ASD-400 Post-Implementation Analysis

Case Studies: URET and TMA

September 30, 2002

Federal Aviation Administration
800 Independence Avenue, SW
Washington, DC 20591
DCN: 4.0-01-02-168

Table of Contents

1.0 INTRODUCTION

1.1 Purpose

The purpose of this document is to present preliminary post-implementation analyses that were recently conducted by the Investment Analysis and Operations Research (ASD-400) Division. ASD-400, an organization with the primary charter of conducting investment analysis, does not promote, sponsor, or endorse FAA programs. Typically, the program office, i.e., an Integrated Product Team (IPT), performs that function. This effort represents a starting point for post-implementation analysis of major capital investments at the FAA[1].

1.2 Scope

This document presents two program evaluations, i.e., User Request Evaluation Tool (URET) and the Traffic Management Advisor Single Center (TMA-SC), that were recently prepared as part of ASD-400's independent assessment of the Free Flight Phase 2 (FFP2) program in support of the Joint Resources Council (JRC) 2b investment decision. The evaluations focused on the flight performance[2] before and after deployment.

The current state of the two evaluated programs is as follows:

1) URET Core Capability Limited Deployment (CCLD) – URET is operational at six Air Route Traffic Control Centers (ARTCCs): Indianapolis Center (ZID), Memphis Center (ZME), Kansas City Center (ZKC), Washington Center (ZDC), Cleveland Center (ZOB), and Chicago Center (ZAU).
2) TMA-SC - TMA has recently gone into Planned Capability Achieved (PCA) at the following FFP1 sites: Miami Center (ZMA), Denver Center (ZDV), and Minneapolis Center (ZMP); and Initial Daily Usage (IDU) status at Oakland Center, Atlanta Center, and Los Angeles Center. It was also implemented at Dallas Fort Worth Center before the establishment of the FFP1 Program.

URET flight performance was evaluated through ZID and ZME since they were the prototype sites. Flight performance through ZMP is examined in this analysis since TMA has been operational with time-based metering significantly longer than at the other operational FFP1 sites. These analyses should be considered preliminary since the evaluation period represents a relatively short timeframe of actual usage. Additional time for the evaluation is needed to reach a more conclusive perspective of the performance of these two programs as well as any program that is early in its deployment.

1.3 State of Post-Implementation Assessment at the FAA

Presently, the FAA does not have a formalized process for conducting post-implementation assessments of deployed acquisitions. In April 1999, a General Accounting Office (GAO) report

[1] This effort does not follow the activities denoted as part of a formal Post-Implementation Review (PIR) laid out by the NAS Configuration Management and Evaluation Office (ACM-1).
[2] Flight performance in this report was measured using airborne times and block times between key city pairs.

titled *FAA's Modernization Investment Management Approach Could Be Strengthened* recommended that the FAA initiate Post-Implementation Reviews (PIRs) on its acquisition program [1]. At this time, it is not clear to the author of this document how the Agency has addressed this issue; however, discussions with managers and literature searches on several FAA web sites revealed no completed or ongoing efforts. Currently, no documentation exists regarding PIRs in the FAA's Acquisition Management System (AMS).

PIRs are excellent vehicles for evaluating the performance of an acquisition. Organizations within an agency that are not directly involved in an acquisition should conduct them. Once a project becomes operational, it is important to understand its impact through a metric(s). Other Federal agencies, i.e., DOT National Highway Traffic Safety and the U.S. Coast Guard, have been performing formal, post-implementation assessments for over 20 years [2]. Similarly, the Department of Energy (DOE) has identified these assessments as a critical component of its management of an acquired capital asset that is part of their agency's portfolio [7].

1.4 Candidate Acquisitions for Post-Implementation Assessment

Several ongoing National Airspace System (NAS)-funded programs lend themselves to conducting post-implementation assessments. Many FAA-funded programs received funding approvals within the last five years and are going operational at this time at various sites. Table 1-1 below lists some of the higher-dollar programs with their dominant metric(s), annotated in the Acquisition Program Baseline (APB) and Investment Analysis Report (IAR) that have recently been deployed in the NAS.

Table 1-1: Recently Deployed FAA Acquisitions

Program	Primary Metric	Deployment Sites	Program
AMASS	Runway incursion reductions	ATL, STL, LAX, ORD	ASDE
ITWS	Airborne, ground and total arrival delay savings	EWR, ATL	Weather Program
PRM	Arrival delay savings	MSP, PHL, JFK, STL, SFO	ATP- 100
TMA	Throughput increase	ZMP, ZTL, ZLA, ZMA, ZDV, ZOA	FFP1
URET	Distance savings	ZID, ZME, ZDC, ZKC, ZOB, ZAU	FFP1

2.0 ASD-400 POST-IMPLEMENTATION INDEPENDENT ANALYSIS

This section describes the details of ASD-400's analyses of TMA and URET. Note: the efforts described below represent a sub-set of the full evaluation that supported ASD-400's independent assessment of the FFP2 program.

2.1 Benefits' Tools and Data Sources

Several data sources were applied for the two analyses (URET and TMA) to measure their pre- and post-implementation benefits. The primary data sources are listed in Table 2-1.

Table 2-1: Tools and Data Sources Used

Models/Tools and Primary Databases	Key Function
Air Traffic Operations Network (OPSNET)	Official air traffic delay database. Captures all reportable delays exceeding 15 minutes. Reports delays by type (i.e., departure, arrival, en route) and cause (i.e., weather, traffic flow, equipment, etc).
Airline Service Performance Metrics (ASPM)	Airborne times and filed-estimated time en route (ETE) for all filed-IFR flight plans (2001 to present) that are flown.
Airline Service Quality Performance (ASQP)	Airborne times, ground times, and delays for 10-12 major carriers submitted monthly to the Department of Transportation. It is used for on-time performance reporting.
Consolidated Operations and Delay Analysis System (CODAS)	Same information as ASPM but has 1998-2000 data. Airborne times and filed-ETE for all filed-IFR flight plans that are flown.
Enhanced Traffic Management System (ETMS)	IFR flight plans that are either "as flown" or "as filed." Trajectories can be built for each flight.
National Climatic Data Center (NCDC) Surface Weather	Hourly surface data that gives ceiling, visibility, and winds by airport. Isolates the weather conditions at the respective airports.
Post Operational Evaluation Tool (POET)	Provides access to ETMS data. Gives detailed flight plan information (includes sectors flown and amendments) and assists analysts with understanding several city pair attributes, e.g, routes flown..
Terminal Area Forecast (TAF)	Historical, current, and forecast data for enplanements, operations, and instrument operations for the majority of the airports in the NAS.

2.1.1 Case 1: URET - Airborne Time Evaluation of Common City Pairs

This part of the assessment compares the airborne times of flights between city pairs that had a significant portion of its flights flying through either the Memphis and Indianapolis Centers (ZME and ZID) or both, before and after, using the URET prototype. Data through the ETMS reveals that several thousand-city pairs spend at least 5 minutes over ZME or ZID. For this analysis, we initially selected 34 city pairs randomly from flights that almost always flew over one of the Centers. The Post Operations Evaluation Tool (POET), developed by METRON, was used to plot one month of flights from these 34 city pairs. Plotting revealed that 28 of the city pairs consistently had flight paths over ZME or ZID. The remaining six city pairs had significant numbers of flights that did not pass through these two Centers and were eliminated from the analysis. Table 2-2 presents the findings on flight paths using POET for the 28 city pairs evaluated.

Table 2-2: 28 Key City Pairs Evaluated

Departure	Arrival	Centers Affected
ATL	CVG	ZID only
ATL	DAY	ZID only (landing in ZID)
ATL	DEN	ZME most of the time
ATL	MCI	ZME
ATL	MSP	ZME or ZID (each half time)
ATL	ORD	Both ZME and ZID
ATL	PIT	ZID only
BWI	MDW	ZID most of the time
CMH	ORD	ZID only
DEN	IAD	ZID only
DTW	CVG	Airport Inside ZID
DTW	MEM	Both ZME and ZID, lands in ZME
IAH	DTW	Both ZME and ZID
IND	STL	ZID only
LAX	IAD	ZID and sometimes ZME
LGA	DFW	Both ZME and ZID
MDW	CMH	ZID only
MEM	ATL	ZME only (starts in the Center)
ORD	ATL	Both ZME and ZID
ORD	CLT	ZID only
ORD	IAH	Mostly through ZME
ORD	MCO	ZID only
ORD	MIA	ZID only
SDF	STL	ZID (airports in ZID)
STL	CLE	ZID to the north of Center
STL	DCA	ZID
STL	IND	ZID only (landing in ZID)
STL	LIT	ZME only (landing in ZME)

2.1.1.1 Methodology

We began by collecting *airborne times* from 1995 to 2000 between all flights from the 28 city pairs. The airborne times from the ASQP data were used. Next, using SPSS Exponential Smoothing Model, frequency distributions were calculated for each city pair's airborne time. Detailed results for each city pair are available upon request. For a given city pair, airborne times that flew through ZME and/or ZID were compared for each timeframe as follows:

- Pre-URET — comparison of airborne performance of 1995 frequency distributions to 1998 distributions
- Post-URET — comparison of airborne performance of 1998 frequency distributions to 2000 distributions

These comparisons show the trends of airborne time distributions after having URET at the Centers when the two-way host interface was implemented. We call these the post-URET airborne time comparisons.

There are two different cases that can be observed. The first case is when a median airborne time increased by 2 minutes between 1995 and 1998, but only increased .5 minutes between 1998 and 2000. One can conclude that having URET might have contributed to the *declining rate* of airborne time increase. The second case is when the mean and median remained the same but the standard deviation fell between 1998 and 2000 when it had an increasing trend prior to 1999; one can conclude that URET may have contributed to a tighter airborne distribution. Predictability metrics may be hard to assign an economic value, but since they impact the airline's bottom line, they should be quantified. Improved airborne time predictability for a city pair may enable the airlines to reduce their scheduled block times. Thus, a tighter airborne time distribution can be assigned an economic value by reduction in block times. In the next section, we discuss changes in scheduled block times for DTW-MEM where the standard deviation fell between 1998 and 2000. Conversely, the opposite effect takes place when the airborne times and standard deviations increase.

In a previous ASD-400 study[3], a preliminary analysis was done to measure changes in airborne distributions for nine airports between 1995 and 2000. When the destination airport of our 28 city pairs were one of the nine airports in that study, we further compared post-URET airborne time comparisons to the results found in the study. This again shows how this city pair performed compared to other city pairs flying into the same airport. Again, if the statistics of the selected city pair were better than the ones found in the study, then one could hypothesize that URET had a positive effect on reducing the increasing airborne time trends. It has been hard to show any significant differences, however, when looking at airborne averages for the entire city pairs' population and those that fly over ZME and ZID. For this reason, we concentrated on entire distributions not just averages. Lastly, we should observe that any improvement might be due to factors other than URET.

2.1.1.2 Data Collection

The ASQP data between 1995 and 2000 was used in this analysis. The ASQP database provides origin, destination, flight time, scheduled, and actual flight and airborne times, as well as taxi-in and taxi-out times. Aircraft speed varies by aircraft type. Propeller aircraft (turboprops and piston-engine) fly at slower speeds than jets. However, for the 28 city pairs, the annual changes in the number of propeller aircraft reported in ASQP were small enough that we did not filter them out of the analysis. ASQP data, which is based on the reporting of the top 10-12 carriers to DOT, comprise about a half million records a month or over an average of 16,000 flights per day.

Manipulating and analyzing several years of data is a time-consuming, difficult task; consequently, we considered sampling to ease handling of this dataset. Sampling plans include using one representative month, or 12 sample days (one day per month of the year) to capture seasonal differences. Since weather plays a significant role in the length of the airborne times, weather effects on delays must be identified and adjusted for all years. This is not an easy task, especially

[3] Nastaran Coleman, "Optimized Flight Time", September 2001.

since our current weather information is limited to hourly surface weather. Furthermore, historical data for convective weather has not been collected for multiple years. Moreover, the effects of different weather scenarios on delay are not understood completely; therefore, more research is required. Thus, we decided to use all days (not segregating the data by weather condition) while working with very large data sets.

2.1.1.3 Results - Comparing Annual Distributions

First, we generated pre- and post-URET distributions. Pre-URET is defined by comparing 1998 airborne distributions for selected city pairs to the corresponding distributions in 1995. Post-URET is defined similarly using 1998 and 2000, respectively. Comparing all three years for a city pair enables the evaluation of the impact of URET on that city pair.

Tables 2-3 and 2-4 provide illustrations of airborne time performance with the relevant statistics for one of the 28 city pairs - DTW to MEM. Table 2-5 shows the DTW to MEM airborne time statistics for each year. In this example, the airborne time's median, mean, and standard deviation increased by 2.0, 1.64, and .23 minutes, respectively, between 1995 (pre-URET) and 1998, but actually decreased between 1998 and 2000 (post-URET) as traffic, defined as number of aircraft handled, increased (see Table 2-6).

Table 2-3: Airborne Time Statistics of DTW to MEM (minutes)

Measure	1995	1998	2000
# of Flights	2569	2795	3100
Min	79	81	81
Median	89	91	90
Mean	90.6	92.2	91.7
Max	120	126.8	122.7
Std. Dev.	8.1	8.4	7.9

Table 2-4: Airborne Time Probability Bands of DTW to MEM (minutes)

Percentiles	1995	1998	2000
0.05	79.1	81.1	81.1
0.25	85.0	86.0	86.0
0.5	89	91	90
0.75	94.9	95.9	95.9
0.95	105.3	107.3	105.4

Table 2-5: DTW-MEM Airborne Time Statistics (Absolute Difference in Minutes)

Measure	Absolute Diff. (95-98)	Absolute Diff. (98-00)
Median	2.00	-1.00
Mean	1.64	-0.64
Std. Dev.	0.23	-0.54

Table 2-6: Number of Aircraft Handled by Center (000s)

ARTCC	1995	1998	2000	Chg. 95-98	Chg. 98-00	Chg. 95-00
ZME	2,003	2,144	2,232	7%	4%	11%
ZID	2,117	2,444	2,685	15%	10%	26%

Source: Administrators Fact Book, January 2002

Looking at scheduled block times for DTW and MEM illustrates how tightening of the airborne distribution impacts the airlines' scheduling (summarized by statistics in Tables 2-7 and 2-8). The scheduled block time follows a similar pattern to the airborne time. The median time increased by over 3 minutes between 1995 and 1998, but remained virtually the same between 1998 and 2000. Thus, despite growth in the ZME and ZID Centers one can conclude that URET might have prevented airlines from increasing their schedule block times, which in turn, may have reduced costs to the airlines.

Table 2-7: DTW-MEM Scheduled Block Time Statistics

Measure	DTW-MEM		
	1995	1998	2000
Min	107	112	113
Median	116.9	120	120.1
Mean	115.9	120.8	121.2
Max	125	129.6	135.7
Std Dev.	5.1	4.5	5.1

Table 2-8: DTW-MEM Scheduled Block Time Probability Bands

Percentile	1995	1998	2000
0.05	107.1	112.2	113.2
0.25	112.1	117	116.1
0.5	116.9	120	120.1
0.75	120	124	125
0.95	123.8	127.9	129.6

As shown in Table 2-9, patterns in the airborne distributions suggest that these distributions widened between 1995 and 1998, but started tightening up in 2000. From 1998 to 2000, the standard deviations tightened in 14 of the 28 city pairs from the 1995 to 1998 difference. These are all positive signs that, despite traffic growth at several of these city pairs, the airborne time distribution is getting tighter and shifting to the left.

In this case, mean and median airborne times are increasing at a much slower rate, and the predictability case is improving (smaller standard deviation). This improvement can result from URET and/or various factors; thus, one can conclude that URET might have contributed to the declining rate of airborne time increase.

Table 2-9: Summary of 28 City Pairs through ZMD and ZID (Standard Variations)

	Difference from 1995 to 1998					Difference from 1998 to 2000			
City Pair	Flights (in 1995)	Median (minutes)	Mean (minutes)	Std. Dev.	Flights (in 1998)	Flights (in 2000)	Median (minutes)	Mean (minutes)	Std. Dev.
ATL-CVG	3091	1	1.95	1.8	3970	3206	0	-0.62	-1.61
ATL-DAY	1475	1	0.66	1.04	1816	1794	0	0.6	-0.5
ATL-DEN	3256	0	-0.49	-0.41	3926	3515	3	3.41	0
ATL-MCI	2195	0	-0.37	-0.57	2515	2377	1	1.82	1.26
ATL-MSP	3118	-2	-1.97	-0.37	3432	3691	0	0.12	-0.33
ATL-ORD	8711	0	0.09	0.52	10120	9261	2	2.53	0.54
ATL-PIT	3920	1	0.81	0.33	3783	3182	0	-0.29	0.05
BWI-MDW	2449	0	0.19	-0.03	2820	4130	2	2.38	1.15
CMH-ORD	2137	-1	-0.78	0.8	2036	1971	2	2.02	0.07
DEN-IAD	2656	2.5	2.7	1.3	2928	3196	3.5	2.97	0.14
DTW-CVG	2527	0	0.63	0.06	1183	367	-0.94	-0.8	-0.28
DTW-MEM	2569	2	1.64	0.23	2795	3100	-1	-0.64	-0.54
IAH-DTW	1443	-1	-1.5	-1.16	1677	3298	2	2.49	1
IND-STL	4288	0.03	0.61	0.98	3785	3754	-0.03	-0.96	-1.28
LAX-IAD	3285	0	0.19	1.35	3941	4434	3	4.23	2.21
LGA-DFW	7153	-1	0.65	-0.49	6747	5916	-1	-1.8	0.89
MDW-CMH	4032	0.03	-0.18	-0.6	3663	2201	0	0.08	0.22
MEM-ATL	4813	0	0.93	0.84	4745	3718	-2	-2.55	-1.48
ORD-ATL	8369	-1	-0.18	1.76	10130	9256	0	-0.68	-1.57
ORD-CLT	3189	-0.97	-0.53	-0.53	3315	4144	1	1.04	0.47
ORD-IAH	6170	1	0.9	1.42	6128	6239	-1	0.12	0.93
ORD-MCO	4016	-1	0.14	0.73	3361	3680	1	0.69	-0.14
ORD-MIA	4486	-2	-1.99	0.74	5114	5149	1	0.39	-0.5
SDF-STL	3785	0.01	0.51	0.28	3750	3246	0	-0.62	-1.39
STL-CLE	4450	-0.94	-0.4	0.22	3287	2832	2	1.95	0.54
STL-DCA	2909	-1.03	-0.8	-0.13	2619	2356	1.09	1.46	0.38
STL-IND	4323	0.94	0.48	0.55	3821	3774	0.06	0.35	-0.36
STL-LIT	3467	0	0.51	0.47	3410	3351	-0.03	-0.62	-0.94

A similar pattern is not present, however, in several of the other city pairs. Tables 2-10 and 2-11 summarize the weighted average statistics over all these city pairs. Table 2-10 details the absolute differences in minutes/seconds, while Table 2-11 shows the percent difference. The changes in mean and median airborne times are about the same for both pre- and post-URET. The average standard deviation of airborne time rose by 7 percent between 1995 and 1998, but stabilized and decreased slightly by .43 percent between 1998 and 2000.

Table 2-10: Weighted Average Statistics for 28 City pairs (Absolute Differences in Minutes/Seconds)

Measure	Absolute Diff. 95-98	Absolute Diff. 98-00
Median	-0.17/-10.5	0.74/44.7
Mean	0.18/11.0	0.74/44.3
Std Dev	0.59/35.2	-0.04/-2.1

Table 2-11: Weighted Average Statistics for 28 City Pairs (Percent Differences)

Measure	% Diff. 95-98	% Diff. 98-00
Median	-0.18%	0.76%
Mean	0.19%	0.75%
Std Dev	7.15%	-0.43%

This suggests that the overall airborne time distributions remained the same while traffic increased between 1998 and 2000. This is a positive improvement to which URET *might* have contributed, but it is hard to assign an economic value because it did not result in reductions in airborne times.

Finally, for flights arriving into DCA, DFW, and ORD, we compared the airborne results to airborne changes at the arrival airports found in the ASD-400 study. This shows how this city pair performed compared to other city pairs flying into the same airport. If the statistics of the selected city pair were better than the ones found in ASD-400 study, then one could hypothesize that URET had a positive effect on reducing the increasing airborne time trends. We analyzed one example closely, Columbus, Ohio International Airport (CMH).

We compared all flights departing from CMH and arriving to ORD. From the ASD-400 study, ORD statistics are based on all arrivals into ORD. Tables 2-12 and 2-13 show the statistics for CMH-ORD and all ORD arrivals. The changes in mean and median airborne times are almost identical between CMH-ORD as well as all the other arrivals into ORD. However, the increase in standard deviation is somewhat less, (.07 minute) for CMH-ORD, compared to all other arrivals into ORD (.93 minute). This shows a similar conclusion, as do all the other ways we look at the distribution.

Table 2-12: CMH-ORD Airborne Time Statistics (Absolute Difference in Minutes)

Measure	Absolute Diff. 95-98	Absolute Diff. 98-00
Median	-1.00	2.00
Mean	-0.78	2.02
Std. Dev.	0.80	0.07
# of Flights	2036	1971

Table 2-13: ORD Airborne Time Statistics (Absolute Difference in Minutes)

Measure	Absolute Diff 95-98	Absolute Diff 98-00
Median	1.60	2.00
Mean	0.40	2.20
Std. Dev.	0.79	0.95

In short, URET might have contributed to stabilizing the actual airborne time distributions, i.e., no increase in standard deviations of airborne time distributions, despite traffic growth. If this is the case, then it may have prevented airlines from increasing their block times for some city pairs flying through the ZME and ZID Centers and, consequently, containing their operating costs. Nevertheless, these savings are not visible when compared to other flights that flew through the rest of the NAS. Therefore, it is difficult to assign an economic value to this possible improvement.

2.1.1.4 URET Summary Evaluation

In the bottoms-up approach that is presented above, we attempted to determine the impact of URET on flights through ZID and ZME since it was operational as a prototype. Using a limited, though representative set of city pairs with many flights, the intent of the evaluation was to identify the behavior of airborne and flight times.

The evaluation provided some indication of improvements in airborne time and the associated variability (standard deviation). These improvements were not compelling or conclusive. Since URET is an enabling technology for other programs and procedural initiatives, as well as being supported by programs such as the Weather Radar Processor (WARP), it is not readily apparent how to allocate the benefits/improvements between programs.

2.1.2 Case 2: TMA - Data Evaluation Perspective

ASD-400 performed this analysis by conducting a comprehensive analysis of 25 dominant origin airports with an emphasis on arrivals during the peak times. This approach evaluates the flight performance data at a TMA (SC) site before and after the implementation. This effort evaluates the post-implementation performance at MSP from a pre-acquisition flight performance baseline.

2.1.2.1 Flight Performance Data

The ASD-400 assessment team extracted performance data from CODAS for 1999-2000 and ASPM for January - July 2001. These two data sources provide a breakdown of flight-by-flight performance metrics including departing/arrival airport, scheduled/actual arrival/departure times, airborne time, and filed-ETE. The core metrics of our analysis are actual airborne time and filed-ETE.

To focus our analysis on consistent flight data, we evaluated 25 airports (by average historical total monthly demand) that had a majority of their flights flying into MSP in 1999 and 2000. The filter ensures that sporadic long- or short-haul flights do not influence the data. In Table 2-14, these 25 airports departing to MSP are listed in order of average monthly demand down, i.e., ORD has the most departures into MSP and IAH has the fewest departures into MSP.

Table 2-14: Departure Airports into MSP

ORD	MDW	MEM	PHL	CLE
DTW	ATL	EWR	LAX	BOS
STL	CVG	SEA	SFO	DCA
DEN	MCI	MKE	OMA	LGA
DFW	PHX	LAS	IND	IAH

The analysis was narrowed to the peak arrival times when a metering tool such as TMA would have the greatest impact. The peak arrival periods were determined using the relative local maxima to retrieve the four highest periods, or "spikes," of the day. Extracting the peak arrival times required querying four days of ETMS arrival (AZ) tables to develop a moving 30-minute arrival count. ETMS was drilled down further to develop flight counts for each 10-minute bin based on the moving 30-minute arrival counts. The following peak arrival times were extended to total three 10-minute bins following the beginning of the peak time (1020, 1210, 1740, and 1930 local times), giving 30-minute peak times (see Figure 2-1). Also, the monthly arrival demands into MSP were extracted from OPSNET.

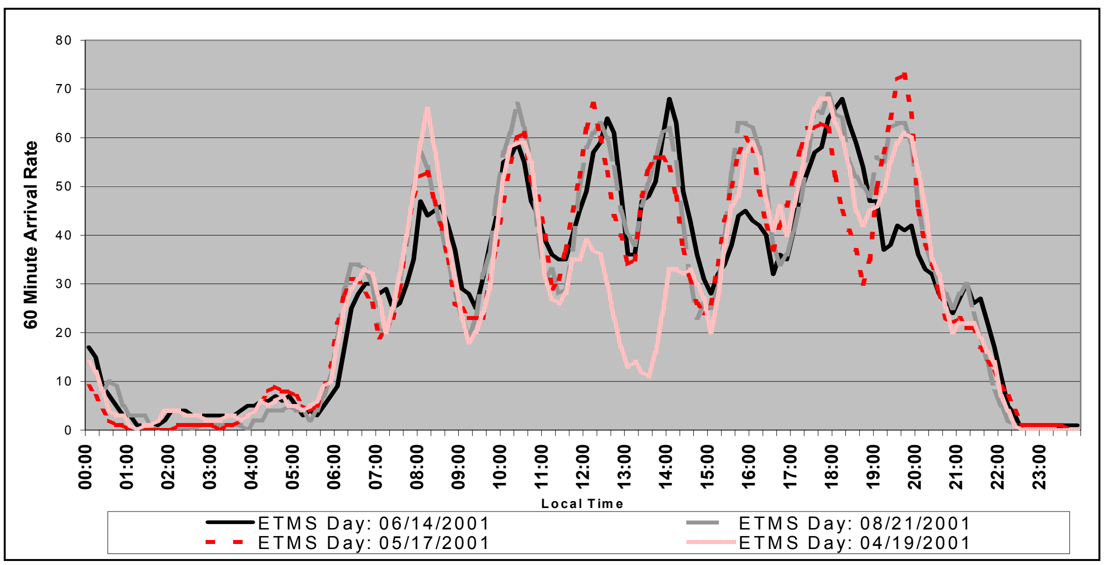

Figure 2-1: Hourly Flight Distribution at MSP

2.1.2.2 Weather Data

This analysis relied upon historical National Climatic Data Center (NCDC) surface hourly weather data to determine VFR and Non-VFR (MVFR and IFR) weather conditions at MSP. This data source provides hourly precipitation, ceiling, and visibility measurements. When weather conditions change, special observations within the hour are reported. The ceiling and visibility measurements provide the necessary metrics to determine whether VFR or Non-VFR conditions existed at the airport (see Table 2-15). The following guidelines were applied[4].

Table 2-15: Weather Criteria

Weather Condition	Minimums	Avg. Frequency
VFR	Ceiling >= 3500 feet AND Visibility >= 8 miles	68.5 %
Non-VFR	Ceiling < 3500 feet OR Visibility < 8 miles	31.5%

These criteria were assigned to each weather event. The weather events, with their respective assignments, were then mapped to those flights arriving during the weather period. Using this information, we determined which flights arrived during VFR or Non-VFR periods.

2.1.2.3 Methodology

2.1.2.3.1 Pre- and Post-Implementation Periods

The pre- and post-implementation analysis involves identifying three stages of the product usage: pre-implementation, transition period, and post-implementation. For the purposes of this study, the pre-implementation period is from January 1999 to the TMA's Initial Daily Usage. The transition period, which is identified by the program office as six months at ZMP, is defined by the onset of IDU to the day of PCA. PCA is defined by effective and fully qualified use of TMA. The post-implementation period is the period following the PCA date to July 2001. Table 2-16 shows the pre- and post-implementation dates for ZMP TMA.

Table 2-16: Key TMA Milestones at MSP

Milestone	Timeframe
Pre-Implementation Period	**01/01/1999 – 06/22/2000**
IDU	06/22/2000
Transition Period	06/22/2000-12/20/2000
PCA	12/20/2000
Post-Implementation Period	**12/21/2000 – 07/01/2001**

2.1.2.3.2 Average Airborne and Filed-ETE Times

Using the January 1999 – July 2001 performance data, we developed average airborne and filed-ETE times for each of the top 25 origin airports arriving at MSP for each month. The city pair averages were further averaged into monthly average airborne and filed-ETE times arriving into MSP. These monthly averages allow us to compare similar periods' historical performance and provide a basis to develop a forecast.

2.1.2.3.3 Forecasting with Exponential Smoothing

We used the SPSS Exponential Smoothing Model to project the pre-implementation period (January 1999 – June 2000) average, monthly airborne times into the post-implementation months for each data set that was categorized by weather condition, i.e., All, VFR and Non-VFR. Filed-ETEs were

[4] Source: MSP tower input from ASD-400/ATP-100 - 2000 Airport Capacity Survey.

evaluated in the All data set. To ensure an accurate assessment, the data within the transitional period before the PCA period was not analyzed. For each data set, the most appropriate model was a Damped curve model with a Least-Sum-of-Square-Error fit test.

2.1.2.4 Analysis

2.1.2.4.1 Demand

The historical arrival demand from OPSNET[5] was summarized within our definition of pre- and post-implementation for similar six-month periods. The six-month demand averages are grouped accordingly as shown in Table 2-17.

Table 2-17: Average Number of Daily Arrivals

Timeframe	Jan. 1999 – Jun. 1999	Jan. 2000 – Jun. 2000	Jan. 2001 – Jun. 2001
Number of Arrivals	685	708	708

2.1.2.4.2 Performance

The resulting data from the methodology provides performance figures for both Airborne and Filed-ETE times. Figures 2-2 and 2-3 illustrates separation by weather conditions and shows the following:

The system state of each graph includes:

- Pre-Implementation period
- IDU and PCA dates
- Transitional period
- Post-Implementation period
- Forecasted Post-Implementation period

Each graph also has performance metrics of:

- Average Daily Demand by month
- Pre/Post-Implementation Actual/Forecast Airborne times
- Pre/Post-Implementation Actual/Forecast-Filed-ETE times (All Case)

Tables 2-18 and 2-19 (following each of the graphs) have additional metrics showing:

- Comparative six-month average airborne and filed-ETE times
- Post-implementation forecast based on all pre-implementation data

[5] The daily reporting is collected from FAA Forms 7230-1, Airport Traffic Record, 7230-12, Instrument Approaches Monthly Summary and 7230-26, Instrument Operations.

Figure 2-2 shows the Airborne and Filed-ETE performance for MSP aggregated for all surface weather conditions.

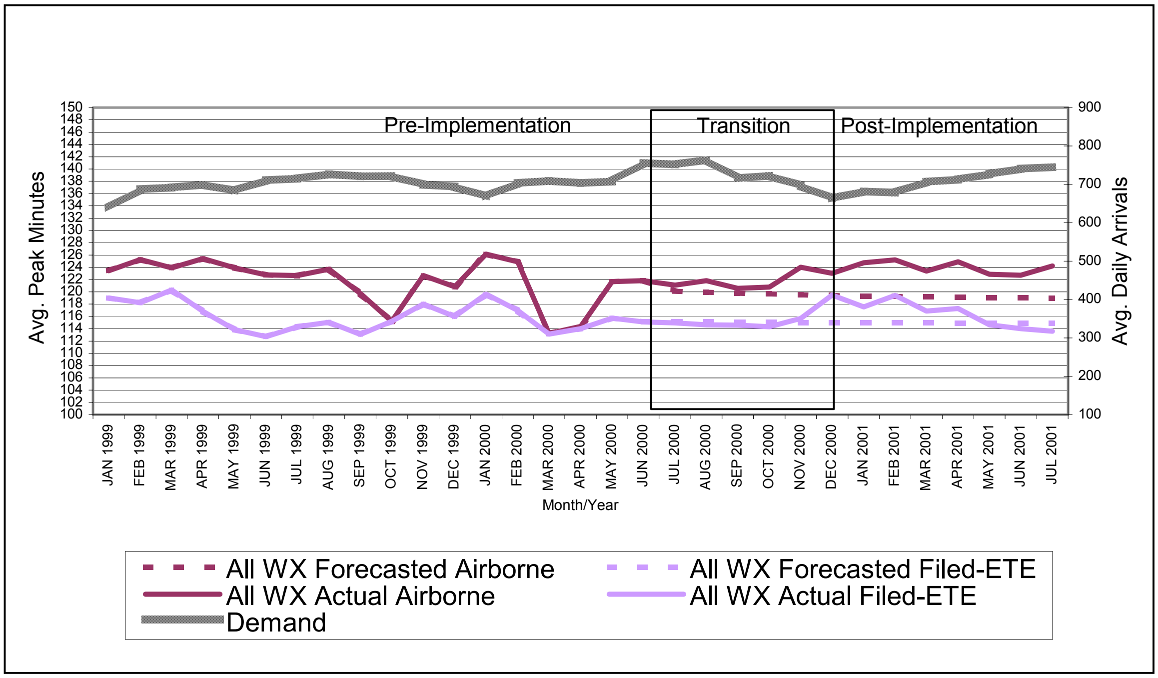

Figure 2-2: Pre/Post-Implementation Results (All Weather)

The data in Table 2-18 represents flight performance times during all weather conditions for similar six-month periods in each year. The resulting data shows a 3.7-minute decrease in average airborne time between the six-month periods in 1999 and 2000. However, the results show a 3.6-minute increase in the average airborne time between the six-month periods in 2000 and 2001. Also, from 1999 to 2001, the results show a .1-minute decrease.

Table 2-18: Summary of Flight Performance (in minutes) – All Weather Conditions

Timeframe	Jan. 1999 - Jun. 1999		Jan. 2000 - Jun. 2000		Jan. 2001 - Jun. 2001		Forecast	
	Airborne	Filed-ETE	Airborne	Filed-ETE	Airborne	Filed-ETE	Airborne	Filed-ETE
Time	124.1	116.8	120.4	115.8	124.0	116.6	119.2	114.9

The forecast produced a trend that showed a decrease in airborne times during the post-implementation period. This information suggests that the airborne times were decreasing over time during the pre-implementation period and were forecasted to continue decreasing. The actual post-implementation period, however, shows a marginal increase for all evaluated flights during all weather conditions.

The resulting data shows a 1.0-minute decrease in average filed-ETE times from the six-month period in 1999 to 2000. The filed-ETE times for the six-month period in 2001 show an increase of .8 minutes from the six-month period in 2000. The filed-ETE times for the six-month period in 2001 shows a marginal decrease of .2 minutes from the six-month period in 1999. The forecasted average filed-ETE times were predicted to continue decreasing through the post-implementation six-month period, however, the actual data show an increase. This data suggests the airlines were decreasing their filed-ETE times through 2000, but increasing them in 2001. Figure 2-3 shows the airborne times during VFR weather conditions for MSP.

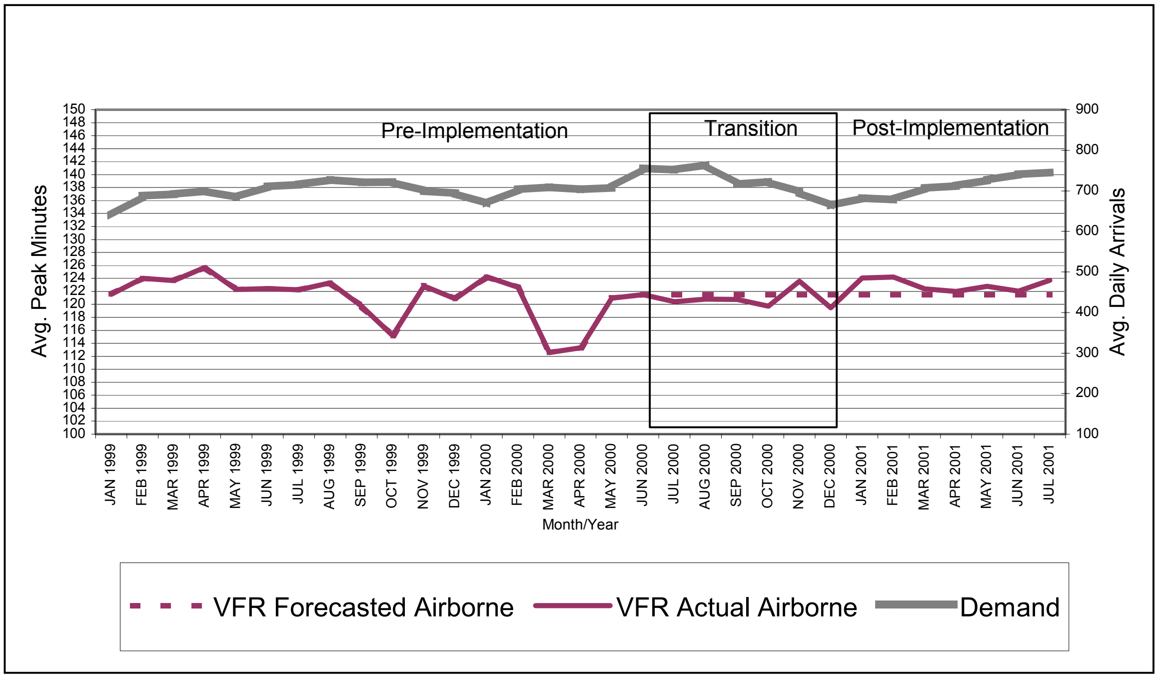

Figure 2-3: Pre/Post-Implementation Results (VFR)

The data in Table 2-19 represents airborne times during VFR weather conditions for similar six-month periods in each year. The resulting data shows a 4.1-minute decrease in average airborne time in the six-month periods from 1999 to 2000. The results show a 3.7-minute increase, however, in the six-month periods from 2000 to 2001. Also, the results show a .4-minute decrease in the six-month periods from 1999 to 2001.

Table 2-19: Forecast Airborne Times (minutes) - VFR Conditions

Timeframe	Jan. 1999 - Jun. 1999	Jan. 2000 - Jun. 2000	Jan. 2001 - Jun. 2001	Forecast
	Airborne	Airborne	Airborne	Airborne
Time	123.3	119.2	122.9	121.5

The forecast produced a trend that showed an increase in airborne times during the post-implementation period. This information suggests that the airborne times were decreasing, followed by an increase during the pre-implementation period and were forecasted to continue increasing. The actual post-implementation period shows a marginal decrease while the forecast times increase.

Figure 2-4 shows the performance of airborne times during Non-VFR weather for MSP.

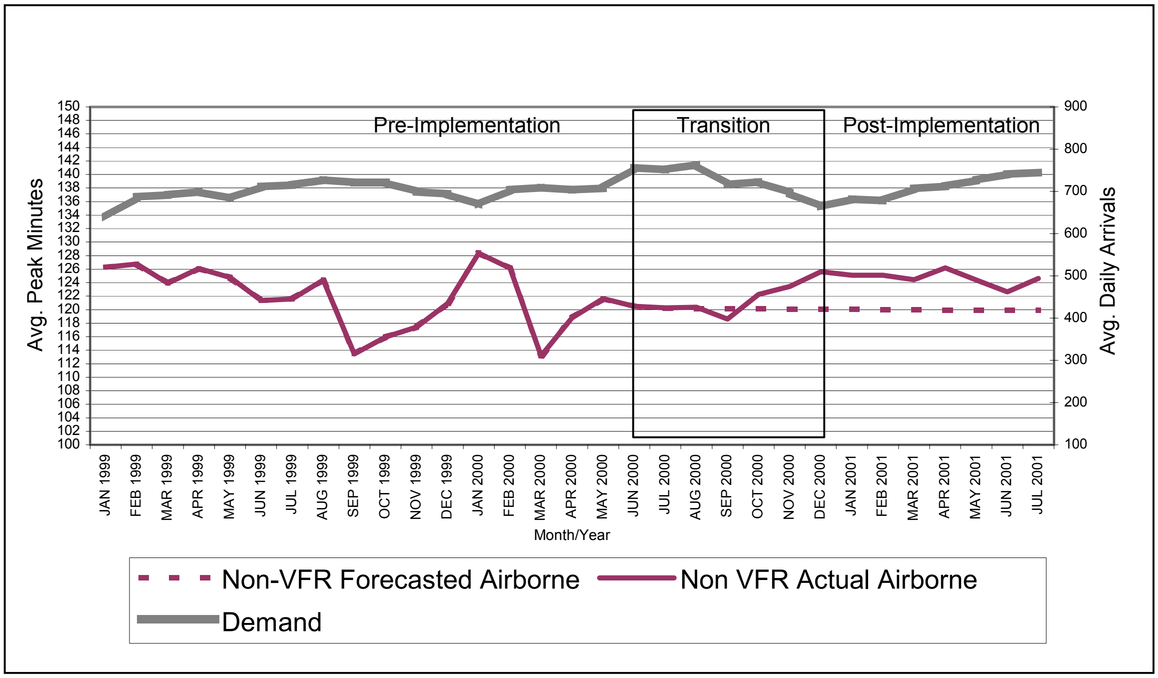

Figure 2-4: Pre/Post-Implementation Results (Non-VFR)

The data in Table 2-20 represents airborne times during Non-VFR weather conditions for similar six-month periods in each year. The resulting data shows a 3.5-minute decrease in average airborne time in the six-month period from 1999 to 2000. The results show a 3.2-minute increase, however, from the six-month period from 2000 to 2001. Also, the results show a .3-minute decrease from the six-month period of 1999 to 2001.

Table 2-20: Forecast Airborne Times (minutes) - Non-VFR Conditions

Timeframe	Jan. 1999 - Jun. 1999	Jan. 2000 - Jun. 2000	Jan. 2001 - Jun. 2001	Forecast
	Airborne	Airborne	Airborne	Airborne
Time	124.9	121.4	124.6	120.0

The forecast produced a trend that showed a decrease in airborne times during the post-implementation period. This information suggests that the airborne times were decreasing during the pre-implementation period and were forecasted to continue decreasing. The actual post-implementation period shows a marginal decrease smaller than the forecasted decrease.

.

2.1.2.5 Conclusions - TMA

The analysis reveals that from the time TMA has been operational, since PCA, the average airborne times from the 25 airports have increased during the six-month evaluation period by 3.6 minutes (3 percent) from the 2000-evaluation period (pre-IDU). With virtually identical daily demand, it is apparent that the implementation of TMA has not improved airborne flight times despite AOZ's (Free Flight Program Office) observations that airport acceptance rates have been increasing during the peak arrival rates. Additionally, examination of post-TMA adherence of carrier filed-ETE time, an indicator of an airline's schedule predictability, reveals that airlines have not adjusted their scheduled filed-ETE times to be consistent with the distribution in airborne times that are occurring with time-based metering.

2.2 Conclusions

This analysis revealed the following preliminary results at the evaluated sites for both URET and TMA:

- URET provided limited indication of improvements in airborne time and the associated variability (standard deviation). These improvements were not compelling or conclusive.
- TMA is not showing flight time reductions (filed flight plans and airborne times) since it has been deployed and declared PCA through from December 2000 through August 2001.[6]

2.3 Recommendations

Given that the appropriate data is accessible, usable, and is of ample quality to evaluate key metrics, e.g., flight time and airborne time, a plan needs to be developed that gives emphasis to formal reviews of key FAA acquisition programs, i.e., the agency needs to define its PIR policy and incorporate the policy into the AMS. If the policy is put into service, the Agency should move forward aggressively with a well-documented plan that describes a sound data collection and evaluation process for conducting these reviews. With the emphasis for better supporting the FAA's focus as a performance-based organization (PBO), it is critical that the resources are made available to follow-up on the flight performance behavior of key programs through a data-driven approach of key programs as they are deployed.

Consideration should be given to several statistical tests for post-implementation benefits analysis described in *The Art of Benefits Prediction and the Statistical Science of Post-Implementation Analysis in Aviation Investment Analysis [2]*. Moreover, with the advent of the preliminary Operational Evolution Plan (OEP) Metrics Plan, which addresses four core quadrants (i.e., arrival departure rates, en route congestion, airport weather conditions, and en route severe weather), core metrics need to be tracked to measure the performance of both the NAS and individual programs, whenever possible. Adherence to related elements of the OEP through post-implementation evaluations will provide the FAA with much needed accountability to its key acquisitions once they are operational.

[6] Note: This finding differs from the conclusions from the *FFP June 2002 Performance Metrics Report* where the primary claimed benefit at ZMP is increased actual peak arrival and operations rates. ASD-400's focus was on the behavior of time for flights that utilized the TMA.

References:

1. An Approach for Developing a Standard Method for Conducting Post-Implementation Reviews, June 2001
2. ASD-400 Independent Benefits Assessment of the Free Flight Phase 2 Program, June 2002
3. The Art of Benefits Prediction and the Statistical Science of Post-Implementation Assessment in Aviation Investment Analysis, June 2002, http://www.faa.gov/asd/ia-or/or-reports.htm
4. Investment Analysis Benefit Guidelines: Quantifying Flight Efficiency Benefits, June 21, 2001
5. OMB Circular No.A-11, Post-Implementation Review http://www.whitehouse.gov/omb/circulars/a11/capman.pdf
6. Pre- and Post- Implementation Analysis of ITWS at EWR (Draft), August 2002
7. Department of Energy document, Chapter 4: Management-in-use-phase