11.1

THE QUALITY ASSESSMENT RESEARCH TEAM (QART) AND NEXTGEN ALIGNMENT †

Jennifer Luppens Mahoney*

NOAA/ESRL/GSD, Boulder, CO

1. INTRODUCTION

The Quality Assessment Research Team (QART) is one of the Research Teams supported by the FAA Aviation Weather Research Program (AWRP). Unlike most of the AWRP aviation weather research teams (e.g., Inflight Icing, Turbulence, etc.), the QART's main function is to provide an independent quality assessment of aviation forecast products that are being transitioned to NWS and FAA operations. Three main activities support this function: 1) the development of forecast assessment techniques that bring together forecast and observation information within a user-specific context utilizing observation datasets that are independent of the forecast product being evaluated, 2) the development of verification tools, specifically the Real-Time Verification System (Mahoney et al. 2002; RTVS), and the Network-Enabled Verification Service (Matheson et al. 2008; NEVS) that provide the infrastructure for QA assessment analyses, an automated verification capability that allows feedback to forecasters, decision makers, managers, and automated decision support tools, and a historical record of aviation forecast performance, and 3) the transition of QA technologies in NWS operations.

The purpose for this paper is to provide a programmatic summary of the main research and development activities underway within the QART. Many of the activities presented here are state of the art and offer significant advances in forecast evaluation techniques for aviation and collaborative verification services. In Section 2, the independent QART evaluation process is presented. Section 3 highlights the user-specific verification techniques that are the foundation for the forecast assessments, while in Section 4, a network-enabled verification service (NEVS), needed to support the QA assessment and future NextGen, is presented. Section 5 summarizes the status of the RTVS transition to NWS operations, and a summary and future efforts are provided in Section 6.

2. INDEPENDENT QART EVALUATION PROCESS

The main QART function is to provide an independent quality assessment of aviation forecast products that are transitioning from a research laboratory environment to fully supporting FAA operations. Figure 1 shows the process for transitioning a forecast product to operations. Within this process, two key decision points are relevant to QART activities (i.e., D3 and D4). Prior to each decision point, the QART assesses the forecast quality of the aviation forecast product.

During the planning and development phase of the verification process and prior to the decision points, the verification plan is being developed. Questions that are being addressed during the verification planning include: the concepts for how the product will be used in operations, the primary user groups that will integrate the product into their operations, and the graphical representation of the product to the end users. These operational attributes are primary drivers for how the assessment evaluation is structured. In addition to the drivers, other important considerations that dictate the evaluation procedures include: the supplemental nature of the product with respect to the operational standard, the expected replacement of an operational standard by the experimental product, and

†This work is sponsored by the National Weather Service and the Federal Aviation Administration.

*Corresponding author address: Jennifer Mahoney
325 Broadway, Boulder, CO 80305-3328
e-mail: Jennifer.Mahoney@noaa.gov

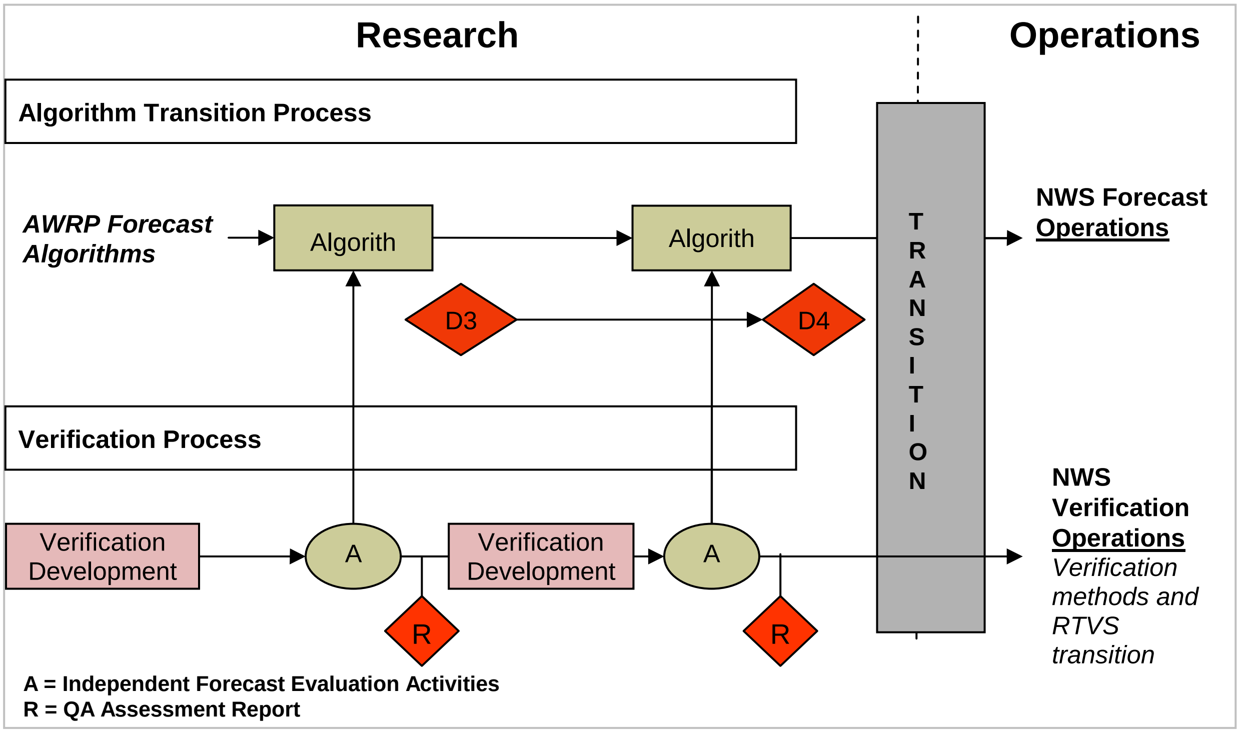

Figure 1. *Diagram describing the aviation product transition process and the supporting QART verification process. For each decision point (D3 and D4), an independent assessment of the forecast product is completed by the QART. At the end of the process, the verification techniques and verification system (RTVS) are transitioned to NWS for on-going operation.*

the overall strengths and weakness of the experimental product itself (e.g., accuracy, timeliness, update rates etc).

Once the planning phase is completed, the evaluation of the product begins. First, implementation of the experimental components determined during the planning phase are developed and integrated into RTVS (and NEVS in the future) to support the analyses phase. Output from the forecast product is collected from the Research Teams and incorporated into RTVS. Statistical results are generated and a detailed analysis of the results within a user-specific context is completed. The findings are summarized in a written report and provided to a Technical Review Panel. The Technical Review Panel evaluates the work (along with other information) and provides recommendations to a joint FAA and NWS transition team for product transition.

Following the evaluation, the verification techniques and the operational components of the RTVS that were used to support the evaluation are transitioned to the NWS so that an official historical forecast performance record can be generated. Information regarding these transition activities is summarized in Section 5.

In 2007, products that were evaluated by the QART for transition to NWS operations included: the Graphical Turbulence Guidance product (GTG2; Kay et al. 2006), the National Ceiling and Visibility Analysis (NCVA; Braid et al. 2007), the Forecast Icing Potential (FIP D3; Chapman and Mahoney 2007). In 2008, products that will be evaluated include: FIP (D4), GTG3, Volcanic Ash, National Ceiling and Visibility Forecast, and Flight Level Winds. All publications are located at the following web site: http://rtvs.noaa.gov, link publications.

The RTVS and the NEVS are the tools that are used to perform the QART assessments and provide the backbone for integrating new datasets (both forecasts and observations), managing the relationships between the databases through a relational database management system, incorporating and integrating

1115 UTC Telecon

Product	INITIAL Time	ISSUE/ AVAILABLE Time	VALID Time MODEL Period (LEAD Period)			
			1300	1500	1700	1900
CCFP - Prelim		1000	1300 (-) (+3)	1500 (-) (+5)	1700 (-) (+7)	
CCFP - Final		1100	1300 (-) (+2)	1500 (-) (+4)	1700 (-) (+6)	
NAM – Refl.	0600	0805	1200 (6) (+4)	1500 (9) (+7)	1800 (12) (+10)	1800 (12) (+10)
RCPF	0900	1100	1300 (4) (+2)	1500 (6) (+4)	1700 (8) (+6)	1900 (10) (+8)
RUC – Refl.	0900	1100	1300 (4) (+2)	1500 (6) (+4)	1700 (8) (+6)	1900 (10) (+8)

Table 1. *Flight planning outlook times are represented by the bold sub-column headings in the main right most column. The lead periods are presented by the green numbers in the sub-columns. The difference in time between the green and blue numbers is the model latency.*

non-meteorological data for user-specific stratifications, and access to and display of the statistical information. The transition of RTVS to NWS is underway, while the research and development of NEVS is beginning to emerge. In the near term, RTVS will continue to provide the main support for QA assessments, while in the longer term, NEVS will replace the RTVS and provide a full data integration capability and a network-enabled way to access and distribute verification information.

3. USER-SPECIFIC VERIFICATION TECHNIQUES

Verification techniques are the foundation for all QART assessments. Therefore, with emphasis on aviation forecast application to FAA planning decisions, the QART is transforming traditional meteorological verification approaches and techniques into a new generation of assessment techniques that are driven by operational aviation constraints and key operational decision points. Two efforts within the QART to develop and test user-specific verification concepts are underway.

First, a convective forecast evaluation (Kay et al. 2007 and 2008), conducted during the summer of 2007, assessed the quality of five convective forecasts with respect to their use in air traffic flow planning. Rather than using traditional verification approaches for verifying forecast quality with respect to meteorological domains or regions, the forecast products were evaluated with respect to key operational strategic planning criteria. Specifically, the forecast quality was evaluated at strategic traffic flow planning periods, with respect to air traffic sectors with high convective coverage, and results were stratified by a weather impact traffic index. This approach gave rise to results that were operationally meaningful for strategic traffic flow management. The relevant contributions to the user-specific approach applied in the convective evaluation are described in the following figures and table. The statistical results from the evaluation are provided in Kay et al. (2007).

Table 1 shows the mappings that were used in the evaluation between the traditional forecast issuances and lead times to strategic flight planning decision periods. For instance, at the 1300UTC flight planning outlook time, the forecasts that were available for decision making were CCFP Preliminary 3-h lead, the CCFP Final 2-h lead, the NAM reflectivity 4-h lead, the RCFP 2-h lead, and the RUC reflectivity 2-h lead. From a traditional verification perspective, comparing the quality of these forecasts for different lead times provides limited meaningful information. However, from a traffic flow planning perspective this verification information is significant since decisions are based only on the available forecast information regardless of

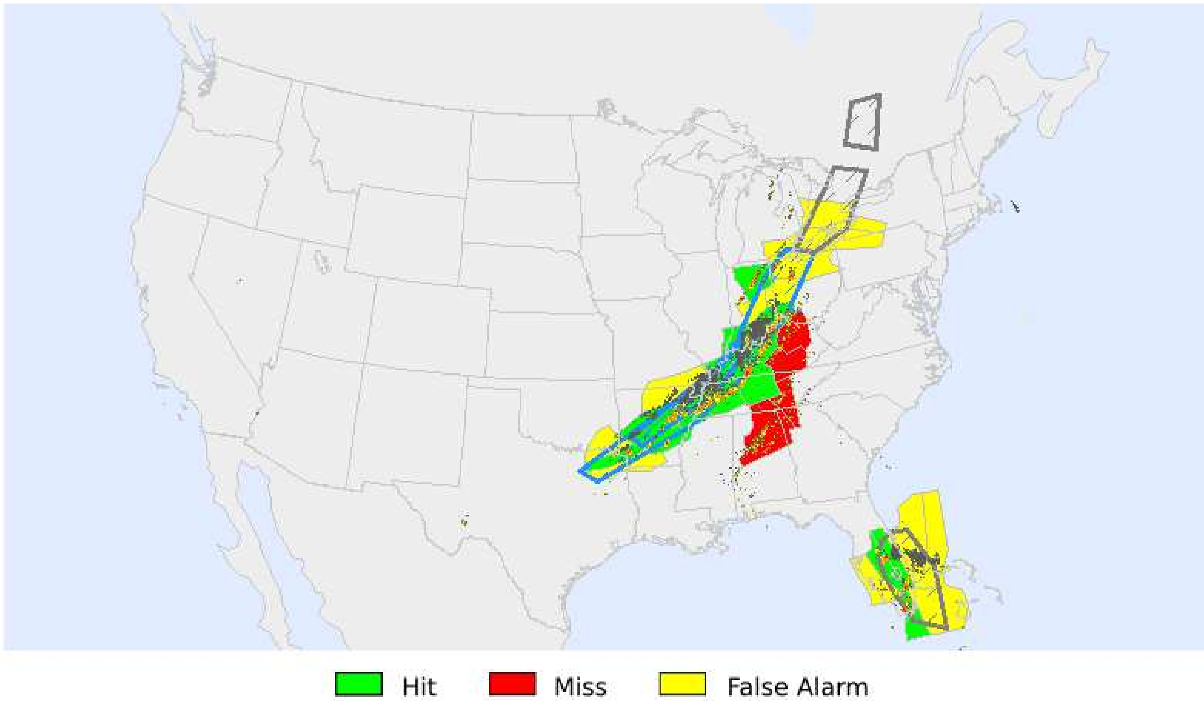

Figure 2. *Sector-based verification of the 2-h CCFP Final forecast from 8 June 2007 issued at 1500 UTC. NCWD observations shown as well. Impacted sectors are color-coded to depict the verification results (from Kay et al. 2007).*

whether the forecast lead times and valid times are in alignment between different forecast products.

Figure 2 represents the flight sector verification approach that was used during the summer convective evaluation. Convective coverage in air traffic sectors was a useful measure for evaluating convective forecast quality because the sectors represented the air space volumes that are used for strategic air traffic flow planning. Figure 2 shows the traditional metrics at the flight sector resolution. For instance, the green sectors indicate good correspondence between the convective forecast and high convective coverage within a sector (i.e., high coverage typically leads to high traffic impact, particularly over the eastern U.S.). The red indicates the sectors that contained a high level of convective coverage (and possible traffic impact), but did not contain a relevant forecast with similar convective coverage.

Figure 3 shows a proxy for the Weather Impact Traffic Index (WITI; Klein 2006) for each day during the summer of 2007. This metric has been adopted by the FAA to measure the health of the National Air Space (NAS) as a function of convective weather coverage. This metric was used in the summer evaluation to stratify the verification scores for days when the NAS impact due to convective weather was large.

The convective coverage, along with the proxy-WITI metric, was used to define operationally relevant regimes as shown in Figure 4. These operationally relevant regimes were used to characterize the quality of the 5 convective forecasts during the summer evaluation. Forecast quality and accuracy within the high coverage, very high impact and the high coverage, high impact regimes are of particular interest to operational flight planning.

A second area under development within the QART with a user-specific focus is the ability to assess skill in forecasting the **onset** of weather events that have a high impact on air traffic flow and safety. For instance, a new lead-time metric is being developed for assessing skill in forecasting the onset of significant

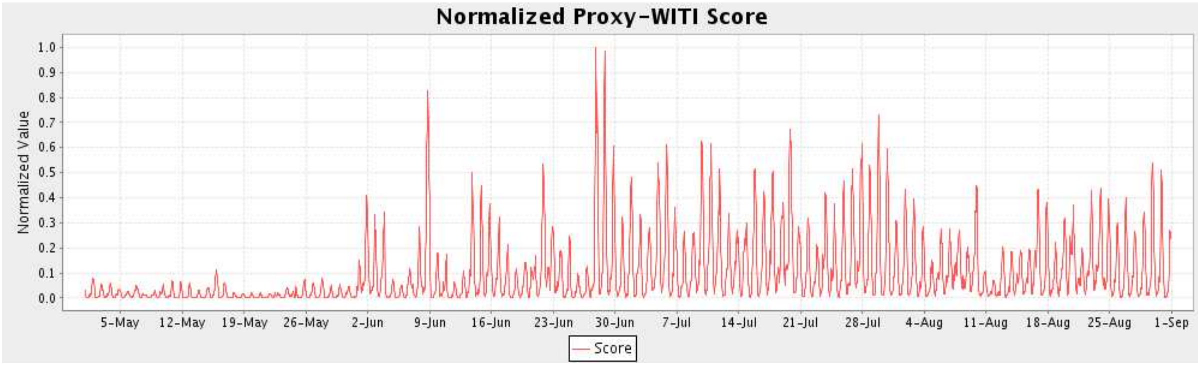

Figure 3. *Normalized proxy WITI score per day from early May to mid Sept 2007 within the NAS.*

ceiling and visibility events (i.e., Instrument Flight Rules events; IFR events) that would impact air travel at airports. The method compares the onset of observed IFR events (via METARs) with the onset from NWS-issued Terminal Aerodrome Forecasts (TAFs) yielding a measure of forecast quality that is directly related to airport strategic planning and operations. Figure 5 shows an example of the lead-time event viewer. This tool is used to visibly show the relationships between the forecast and the observed weather events and the associated lead time. Loughe et al. (2008) describes in detail this lead time verification approach and the tools used through the Real-Time Verification System (RTVS) to analyze the verification results by station, region, weather category, forecast issuance hour, or other TAF attributes. This technique will be used to evaluate ceiling and visibility forecast products transitioning to NWS operations and will be provided to NWS forecasters as feedback into the TAF performance and improvement.

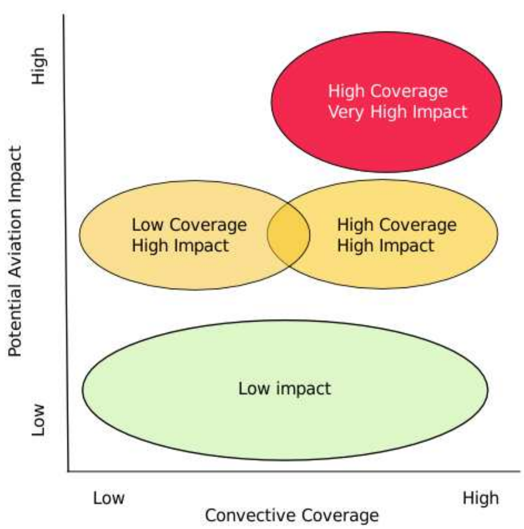

Figure. 4. *Schematic depiction of regimes that are of differing importance to aviation planning as a function of location of convection (as measured by potential aviation impact) and overall convective coverage over CONUS (Kay et al. 2007).*

4. NETWORK-ENABLED VERIFICATION SERVICE (NEVS)

The Network-Enabled Verification Service (NEVS) is the evaluation tool that brings the user-specific verification capabilities, previously discussed, together in a verification service. As part of the QART alignment toward NexGen, the network-enabled verification service would provide the right verification information to the right users at the right time. To accomplish this effort, the QART staff is taking advantage of 10 years of experience in building verification systems (through the development of RTVS) to construct a Network-Enable Verification Service (NEVS) that can become an integral part of the NextGen System Wide Information Managment (SWIM) infrastructure of the future. The emphasis of NEVS is on collaborative data integration, forecast comparison, detailed statistical analysis, and network-enabled dissemination of specific verification information to a variety of end users (including automated decision support tools and human

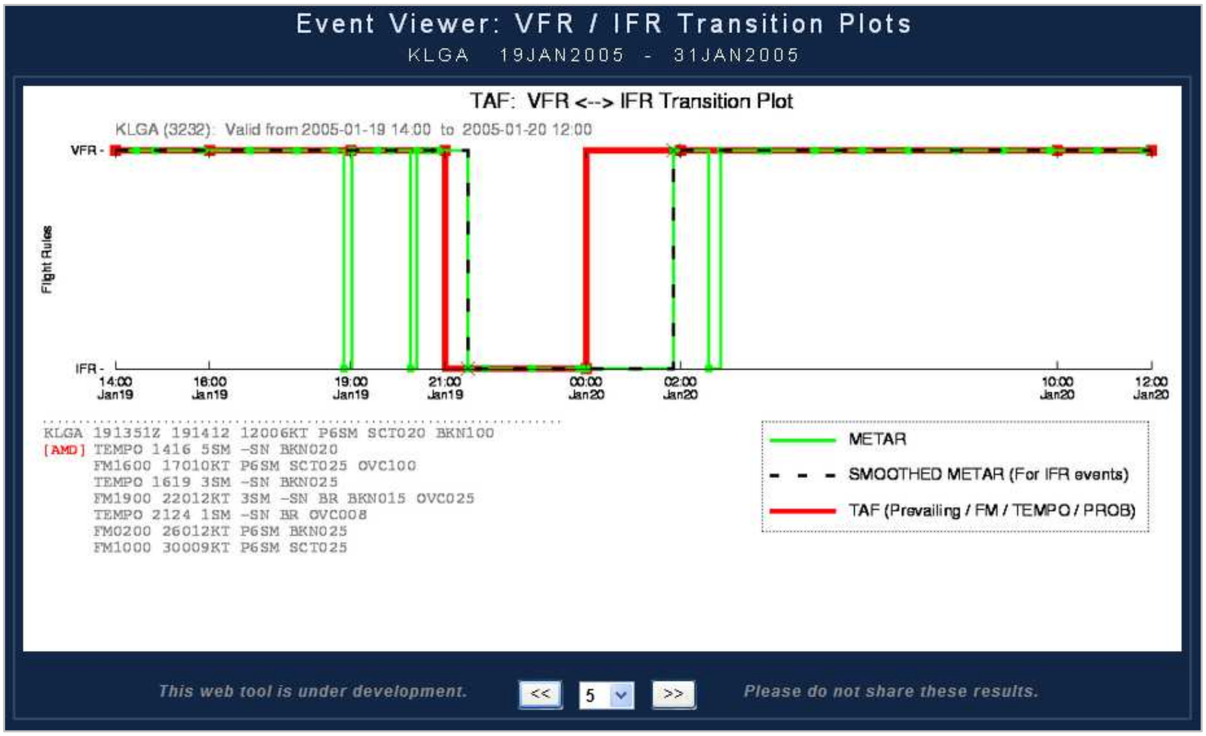

Figure 5. *TAF lead-time metric event viewer. Green lines represent transition periods between observed IFR and VFR events. Red lines represent the transition periods as forecast by the TAF.*

forecasters). NEVS takes advantage of relational database management systems and web-based technologies for access to, and dissemination of ,statistical information. The NEVS framework will help to unify resources from various organizations, will allow for data sharing and rapid transition of verification information to automated decision support tools and human forecast processes, and will ultimately provide a more complete assessment of forecast quality for user-specific applications. Figures 6 and 7 show the graphical user interface to the NEVS proof of concept. Figure 6 highlights the generation of statistical metrics that are stratified for operational meaningfulness (i.e., strategic planning periods, weather traffic impact (WITI), and effective valid time. The traditional grid to grid verification approaches and the newly developed sector approach are also introduced within NEVS. Figure 7 shows the probabilistic verification capabilities of NEVS as well as introducing an object-based verification technique capability.

5. TRANSITION ACTIVITIES

The goals for transitioning aviation verification capabilities to the NWS are to provide:

- An automated operational verification system that will make available a historical record of aviation forecast quality.
- Automated verification information to forecasters and end users in near real time to
- aid improvements in the operational forecast quality and end user decisions.
- A system infrastructure that would support the transition of new verification capabilities, such the new forecast lead time GPRA metric, and user-specific verification information.
- A verification infrastructure that will support future verification efforts within the NextGen environment and for evaluation of the 4-D aviation data cube.

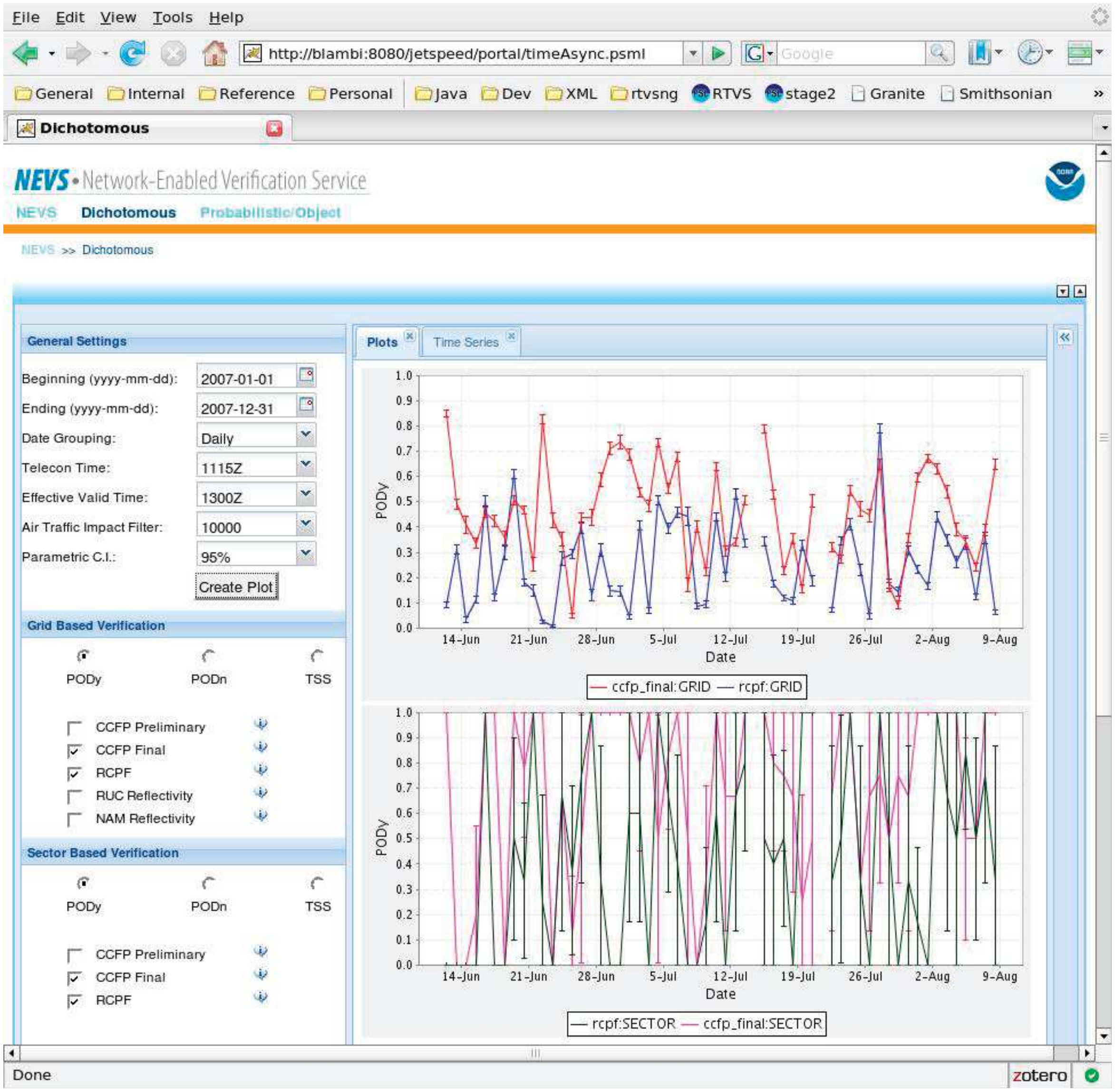

Figure 6. *NEVS Use Case I - Extending traditional verification applications. Traditional verification features include sub-setting of data based on date and time, intercomparison of multiple products, selection among various statistics for comparison, and confidence intervals on statistics of interest. In addition, NEVS allows for easy extension of the traditional verification paradigm, incorporating the concept of air traffic planning periods and effective valid times, equalizing and comparing products based on user application not forecast production attributes, filtering the data by a user domain attribute, in this case a weather/traffic impact measure, and the introduction of the sector based dichotomous statistical approach (Matheson et al. 2008).*

To accomplish these goals the QART is in the process of formally transitioning the operational aspects of the Real-Time Verification System (RTVS) to the NOAA Telecommunication Operations Center (TOC) with user support for the NWS application of RTVS provided by staff from the NWS Performance and Aviation Services Branches.

Current status of the transition includes: RTVS transition plan coordinated between NWS and OAR,g Gate 1 approval through the NWS OSIP transition process (https://osip.weather.gov), RTVS help desk established for addressing system user questions, and funding from FAA and NWS to support transition

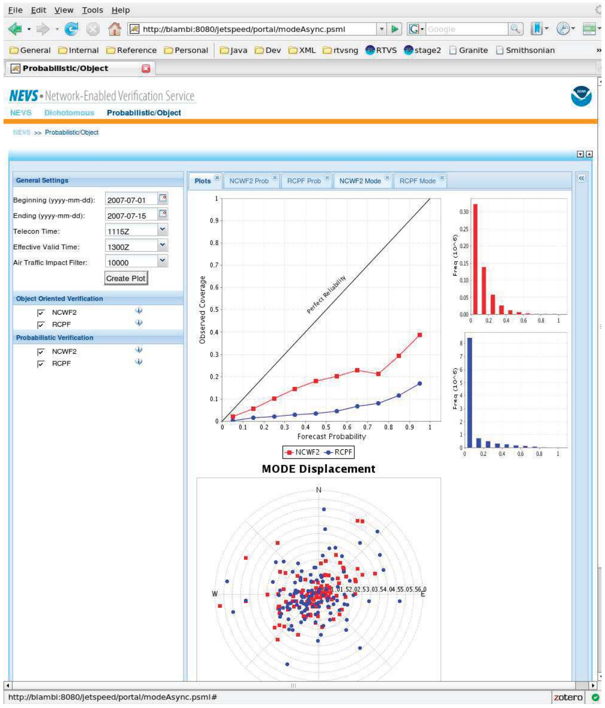

__Figure 7.__ NEVS Use Case II - Facilitating new verification applications. The NEVS concept can also be utilized to build emerging verification applications. This graphic presents probabilistic and object oriented verification methods in a unified view that ensures the datasets: conform to the user viewpoint, are event equalized, are filtered along the same user-specific criteria (Matheson et al. 2008).

activities. We expect a full transition of the RTVS framework to be completed by FY11 with verification technique and dataset upgrades continuing thereafter. By FY12, NEVS will begin to replace RTVS in NWS operations, so that the NWS is well positioned to move into the NextGen era.

6. SUMMARY AND FUTURE QART EFFORTS

The QART is an independent assessment group that provides support for aviation forecast product transition to NWS operations. Three main activities support this function, which include the development of user-specific verification techniques, NEVS development, and RTVS transition to NWS. These research and development activities are providing the mechanism for the QART to directly align with the goals of NextGen by providing user-specific verification information that is distributed through a network-enabled distribution service to automated decision support tools and end-user applications. Future efforts include a realization of NEVS to support FY08 convective evaluations, complete transition of RTVS to NWS, and enhancements to the user-specific verification techniques.

7. ACKNOWLEDGEMENTS

This research has been sponsored by the National Weather Service and 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 NWS or the FAA. The author would like to thank the QART staff (Mike Chapman (CIRES), Joan Hart (CIRES), Judy Henderson (NOAA), Mike Kay (CIRES), Brice Lambie (CIRES), Steven Lack (CIRES), Geary Layne (SRG), Andy Loughe (CIRES), Sean Madine (CIRA), Nick Matheson (CIRES), Missy Petty (CIRA), and Dan Schaffer (CIRA)) for their excellent work and dedication to these efforts.

8. REFERENCES

Braid, J.T., A. Holmes, M.J. Pocernich, L.D. Holland, P.A. Kucera, B.G. Brown, and J.L. Mahoney, 2007: Quality Assessment Report: National Ceiling and Visibility Analysis Product (D4). Submitted to FAA Aviation Weather Technology Transfer (AWTT) Technical Review Panel. Available at http://rtvs.noaa.gov, link publications.

Chapman, M.B., and J.L. Mahoney, 2007: Forecast Icing Product - Icing Probability: Quality Assessment Report. Submitted to Aviation Weather Technology Transfer (AWTT) Technical Review Panel. Available at http://rtvs.noaa.gov, link publications.

Kay, M.P., S. Madine, J.L. Mahoney, J.E. Hart, 2008: Aligning forecast verification with user-specific needs: An example for aviation. Preprints: *13th Conf.Aviation, Range, and Aerospace Meteor.*, New Orleans, LA.

Kay, M.P., S. Madine, J.L. Mahoney, J.E. Hart, 2007: The 2007 Convective Forecast Scientific Evaluation. Submitted to FAA System

Operations. Available at http://rtvs.noaa.gov, link publications, 33 pp.

Kay, M.P., J.K. Henderson, S.A. Krieger, J.L. Mahoney, L.D. Holland, and B.G. Brown, 2006: Quality Assessment Report - Graphical Turbulence Guidance (GTG) Version 2.3. Submitted to FAA Aviation Weather Technology Transfer (AWTT) Technical Review Panel. Available at http://rtvs.noaa.gov, link publications.

Klein, A, 2006.: Normalized Delay-and Cost-Derived National Airspace System Performance Indexes. *American Institute of Aeronautics and Astronautics*, 13pp.

Loughe, A. F., S. Madine, J.L. Mahoney, M. Graf, 2008: A lead-time metric for assessing skill in forecasting the onset of IFR events. Preprint: *13th Conf.Aviation, Range, and Aerospace Meteor.*, New Orleans, LA.

Matheson, N.D., M.A. Petty, D.S. Schaffer, S. Madine, and J. Mahoney, 2008: The Network-Enabled Verification Service (NEVS): Integrating data to support user-specific verification. Preprints: *13th Conf. Aviation, Range, and Aerospace Meteor.*, New Orleans, LA.

Mahoney, J.L., Judy K. Henderson, Barbara G. Brown, Joan E. Hart, Andrew Loughe, Christopher Fischer, and Beth Sigren, 2002: The Real-Time Verification System (RTVS) and its Application to Aviation Weather Forecast. Preprints: *10th Conference on Aviation, Range, and Aerospace Meteorology, 13-16 May, Portland, OR.*