The FAA Aviation Weather Research Program Quality Assessment Product Development Team

Jennifer Luppens Mahoney
NOAA Research-Earth System Research Laboratory/Global Systems Division

Barbara Brown
National Center for Atmospheric Research/Research Applications Laboratory

1. Introduction

The FAA Aviation Weather Research Program is engaged in research that is directed toward improving weather forecasts for the aviation community. Much of this research, in the form of automated algorithms to predict aviation weather phenomena such as icing and turbulence, is transferred from research laboratories to the National Weather Service (NWS) through the Aviation Weather Technology Transfer (AWTT) process. The main purpose of the Quality Assessment Product Development Team (QA PDT) is to objectively evaluate the forecasting performance of these automated algorithms and ensure that the algorithms provide improved forecasting capabilities.

Since its inception 4 years ago, the QA PDT staff has supported the transition of the National Convective Forecast Product (NCWF), Graphical Turbulence Guidance (GTG), Current Icing Potential (CIP), and the Forecast Icing Potential (FIP) to full operational status at the NWS Aviation Weather Center (AWC) for use as forecast guidance to AWC forecasters.

Corresponding Author:
Jennifer Mahoney
NOAA-Research/ESRL/GSD
325 Broadway
Boulder, CO 80305
Jennifer.mahoney@noaa.gov

In support of the QA PDT evaluation responsibility, the PDT staff work diligently to advance verification methodologies to address the complexities of aviation weather forecasts. For instance, an effort to develop approaches that incorporate the operational significance or impact of an aviation forecast is underway. Also, new observation datasets, such those produced by satellites, are being investigated by the PDT staff for use in assessing the quality of global aviation forecasts.

This paper summarizes the responsibilities of the QA PDT and highlights some of the recent accomplishments. The report is organized as follows: Section 2 summarizes the methodology used by the PDT to assess the quality of forecasts transitioning through the AWTT process; Section 3 highlights many of the PDT's recent accomplishments; and Section 4 outlines the future goals of the PDT.

2. Methodology

The QA PDT's approach for evaluating the algorithms' performance is to conduct intensive, independent, assessment exercises. Each algorithm transitioning through the AWTT process is evaluated for a three-month period, with its performance compared to an operational standard. The results are summarized in a written report and provided to the AWTT

Technical Review Panel (TRP) as input to the transition process.

The QA PDT establishes objective procedures for evaluating forecast quality and accuracy, utilizing and applying *in situ* observations, advanced weather observations, and measurements from remote-sensing instruments. These observations form the basis for the verification methodologies and estimates of forecast accuracy. Often, the available observational datasets do not directly represent the forecast attributes, so inferences and comparisons among a variety of observational datasets are used to establish forecast quality. The statistical verification methodologies used in these studies are designed to represent the operational use of the forecasts. To evaluate the aviation weather forecasts and algorithms, the QA PDT often must develop new advanced verification methodologies that provide a better measure of forecast accuracy than the standard approaches.

3. Recent Accomplishments

Many of the accomplishments summarized in this Section have supported the transition of AWRP algorithms to operations at the NWS. In addition, many of the new verification techniques and capabilities developed by the QA PDT staff have been implemented into the Real Time Verification System (RTVS; Mahoney et al 2002?) for long-term tracking of forecast quality and for short-term intercomparison evaluations.

The QA PDT staff completed evaluations for the National Ceiling and Visibility Analysis (NCV-A; reference) and the Oceanic Weather Cloud Top Height (CTOP; Quality Assessment PDT 2005), and to support the transition of the Current Icing Potential (CIP) to a 20-km grid (reference). The results were summarized in written reports and provided to the AWTT Technical Review Panel as supporting evidence regarding the scientific validity of the products. In addition, the QA PDT completed a real-time evaluation of the National Convective Weather Forecast (NCWF) product, version 2, and continues to prepare for FY06 evaluations of the NCV Forecast; the CIP and Forecast Icing Potential (FIP) severity forecasts; and the Graphical Turbulence Guidance product, version 2.

In support of these AWTT activities, the QA PDT developed verification approaches that made use of satellite-based observations and satellite-derived algorithms to infer forecast quality over oceanic domains. Figure 1 shows one result from a comparison between the Oceanic Weather PDT (reference) Cloud Top Height (CTOP) algorithm and a cloud top pressure product produced by NOAA National Environmental Satellite Data and Information Service (NESDIS). The intent of the comparison was to infer forecast quality for the CTOP algorithm as it was compared to other similar operational products. The results indicated that for opaque deep convective clouds that could be an impact to aviation, the bias comparing the two products was small for 3 out of 4 domains (Fig. 1a). However, the bias between the two products was quite large for clouds of less opacity (Fig. 1b). The QA PDT concluded that the CTOP algorithm was quite accurate at predicting the height of deep convective clouds, but was not very accurate at predicting the height of other cloud types (see reference for further discussion or RTVS for on-going verification statistics; http://www-ad.fsl.noaa.gov/fvb/rtvs/, link CTOP).

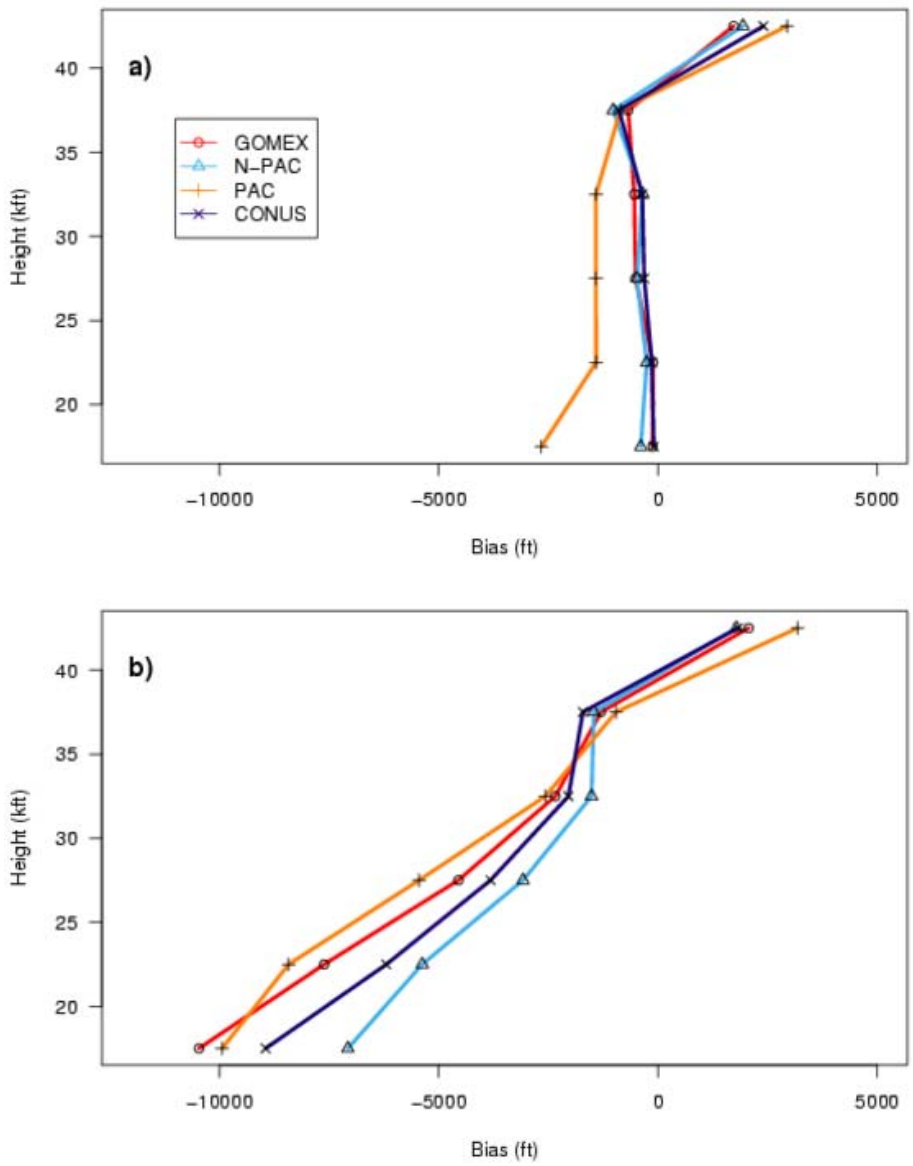


Figure 1. *Height series of bias values stratified by CTOP region. Results are presented for (a) opaque (ECA >= 95%) and (b) non-opaque (ECA < 95%). (Quality Assessment PDT 2005)*

In a second evaluation, the QA PDT assessed the quality of the National Ceiling and Visibility Analysis (NCV-A) algorithm produced by the National Ceiling and Visibility PDT using a cross-validation technique (reference). Figure 2 shows the distribution of the ceiling height errors produced by the NCV-A at locations between the METAR stations that were used to create the analysis, as compared to the ceiling height observed by the surface observation. The great majority of the errors shown in Fig. 2 are small,

typically less than one thousand feet. However, the negative

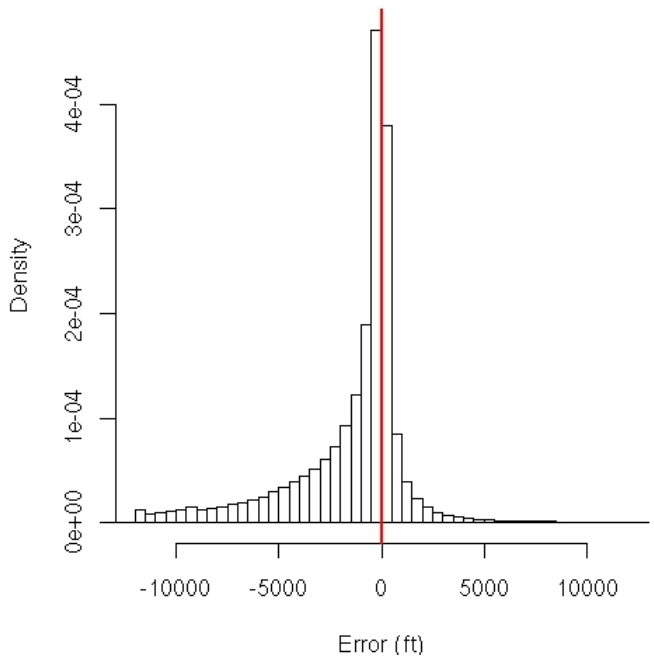


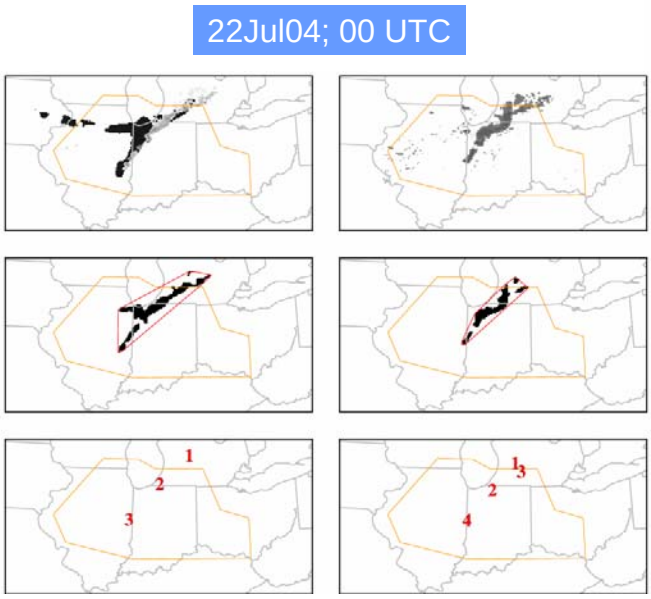
Figure 2. *Histogram showing error in ceiling height (METAR-NCV-A). (Fowler et. al, 2005)*

skew in the histogram indicates that the NCV-A is more likely to indicate that the ceiling is too high rather than too low. In other words, the NCV-A product is somewhat biased toward higher ceilings than are observed. The QA PDT concluded that the NCV-A visibility field closely matched the observed METAR visibility at all levels and that the NCV-A ceilings matched well with the METAR ceilings, especially when ceilings were unlimited or below 10K ft (see reference for further discussion of the results).

In addition, the QA PDT is developing object-based assessment techniques that provide measures of forecast performance in terms of errors in location, timing, and other user-relevant attributes. Figure 3 shows an example of an application of this approach to output of the Autonowcaster (a 1-h convective forecasting system, developed by the Convective Weather Product Development Team). This verification approach evaluates various attributes of the forecasts (e.g., forecast size, shape, location) as a complement to the standard verification statistics (e.g., POD, FAR, CSI). In the case shown, a single composite forecast object was identified, along with a single composite observed object. Applications of this approach are expected to lead to evaluations of convective forecasts that provide information that is more useful to operational users of the forecasts.

4. Future Goals

In the future, the QA PDT will continue AWTT evaluations for the GTG2, NCWF2, NCV Forecast algorithm, Oceanic Weather convective and cloud top forecast products, and CIP and FIP severity products. The QA PDT plans to collaborate with the Developmental Testbed Center (DTC) to evaluate the impact of numerical model enhancements on the forecast performance of the AWRP developed algorithms. Finally, the QA PDT plans to: 1) continue its research on object-based verification approaches and 2) incorporate operational considerations into the forecast verification methodologies.



	Diff
Centroid X	-14.92
Centroid Y	2.41
Axis angle	-7
Area	5931
Inten 10	-3.4
Inten 25	-8
Inten 50	-8.9
Inten 75	-3.3
Inten 90	1.2
Ugliness	0.8
Intersection	0.22
SD	0.78
POD	0.50
FAR	0.67
Bias	1.49
CSI	0.25

Figure 3. Example of new object-based forecast evaluation methodology, applied to convective nowcasts produced by the NCAR Autonowcaster.

References

Fowler, Tressa, Anne Holmes, Matthew J. Pocernich, Jamie T. Braid, Barbara G. Brown, Jennifer Luppens Mahoney, 2005: National Ceiling and Visibility Analysis Product Quality Assessment Report, FAA Aviation Weather Research Program Quality Assessment Product Development Team (available from author, Tressa@ucar.edu).

Mahoney, J.L., J. K. Henderson, B.G. Brown, J.E. Hart, A. Loughe, C. Fischer, and B. Sigren, 2002: The Real-Time Verification System (RTVS) and its application to aviation weather forecasts. *Preprints, 10th Conference on Aviation, Range, and Aerospace Meteorology*, 13-16 May, Portland, OR, American Meteorological Society (Boston), 323-326.

Quality Assessment PDT, contributors; Brown, Barbara G., Eric Gilleland , Robert Hueftle, Lacey Holland, Mike Kay, Sean Madine, Jennifer Mahoney, Agnes Takacs, and Jamie Wolff, 2005: *Oceanic Weather Cloud-Top Height.* Quality Assessment Report, FAA Aviation Weather Research Program Quality Assessment Product Development Team (available from author, bgb@ucar.edu).