P1.17 DEVELOPMENT OF FAA NATIONAL CEILING AND VISIBILITY PRODUCTS: CHALLENGES, STRATEGIES AND PROGRESS

Paul Herzegh*[1], Gerry Wiener[1], Richard Bankert[2], Richard Bateman[1], Stan Benjamin[3], John Brown[3], James Cowie[1], Michael Hadjimichael[2], Matthew Tryhane[1] and Beth Weekley[1]

[1]National Center for Atmospheric Research, Boulder, CO 80307
[2]Naval Research Laboratory, Monterey, CA 93943
[3]Global Systems Division, NOAA Earth Sciences Research Division, Boulder, CO, Canada

1. INTRODUCTION

The complex physical processes controlling ceiling and visibility (for example, the formation, evolution and motion of low cloud, precipitation and fog) and the diverse seasonal and geographic influences that modulate these controls across the continental U.S. and Alaska yield an extremely difficult analysis and forecast problem. This same phenomenology significantly impacts the safety and continuity of aviation operations, making VFR (visual flight rules) flight into impacted conditions associated with IFR (instrument flight rules) the leading cause among weather-related aviation accidents in the U.S.

Recent development of an automated ceiling and visibility (C&V) forecast system for the continental U.S. has utilized expert system methodology to blend numerical and observational inputs in the synthesis of ceiling and visibility analyses and forecasts out to 10 hours. This experimental system (posted at www.rap.ucar.edu/projects/cvis) makes use of current and historical METAR data, GOES satellite observations, numerical models, MOS forecasts and observations-based ruleset forecasts. As this system matures, a similar approach will be used to provide corresponding analyses and forecasts in Alaska.

This paper outlines recent progress and current functionality of the continental U.S. system. This work is carried out by the National Ceiling and Visibility (NCV) product development team under funding from the FAA's Aviation Weather Research Program.

* *Corresponding author address*: Paul Herzegh, NCAR, P.O. Box 3000, Boulder, CO 80307;
Email: herzegh@ucar.edu

2. OPERATIONS STRATEGY

NCV systems are being developed to provide high quality automated input to the National Weather Service (NWS) forecast process. As shown in Figure 1, the NCV grids are conceived to flow into the forecast process and undergo further processing under forecaster oversight. The processed grids are then planned to populate the aviation parameters within the NWS' National Digital Forecast Database, from which a unified, self-consistent set of aviation products and warnings will flow. The NCV system is expected to begin testing this operational role in stages over the next several years.

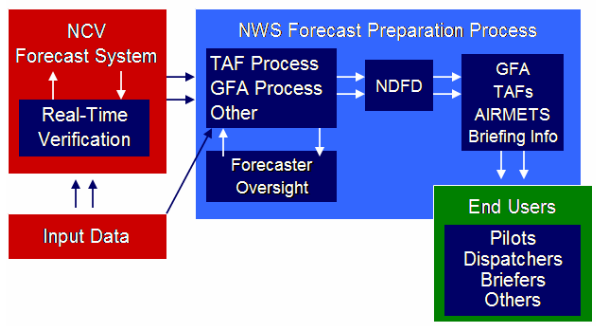


Figure 1. Conceptual view of the relationship between the NCV system (red) and the NWS forecast preparation process (blue). NCV analysis and forecast grids flow to NWS as initial fields for the TAF and GFA forecast process. Following forecaster input, modified grids populate the National Digital Forecast Database. Derived warning information flows to end users (green).

3. ANALYSIS PRODUCT

Nearest-neighbor interpolation of real-time observations from approximately 1524 operational CONUS C&V reporting sites are selectively combined with satellite data and high-resolution terrain data to formulate a continu- ously updating (15 minute) analysis of ceiling, visibility, flight category and terrain obscuration on the RUC 20 km grid. Higher resolution analysis grids (to 5 or 2.5 km resolution) will be implemented as soon as resources allow. The grids for ceiling, visibility and flight category are found to be accurate for local conditions and have good reliability toward representing the most probable conditions *between* reporting sites. Incorporation of high-resolution terrain data has allowed realistic representation of terrain obscuration, a major hazard under IFR conditions.

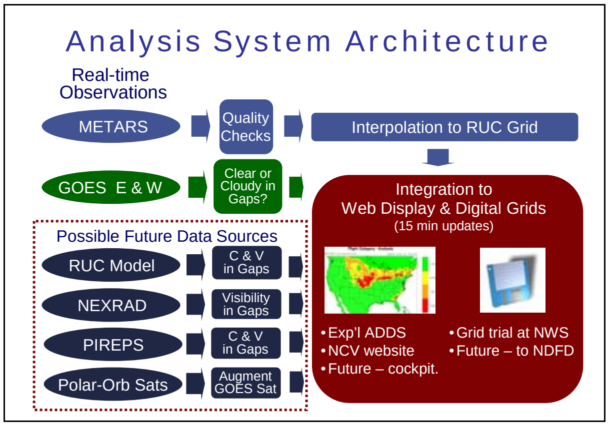


Figure 2. Schematic view of the NCV analysis processing system. Input data are shown in ovals at left. Output products are at lower right in red.

Key elements of the analysis system shown in Fig. 2 are outlined below.

- The light blue elements highlight the real-time ingest of ceiling, visibility and related data from each of the METAR sites operating within the RUC domain. Processing steps include quality checking, interpolation of ceiling and visibility observations to the RUC grid, and output to the integration process (red) that prepares a finalized analysis grid and corresponding real-time display.
- The green elements highlight the ingest of GOES-10 and GOES-12 real-time data, subsequent processing to distinguish cloudy from cloud-free regions, and application of the resulting cloud information to detect any clear (no ceiling) regions in the gap areas between METAR sites. Cloud-free areas are determined using the cloud mask developed by Jedlovec et al. (2003). This information on clear conditions in gap areas is then incorporated into the final analysis grids and displays in the integration step (red).
- The dark blue elements outline a variety of future sources of observational data not used today (*e.g.*, POES satellite data for improved cloud masking, PIREPS and radar data), the intermediary processing required to prepare these data for use, and the transfer of the resultant data products for use in the final integration step. Of course, the intermediary processing will vary from case to case depending upon the input data type and the output information or product desired.

A real-time NCV ceiling analysis demonstrating conditions ranging from clear to terrain-obscurred is shown in Fig. 3. Additional real-time experimental analyses can be viewed for evaluation (not operational use) at www.rap.ucar.edu/projects/cvis.

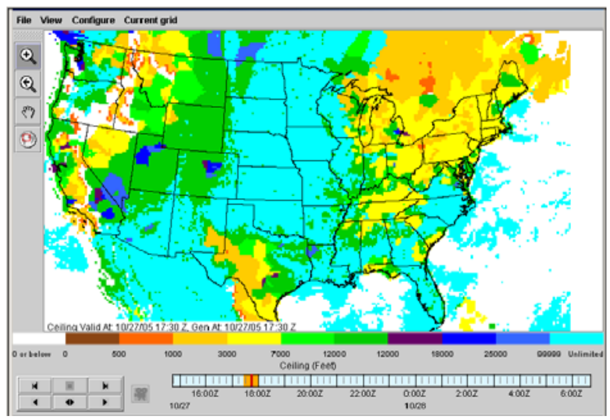


Figure 3. NCV real-time ceiling analysis showing clear conditions (cyan) through the midwest and lowered ceilings in the northeast, northwest and Texas. Significant terrain obscuration (white) is shown in the northwest.

4. IMPROVING CEILING ANALYSIS USING KNOWLEDGE DISCOVERY FROM DATABASES

Knowledge Discovery from Databases (KDD) methodology is used to develop algorithms to estimate ceiling and visibility where direct observations are not available. As part of the KDD procedure, data mining of historical satellite, numerical model, and METAR data will uncover the data relationships used to estimate ceiling and visibility through satellite and/or numerical model data only.

Filling in the spatial or temporal holes of ceiling and visibility observations is a key challenge. Previous work (Bankert, et al, 2004) demonstrated the potential usefulness of cloud ceiling height estimation algorithms developed through the KDD approach. Data collection, data processing, data mining, and algorithm development are all part of the KDD methodology.

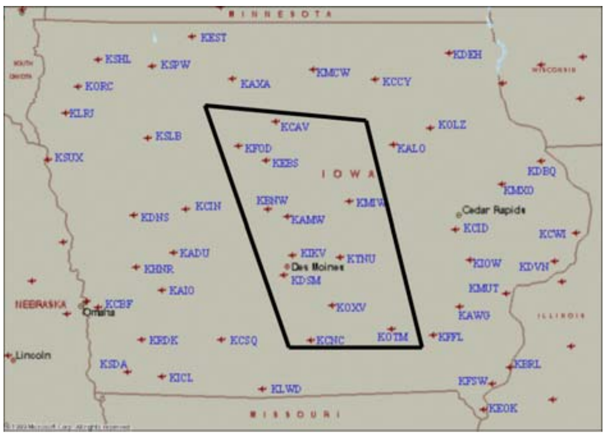


Figure 4. Iowa METAR station locations with testing set of METARS enclosed by black border. All other station data are used to train the specific algorithm through the data mining process.

Hourly Rapid Update Cycle (RUC) model and GOES-12 data from 2004 are being used to create a database of model and satellite parameters. Four geographic regions within the continental U.S. are being studied separately: Iowa, Northeast Texas, Gulf Coast (Texas to Florida), and Mid-Atlantic (Connecticut to Virginia) regions. Each of these areas provides a sufficient quantity and density of METAR stations to train and test the KDD-developed algorithms.

Each record in the database is a set of RUC, GOES-12, and METAR parameter values for a given time at a selected METAR location. A list of RUC parameters can be found in Appendix A. GOES-12 parameters include spectral channel (1 visible and 4 infrared channels) and channel-differencing data. In the near future cloud property parameters computed from a combination of GOES-12 and RUC data will be added.

To evaluate the performance of these KDD-generated algorithms, the data is split into training and testing sets for a given region with all records for a specific METAR station in either the training or testing set. The cloud ceiling algorithm (developed from the training set) is a three-step procedure to, ultimately, identify and estimate the height of low cloud ceilings at a specific location:

Step 1: Yes/No Ceiling Classification
If ceiling exists (yes class), proceed to step 2.

Step 2: Low/High Ceiling Classification
If ceiling is low (below 1000 m), proceed to step 3.

Step 3: Compute Ceiling Height.

Similarly, the visibility algorithm is currently a two-step process to identify and estimate low visibilities at the surface:

Step 1: Low/High Visibility Classification.
If visibility is low (less than 5 mi), proceed to step 2

Step 2: Compute Visibility Distance.

As mentioned earlier, data from selected METAR stations are used for testing and not involved in the data mining – development of the algorithm - for each region. As an example, the selected Iowa stations are shown in Figure 4. Each of the three steps within the cloud ceiling algorithm can be evaluated. Day and night records are considered separately.

Figure 5 is a graph representing the performance for each algorithm in Step 1 (yes/no ceiling) using Iowa daytime data. Three algorithms are compared: GOES-only, RUC-only, and GOES/RUC combined. As expected, a combination of satellite and numerical model data provide the best performance; however, algorithms using single data source (GOES-12 or RUC) also performed reasonably well.

Performance measures (Iowa daytime data) for Step 2 – low/high ceiling classification are displayed in Figure 6. POD, FAR, CSI, TSS are measured with respect to low ceilings – which is of most interest in the aviation community. These results demonstrate fairly high accuracy and skill for the RUC-only algorithm estimating whether a

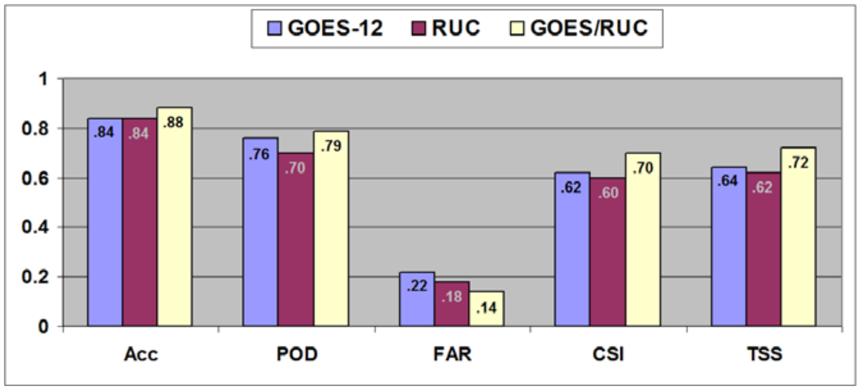


Figure 5: *For each data set – performance measures for cloud ceiling yes/no (Step 1) for Iowa daytime algorithm. Acc: Overall Accuracy, POD: Probability of Detection, FAR: False Alarm Ratio, CSI: Critical Success Index, TSS: True Skill Score.*

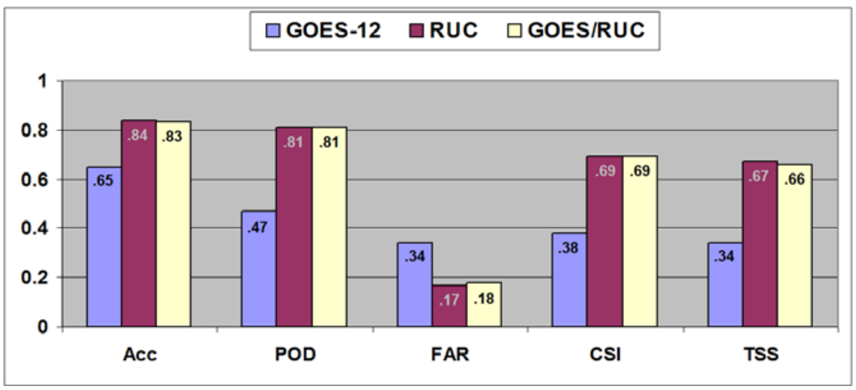


Figure 6: *For each data set – performance measures for cloud ceiling low/high (Step 2) for Iowa daytime algorithm. Acc: Overall Accuracy, POD: Probability of Detection, FAR: False Alarm Ratio, CSI: Critical Success Index, TSS: True Skill Score.*

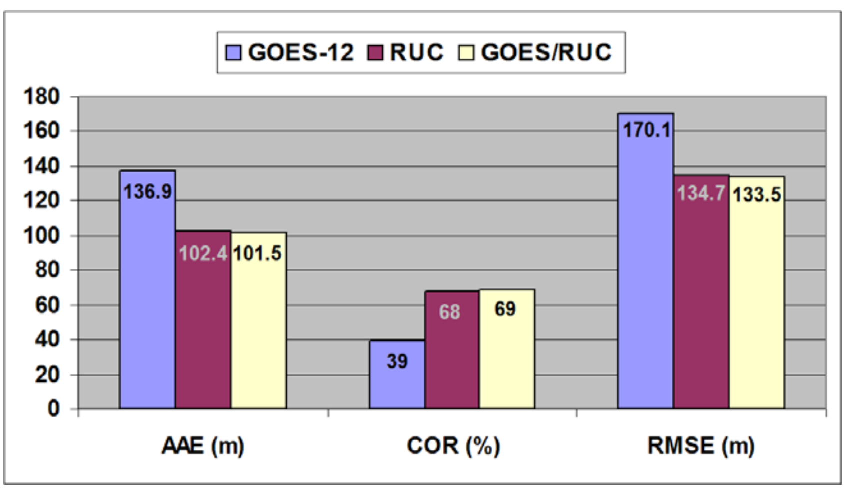


Figure 7: *For each data set – performance measures for low ceiling height estimation (Step 3) for Iowa daytime algorithm. AAE: Average Absolute Error (meters), COR: Correlation Coefficient (%), RMSE: Root Mean Square Error (meters).*

ceiling is low or high. However, the addition of GOES-12 data provides no positive contribution.

Figure 7 is a graphical display of the performance measures for Step 3 – low ceiling height estimation. A combination of data types (GOES-12 and RUC) produced best results, but GOES-12 provided only a small contribution.

Similar results for all three steps were obtained for Iowa night data. Further analysis of these results and additional cloud ceiling studies will be performed for the other three regions. Visibility algorithms for all four regions will also be produced and their performance analyzed.

5. FORECAST PRODUCT

The premise underlying the functionality of the NCV forecast system is that five or more independent forecast modules (numerical models, NWS forecaster guidance products and observations-based methods) can be adaptively combined to yield an integrated forecast that is significantly more skillful than any of its input modules acting alone. The system to do this has been developed and is in testing. Forecast output quantities are ceiling, visibility, flight category and terrain obscuration at lead times from 2 to 10 h. Forecasts are made for each of approximately 1524 METAR sites in the ConUS. These site-specific forecasts are then generalized to the surrounding area through nearest neighbor interpolation. Terrain effects influencing ceiling height (above ground level) are taken into account.

A schematic overview of the forecast system is given in Fig. 8. A key step in the forecast process is the selection of an optimum forecast for a given METAR site. That selection process, which is evolving as this is written, utilizes past performance statistics for each input module, each target variable, each forecast lead time, each forecast initiation time and each forecast site. The performance statistics are automatically generated for a variety of past periods ranging from as little as a few days to as long as several months. Automated processing utilizes the statistics to choose an optimum forecast module, and the output of that module then comprises the forecast for that site, target parameter and lead time. The forecast modules in use and the number sites at which each is available is summarized in Table 1 below.

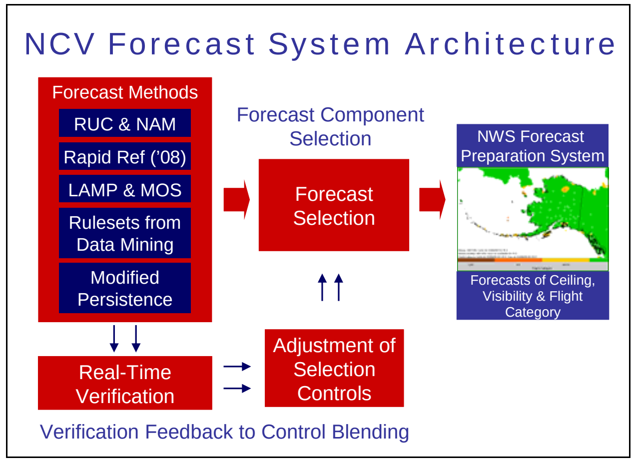
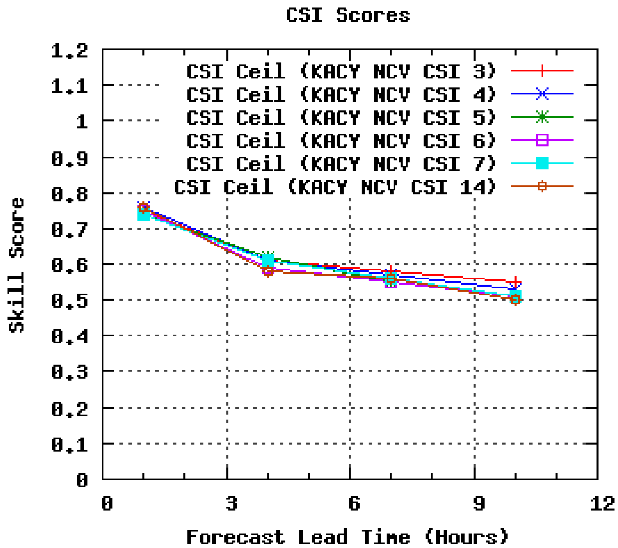


Figure 8. Schematic view of the NCV forecast processing system. Input forecast modules are shown in blue at left. Output products are at lower right in red.

Early testing results indicate that the adaptive selection approach yields forecasts whose skill typically exceeds that of the operational forecast modules used as input. Ceiling forecast test results for the November 2004 through April 2005 period for Atlantic City (KACY) are shown in Fig. 9. While significantly more results will be available at the time of the conference (January, 2006), the current data from KACY and other sites are highly encouraging in that they are competitive with persistence at 1 h lead time and demonstrate improvement over the available forecast guidance products used as input at 4, 7 and 10 h lead times.

Table 1. Forecast Inputs to NCV ConUS Forecast System

Forecast Modules in Use	Number of Sites (out of 1524 total)
RUC 13 Persistence GFS MOS	1524
RUC 13 Persistence GFS MOS NGM LAMP	~ 950
RUC 13 Persistence GFS MOS NGM LAMP Data Mining Rulesets	51



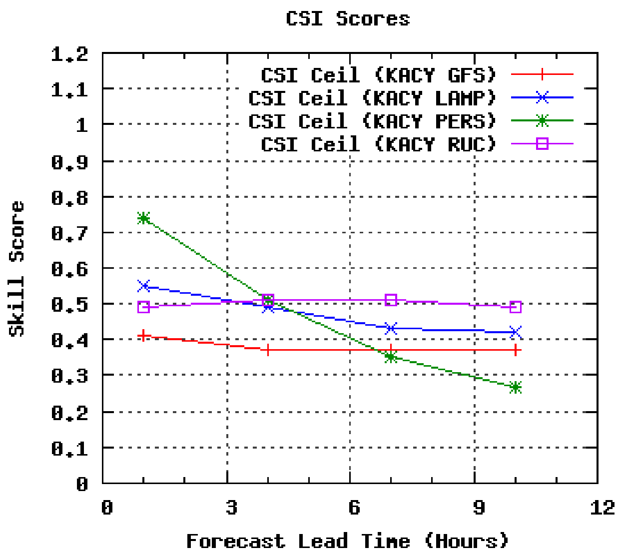


Figure 9. *<u>Top</u>: Critical success index (CSI) for NCV forecasts of ceiling at Atlantic City over the period November 2004 through April 2005. Individual curves apply to tests of varying forecast control settings. <u>Bottom</u>: CSI scores as above for four operational forecast guidance products (GFS MOS, ETA LAMP, Persistence and RUC 13). Scores for the NCV system (top) show significant improvement over the guidance products.*

6. SUMMARY

This paper outlines the data sources and methods used within the automated ConUS analysis and forecasting systems currently under development within the FAA's National Ceiling and Visibility development team. Encouraging results are obtained in work toward use of KDD methods for estimation of ceiling height in gap areas within the real-time analysis system and in the use of a new forecast selection scheme to guide synthesis of the forecast product from 2 to 10 h using a variety of operational guidance products as input. These and other techniques are used to improve the performance of the analysis and forecast systems, which are targeted toward production of improved automated forecast input grids and forecaster guidance for operational use.

Additional results will be available at the time of the conference.

Acknowledgements

This research is in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA.

REFERENCES

Bankert, R.L., M. Hadjimichael, A.P. Kuciauskas, W.T. Thompson, and K. Richardson, 2004: Remote cloud ceiling assessment using data-mining methods. J. Appl. Meteor., **43**, 1929-1946.

APPENDIX A

Table A.1. RUC output fields used in KDD processing. Parameters are stored in hourly, location dependent database records and used in finding relationships between them and observed cloud ceiling (C) and visibility (V).

RH at lowest model level (C&V)
Dewpoint temperature at lowest model level (C&V)
Empirical ceiling (using lowest level T and Td) (C)
LCL (C)
Temperature at lowest model level (C&V)
u-wind component at lowest model level (C&V)
v-wind component at lowest model level (C&V)
Total wind speed at lowest model level (C&V)
Sensible heat flux at surface (C&V)
Latent heat flux at surface (C&V)
Bowen ratio (C&V)
Height of lowest model level with RH > 90% (C)
Terrain height (C&V)
Cloud base height (C)
Cloud top height (C&V)
Cloud top temperature (C&V)
Cloud / No cloud (yes/no) (C&V)
Snow cover depth (V)
Snow cover / No snow cover (yes/no) (C&V)
... Continued on next page..

Height model level with max vapor mixing ratio (C)
u-wind component in 0-30 mb AGL layer (C&V)
v-wind component in 0-30 mb AGL layer (C&V)
Total wind speed in 0-30 mb AGL layer (C&V)
Average RH in lowest 150 mb (C&V)
Temp diff (top and bottom) in lowest 150 mb (C&V)
Precipitable water (C)
Precipitable water ratio (C&V)
Richardson number (lowest 4 levels) (C&V)
PBL depth (C&V)
Vertically-averaged TKE in PBL (C&V)
Stoelinga-Warner ceiling (C)
Soil temperature (C&V)
Net longwave radiation (C&V)
Net shortwave radiation (C&V)
Stoelinga-Warner visibility (V)
Categorical rain (V)
Categorical snow (V)
Categorical freezing rain (V)
Categorical ice pellets (V)
Ground moisture availability (C&V)