3rd USA/Europe Air Traffic Management R&D Seminar Napoli, 13-16 June 2000

The Operational Assessment of Free Flight Phase 1 ATM Capabilities

David Knorr
Manager, Performance Metrics, Free Flight Phase 1 Program Office
Federal Aviation Administration, Washington, DC
dave.knorr@faa.gov

Joseph Post
Senior Operations Research Analyst
TRW Inc., Washington, DC
joseph.ctr.post@faa.gov

Jeff Biros
Operations Research Analyst
TRW Inc., Washington, DC
jeffrey.ctr.biros@faa.gov

Michelle Blucher
Project Team Manager, Systems Analysis
Mitre/CAASD., McLean, VA
mblucher@mitre.org

Abstract

The Federal Aviation Administration's Free Flight Phase 1 (FFP1) program is fielding five air traffic control automation systems that are intended to assist controllers and airlines with decision-making, thereby increasing the efficiency of operations. FFP1 tools will be deployed at a limited number of sites between 1998 and 2002, and evaluated to determine their operational effectiveness. The paper briefly describes the five FFP1 tools, and then discusses the collaborative approach being used to assess the operational impact of the program on airspace system users and service providers. The metrics chosen for each tool, and the categorization method for these metrics, are presented. The methodology being used to quantify and assess these metrics is then discussed. The paper concludes with a case study of the passive Final Approach Spacing Tool (pFAST) at Dallas-Ft. Worth International Airport. The quantified effects of pFAST usage on airport acceptance rates, actual peak arrival rates, runway balancing, and peak operation rates are presented.

Introduction

The Federal Aviation Administration's (FAA) Free Flight Phase 1 (FFP1) program will deploy Air Traffic Management (ATM) capabilities that can provide early benefits to National Airspace System (NAS) users and service providers, leveraging proven technologies with needed procedural enhancements and appropriate standards. FFP1 capabilities have been developed by the FAA, in concert with the user community, and are intended to assist controllers and airlines with decision-making, thereby increasing the efficiency of operations. FFP1 tools will be deployed at a limited number of sites between 1998 and 2002 and evaluated to determine their operational effectiveness, allowing for informed decisions regarding future system development and acquisitions.

Evaluating the operational effectiveness of FFP1 capabilities will be accomplished collaboratively by the FAA and industry stakeholders. This paper describes the collaborative approach used to develop FFP1 performance metrics, outlines the proposed measurement process, and presents some sample evaluation results for one of the capabilities. We believe a joint FAA/industry approach for gauging the success of FFP1 capabilities is an important step in maintaining the original FFP1 consensus and preparing for future deployment decisions.

System Description

The FFP1 program will field five different systems between 1998 and 2002 at a limited number of sites. The operational evaluation of this Core Capability Limited Deployment (CCLD) will provide the basis for decisions regarding national deployment and further development of these systems. The flight domains in which the tools operate are depicted in Figure 1.

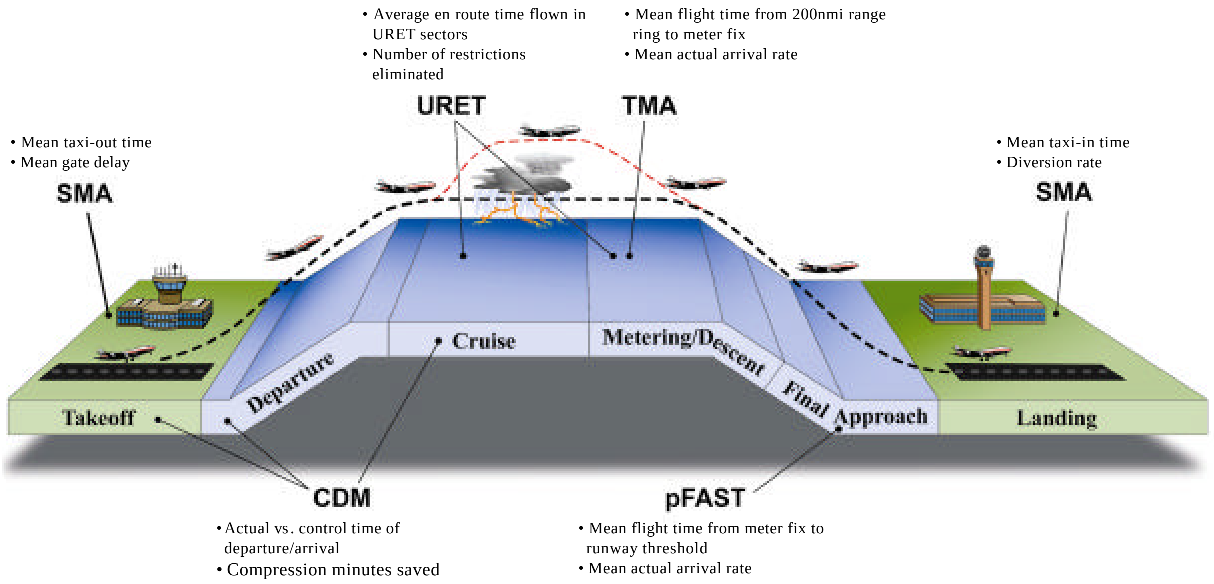

Figure 1. FFP1 Tools, Associated Flight Domains, and Principal Metrics

Collaborative Decision Making

Collaborative Decision Making (CDM) is a set of tools and procedures that allows the airlines and FAA to improve operations through information sharing. Ground Delay Program Enhancements (GDP-E), the initial focus of CDM, started prototype operations at San Francisco and Newark airports in January 1998. Under GDP-E, participating airlines send operational schedules and changes to schedules to the Air Traffic Control Systems Command Center (ATCSCC) on a continuous basis. This schedule information includes, but is not limited to, flight delay information, cancellations, and newly created flights. The ATCSCC uses this information to better implement and manage ground delay programs (GDPs).

In addition to improving the execution of GDPs, CDM has been found to have application to other air traffic management problems, such as airspace congestion due to heavy traffic or en route weather. CDM's Collaborative Routing (CR) function is intended to provide better information to airspace users about potential flow problems that are likely to require rerouting or other flow management actions. This may allow users to prepare for possible effects on their operations in advance. The NAS Status Information (NASSI) function will provide a mechanism to share critical safety and efficiency data with NAS users. A recently formed group has been tasked to determine what these data are and how to set priorities for getting the data distributed.

User Request Evaluation Tool

The User Request Evaluation Tool (URET) is a decision support system developed by the MITRE Center for Advanced Aviation System Development (CAASD) for use by en route center controllers. URET provides aircraft-to-aircraft and aircraft-to-airspace conflict detection and trial planning of proposed Air Traffic Control (ATC) solutions to ensure that they are conflict free. Capabilities will be used primarily by D-Controllers for strategic problem detection. The basis for URET strategic planning capabilities is aircraft flight plan information, track data, forecasted winds and temperatures, aircraft performance characteristics, and facility information. Using this information, the progress of an aircraft is continuously monitored, problems are detected, and controllers are notified of possible conflicts between the current flight and other aircraft and/or airspace. In addition, when a pilot requests a new clearance, the controller can use URET to identify any possible conflicts.

Center-TRACON Automation System

The Center-TRACON Automation System (CTAS) is a set of decision support tools that assist air traffic managers and controllers by generating aircraft schedules and advisories to regulate aircraft arrivals to a runway complex. Traffic Management Advisor (TMA) assists controllers in the en route cruise and transition airspace managed by Air Route Traffic Control Centers (ARTCCs). TMA provides ARTCC personnel with a means of optimizing the arrival throughput of capacity constrained airports. Inputs to the system include real-time radar track data, flight plan data, and local weather conditions. TMA's trajectory models use this information, updated every 12 seconds, to compute routes and optimal schedules to the meter fixes for all arriving aircraft which have filed IFR flight plans, with consideration given to separation, airspace, and airport constraints.

The passive Final Approach Spacing Tool (pFAST), the other CTAS tool being fielded, is used by controllers and air traffic managers to manage the flow of arrivals in the terminal airspace. pFAST computes a relative sequence for each arrival aircraft for each runway at a particular airport. The system calculates a runway assignment for each aircraft in such a way as to minimize overall flight delay, with consideration given to aircraft type, speed, and trajectory. Runway advisories are displayed to the controller on the ARTS display. The controller may manually override both the relative sequence number and the runway advisory displayed by pFAST, and the system automatically adjusts to sequence number changes.

Surface Movement Advisor

The concept for Surface Movement Advisor (SMA) sprung from development work by the National Aeronautics and Space Administration (NASA) on a tool to aid ramp controllers with gate management. This tool was prototyped at Atlanta's Hartsfield International Airport, and continues to be used there. For the SMA CCLD, the FAA will distribute filtered Automated Radar Terminal System (ARTS) data to participating airlines and service providers via the Collaborative Decision Making Network (CDMNet). FAA is also providing software that will present a planform display of arriving and departing aircraft (similar to a radar display). Several airlines (including Northwest Airlines, Southwest Airlines, and USAirways) have begun using this display in their operations centers to better manage operations at the SMA airports. These airlines also intend to integrate this data into their respective gate management tools, so that ramp controllers will benefit from timely information on arriving aircraft.

Collaborative Approach

FFP1 capabilities and sites were originally selected by the RTCA Free Flight Steering Committee [1].[1] The collaboration on FFP1 between industry stakeholders (as represented by the RTCA) and the FAA will continue through deployment and include an evaluation of FFP1's operational impact. The RTCA and FAA developed a joint workgroup to define a core set of metrics to assess the impact of each capability. This metrics workgroup included expert representatives from the airspace user community (airlines, cargo carriers, and general aviation), as well as service providers and CAASD. Quantitative measures were derived from the desired impacts outlined in the operations concept and aligned with the FAA's operational goals [2]. Mapping the FFP1 metrics to the RTCA's operational concept and the FAA's goals was a key step in reaching an FAA/industry agreement on the metrics.

In addition, these metrics were assessed for their measurability (practicality of data collection) and meaningfulness (interpretability in terms of a positive or negative value). Since approval of the core set in April 1999, the FFP1 Metrics Team has continued to work with the RTCA, developing a detailed operational evaluation plan to assess the metrics [3]. This evaluation plan further refines the set of metrics following more detailed discussions with air traffic controllers and managers at the FFP1 prototype sites, discussions with other FAA analysts, preliminary analyses of prototype data, and continued interface with industry stakeholders. While the plan defines metrics for each tool, we recognize that the metrics may need to be adjusted as we gain more experience. The plan specifically identifies data sources and methodologies used to calculate each metric.

Stakeholders' data inputs, and their interpretation and validation of impacts, are a vital link to operational impacts. The FFP1 Program Office (PO) will make data available for stakeholder review, and will provide stakeholders with consistent information from reliable sources to assess operational performance. This information will also facilitate future decisions about system enhancements, site proliferation, and funding. The FFP1 PO will implement both formal and informal reporting mechanisms to share the results of operational evaluations with stakeholders. Formal mechanisms will include quarterly reporting to the RTCA Free Flight Steering Committee, which will be coordinated with the RTCA Select Committee on Free Flight Implementation.

Metrics Definitions

The FFP1 metrics have been grouped into five categories aligned with the FAA's operational goals: safety, access, delay/efficiency, predictability, and flexibility. All of the revised FFP1 performance metrics are tabulated in Table 1.

Safety – Safety may be defined as the ability to maintain the standards specifying safe spacing distances, both between aircraft and between aircraft and terrain or obstructions. FFP1 safety metrics are the changes in the number of operational errors and operational deviations.

Access - Access can be defined as the ability of users to enter the ATC system and obtain services on demand. For FFP1, access focuses on maximizing the use of existing runways for arrivals and departures. For CDM access also includes system throughput related to improved information. Clearly, improved access to airspace and runways will have a direct relationship to delays. However, as demand increases runway throughput may increase while delays remain constant or even increase. This phenomenon is well known in surface transportation; when a lane is added to a highway drive times initially decrease, but then increases as traffic increases. For this reason it is important to have specific measures for access (throughput) and delay.

Specific FFP1 access metrics include mean peak period actual arrival rates per airport and per runway, diversion rate, and the number of unused slots.

Table 1. FFP1 Performance Metrics

Outcome Category	CDM	URET	TMA	pFAST	SMA
Safety	• Change in operational errors • Change in operational deviations				
User Access	• # of unused slots		• Actual arrival rate	• Actual arrival rate • Actual arrival rate for each runway	• Diversion rate
Delay/ Efficiency	• Mean flight time • Compression minutes saved	• Mean enroute time • Mean distance flown • Mean air distance flown • Mean fuel usage • % time spent at or near desired altitude • # restrictions eliminated • Aggregate degrees turned	• Mean flight time from 200nmi range ring to meter fix • Mean arrival delay • Mean fuel usage from 200nmi range ring to meter fix • Variability of fuel usage	• Mean flight time from meter fix to runway threshold • Mean fuel usage from meter fix to threshold • Variability of fuel usage from meter fix to threshold	• Mean taxi-in time • Mean taxi-out time • Mean gate delay
Predictability	• Integrated Predictive Error • Rate Control Index • EDCT compliance ratio • # of GDPs canceled near start • # of GDP revisions	• Planned vs. actual enroute time • Planned vs. actual distance flown	• Mean error in predicted meter fix arrival time • Variability in error • Variability of actual arrival rate • Mean difference between AAR and actual arrival rate • Variability of time from 200nmi range ring to meter fix	• Mean difference between AAR and actual arrival rate • Variability of flight time from meter fix to threshold	• Variability of taxi-in time • Variability of taxi-out time • Variability of gate delay • Gate reassignment rate
Flexibility	• Mean distance flown • Control time of arrival	• Mean enroute time (late departures) • Mean enroute distance flown (late departures)			
Productivity		• # of a/c per sector per unit time • Change in monitor alert threshold	• Mean actual arrival rate/throughput per sector or position	• Distribution & throughput of operations per runway/position	

EDCT = Estimated Departure Clearance Time AAR = Airport Acceptance Rate

Delay/Efficiency - We have elected to combine delay and efficiency metrics into one category, since some metrics in these areas are closely related or even equivalent. In the past delay has been defined in three different ways:

- the amount of time beyond expectations that it takes to complete a flight or flight segment,
- the difference between actual and scheduled arrival times, and
- the additional transit time above the “optimal” or unimpeded time.

We intend to consider all of these definitions of delay, since each has a unique impact on the NAS user’s value function.

Definitions of efficiency have centered around fuel efficiency for a given flight as well as reductions in flight times. Efficiency has often incorporated all reductions in delay. From our perspective, both definitions have unique components which are valued separately as well as a common component. Both definitions will be captured under this metric category.

FFP1 delay/efficiency metrics include mean gate delay, mean flight time, mean taxi out time, mean flight time in URET sectors, mean flight time from the 200 nmi range ring to the meter fix, mean flight time from the meter fix to the runway threshold, mean taxi-in time, mean arrival delay, GDP compression minutes saved, and the mean and variance of fuel usage per flight segment by aircraft type.

Predictability - Predictability measures the variation in the ATM system as experienced by the user. Our definition of predictability focuses on the dispersion (specifically, the variance) associated with flight segment times. Commercial airlines may benefit as much from a reduction in the variance (or an improvement in the consistency) of flight/taxi times as they would from a reduction in average flight times. System predictability allows for improved scheduling and more efficient bank operations.

Specific predictability metrics include variability of flight time from the 200 nmi range ring to the meter fix, variability of flight time from the meter fix to the runway threshold, mean difference between actual arrival rate and AAR, variability of actual arrival rate, variability of taxi-in and taxi-out times, variability of gate delay, number of GDPs cancelled, and time spent at or near desired altitude for specific city pairs.

Flexibility - Ultimately flexibility measures the ability of the ATC system to meet users changing needs in their efforts to optimize daily operations. For example, commercial air carriers may prefer increased delay in exchange for an on-time arrival on a specific flight. We have focused flexibility metrics on capturing anything an airline would like to accomplish on an individual flight not already captured in the previous metrics, including faster routes (flight times) to make up lost schedule, slower routes to reduce taxi-in delay, and requests for altitude changes for passenger comfort.

In practice, it is extremely difficult to establish airline intent on an individual flight basis. In fact, within an airline the pilot and dispatcher may have different objectives. For this reason, our approach to measuring flexibility is to segregate flights delayed at departure from those departing on time. Our supposition is that those aircraft leaving late will desire to make up time en route. Flexibility will be measured by an airline's ability to make up time (i.e., to keep to schedule). Other measures of flexibility will be developed after obtaining feedback from users on perceived changes or improvements in service.

Measurement Process

An extensive data collection effort is planned in order to fully assess the impact of the FFP1 tools at each of the CCLD sites. Various data sources from the FAA, Department of Transportation, NASA, the National Oceanic and Atmospheric Administration (NOAA), and the airlines will be combined to provide a complete picture of NAS operations and FFP1 tool performance, both prior and subsequent to fielding of the tools at each of the CCLD sites.

Data for each metric will be collected for at least one year prior to Initial Daily Use (IDU) at each site so that a robust baseline can be established (for CDM metrics data are collected nationally). Data for each metric will then be collected for a period of at least one year following Planned Capability Available (PCA). In this way seasonal factors may be fully removed. Between IDU and PCA, operations will be observed and trends in the metrics reported in order to understand any learning-curve effects, and to provide feedback to system developers as to local adaptation. Local environmental, airport configuration, and airport demand data will also be collected for this length of time so that we may better isolate the effects of the particular FFP1 tool from those of changes in these conditions.

For each capability, the evaluation (and consequently the data collection effort) will focus on the flight segments that are expected to be most affected by the particular tool (see Figure 1). For example, the pFAST evaluation will focus on flight times in the terminal area (specifically, flight times from the meter fix to the runway threshold) and airport arrival rates. However, since NAS operations are tightly coupled, the evaluation will also consider upstream and downstream effects where appropriate. As an example, one of the primary metrics for SMA will be mean taxi-out time. Additionally, we intend to examine upstream gate delay for SMA (gate delay is defined here as the difference between scheduled and actual gate departure times).

Wide-ranging data relating to the local environment and conditions at each FFP1 site will be collected for the same time periods that performance metric data are collected. This data includes, but will not be limited to, airport configuration; surface weather, including visibility, ceiling, and precipitation rate; winds aloft; arrival demand (i.e., actual arrivals per unit time); and departure demand. This data is essential to ensure that "apples to apples" comparisons of system performance are made before and after the deployment of a capability. For tools which operate in the extended terminal area, arrival demand data is particularly critical, since flight times can be expected to be larger when there is a high level of arriving traffic.

In order to "normalize" for differing distributions of these conditions pre- and post-deployment, two different analytical techniques will be employed. The first and most transparent approach will be to group data into "bins" with similar local conditions. This approach is simplest, but does not take full advantage of all the information content in the data. To remedy this limitation, multivariate regression techniques will also be used. These statistical techniques are well suited to problems where an output variable (e.g., flight time) may be influenced by many "exogenous" factors.

A number of data sources will be used to compute the metrics and obtain the associated environmental data, including existing FAA databases, airline data, and new data sources. Actual arrival rates, and coarse flight time and distance information will be collected from the Enhanced Traffic Management System (ETMS), Airline Service Quality Performance (ASQP), and Consolidated Operations and Delay Analysis System (CODAS). More precise flight times and distances will be obtained from local ARTSs for the terminal domain, and from the Host computer for the en route domain. In some cases flight tracks and associated data will be obtained from log files produced by the various FFP1 tools. AARs will be obtained from facility and ATCSCC logs. Fuel usage will be modeled using airline-provided equipment-specific fuel consumption data and observed flight trajectories. Finally, safety data will be obtained from the National Airspace Incident Monitoring System (NAIMS) databases.

Preliminary pFAST Results

In February 1999 air traffic controllers at the Dallas-Ft. Worth International Airport (DFW) Terminal Radar Approach Control (TRACON) began using pFAST to help sequence and assign runways to arriving aircraft. Initially pFAST was only used by a cadre of controllers, but over the past year pFAST usage has gradually increased, to the point now that all controllers in the TRACON are using the tool.

We have conducted a preliminary analysis of the impact of pFAST on operations at DFW, examining airport acceptance rates, actual arrival rates and operations (i.e., arrivals plus departures) rates, TRACON flight times, taxi times, and runway balancing. Our results suggest that pFAST usage has indeed led to an increase in airport acceptance rates, which has thereby led to an increase in peak arrival rates. An observed improvement in runway balancing has also resulted in an increase in total operations rates. While taxi-in times have slightly increased, there has been an offsetting and larger decrease in taxi-out times. Finally, these improvements have occurred without any measurable change in TRACON flying times. We present here summary results for acceptance rates, actual arrival rates, runway balance, and operation rates.

In order to determine if pFAST usage has led to an increase in Airport Acceptance Rates (AARs) at DFW, we performed a regression analysis of the airport acceptance rate and various environmental variables that, in the judgement of experienced air traffic controllers, should affect the AAR. Specifically, we regressed the number of arrival runways in use, the type of approaches being used (visual or instrument), the natural logarithm of ceiling, the square of crosswind component, and a pFAST dummy variable on AAR.[2] We also included a dummy variable that accounts for a procedural change implemented in July 1999 whereby the TRACON accepts an unlimited number of aircraft for specified times on a fifth arrival route (a so-called "unlimited dual" route). We included data in ten minute increments from February 20 through December 31 1999, for a total of approximately 220,000 observations.

The results of this regression analysis are presented in Table 2. All of the variables included in this model were found to be significant at the five percent level, and the signs of the coefficients were all as we would expect. For example, when DFW uses three arrival runways (rather than four) the acceptance rate is reduced *ceteris paribus* by approximately 22 aircraft per hour. Similarly, when the ceiling increases by 10,000 feet the acceptance rate increases by $0.97 \cdot ln$ (10,000 ft.) $\approx$ 8.9 aircraft per hour. After controlling for all of these factors, we found that pFAST usage resulted in an increase in acceptance rates of approximately 2.5 aircraft per hour.

Next we examined actual airport arrival rates, to see if controllers were able to land aircraft at the higher rates that the TRACON is now requesting from the en route center. We collected aircraft arrival counts in ten minute intervals for DFW from 20 February 1999 through 29 February 2000. In order to diminish the impact of any possible sampling error in these rates, we used a sliding 30 minute window to calculate the arrival rates (thus we calculated the number of aircraft arriving in a 30 minute period every 10 minutes). Once we had "filtered" the arrival rate data in this manner, we used an algorithm to identify the eight highest arrival peaks per day.[3] Figure 2 presents a box plot of these peak arrival rates, segregated into instrument and visual arrival conditions, with and without pFAST in use. The dark central line within each box indicates the median flight time for that particular set of conditions.[4] The sample size for each box plot is indicated at the bottom of the plots.

Figure 2 suggests that there has been a measurable increase in median peak arrival rates at DFW associated with pFAST usage. The median peak arrival rate has increased from 58 arrivals per 30 minute period to 59.5 when running instrument approaches. When using visual approaches, the median arrival rate has increased from 61 to 62 arrivals per 30 minute period. A chi-square-type test of these medians indicates that the differences are statistically significant at the five percent level [5].

We also examined the "balancing" of the runways at DFW. Runways are considered to be balanced if the arrival rates on the individual runways are approximately equal. By balancing the runways the overall arrival rates should be able to be increased, and surface congestion reduced. Our relatively simple measure of the degree to which the arrival runways are balanced is the difference in the percent of arrivals handled by the most used and least used runways. The data used for this calculation are the same data used above, namely the 10 minute arrival counts. We limit the data sample to periods when the airport was in a south flow configuration, since for most of the period under examination pFAST was used only in this configuration. Additionally, we only included 10 minute time periods when there were at least four total arrivals. All of the arrivals are then summed by month, and the difference between the most used and least used runways is expressed as a percentage of total arrivals. The result of this calculation, displayed in Figure 3, indicates that the difference between the most and least used arrival runways is reduced when pFAST is in use.

Table 2. DFW Acceptance Rate Regression Analysis

Dependent Variable:	
AAR	Airport Acceptance Rate (arrivals/hr)
Independent Variables:	
3_Runways	0 - four arrival runways
	1 - three arrival runways
IFR	0 - visual approaches
	1 - instrument approaches
UnltdDuals	0 - FEB 20 - JUN 30 1999
	1 - JUL 1 - DEC 31 1999
Ln_Ceiling	natural logarithm of ceiling in feet
NorthFlow	0 - south flow
	1 - north flow
CrosswindCompSq	square of crosswind component in knots
pFAST	0 - pFAST off
	1 - pFAST on

	Unstandardized Coefficients		Standard-ized Coeffic-ients		
Independent Variables	B	Std. Error	Beta	t	Sig.
(Constant)	127.936	.173		739.002	.000
3_Runways	-21.871	.060	-.444	-366.191	.000
IFR	-13.978	.038	-.540	-368.845	.000
UnltdDuals	1.382	.030	.058	45.793	.000
Ln_Ceiling	.970	.017	.085	58.499	.000
NorthFlow	-.936	.031	-.036	-29.775	.000
CrosswindCompSq	-.01196	.000	-.049	-40.816	.000
pFAST	2.486	.030	.098	82.543	.000

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.859	.739	.739	6.0243

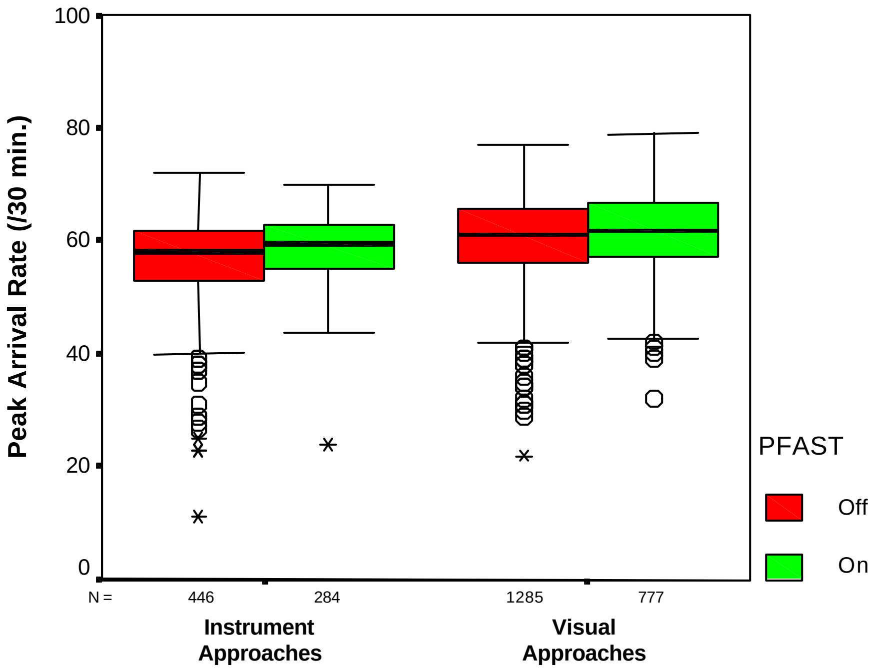

Figure 2. DFW Peak Arrival Rates

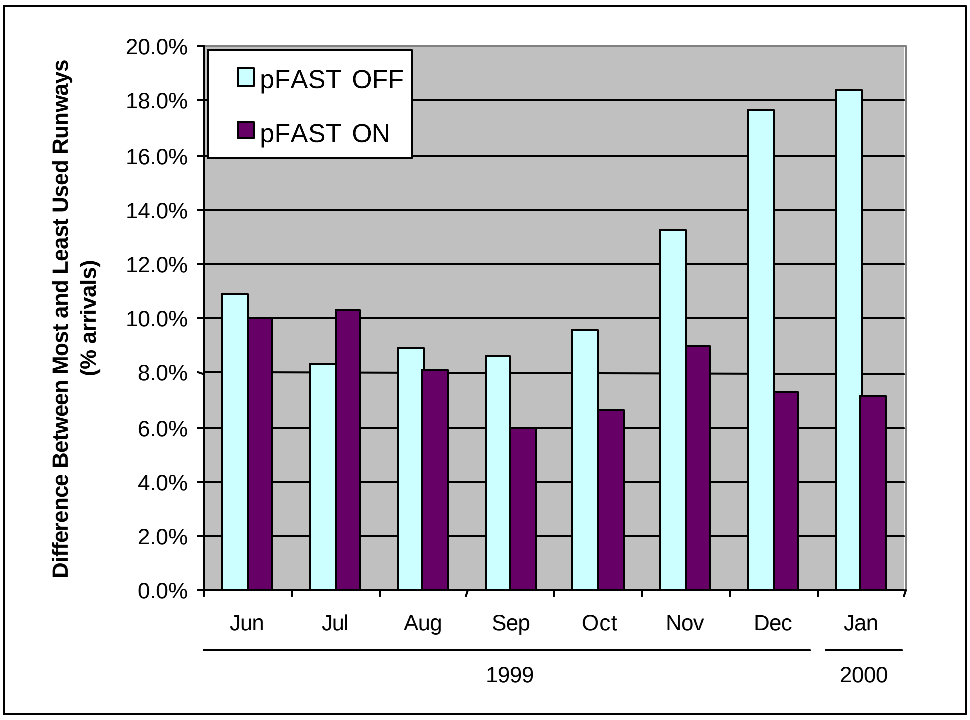

Figure 3. DFW Arrival Runway Balance

The final metric examined here is the peak rate of operations. The improvement in runway balancing made possible with pFAST has resulted in less surface congestion on DFW taxiways at peak times. This decrease in congestion has led to a corresponding increase in departure rates. Figure 4 illustrates the distributions of peak operation rates (arrivals plus departures) from 22 April 1999 through 29 February 2000. We filtered the operation rates using a 30 minute window, selected the eight highest peak operation rates for each day, and segmented these peak rates by type of approaches being flown and pFAST status (the identical procedure as that used for arrival rates above). The median peak rate of operations has increased from 105 to 109 operations per 30 minute period under instrument approaches, and from 111 to 114 operations per 30 minutes under visual approaches. These differences are statistically significant at the five percent level.

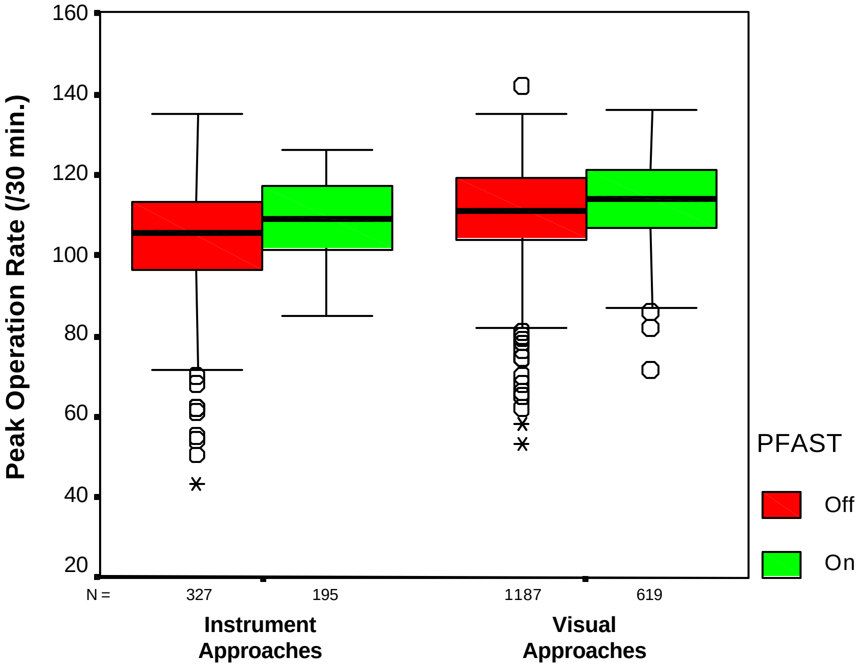

Figure 4. DFW Peak Operation Rates

Conclusions

The FFP1 operational evaluation represents a departure from the FAA's usual approach to system acquisition, where impact studies are usually limited to a specified operational evaluation period. We recognize the manifold challenges inherent in this approach. In order to mitigate these risks, we have assembled an expert team of engineers, analysts, and air traffic controllers with considerable experience in conducting operational evaluations. In addition, we will draw on the experience of our industry partners to help guide our efforts and interpret our results. By studying data and experiences from the FFP1 prototype sites we will be well positioned to accommodate the operational data. We are confident that we will be able to conduct an objective assessment of FFP1 capabilities which will consider the interests of NAS users and service providers alike, and will aid developers in refining their products and facilitate informed national deployment decisions.

References

[1] *Government/Industry Operational Concept for the Evolution of Free Flight, Addendum 1: Free Flight Phase 1 Core Capabilities Limited Deployment*, RTCA Select Committee for Free Flight Implementation, August 1998.

[2] *Air Traffic Services Performance Plan for Fiscal Years 1998-2000*, FAA Office of Air Traffic Services.

[3] *FFP1 Performance Metrics: An Operational Impact Evaluation Plan*, FAA Free Flight Phase 1 Program Office, 12 August 1999.

[4] Tukey, John W., *Exploratory Data Analysis*, Reading, Mass.:Addison-Wesley, 1977.

[5] Conover, W.J., *Practical Nonparametric Statistics*, New York:Wiley, 1980.

---

[1] RTCA, Inc. is a private, not-for-profit corporation that advises the FAA by developing consensus-based recommendations on communications, navigation, surveillance, and air traffic management issues.

[2] In actuality we used the Airport Landing Rate (ALR) for this analysis. The ALR adds to the AAR certain additional arrival traffic, and is a more accurate representation of the total arrival rate being specified. The DFW TRACON is the only approach control facility in the United States that makes such a distinction.

[3] The choice of eight peaks per day is somewhat arbitrary. There are typically nine arrival and departure peaks per day at DFW. We have elected to discard the smallest of these peaks each day.

[4] The bottom and top of the central box represent the 25th and 75th percentiles of the data, respectively. The lower whisker (the line distending downward from the box) extends from the 25th percentile to the value equal to or greater than the 25th percentile minus 1.5 times the interquartile range. Similarly, the upper whisker extends upward from the box to the value equal to or less than the 75th percentile plus 1.5 times the interquartile range. Values more than 1.5 but less than 3 times the interquartile range from the box are represented by circles, while those more than 3 times the interquartile range from the box are depicted by asterisks [4].