5.4 VALIDATION OF FAA WARP SYSTEM RADAR MOSAIC GENERATION ALGORITHMS

Joseph Lang *
Unisys Weather Information Service, Kennett Square, Pennsylvania

James Stobie
SAIC, Washington, DC

Kerryaine Yarber
FAA, ATO-D, Washington, DC

1.0 INTRODUCTION

The FAA WARP program recently completed development and implementation of a set of improved radar mosaic generation algorithms. The Unisys Weather Information Services Group performed the work under sub-contract to the Harris Corporation, the prime contractor on the WARP program. This development effort, which began in March 2003, included an initial study phase to define and prototype the mosaic generation algorithms. A paper presented at 20th IIPS (Lang 2003) described the mosaic generation algorithm and the initial effort to validate the performance of the algorithm. Implementation was completed in August, 2004. At the time this paper is being prepared (October 2004), the System Acceptance Test has been successfully completed. Key site testing is scheduled for December 2004, with deployment in the first quarter 2005

The radar mosaic products generated by the WARP system are used as weather backgrounds on en-route air traffic controller aircraft situation display screens. Because of the mission critical nature of this application, the performance of the new algorithms must be formally validated before the new mosaic products can be displayed on the controller screens. While algorithm validation was part of the original scope of the overall algorithm development effort, early on in the validation effort a situation occurred that sharpened the focus on the validation effort and resulted in a more aggressive validation program. In the summer of 2003, deficiencies in the Nexrad AP-mitigated low layer composite reflectivity product (product 67) came to light that resulted in grave concerns in the controller community about the validity of the WARP mosaic products generated from this radar product. An emergency WARP software release was required to remove these mosaics from the controller display screens. There was a general perception that the product had been fielded without adequate validation of the Nexrad AP-mitigated product data. This event effectively raised the bar for the mosaic generation algorithm validation effort.

* Corresponding author address: Joseph C. Lang, Unisys Corporation, weather Information Services, 221 Gale Lane, Kennett Square, PA 19348; e-mail: joseph.lang@unisys.com

This paper describes the validation program conducted to validate the performance of the WARP mosaic generation algorithms. This validation effort, which was effectively conducted in parallel with the software development effort, required close cooperation between the Harris/Unisys development team, the WARP program office, controller organizations, and other FAA and NWS organizations.

2.0 PROJECT CHRONOLOGY

Figure 1 shows the algorithm validation activities in the context of the overall software development effort. Development began in March 2003. The contract included an initial four-month study phase for developing the optimal mosaic generation algorithms. In addition to the normal software development cycle activities and milestones (SRR, CDR, FAT, SAT, etc.) the contract provided for two Technical Interchange Meetings (TIM). The purpose of the first TIM, scheduled in the middle of the algorithm definition study, was to assess the progress of the algorithm definition effort. The second TIM, scheduled mid-way through the implementation phase, was intended to address testing issues, including validation of the new mosaic generation algorithms.

The algorithm validation issue surfaced during the first TIM in early MAY 2003. At contract startup the development group had started to collect radar product data sets for assessing prototype algorithm performance during the algorithm definition phase. As a result of the first TIM, the development group decided to make the data being used by the development team to assess algorithm performance during the algorithm definition phase available to selected FAA personnel via a limited access website. The website was established in mid-June, 2003, toward the end of the algorithm definition phase. At the SRR in June, 2003 it was decided to continue the collection of test case data sets beyond the algorithm definition phase for use by the FAA in a formal validation of mosaic algorithm performance. The algorithm prototype test bed was frozen in July, 2003. From that point, as test case data sets were collected, they were processed on the prototype test bed and the resulting mosaic products and supporting data were posted on the website.

The website was initially used by the WARP Program Office for a qualitative assessment of the algorithm

performance, and then as the basis for defining and implementing the formal algorithm validation plan. The WARP Program Office in close cooperation developed the validation plan with the development group and other FAA and NWS organizations. The validation plan was reviewed and approved at the second TIM in February 2004. The formal validation proceeded of two phases. In the first phase, test cases processed on the prototype algorithm test bed were evaluated. A preliminary validation report based on these cases was published in June 2004 and submitted to a Technical Review Panel (TRP) convened by the joint FAA-NWS Aviation Weather Technology Transfer (AWTT) Board (see section 2.6). On the basis of this preliminary report, the TRP gave conditional approval for operational use of the new mosaic products. In the second validation phase, mosaic products generated by the delivered operational software for the set of test cases selected by the FAA were evaluated. A final validation report based on these cases was published in September 2004 and submitted to the AWTT TRP. On the basis of this final report, the TRP sent a recommendation to the AWTT Executive Board to approve the new WARP mosaic products for operational use. The AWTT Executive Board is scheduled to consider this recommendation at its November meeting.

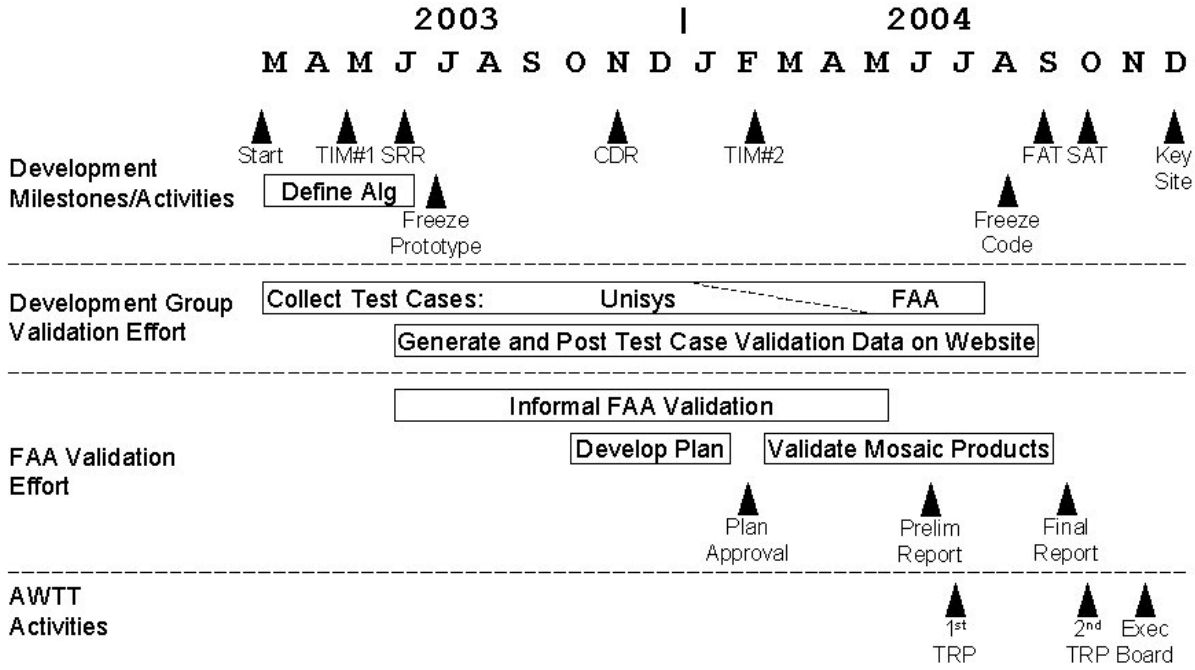

Figure 1. Chronology of algorithm validation effort

2.1 Test Case Data Collection

Collection of test case data sets began in the fall of 2002. Each data set consisted of approximately one hour of the radar product data used by the WARP mosaic generators for 40 radars in one of seven CONUS geographic areas. The data sets also included visible and IR satellite images for the data collection interval. The data was acquired from the Unisys commercial data feed. Initially, the objective was to collect approximately 20 representative test case data sets to be used by the development team during the algorithm design phase of the development effort. After the first TIM in May 2003, the focus shifted toward collection of data sets to be used by the FAA for formally validating the performance of the mosaic generation algorithms.

The Unisys development team initiated collection of 107 test case data sets. Data collection initiated by Unisys had two objectives: to assemble data sets for a variety of synoptic weather scenarios, for different geographic areas, and in different seasons of the year; and to collect data sets for the types of weather and in the areas of the country that have the greatest potential negative impact on air traffic control operations. The FAA initiated collection of 50 data sets. The validation plan identified two CWSU meteorologists and two air traffic controllers as FAA test case selectors. These selectors based their decisions to initiate data collection on real-time observations of developing weather situations in their respective ARTCC coverage areas. The focus of the FAA selected cases was severe weather in the spring and summer months, and also on cases which included radar anomalies. Tables 1 and 2 are a breakdown of the Unisys and FAA selected test cases by region and season.

It is noteworthy that there was no post-selection of data sets. All of the test case data sets collected were used in the algorithm validation process.

Table 1. Unisys selected test cases by season and region

	North East	Central Plains	Great Lakes	North West	Rockies	South East	Southern Plains
Winter		1	1	2		2	
Spring	6	2		1	2	7	7
Summer	17	1	2	3	10	10	12
Fall	6	1		2		6	6

Table 2. FAA selected test cases by season and region

	North East	Central Plains	Great Lakes	North West	Rockies	South East	Southern Plains
Winter	3						
Spring	8				1	5	10
Summer	6	3		1		6	5
Fall							2

2.2 Test Case Website

Initially the website was used to provide the WARP Program Office visibility into the work-in-progress by the development group during the later stages of the algorithm design effort. After the algorithm definition phase was completed, the website was used to make the large volume of test case data available to the FAA validation team in a timely and cost-effective manner. After the prototype algorithm was frozen in July 2003, all previously collected test cases were regenerated using the frozen prototype algorithm and reposted on the website. As new test case data sets were collected, they were processed on the prototype system and posted on the website for use by the FAA validation team. The FAA's validation plan evolved around the website. As the validation plan was being developed, adjustments were made to the website content to tailor it to the needs of the validation effort. The preliminary FAA evaluation was based on the prototype algorithm results. After the operational code was frozen in August 2004, the FAA-selected test cases were rerun on an operational test bed and the results posted on the website. This information was used for the final FAA evaluation.

2.3 Algorithm Validation Plan

The FAA Program Office developed a comprehensive algorithm validation plan in close cooperation with the Harris/Unisys development group. The validation effort was structured to take maximum advantage of the development group assets and to minimize the cost and schedule impacts on the overall development effort. The significant features of the validation plan that contributed to the overall success of the validation effort were:

1. It defined a schedule for the validation effort that was closely coordinated with the detailed software development schedule.
2. It defined roles and responsibilities of specific FAA and NWS organizations, and the WARP development group.
3. It identified the test case data set information to be provided by the development group to support the validation effort.
4. It identified a five-member validation team.
5. It provided a mechanism for the FAA to select test cases to be used in the validation.
6. It defined specific procedures and included a set of guidelines to be used by the validation team for evaluating and scoring test cases.
7. It provided for training the validation team to ensure consistent evaluation by the five-member validation team.
8. It defined the acceptance criteria for the new mosaic products.
9. The plan defined a phased approach that allowed an early start to the validation effort using the prototype algorithm.
10. The plan included steps to directly compare the performance of the prototype and operational algorithms.
11. It provided for independent review of the algorithm validation process by the AWTT.

2.4 Validation Team

The performance of the mosaic generation algorithms was evaluated by a team of five meteorologists familiar with weather operations in the en-route air traffic control environment. The team consisted of two CWSU meteorologists, two meteorologists affiliated with the Weather Group at the FAA Technical Center, and an air

traffic controller with a background in meteorology. Each team member independently evaluated all of the test cases and quantitatively scored each case using the predefined methodology.

2.5 Validation Effort

The validation plan defined a two-stage validation effort: a preliminary evaluation based on the prototype algorithms, and a final evaluation based on the delivered operational software. The primary motivation for using a two-stage approach was to complete the validation effort in the same time frame as the development effort to permit operational use of the new mosaic products as soon as they were deployed. The two-stage approach contributed significantly to achieving this objective. Training of the validation team and fine-tuning of the evaluation procedures occurred during the preliminary evaluation phase. The preliminary evaluation effectively provided for a calibration of the five-member validation team to ensure consistency of the results in the final evaluation. The lessons learned during the preliminary evaluation facilitated the rapid completion of the final evaluation after the operational software was delivered.

An additional significant benefit of the two-phase approach was that the preliminary evaluation identified a deficiency in the prototype algorithm. The deficiency was detected in time for changes to the operational algorithms to be implemented, tested, and informally validated by the FAA before the operational baseline was frozen prior to the start of the FAT/SAT tests. As a result of these changes the operational products scored higher in the final validation than the prototype products in the preliminary evaluation.

The results of the validation team evaluations were compiled and published in two reports. A preliminary report based on mosaic products generated using the prototype algorithm was issued in June 2004. A final report based on mosaic products generated by operational software was issued in September 2004. In both the preliminary and final validation reports, the individual and cumulative evaluation scores of the five evaluators exceeded the acceptance criteria defined by the validation plan.

2.6 AWTT Board Approval

The FAA's validation of the new mosaic generation algorithms is being reviewed by the joint FAA-NWS AWTT Board to determine if the new mosaic products are suitable for operational use. A Technical Review Panel (TRP) was convened in July 2004 to evaluate the scientific validity and the FAA's verification process of the optimal mosaic product, and to provide an assessment to the AWTT Board. Based on the findings in the FAA's evaluation of the prototype algorithm using the Unisys selected test cases, the TRP approved the optimal mosaic generation algorithm contingent on the results of the FAA's evaluation of the operational algorithm using the FAA selected test cases due in September 2004. Following a supplementary meeting of the TRP in October 2004 to review the final validation report, the TRP issued a recommendation to the AWTT Executive Board to approve the WARP optimal mosaic products for operational use. The TRP recommendation is scheduled for review by the AWTT Executive Board at its November 2004 meeting.

3.0 SUMMARY AND CONCLUSION

Over the course of the project the website proved to be an efficient and effective link between the algorithm development and validation activities. The website allowed the development group to provide the required support to the FAA's validation program without significantly impacting the on-going development effort.

The two phased validation approach (a preliminary evaluation based on prototype algorithm results, and a final evaluation based on results using delivered software) allowed the validation effort to be completed in the same time frame as the software development effort.

Through this validation process the FAA was able to measure optimal mosaic's performance and its suitability for operational use by air traffic controllers. According to the 785 case reviews (5 reviewers X 157 cases), the optimal mosaic provides a significant improvement over the current highest reflectivity algorithm. Using the reviewers' estimates and pixel count data, we estimate the net improvements to the mosaic products generated for the controller displays to be as follows:

- Using the current highest reflectivity rule, 88% of the radar echoes on the controllers' displays are real weather.
- Using the optimal mosaic, 98% of the echoes on the controllers' displays are real weather.
- To achieve this improvement, the optimal mosaic reduces 3% of the real weather echoes by 10 dBZ or less and reduces 1% of the real weather echoes by more than 10 dBZ.

4.0 REFERENCES

Lang, J., 2003: Radar Mosaic Generation Algorithms Being Developed for FAA WARP System. Preprints 20th International Conference on Interactive Information Processing Systems, Seattle, WA, Amer. Meteor. Soc., paper 12.10.