 2008 and a wintertime month in 2009, we find that:

- The linear correlation between an NCVA issuance and its successor one hour later is ~0.8, representing a significant change in flight conditions over the CONUS.
- The five-minute updates that occur between hourly issuances appear to effectively capture the incremental changes over the CONUS, as indicated by linear changes in correlation between the initial and intermediate issuances.

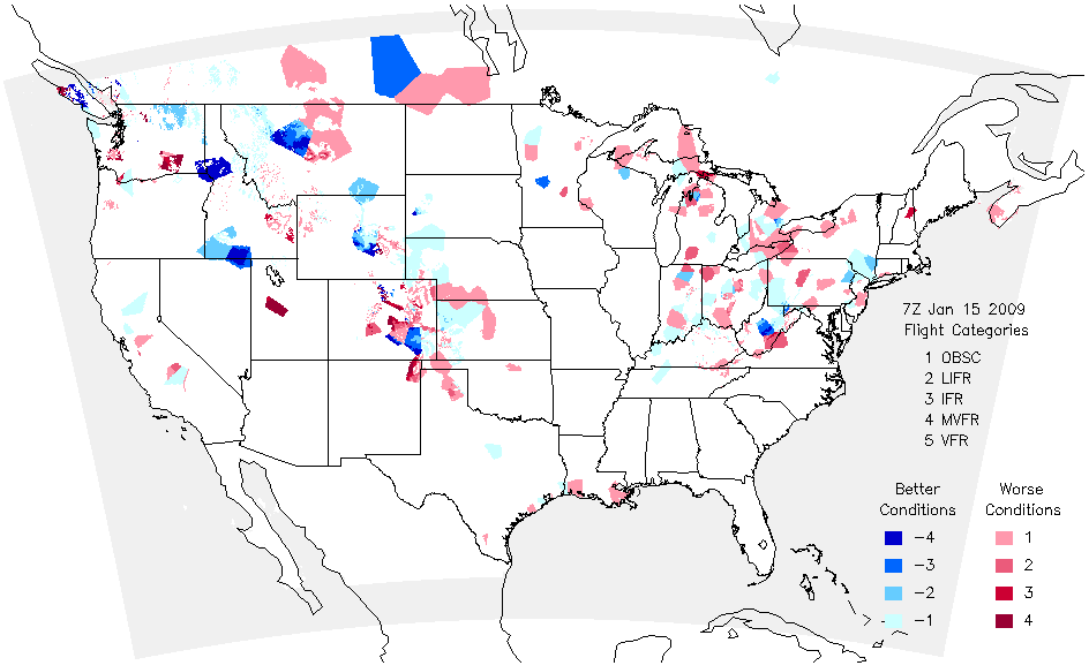

Figure 5.8: Difference field illustrating change in flight categories from 0700 to 0800 UTC 15 Jan 2009. Shades of red reveal worsening conditions while shades of blue reveal improving conditions. The correlation between the 0700 and 0800 UTC analysis grids was measured to be 0.85.

An example of correlation decay over the computational NDFD grid for the NCVA is illustrated in **Figure 5.8**. This difference field illustrates the change in flight categories from 0700 to 0800 UTC on 15 January 2009. Note, for this particular example, that the correlation between the two grids is measured to be 0.85.

5.4.1 Findings from the Temporal Correlation Study

We will now consider specific results illustrated in **Figure 5.9**. When comparing adjacent-in-time analyses, there is a steady reduction in correlation throughout the time period of an hour down to the more abrupt change in correlation at 60 minutes, which is the top of the next hour (**Figure 5.9a-c**). The decrease in correlation measured between 55 minutes past the hour (0.85) and 60 minutes past the hour (0.80) represents change in flight category conditions not captured by the intermediate 5-minute updates. The frequent updates provide a consistent and steady change with infusion of new METAR reports throughout the course of the hour.

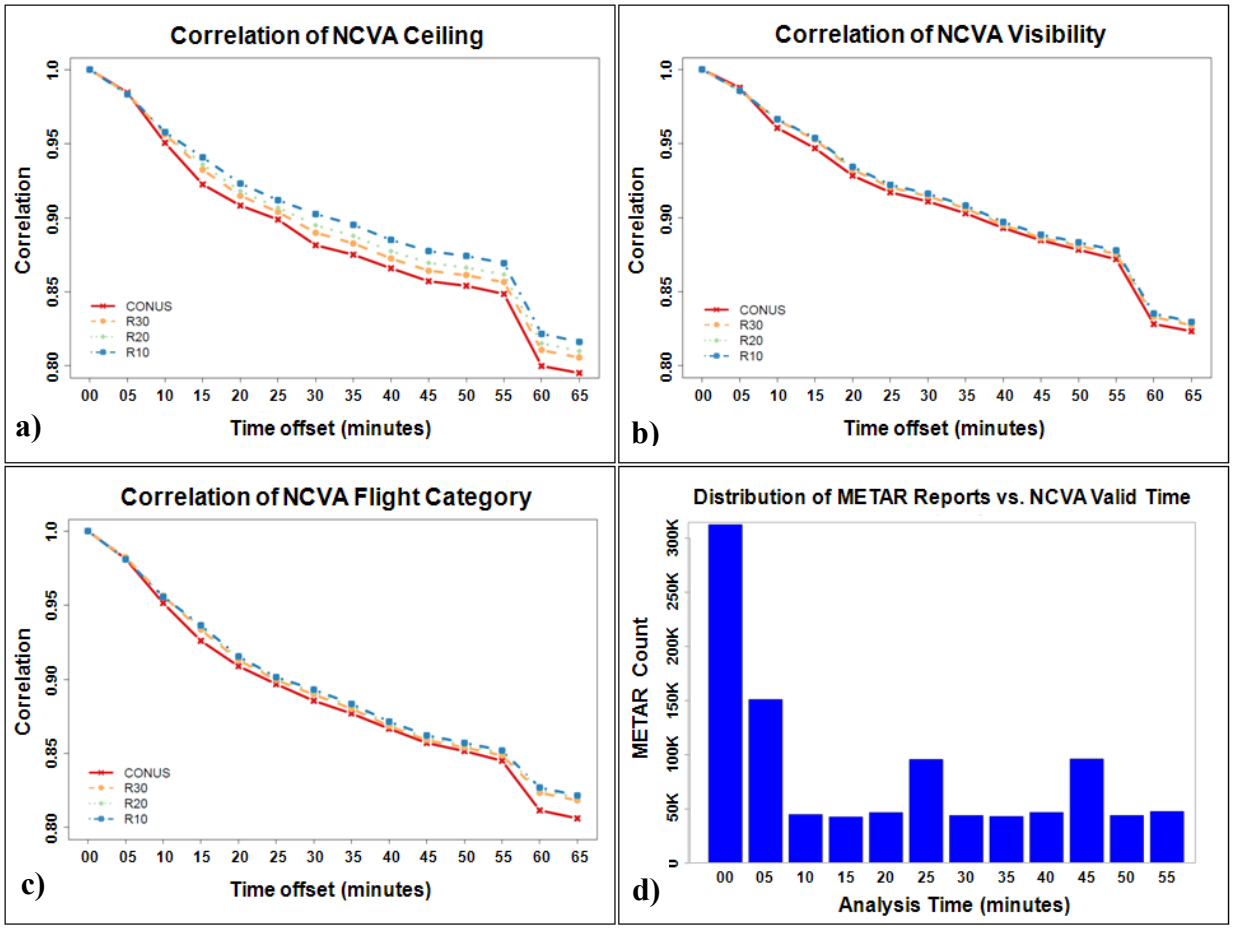

Figure 5.9: Decrease in correlation over the course of one hour (and just slightly beyond) from reference analyses taken from the top of the hour during July 2008 and January 2009. Note that the inclusion of new METAR data every five minutes leads to a significant and steady decline in correlation over the course of one hour, with a larger drop in correlation noted near the time of the hourly infusion of a sizeable number of new METAR reports (60 minutes past the top of the hour).

The drop in correlation from the zero-minute analysis is greater during the first half hour (steeper decline) than during the last half hour (slower decline) for each of the three fields shown: ceiling, visibility, and flight category. As shown by the METAR analysis bar chart (**Figure 5.9d**), the steeper decline (greater change) in correlation during the first half hour is primarily due to the receipt of a large number of METAR reports during that time period. The minimal decrease in correlation during the final 30 minutes corresponds to a time period when fewer updated METAR reports are ingested into the analysis system. Operationally, numerous METARs are issued near the time of 53 minutes past the hour, but are delayed due to communications and decoding issues, and do not arrive for analysis processing until nearer the top-of-the hour.

There is little difference between full-CONUS results and results derived using only grid points that are located within close proximity of a METAR site (from 10-, 20-, and 30-km). Regarding the three fields of measurement: ceiling, visibility, and flight category, all exhibit similar reductions in correlation over time, with some slight variation in results by proximity to a METAR (10-, 20-, 30-km), evident from the line separation shown in the ceiling results.

5.4.2 Analysis of Results from the Temporal Correlation Study

The steady reduction in correlation between analysis times indicates that the 5-minute cycle is achieving its goal of providing suitable updates throughout the course of an hour. It also demonstrates that the analysis system is stable, and not prone to dramatic changes in results between fixed time increments. The decrease in correlation is steeper during the first 30 minutes of the hour as a result of the larger number of new METAR reports that are utilized during that time period than in the latter half hour. It is not surprising that the largest decrease in correlation is associated with the largest integration of new data at the top of the hour, as shown by the large drop in correlation from 55 to 60 minutes past the reference zero-minute analysis (correlation of 0.85 to 0.80).

Results were computed over the entire CONUS as well as for those analysis grid points that are within 10-, 20-, and 30-km of a METAR station. There is little difference in results for the visibility and flight category calculations of correlation differences between adjacent analysis times, but a greater separation in results between correlations of ceiling. This is representative of the physical nature of frequent changes in the overhead cloud cover, compared to visibility. The categorical nature of flight category acts to dampen fluctuations in this field, since particular thresholds of ceiling and visibility must be crossed to record a change in category. For ceiling, the further out in time from the reference analysis, out to 55 minutes, the greater the separation in correlation results by proximity to the analysis points. As noted in the data section of this report, the reported sky cover is representative of a broader horizontal area than is the reported visibility value.

5.5 Correlation to Manual Analysis Product

The primary question with regard to correlation of NCVA and the operational hourly Weather Depiction Analysis is to examine if the automated product performs at least as well as manually generated products, and that it is consistent with those products (FAA/ATO 2007). During the course of this analysis, it was found that:

- NCVA and the Weather Depiction Analysis have an overall correlation greater than 0.6, indicating that these disparate products may perform in a consistent manner.
- With no suitable way to directly measure the quality of the Weather Depiction Analysis, the study finds indications that NCVA performs at least as well as the Weather Depiction Analysis.

The correlation between the gridded NCVA product and the operational hourly Weather Depiction Analysis was calculated for a period of 2 months (July 2008 and January 2009) through a grid-to-grid analysis. The overall conclusion derived from the correlation of weather depiction analyses and NCVA is that the automated NCVA product performs at least as well as and is consistent with the Weather Depiction Analysis. The skill of the NCVA, using CSI, is plotted against its correlation with the Weather Depiction Analysis, color coded by season and distributions calculated. The numbers in the center of each quadrant reveals the relative distribution of points. Distribution results show that greater than 90% of compared results have a significant positive correlation (>0.5). Only 45% of those comparisons show lower skill (**Figure 5.10**). A majority of the comparisons where NCVA showed better skill (CSI > 0.5) occur during the month of January, a month shown to have a greater frequency of restricted flight category events. Further analysis revealed that on average, the NCVA grid tended to analyze a greater spatial quantity of events MVFR or worse (**Figure 5.11**).

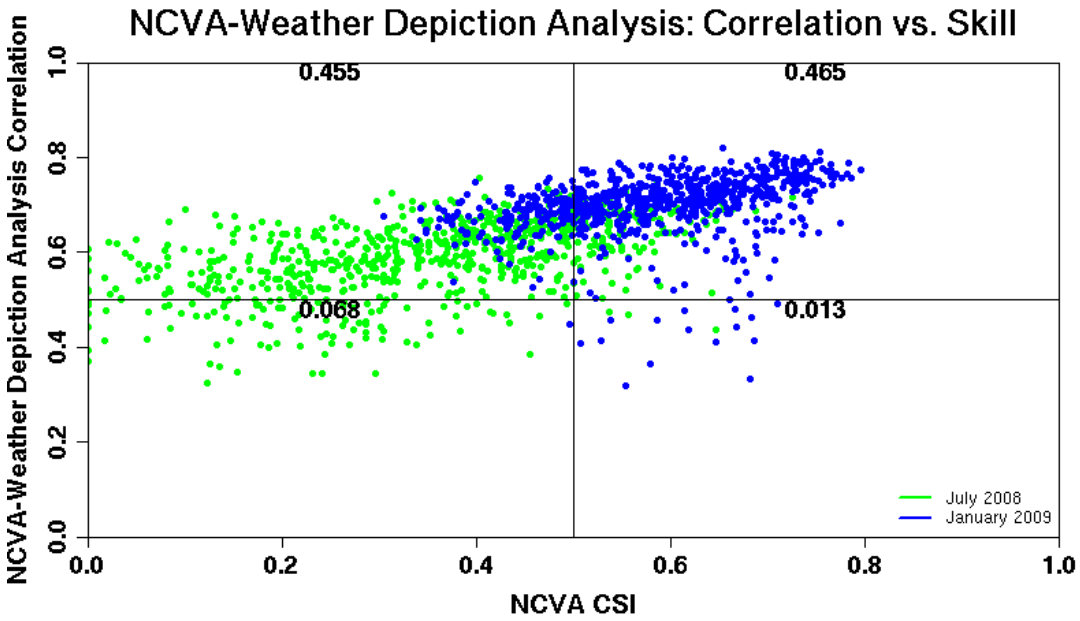

Figure 5.10: Scatter plot of NCVA-Weather Depiction Analysis correlation vs. NCVA skill, color coded by season (January 2009 in blue and July 2008 in green) with quadrant ratio in the upper center of each quadrant.

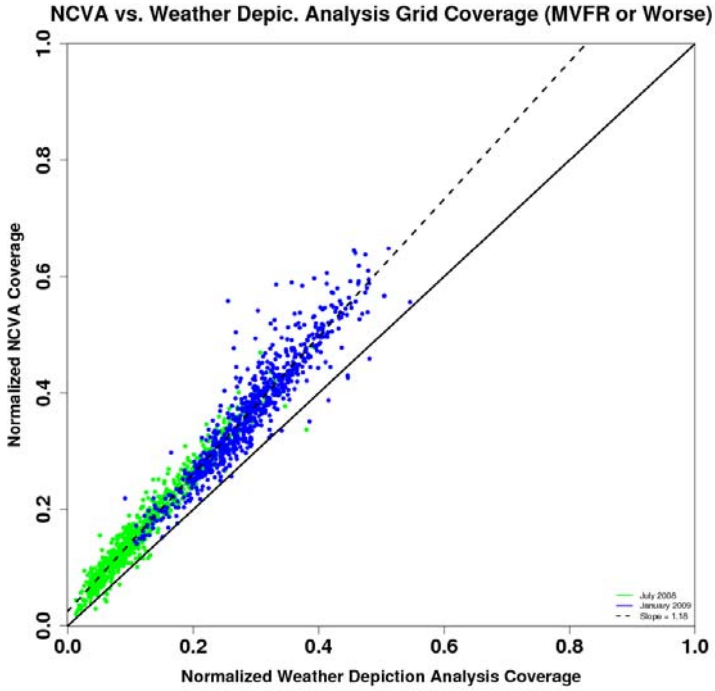

Figure 5.11: Scatter plot diagram of NCVA grid coverage vs. Weather Depiction Analysis grid coverage. Solid line represents a slope=1 line and the dashed line represents the linear regression of points (slope=1.18).

Summary statistics of the direct grid-to-grid comparison between NCVA and the Weather Depiction Analysis revealed a bias between the two products of less than 0.1 (not shown). This positive bias denotes a tendency for the NCVA to analyze MVFR or worse conditions over a larger area than the Weather Depiction Analysis. The RMSE between the two products simply denotes that on average, each grid box analysis is different by approximately 0.3 categories (**Figure 5.12b**). Categories for the matched NCVA and Weather Depiction Analysis are 0 (VFR conditions), 1 (MVFR conditions), and 2 (IFR or worse conditions).

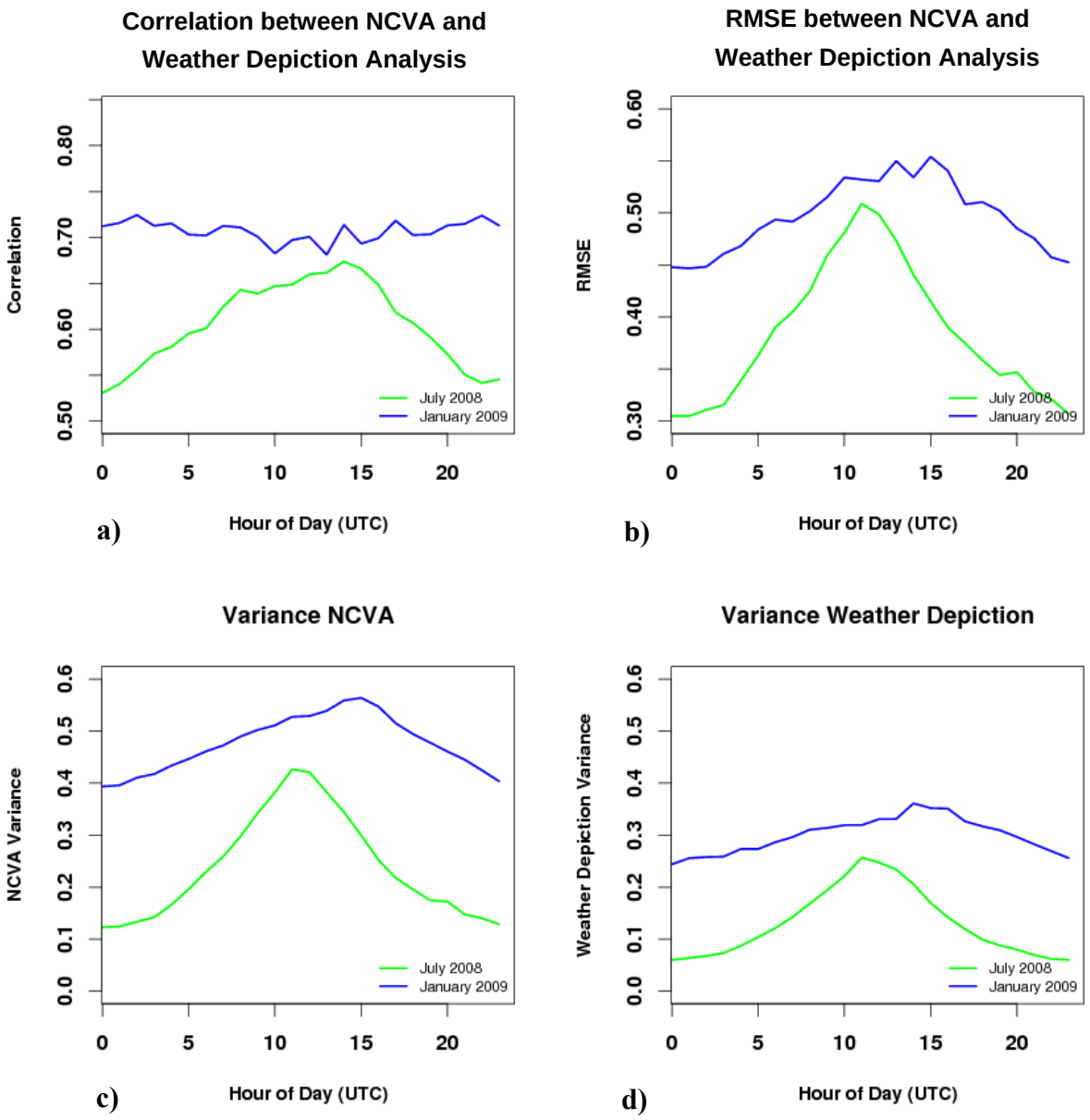

Figure 5.12: Continuous statistics for NCVA vs. Weather Depiction Analysis separated by analysis time (July 2008 in green and January 2009 in blue): **a)** Correlation between NCVA and Weather Depiction Analysis, **b)** Root Mean Square Error between NCVA and Weather Depiction Analysis, **c)** Variance of the NCVA product and **d)** Variance of the Weather Depiction Analysis.

The variance of each matched grid was calculated. The variance was elevated in both analyses in January, but overall, the NCVA product had a higher variance than did the Weather Depiction Analysis (**Figure 5.12c,d**). Given the equalization of analysis categories between grids in this evaluation, the reduction in variance in the Weather Depiction Analysis is most likely the cause of a smaller sampling of METAR observations from which its analysis is created. Seasonally, these summary statistics are all elevated for the month of January given the higher

frequency of events for analysis. Further, diurnal signals in these statistics are more defined during the summer months and are most likely an artifact of a climatological signal.

6. Conclusions

This report summarizes a formal quality assessment of the National Ceiling and Visibility Analysis product (NCVA), a gridded analysis that evaluates reported ceiling and visibility information for the purpose of improving the flight planning process. On behalf of the Federal Aviation Administration's Aviation Weather Research Program, and in support of an Aviation Weather Technology Transfer (AWTT) D4 (operational) decision point, this study was carried out to examine the following:

- The quality of the NCVA product with respect to a baseline analysis.
- The effect on the NCVA of utilizing a satellite-based cloud mask.
- The potential value of NCVA's frequent update-cycle to the flight planning process.
- NCVA's performance compared to the operational Weather Depiction Analysis, a product specifically referenced in the NCVA Concept of Use.

Constrained to the use of METARs for verification, this assessment employs a cross-validation technique to create independence between the observational set and the input set that is utilized by the NCVA algorithm. Cross-validation statistics produced from METAR data that are withheld from the NCVA should not be interpreted as measures of the algorithm's skill *at* METAR sites, but rather as a measure of performance of the NCVA grid at points *away from* METAR sites. The NCVA algorithm retains information from all operationally available METARs used as input, never altering data at grid points associated with available METAR reports.

In the absence of a secondary gridded product that represents current operational planning information, the analysis team created a proxy baseline product for comparison. This product, referred to in this report as the Nearest Neighbor Analysis (NN-A), simply uses the METAR closest to each evaluation point in determining ceiling, visibility, and flight category values. The study period consists of the summer months of 2008 and the winter months of 2008-09, chosen to provide significant data for stratification of results by season and time of day.

6.1 Quality of the NCVA Product

Overall, the NCVA could add significant value to the planning process compared with the baseline analysis (NN-A):

- By more effectively detecting IFR events and reducing risk throughout the airspace (NCVA Probability of Detection of 0.71 vs. 0.60 for the NN-A).
- By more effectively reducing false alarms of IFR events, resulting in more efficient use of the airspace (NCVA False Alarm Ratio of 0.25 vs. 0.39 for the NN-A), with a lower False Alarm Ratio being more favorable.
- Continuous measures of the error in ceiling and visibility attributes concur with the observations stated above.

The quality of the NCVA and the NN-A differ during the wintertime and summertime:

- NCVA has a significantly higher detection rate of IFR events in the wintertime than in the summertime (0.77 vs. 0.60).
- The NCVA has a slightly lower (better) False Alarm Ratio in the wintertime than in the summertime (0.22 vs. 0.31).
- Continuous measures of the error in ceiling and visibility attributes concur with the observations stated above.

Quality measures differ between normal and low confidences:

- While the IFR detection rates for *normal* and *low* flight category confidences are similar (0.69 vs. 0.71), the False Alarm Ratio for the *normal* confidences is lower (0.17 vs. 0.25).
- For ceiling and visibility values, *normal* confidence analysis points are more accurate than *low* confidence analysis points: Root Mean Square Errors of 1527 ft. and 1.88 SM for *normal* confidences vs. 1821 ft. and 2.41 SM for *low* confidences.

NCVA performance was found to vary greatly by region, sub-region, and weather regime:

- Performance of the NCVA was found to be most favorable along the East Coast, the Southwest Coast, and in the Midwest. It performed much less favorably in the Intermountain West, High Plains region, and during the summertime along the Gulf Coast. Consistent among the sub-regions is demonstrated superiority of the NCVA over the NN-A.

- Skill was found to be greater in sub-regions possessing high METAR station density and where long-lived, large-scale IFR events occur frequently. This is evident along the West Coast in the summertime (sub-regional detection rate of 0.89) and in the Northeast in the wintertime (sub-regional detection rate of 0.91).

6.2 Effect of the Satellite-based Cloud Mask

The study indicates that when the cloud mask is utilized there is**:**

- An overall decrease in false alarms (5.9%) that outweighs an increase in misses (0.5%), resulting in more efficient use of the airspace while only slightly increasing the risk; the NCVA risk is still significantly lower than that of the baseline.
- A measurable difference between the daytime and nighttime, as a large number of analysis grid points possessing *low* flight category confidence are actively re-assigned to *normal* flight category confidence during the daytime (24.3%), but only a negligible change during the nighttime.

6.3 Potential Value of the Frequent Update Rate

Incremental changes in the updates of NCVA every five minutes appear to contain information useful to planners:

- The linear correlation between an NCVA issuance and its successor one hour later is ~0.8, representing a significant change in flight conditions over the CONUS.
- The five-minute updates that occur between hourly issuances appear to effectively capture the incremental changes over the CONUS, as indicated by linear changes in correlation between the initial and intermediate issuances.

6.4 Comparison to the Weather Depiction Analysis

This study shows that the NCVA, in accordance with its Concept of Use, performs as well as and is consistent with the Weather Depiction Analysis:

- NCVA and the Weather Depiction Analysis have an overall correlation greater than 0.6, indicating that the two products may perform in a consistent manner.
- With no suitable way to directly measure the quality of the Weather Depiction Analysis, the study found indications that NCVA performs at least as well as the Weather Depiction Analysis.

Acknowledgements

This research is in response to requirements and funding provided by the Federal Aviation Administration. The views expressed are those of the authors and do not necessarily represent the official policy and position of the U.S. Government.

7. References

Carrin, G., J. DeBerry, J. Hansford, 2008: Climate of Shreveport, Louisiana. National Weather Service Office Shreveport, LA, http://www.srh.noaa.gov/ssd/techmemo/sr230.pdf.

COMET, 2004: Fog and Stratus Forecast Approaches. University Corporation for Atmospheric Research, http://www.meted.ucar.edu/dlac/lesson1/print.htm.

Federal Aviation Administration, 1975: Aviation Weather for Pilots and Flight Operations Personnel. Department of Transportation and Department of Commerce, ASA publications, http://rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/0/C2DF8D9D7471617786256A020078083A?OpenDocument.

Federal Aviation Administration, 2008: Aeronautical Information Manual (section 7-1-7). http://www.faa.gov/air_traffic/publications/ATpubs/AIM.

Federal Aviation Administration Air Traffic Organization, 2007: National Ceiling and Visibility (NCV) CONUS Concept of Use, version 2.0.

Glahn, H.R. and D. P. Ruth, 2003: The New Digital Forecast Database of the National Weather Service, Bull. Amer. Meteor. Soc., 84, 195-201.

Herzegh, P., G. Wiener, R. Bateman, J. Cowie, J.Black, 2009: NCVA – The NCV CONUS Analysis Product, available from herzegh@ucar.edu.

Jedlovec, G. J. and K. Laws, 2003: GOES cloud detection at the Global Hydrology and Climate Center. Preprints, 12th Conf. on Satellite Meteorology and Oceanography, Long Beach, CA, Amer. Meteor. Soc., CD-ROM, P1.21.

Leipper, D., 1994: Fog on the U.S. West Coast: A Review. *Bull. Amer. Meteor. Soc.*, **75**, 229-240.

National Weather Service, 2000: National Weather Service Operations Manual. National Weather Service, http://www.nws.noaa.gov/wsom/manual/archives/NC750006.pdf.

National Weather Service Instruction 10-813, 2005: Aviation Weather Services, NWSPD 10-8, Terminal Aerodrome Forecasts, 57 pp.

Neter, J., M. H. Kunter, C. J. Nachtsheim, and W. Wasserman, 1996: Applied Linear Statistical Models. McGraw Hill / Irwin, Chicago, Ill.

Skiena, Steven S., The Algorithm Design Manual, Telos, 1997.

Tardif, Robert, Roy M. Rasmussen, 2007: Event-Based Climatology and Typology of Fog in the New York City Region. *J. Appl. Meteor. Climatol.*, **46**, 1141–1168.

Vines, Andrew, 2010: Low-level stratus and fog development in Houston, Texas. http://www.theweatherprediction.com/weatherpapers/086/index.html.

Wilks, Daniel S., 2006: Statistical Methods in the Atmospheric Sciences, Vol. 91, 2d ed. Academic Press, 627 pp.

Appendix A

Table A.1: Performance measures from a matching method using radius 10-km and "best flight category" mechanics. Measures are for all seasons, summer, and winter, and for the entire CONUS (in bold) and the northeast. The NN-A analysis is marked with an asterisk. The NCVA results are for *all*, *low*, and *normal* flight category confidences. HSS is the Heidke Skill Score, and ETS is the Equitable Threat Score (Wilks, 2006).

Analysis	F.C. Confidence Level	Season	Region	PODy	CSI	HSS	ETS	TSS	PODn	FARatio	Bias
NN-A*	ALL	ALL	CONUS	0.60	0.43	0.57	0.40	0.57	0.97	0.39	0.99
NCV-A	ALL	ALL	CONUS	0.71	0.58	0.71	0.55	0.69	0.98	0.25	0.95
NCV-A	LOW	ALL	CONUS	0.71	0.58	0.70	0.54	0.69	0.97	0.25	0.95
NCV-A	NORMAL	ALL	CONUS	0.69	0.61	0.76	0.61	0.69	1.00	0.17	0.83
NN-A*	ALL	ALL	NE	0.64	0.46	0.58	0.40	0.59	0.95	0.39	1.04
NCV-A	ALL	ALL	NE	0.76	0.62	0.74	0.58	0.73	0.97	0.23	0.99
NCV-A	LOW	ALL	NE	0.77	0.63	0.73	0.57	0.73	0.96	0.23	0.99
NCV-A	NORMAL	ALL	NE	0.59	0.54	0.70	0.54	0.59	1.00	0.13	0.69
NN-A*	ALL	S	CONUS	0.47	0.32	0.46	0.30	0.45	0.97	0.51	0.97
NCV-A	ALL	S	CONUS	0.60	0.47	0.62	0.45	0.59	0.99	0.31	0.87
NCV-A	LOW	S	CONUS	0.60	0.47	0.62	0.45	0.58	0.98	0.32	0.88
NCV-A	NORMAL	S	CONUS	0.54	0.47	0.64	0.47	0.54	1.00	0.20	0.67
NN-A*	ALL	S	NE	0.58	0.40	0.51	0.34	0.51	0.94	0.43	1.02
NCV-A	ALL	S	NE	0.71	0.58	0.70	0.54	0.68	0.97	0.24	0.94
NCV-A	LOW	S	NE	0.72	0.58	0.69	0.52	0.67	0.95	0.24	0.95
NCV-A	NORMAL	S	NE	0.32	0.30	0.46	0.30	0.32	1.00	0.21	0.41
NN-A*	ALL	W	CONUS	0.66	0.49	0.62	0.45	0.62	0.96	0.34	1.00
NCV-A	ALL	W	CONUS	0.77	0.63	0.75	0.60	0.74	0.97	0.22	0.98
NCV-A	LOW	W	CONUS	0.77	0.63	0.74	0.59	0.74	0.97	0.22	0.99
NCV-A	NORMAL	W	CONUS	0.81	0.71	0.83	0.71	0.81	1.00	0.15	0.95
NN-A*	ALL	W	NE	0.72	0.53	0.65	0.49	0.67	0.96	0.33	1.07
NCV-A	ALL	W	NE	0.83	0.68	0.78	0.65	0.80	0.97	0.21	1.05
NCV-A	LOW	W	NE	0.83	0.68	0.78	0.63	0.79	0.96	0.21	1.05
NCV-A	NORMAL	W	NE	0.85	0.77	0.87	0.77	0.85	1.00	0.11	0.95

Table A.2: Performance measures from a matching method using 4pt and "best flight category" mechanics. Measures are for all seasons, summer, and winter, and for the entire CONUS (in bold) and the northeast. The NN-A analysis is marked with an asterisk. The NCVA results are for *all*, *low*, and *normal* flight category confidences. HSS is the Heidke Skill Score, and ETS is the Equitable Threat Score (Wilks, 2006).

Analysis	F.C. Confidence Level	Season	Region	PODy	CSI	HSS	ETS	TSS	PODn	FARatio	Bias
NN-A*	ALL	ALL	CONUS	0.60	0.43	0.57	0.40	0.57	0.97	0.39	0.99
NCV-A	ALL	ALL	CONUS	0.66	0.50	0.64	0.47	0.63	0.97	0.33	0.98
NCV-A	LOW	ALL	CONUS	0.66	0.50	0.63	0.46	0.62	0.96	0.33	0.98
NCV-A	NORMAL	ALL	CONUS	0.64	0.52	0.68	0.52	0.64	1.00	0.26	0.86
NN-A*	ALL	ALL	NE	0.64	0.46	0.58	0.40	0.59	0.95	0.39	1.04
NCV-A	ALL	ALL	NE	0.70	0.52	0.64	0.47	0.65	0.96	0.33	1.03
NCV-A	LOW	ALL	NE	0.70	0.52	0.63	0.46	0.64	0.94	0.33	1.04
NCV-A	NORMAL	ALL	NE	0.54	0.43	0.60	0.43	0.54	1.00	0.31	0.77
NN-A*	ALL	S	CONUS	0.48	0.32	0.46	0.30	0.45	0.97	0.51	0.97
NCV-A	ALL	S	CONUS	0.53	0.38	0.53	0.36	0.51	0.98	0.43	0.93
NCV-A	LOW	S	CONUS	0.54	0.38	0.52	0.35	0.51	0.97	0.43	0.94
NCV-A	NORMAL	S	CONUS	0.46	0.37	0.53	0.36	0.46	1.00	0.36	0.72
NN-A*	ALL	S	NE	0.58	0.40	0.51	0.34	0.51	0.94	0.43	1.02
NCV-A	ALL	S	NE	0.63	0.47	0.58	0.41	0.58	0.95	0.36	0.99
NCV-A	LOW	S	NE	0.64	0.47	0.57	0.40	0.57	0.93	0.36	1.00
NCV-A	NORMAL	S	NE	0.26	0.19	0.32	0.19	0.25	1.00	0.57	0.59
NN-A*	ALL	W	CONUS	0.66	0.49	0.62	0.45	0.62	0.96	0.34	1.00
NCV-A	ALL	W	CONUS	0.72	0.56	0.68	0.52	0.68	0.97	0.28	1.00
NCV-A	LOW	W	CONUS	0.72	0.56	0.67	0.51	0.67	0.96	0.29	1.00
NCV-A	NORMAL	W	CONUS	0.77	0.64	0.78	0.64	0.77	1.00	0.20	0.96
NN-A*	ALL	W	NE	0.72	0.53	0.65	0.49	0.67	0.96	0.33	1.07
NCV-A	ALL	W	NE	0.77	0.59	0.71	0.55	0.73	0.96	0.29	1.08
NCV-A	LOW	W	NE	0.77	0.59	0.70	0.53	0.72	0.95	0.29	1.08
NCV-A	NORMAL	W	NE	0.81	0.71	0.83	0.71	0.81	1.00	0.15	0.95

Table A.3: Sub-regional performance measures from a matching method using radius 10-km and "best flight category" mechanics. Measures are for NCVA from all seasons, summer, and winter. These NCVA results are for ***all*** flight category confidences. HSS is the Heidke Skill Score, and ETS is the Equitable Threat Score (Wilks, 2006).

Analysis	Season	Region	PODy	CSI	HSS	ETS	TSS	PODn	FARatio	Bias
NCV-A	ALL	Chicago Area	0.82	0.71	0.81	0.69	0.80	0.99	0.16	0.97
NCV-A	S	Chicago Area	0.68	0.57	0.72	0.56	0.67	0.99	0.22	0.88
NCV-A	W	Chicago Area	0.86	0.76	0.84	0.72	0.84	0.98	0.14	1.00
NCV-A	ALL	Denver Area	0.70	0.54	0.68	0.52	0.68	0.98	0.30	0.99
NCV-A	S	Denver Area	0.58	0.45	0.61	0.44	0.57	0.99	0.33	0.86
NCV-A	W	Denver Area	0.73	0.56	0.69	0.53	0.71	0.97	0.29	1.03
NCV-A	ALL	Seattle Area	0.70	0.58	0.69	0.53	0.66	0.97	0.23	0.91
NCV-A	S	Seattle Area	0.44	0.39	0.54	0.37	0.44	0.99	0.23	0.58
NCV-A	W	Seattle Area	0.77	0.63	0.71	0.54	0.71	0.94	0.23	1.00
NCV-A	ALL	Houston - Pensacola	0.71	0.57	0.71	0.55	0.69	0.98	0.25	0.94
NCV-A	S	Houston - Pensacola	0.23	0.16	0.27	0.16	0.22	0.99	0.63	0.61
NCV-A	W	Houston - Pensacola	0.82	0.68	0.78	0.65	0.79	0.97	0.19	1.01
NCV-A	ALL	Los Angeles – San Francisco	0.82	0.72	0.82	0.69	0.81	0.98	0.15	0.97
NCV-A	S	Los Angeles – San Francisco	0.89	0.77	0.85	0.74	0.87	0.98	0.14	1.03
NCV-A	W	Los Angeles – San Francisco	0.71	0.62	0.75	0.60	0.69	0.99	0.16	0.84
NCV-A	ALL	Memphis - Atlanta	0.72	0.58	0.71	0.55	0.70	0.98	0.26	0.98
NCV-A	S	Memphis - Atlanta	0.49	0.36	0.50	0.34	0.47	0.98	0.44	0.87
NCV-A	W	Memphis - Atlanta	0.82	0.68	0.79	0.65	0.79	0.97	0.20	1.02
NCV-A	ALL	New York - Boston	0.83	0.76	0.84	0.73	0.82	0.99	0.11	0.93
NCV-A	S	New York - Boston	0.77	0.70	0.80	0.67	0.76	0.98	0.12	0.88
NCV-A	W	New York - Boston	0.91	0.83	0.89	0.81	0.90	0.99	0.09	1.00
NCV-A	ALL	Oklahoma City – San Antonio	0.83	0.72	0.83	0.71	0.82	0.99	0.16	0.99
NCV-A	S	Oklahoma City – San Antonio	0.67	0.54	0.69	0.53	0.66	0.99	0.27	0.91
NCV-A	W	Oklahoma City – San Antonio	0.89	0.79	0.87	0.78	0.88	0.99	0.12	1.02

Table A.4: Sub-regional performance measures from a matching method using radius 10-km and "best flight category" mechanics. Measures are for NN-A from all seasons, summer, and winter. HSS is the Heidke Skill Score, and ETS is the Equitable Threat Score (Wilks, 2006).

Analysis	Season	Region	PODy	CSI	HSS	ETS	TSS	PODn	FARatio	Bias
NN-A	ALL	Chicago Area	0.72	0.55	0.69	0.52	0.69	0.97	0.29	1.02
NN-A	S	Chicago Area	0.54	0.38	0.53	0.36	0.52	0.98	0.44	0.97
NN-A	W	Chicago Area	0.78	0.62	0.73	0.57	0.74	0.96	0.25	1.04
NN-A	ALL	Denver Area	0.51	0.34	0.48	0.31	0.48	0.97	0.50	1.02
NN-A	S	Denver Area	0.43	0.29	0.44	0.28	0.42	0.99	0.52	0.90
NN-A	W	Denver Area	0.53	0.35	0.48	0.31	0.49	0.96	0.49	1.05
NN-A	ALL	Seattle Area	0.54	0.40	0.50	0.34	0.48	0.94	0.40	0.90
NN-A	S	Seattle Area	0.30	0.23	0.35	0.21	0.28	0.98	0.48	0.57
NN-A	W	Seattle Area	0.61	0.44	0.51	0.34	0.50	0.89	0.39	1.00
NN-A	ALL	Houston - Pensacola	0.66	0.49	0.63	0.46	0.63	0.97	0.34	1.00
NN-A	S	Houston - Pensacola	0.19	0.12	0.20	0.11	0.18	0.99	0.74	0.75
NN-A	W	Houston - Pensacola	0.76	0.59	0.71	0.55	0.73	0.96	0.28	1.06
NN-A	ALL	Los Angeles – San Francisco	0.66	0.47	0.60	0.43	0.61	0.95	0.38	1.06
NN-A	S	Los Angeles – San Francisco	0.74	0.51	0.62	0.45	0.67	0.93	0.38	1.18
NN-A	W	Los Angeles – San Francisco	0.52	0.39	0.53	0.36	0.49	0.97	0.38	0.83
NN-A	ALL	Memphis - Atlanta	0.62	0.45	0.59	0.42	0.59	0.97	0.37	0.98
NN-A	S	Memphis - Atlanta	0.39	0.25	0.38	0.23	0.37	0.97	0.58	0.93
NN-A	W	Memphis - Atlanta	0.71	0.55	0.67	0.51	0.68	0.96	0.29	1.00
NN-A	ALL	New York - Boston	0.68	0.53	0.65	0.48	0.65	0.96	0.30	0.98
NN-A	S	New York - Boston	0.60	0.44	0.56	0.39	0.55	0.95	0.37	0.96
NN-A	W	New York - Boston	0.79	0.65	0.76	0.62	0.76	0.98	0.21	1.00
NN-A	ALL	Oklahoma City – San Antonio	0.72	0.55	0.70	0.54	0.71	0.98	0.29	1.02
NN-A	S	Oklahoma City – San Antonio	0.48	0.32	0.47	0.31	0.47	0.99	0.52	1.00
NN-A	W	Oklahoma City – San Antonio	0.81	0.66	0.78	0.64	0.79	0.98	0.22	1.03

Appendix B

Table B.1: Root Mean Square Error (RMSE) for NCVA and NN-A ceiling along with ΔRMSE (NCVA minus NN-A) for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all analyzed spatial domains.

CEILING RMSE (ft.)		All			Summer			Winter		
		NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE
All Conf	NDFD	1859.04	2012.34	-153.3	2312.44	2465.28	-152.84	1780.61	1934.79	-154.2
	CONUS	1819.87	1980.72	-160.84	2282.21	2439.6	-157.38	1739.69	1902.1	-162.4
	NE US	1959.6	2160.32	-200.72	2332.79	2549.99	-217.2	1750.33	1943.55	-193.2
Low Conf	NDFD	1859.81	2013.36	-153.56	2314.66	2467.92	-153.26	1781.27	1935.68	-154.4
	CONUS	1820.62	1981.76	-161.14	2284.51	2442.37	-157.87	1740.33	1903	-162.7
	NE US	1960.29	2161.44	-201.15	2333.04	2550.48	-217.44	1750.94	1944.7	-193.8
Norm Conf	NDFD	1528.23	1559.8	-31.57	1744.24	1774.22	-29.98	1458.27	1490.53	-32.25
	CONUS	1527.23	1558.83	-31.59	1745.36	1775.33	-29.97	1456.53	1488.83	-32.3
	NE US	1654.02	1640.83	13.19	2131.84	2141.61	-9.78	1534.76	1514.63	20.13

Table B.2: Root Mean Square Error (RMSE) for NCVA and NN-A ceiling along with ΔRMSE (NCVA minus NN-A) in the NDFD domain for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all METAR flight category conditions.

CEILING RMSE by FLIGHT CATEGORY (ft.); NDFD DOMAIN

		All			Low Conf			Normal Conf		
		NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE
All	LIFR Condition	1516.29	1750.8	-234.5	1517	1751.66	-234.7	666.23	688.36	-22.1
	IFR Condition	1422.09	1645.77	-223.7	1423.19	1647.08	-223.9	745.26	832.25	-87
	MVFR Condition	1531.1	1658.77	-127.7	1531.41	1659.19	-127.8	1332.03	1380.16	-48.1
	VFR Condition	2261.49	2406.78	-145.3	2263.68	2409.44	-145.8	1745.53	1767.01	-21.5
Winter	LIFR Condition	1405.11	1647.51	-242.4	1405.92	1648.51	-242.6	593.04	616.87	-23.8
	IFR Condition	1361.61	1575.14	-213.5	1362.08	1575.69	-213.6	829.35	945.31	-116
	MVFR Condition	1483.78	1614.16	-130.4	1483.93	1614.4	-130.5	1344.27	1396.32	-52.1
	VFR Condition	2165.68	2313.59	-147.9	2167.8	2316.17	-148.4	1638.04	1657.23	-19.2
Summer	LIFR Condition	1837.46	2056.55	-219.1	1837.62	2056.77	-219.1	1294.03	1311.35	-17.3
	IFR Condition	1677.85	1944.19	-266.3	1682.78	1950.04	-267.3	677	738.44	-61.4
	MVFR Condition	1915.37	2026.51	-111.1	1918.48	2029.95	-111.5	1310.32	1351.41	-41.1
	VFR Condition	2718.64	2855.46	-136.8	2721.72	2859.06	-137.3	2135.2	2164.55	-29.4

Table B.3: Root Mean Square Error (RMSE) for NCVA and NN-A ceiling along with ΔRMSE (NCVA minus NN-A) in the CONUS domain for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all METAR flight category conditions.

CEILING RMSE by FLIGHT CATEGORY (ft.); CONUS DOMAIN										
		All			Low Conf			Normal Conf		
		NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE
All	LIFR Condition	1471.6	1725.77	-254.2	1472.34	1726.69	-254.4	666.23	688.36	-22.1
	IFR Condition	1373.8	1612.46	-238.7	1374.9	1613.82	-238.9	749.33	836.58	-87.3
	MVFR Condition	1486.12	1625.03	-138.9	1486.39	1625.44	-139.1	1332.49	1380.74	-48.2
	VFR Condition	2220.01	2367.56	-147.5	2222.18	2370.24	-148.1	1743.86	1765.33	-21.5
Winter	LIFR Condition	1360.12	1622.45	-262.3	1360.97	1623.52	-262.5	593.04	616.87	-23.8
	IFR Condition	1311.84	1542.03	-230.2	1312.31	1542.6	-230.3	831.22	947.4	-116.2
	MVFR Condition	1437.4	1580.05	-142.7	1437.51	1580.27	-142.8	1344.6	1396.7	-52.1
	VFR Condition	2122.11	2272.66	-150.6	2124.2	2275.25	-151.1	1635.55	1654.73	-19.2
Summer	LIFR Condition	1795.51	2034.21	-238.7	1795.68	2034.44	-238.8	1294.03	1311.35	-17.3
	IFR Condition	1634.75	1910.11	-275.4	1639.9	1916.29	-276.4	682.12	743.59	-61.5
	MVFR Condition	1874.04	1989.91	-115.9	1877.25	1993.51	-116.3	1310.94	1352.21	-41.3
	VFR Condition	2692.41	2830.02	-137.6	2695.62	2833.8	-138.2	2134.15	2163.45	-29.3

Table B.4: Root Mean Square Error (RMSE) for NCVA and NN-A ceiling along with ΔRMSE (NCVA minus NN-A) in the northeast US domain for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all METAR flight category conditions.

CEILING RMSE by FLIGHT CATEGORY (ft.); NE DOMAIN

		All			Low Conf			Normal Conf		
		NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE	NCVA/METAR	NN-A/METAR	Δ RMSE
All	LIFR Condition	1822.35	1945.48	-123.1	1822.38	1945.51	-123.1	1211.33	1252.82	-41.5
	IFR Condition	1357.77	1589.42	-231.7	1358.01	1589.85	-231.8	1142.89	1178.36	-35.5
	MVFR Condition	1707.76	1847.28	-139.5	1707.86	1847.51	-139.7	1572.14	1495.56	76.6
	VFR Condition	2229.63	2470.11	-240.5	2232.61	2474.35	-241.7	1726.96	1719.87	7.1
Winter	LIFR Condition	1232.63	1434.48	-201.8	1232.72	1434.61	-201.9	939.68	977.68	-38
	IFR Condition	1148.08	1313.55	-165.5	1148.44	1314.14	-165.7	939.06	956.49	-17.4
	MVFR Condition	1582.08	1755.1	-173	1582.13	1755.3	-173.2	1519.91	1452.34	67.6
	VFR Condition	2007.29	2220.37	-213.1	2009.81	2224.14	-214.3	1602.19	1586.04	16.1
Summer	LIFR Condition	1900.15	2015.57	-115.4	1900.15	2015.57	-115.4	1973.02	2029.03	-56
	IFR Condition	1554.59	1843.88	-289.3	1554.46	1843.85	-289.4	1832.22	1917.73	-85.5
	MVFR Condition	1990.57	2062.14	-71.6	1990.76	2062.43	-71.7	1699.4	1601.6	97.8
	VFR Condition	2909.22	3231.72	-322.5	2912.76	3236.56	-323.8	2203.13	2225.09	-22

Appendix C

Table C.1: Root Mean Square Error (RMSE) for NCVA and NN-A visibility along with ΔRMSE (NCVA minus NN-A) for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all analyzed spatial domains.

VISIBILITY RMSE (SM)

		All			Summer			Winter		
		NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE
All Conf	NDFD	2.44	2.59	-0.15	2.47	2.65	-0.18	2.43	2.55	-0.12
	CONUS	2.41	2.57	-0.16	2.45	2.63	-0.18	2.38	2.51	-0.13
	NE US	2.54	2.75	-0.21	2.63	2.86	-0.23	2.31	2.46	-0.15
Low Conf	NDFD	2.45	2.6	-0.15	2.48	2.66	-0.18	2.43	2.55	-0.12
	CONUS	2.41	2.57	-0.16	2.45	2.64	-0.19	2.38	2.51	-0.13
	NE US	2.54	2.75	-0.21	2.63	2.86	-0.23	2.31	2.46	-0.15
Normal Conf	NDFD	1.87	1.95	-0.08	1.82	1.89	-0.07	2.11	2.19	-0.08
	CONUS	1.88	1.96	-0.08	1.83	1.9	-0.07	2.1	2.18	-0.08
	NE US	1.65	1.86	-0.21	1.64	1.87	-0.23	1.73	1.82	-0.09

Table C.2: Root Mean Square Error (RMSE) for NCVA and NN-A visibility along with ΔRMSE (NCVA minus NN-A) in the NDFD domain for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all METAR flight category conditions.

VISIBILITY RMSE by FLIGHT CATEGORY (SM); NDFD DOMAIN										
		All			Low Conf			Normal Conf		
		NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE
All	LIFR Condition	2.56	2.79	-0.23	2.56	2.79	-0.23	2.06	2.18	-0.12
	IFR Condition	2.24	2.44	-0.2	2.24	2.44	-0.2	2.39	2.49	-0.1
	MVFR Condition	2.2	2.3	-0.1	2.2	2.3	-0.1	2.34	2.38	-0.04
	VFR Condition	2.57	2.71	-0.14	2.6	2.73	-0.13	1.56	1.63	-0.07
Winter	LIFR Condition	2.35	2.56	-0.2	2.36	2.56	-0.2	1.9	2.03	-0.13
	IFR Condition	2.15	2.33	-0.18	2.15	2.33	-0.18	2.23	2.34	-0.1
	MVFR Condition	2.21	2.26	-0.05	2.21	2.26	-0.05	2.28	2.31	-0.03
	VFR Condition	2.84	2.96	-0.12	2.84	2.96	-0.12	1.91	2.01	-0.1
Summer	LIFR Condition	3.19	3.51	-0.32	3.19	3.51	-0.32	4.48	4.5	-0.02
	IFR Condition	2.7	2.98	-0.28	2.69	2.98	-0.28	3.15	3.26	-0.11
	MVFR Condition	2.2	2.33	-0.13	2.2	2.33	-0.13	2.37	2.41	-0.04
	VFR Condition	2.38	2.53	-0.15	2.41	2.56	-0.15	1.5	1.59	-0.09

Table C.3: Root Mean Square Error (RMSE) for NCVA and NN-A visibility along with ΔRMSE (NCVA minus NN-A) in the CONUS domain for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all METAR flight category conditions.

VISIBILITY RMSE by FLIGHT CATEGORY (SM); CONUS DOMAIN

		All			Low Conf			Normal Conf		
		NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE
All	LIFR Condition	2.54	2.78	-0.24	2.54	2.78	-0.24	2.05	2.17	-0.12
	IFR Condition	2.2	2.41	-0.21	2.2	2.41	-0.21	2.35	2.45	-0.1
	MVFR Condition	2.18	2.28	-0.1	2.18	2.28	-0.1	2.33	2.37	-0.04
	VFR Condition	2.54	2.68	-0.14	2.56	2.7	-0.14	1.55	1.64	-0.09
Winter	LIFR Condition	2.32	2.53	-0.21	2.32	2.53	-0.21	1.9	2.03	-0.13
	IFR Condition	2.1	2.28	-0.18	2.1	2.28	-0.18	2.18	2.29	-0.11
	MVFR Condition	2.17	2.23	-0.06	2.17	2.23	-0.06	2.27	2.3	-0.03
	VFR Condition	2.81	2.93	-0.12	2.81	2.93	-0.12	1.93	2.03	-0.1
Summer	LIFR Condition	3.17	3.49	-0.32	3.17	3.49	-0.32	4.46	4.47	-0.01
	IFR Condition	2.67	2.96	-0.29	2.67	2.96	-0.29	3.14	3.24	-0.1
	MVFR Condition	2.19	2.32	-0.13	2.19	2.32	-0.13	2.36	2.4	-0.04
	VFR Condition	2.36	2.52	-0.16	2.39	2.55	-0.16	1.52	1.6	-0.08

Table C.4: Root Mean Square Error (RMSE) for NCVA and NN-A visibility along with ΔRMSE (NCVA minus NN-A) in the northeast US domain for the summer, winter, and combined seasons against low, normal, and all confidences. RMSE information is shown for all METAR flight category conditions.

VISIBILITY RMSE by FLIGHT CATEGORY (SM); NE DOMAIN

		All			Low Conf			Normal Conf		
		NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE	NCVA/METAR	NN-A/METAR	ΔRMSE
All	LIFR Condition	2.57	2.86	-0.29	2.57	2.86	-0.29	2.71	2.72	-0.01
	IFR Condition	2.31	2.55	-0.24	2.31	2.55	-0.24	1.73	1.87	-0.14
	MVFR Condition	2.29	2.41	-0.12	2.28	2.41	-0.12	2.06	2.19	-0.13
	VFR Condition	2.73	2.96	-0.23	2.75	2.98	-0.23	1.42	1.69	-0.27
Winter	LIFR Condition	1.6	1.82	-0.22	1.6	1.82	-0.22	1.77	1.79	-0.02
	IFR Condition	2.01	2.23	-0.22	2.01	2.23	-0.22	1.41	1.54	-0.13
	MVFR Condition	2.26	2.3	-0.04	2.26	2.3	-0.04	1.76	1.76	0
	VFR Condition	2.77	2.93	-0.16	2.77	2.94	-0.17	1.74	1.86	-0.12
Summer	LIFR Condition	2.88	3.19	-0.31	2.88	3.19	-0.31	4.67	4.67	0
	IFR Condition	2.58	2.84	-0.26	2.58	2.84	-0.26	2.3	2.46	-0.16
	MVFR Condition	2.29	2.44	-0.15	2.29	2.44	-0.15	2.14	2.33	-0.19
	VFR Condition	2.71	2.96	-0.25	2.74	3	-0.26	1.41	1.69	-0.28