

MTR 97W0000035

MITRE TECHNICAL REPORT

National Airspace System Performance Measurement: Overview

August 1997

Catherine N. Bolczak
Dr. Jonathan H. Hoffman
Anne J. Jensen
Dr. William W. Trigeiro

Sponsor: Federal Aviation Administration
Dept. No.: F066

Contract No.: DTFA-01-C-00001
Project No.: 0297012K-AA

MITRE

Center for Advanced Aviation System Development
McLean, Virginia

Abstract

One of the major goals of the Federal Aviation Administration (FAA) is to ensure the safe, efficient, and orderly movement of air transportation through the air traffic management (ATM) system, including both domestic and oceanic airspace. Historically, the performance of the ATM system and its impacts on the operating efficiency of its users have been measured primarily with delay metrics. While delay metrics provide insights into some aspects of ATM system performance, they do not provide a complete picture. Many other aspects of performance must also be assessed if the FAA is to evolve the ATM system to meet changing user expectations and needs. Recognizing this, the FAA, through the Office of System Capacity (ASC) the Oceanic and Offshore Integrated Product Team (IPT), has been working to enhance its ATM system performance measurement capabilities. The initial set of ATM system performance indicators and metrics discussed in this paper have been and will continue to be refined, more efficient metrics calculation capabilities are being developed, and efforts are underway to embed the metrics throughout the FAA's business processes, such as budgeting and strategic planning. The report provides an overview of the ATM system performance measurement approach and associated tasks.

KEYWORDS: System Performance, Performance Measurement, Indicators, Metrics

Executive Summary

One of the goals of the Federal Aviation Administration (FAA) is to support the safe and efficient movement of air transportation. The FAA is striving to provide a high level of service by continually improving its operational efficiency and effectiveness while maintaining or improving the current level of safety.

Historically, throughput and delay have been used to measure the effectiveness of the air traffic management (ATM) system and the impact that the system has on the operating efficiency of its customers. While delay is an adequate measure of operational effectiveness in some instances, it does not present a complete picture of the many aspects of performance that determine the quality and level of service the users receive. Recognizing this, the FAA is working to improve the approach and metrics it uses to assess its performance and the level of service that domestic and oceanic airspace users receive.

At the highest level, the objective of this project is to pioneer the use of a new set of user-oriented performance metrics by the FAA. This new set presents a more composite view of operational effectiveness and customer outcomes encompassing the need for increased flexibility, predictability, access, and delay, as discussed below. The immediate objective of this project is to establish a set of indicators through identification of specific metrics, exploration of data and assessment techniques, and investigation of the utility and viability of resulting measurements. The focus of this paper is primarily on the establishment of the indicators and metrics with a brief overview of the analytical work performed to date.

This emphasis on serving the customer is summarized in the Air Traffic Services (ATS) Performance Plan [7] as follows:

ATS has focused on the aviation community as its direct customer and recognizes that these customers desire to be better served, not controlled. These customers need different and better types of services if they are to remain economically viable, and serve the ultimate public interest in safe, affordable air travel. ATS services must enable our customers to accomplish their operational objectives in a safe and efficient manner and must also be mutually agreeable to all operational components of the ATS organization. ATS also bears a responsibility to the American taxpayers to provide these services to the industry and the flying public at a reasonable cost.

A key to this performance measurement is a focus on user-oriented outcomes rather than outputs. As the terms are used in this work:

- An *outcome* metric is used to measure the extent to which an organization's broader goals are being met. The focus with outcome measurement is not on the work that is being done, but the results of that work. For the FAA, outcomes represent the *externally-focused*, bottom-line results of the delivery of services to the ATM system

by the FAA *as perceived by customers*. Thus, they are user-oriented measures of ATM system performance.

- An **output** metric is used to measure the results of an organization's immediate activities. Outputs are measures of ATM system performance from the *internally-focused view of the service provider*, namely the FAA.

The FAA has identified four categories of system performance indicators to measure capacity and efficiency: flexibility, predictability, access, and delay. Each of the categories is described briefly in detail in this document. It must be emphasized that safety will not be compromised as the FAA strives to make the ATM system more flexible, more predictable, and more accessible, and to reduce delays. The FAA also recognizes that it may not always be possible to improve flexibility, predictability, access, and delay simultaneously. There may be times when it will be necessary to balance one against another. Moreover, the appropriate balance may change as the FAA gains experience with these indicators and the needs of the user community change.

The **flexibility** indicators are measures of the ability of the system to meet users' changing needs and to permit users to adapt their operations to changing conditions. In the past, the FAA has assumed that it could estimate what values to optimize with respect to user operations. Delays on the ground were assumed to be preferred to airborne delays, to minimize fuel costs. In reality, users may desire to optimize their operations based on time-efficiency, fuel-efficiency, or distance-efficiency (flight trajectory), depending on their needs. For example, commercial air carriers may desire a flight to delay arrival in order to ensure the availability of a gate, or to allow another flight to arrive first.

The **predictability** indicators are measures of the variation in the ATM system as experienced by the user. This variability is a result of the inherent uncertainty that accompanies the operation of the complex aviation system caused by such factors as changes in meteorological conditions and traffic demand (especially visual flight rules (VFR) weather). For users that rely on schedules, this variability is a key parameter used in planning and operating flights. Statistics are used to describe the variation of system performance from day to day.

The **access** indicators are measures of the ability of users to enter the ATC system and obtain services on demand. The indicators reflect the quality and level of service, as well as the availability of system resources. These resources include physical resources such as airspace, airports, and other facilities, and information resources, such as Flight Services.

The **delay** indicators are measures of the amount of time beyond expectations that it takes to complete an operation. Delay has traditionally been one of the key indicators for ATM system performance. Although a wide variety of delay metrics have been used, the delay metric currently used for reporting performance is a count of the number of flights delayed more than

15 minutes. Delay continues to be an important indicator. However, more sophisticated delay metrics will provide a more complete picture of ATC system performance.

The new delay indicators are not meant to necessarily be consistent with current delay metrics. In fact, because user expectations are a part of the definition of delay as presented above, one should not expect perfect consistency. In addition, note that although in general one can say that more flexibility, predictability, and access can probably be assumed to be indicators of better performance, less delay (in the traditional sense) may not always be better. Delay does not necessarily have a negative impact. If a small increase in delay should be the price of free routing through airspace, for example, many users might be willing to accept an increase in delay, especially if taken on the ground. Therefore, delay in this baseline is compared to user expectations, not to some absolute “minimum time” standard.

The FAA’s measurement system is focused on the parts of the ATM system which directly affect user services and which the FAA can control. This focus is intended to help identify problems and opportunities and, when they are identified, to enable the FAA to make worthwhile, targeted, performance improvement changes in the ATM system. The measurement approach and metrics used to calculate performance have been, and will continue to be, developed in consultation with the user community. It is essential that the metrics represent the operational outcomes desired by the users and that they reflect the many different aspects of ATM system performance of interest to them. In addition, it is essential that the metrics evolve as the ATM system, user needs, and financial, technical, and political environments change.

The measurement process is focused primarily on the system-wide, high-level efficiencies and inefficiencies that are dictated by ATM system capacity. The basic ideas can be applied more broadly, as well as to the measurement of performance on a localized basis. In addition, it is important to note that there are many other aspects of ATM system performance that could be measured, but which are not within the scope of this project. In a sense, maintaining or improving safety acts as a constraint that leads to tradeoffs among performance indicators, since the FAA has repeatedly and explicitly stated that ATM system efficiency performance will not be improved at the expense of safety. Other important performance aspects such as project management, cost control, and innovation are also not addressed by this initiative because they are not part of operational ATM system performance. In addition, the impacts of ATM system performance on factors such as the United States’ gross domestic product and the investment infrastructure are not addressed because they are not directly related to the level of services users receive.

Besides producing performance information, the act of measuring performance almost always has residual and often unintended results. The following considerations were developed to help guide development of the ATM system performance measurement approach:

- Users’ current demands on the ATM system do not necessarily reflect their true needs.

- You get what you measure.
- If you measure something, people will expect that you have the authority and resources to improve it.
- Regardless of their initial intent, metrics often inadvertently turn into goals.
- The metrics set conveys information about what is important to the organization.
- Measurement teams must solicit and receive inputs to, and feedback on, metrics sets prior to their implementation.

Section 3 contains descriptions of the completed, on-going, and planned ATM system performance measurement tasks. In particular, this includes a summary of the initial results of this project, which have become known as the “January 1996 Baseline,” as well as a summary of selected work completed since then.

Acknowledgments

The authors would like to thank the numerous individuals who supported the FAA's ATM system performance measurement project. Our sponsors in the Office of System Capacity (ASC), Mr. Bill Voss (now AUA-501) and Mr. Carl Schellenberg, provided invaluable insights and feedback throughout multiple phases of the project. They were supported by other members of ASC including Ms. Paula Lewis, Mr. Ken Peppard, and Mr. Rich Nehl. In addition, Ms. Lee Brown, CAASD's representative to Air Traffic, was instrumental in taking the initial list of indicators and developing the approach and metrics required to assess them. Other CAASD staff members, including Mr. Sam Bowden, Mr. David Burke, Mr. Elza Hisel, Mr. Satish Mohleji, Mr. Michael White, and Mr. Joshua Hochberg (formerly of CAASD), provided additional technical support.

Table of Contents

Section	Page
1 INTRODUCTION	1
1.1 Outcomes Versus Outputs	2
1.2 Background	2
1.2.1 Performance Measurement Needs	2
1.2.2 Project Objectives	3
1.2.3 Definitions	3
1.3 Flexibility, Predictability, Access, and Delay	4
1.4 ATM System Performance Measurement Focus	6
2 PERFORMANCE MEASUREMENT APPROACH	9
2.1 Philosophy and Tenets	9
2.2 Factors to Consider When Developing Metrics	10
2.3 Planned Measurement Process	12
3 PROGRESS AND PLANS	13
3.1 Metrics Project History	14
3.2 Initial Performance Categories and Indicators	16
3.3 January 1996 Baseline Metrics	18
3.4 Metrics for Oceanic Operations	19
3.5 Additions and Refinements to the Baseline Metrics	20
3.5.1 ARTCC Outages	21
3.5.2 Military Aviation and Military Airspace	21
3.5.3 Pilot's Opinion Survey	23
3.5.4 Workshop on FAA System Performance Measures	24
3.5.5 Metrics Calculation Enhancements	24
3.5.6 Indicators and Metrics Addressed Outside CAASD	25
3.6 Current Set of Performance Indicators and Metrics	25
3.7 Next Steps	28
List of References	31
Glossary	33

List of Figures

Figure		Page
1	Metrics Identification, Refinement and Usage Process	12
2	Project Flow and Tasks	13

List of Tables

Table		Page
1	FAA's Initial Performance Categories and Indicators	17
2	January 1996 Baseline Indicators and Metrics	19
3	Oceanic Baseline Indicators and Metrics	20
4	Indicators to be Supported by the Pilot's Opinion User Survey	23
5	Current Indicators and Metrics	26

Section 1

INTRODUCTION

One of the goals of the Federal Aviation Administration (FAA) is to support the safe and efficient movement of air transportation. The FAA is striving to provide a high level of service by continually improving its operational efficiency and effectiveness while maintaining or improving the current level of safety.

Historically, throughput and delay have been used to measure the effectiveness of the air traffic management (ATM) system and the impact that the system has on the operating efficiency of its customers. While delay is an adequate measure of operational effectiveness in some instances, it does not present a complete picture of the many aspects of performance that determine the quality and level of service the users receive. Recognizing this, the FAA is working to improve the approach and metrics it uses to assess its performance and the level of service that domestic and oceanic airspace users receive. Ultimately, the FAA desires to make the system more responsive to the needs of users.

Underlying the measurement of system efficiency is the assumption that safety will be maintained or improved throughout any attempt to monitor and improve system performance. The responsibility for measuring safety belongs to the FAA Safety Office and is not addressed with this effort.

The MITRE Corporation's Center for Advanced Aviation System Development (CAASD), in partnership with FAA's Office of System Capacity (ASC), has been working to develop a new ATM system performance measurement system. The objective of this report is to describe the philosophy, definitions and approach, progress made to date, and current and future plans for the use of a new set of user-oriented performance metrics by the FAA that relates to its overall operations in domestic and oceanic airspace.

Some of the material included in this document was provided to FAA's Air Traffic Services (ATS) for use in the development of the *ATS Performance Plan* [7]. The efforts which led to these two documents have been closely coordinated. The *ATS Performance Plan* states the focus on user outcomes very clearly:

ATS has focused on the aviation community as its direct customer and recognizes that these customers desire to be better served, not controlled. These customers need different and better types of services if they are to remain economically viable, and serve the ultimate public interest in safe, affordable air travel. ATS services must enable our customers to accomplish their operational objectives in a safe and efficient manner and must also be mutually agreeable to all operational components of the ATS organization. ATS also bears a responsibility to the American taxpayers to provide these services to the industry and the flying public at a reasonable cost.

1.1 Outcomes Versus Outputs

The terms “outcomes” and “outputs” distinguish between measuring what an organization does (outputs) and the results it wishes to achieve (outcomes). These terms are borrowed from the language of the Government Performance and Results Act of 1993 [Public Law 103-62] (GPRA), which mandates that federal government organizations develop strategic and annual plans which include their mission statements, long and short term goals, and metrics for measuring progress toward those goals. The distinction between these two terms is critical to understanding the motivation for this metrics effort. As the terms are used in this work:

- An *outcome* metric is used to measure the extent to which an organization’s broader goals are being met. The focus with outcome measurement is not on the work that is being done, but the results of that work. For the FAA, outcomes represent the *externally-focused*, bottom-line results of the delivery of services to the ATM system by the FAA *as perceived by customers*. Thus, they are user-oriented measures of ATM system performance.
- An *output* metric is used to measure the results of an organization’s immediate activities. Outputs are measures of ATM system performance from the *internally-focused view of the service provider*, namely the FAA.

An example of an aviation-related output measure would be the number of aircraft handled by the system. An associated outcome measure would evaluate the level of service received by those flights, such as the level to which the user preferences in schedule and route were granted.

1.2 Background

1.2.1 Performance Measurement Needs

The FAA’s ATM system performance measurement effort is being driven by the need to allocate scarce resources where they will be most beneficial, the need to continually improve ATM services, and the need to abide by various government requirements. GPRA mandates that federal government organizations develop multi-year strategic plans and annual plans that include their mission statements, long and short term goals, and metrics for measuring progress toward those goals.

The plans called for by GPRA must include planned performance improvement projects and the budget allocation requests needed to fund them. These plans must be periodically updated, identifying the goals that were achieved and explaining why any goals were not achieved. The philosophy behind this approach is to manage for results by focusing attention and resources on those activities which support the organization’s ultimate mission and goals. Of importance here, according to GPRA, output measurement alone is not sufficient for performance planning—the measurement of outcomes is essential.

In accordance with GPRA, the FAA's Oceanic and Offshore Integrated Product Team (IPT) developed oceanic service, acquisition, program cost, and team performance metrics which were approved by the General Services Administration (GSA), the Department of Transportation (DOT), and the FAA's Office of Information Technology (AIT) in February 1995. Some of these metrics are intended to measure the benefits to the users and operators resulting from procedure and system modifications, and help the FAA in decision-making to prioritize future system requirements cost effectively.

1.2.2 Project Objectives

At the highest level, the objective of this project is to pioneer the use of a new set of user-oriented performance metrics by the FAA. This new set presents a more composite view of operational effectiveness and customer outcomes encompassing the need for increased flexibility, predictability, access, and delay, as discussed below. The immediate objective of this project is to establish a set of indicators through identification of specific metrics, exploration of data and assessment techniques, and investigation of the utility and viability of resulting measurements. The focus of this paper is primarily on the establishment of the indicators and metrics with a brief overview of the analytical work performed to date.

Key to performance improvement are the metrics for use in overall operational improvement as well as metrics which focus on geographically localized or event-specific details. These include daily reports for use in tactical review of a previous day's performance at a specific location (e.g., airport or sector of airspace) in order to diagnose potential problems and their resolution. As important as these other types of metrics are, they are not a part of the initial metrics effort reviewed in this document, which concentrates primarily on improving operational efficiency of the ATM system by monitoring long-term patterns of performance throughout the ATM system. Work is underway within the FAA, CAASD, and elsewhere to develop these. The efforts represented here are being coordinated broadly, but are distinct tasks.

1.2.3 Definitions

Throughout this paper there are references to various measurement concepts. These concepts are key to understanding the content of this paper and are defined here for clarity.

Indicators and metrics. These terms are not interchangeable, and it is imperative that the reader be able to distinguish between them.

An *indicator* is a high level concept which describes how the FAA will meet its long term strategic goals. Each indicator captures some aspect of system capacity or performance, and may be targeted towards a specific user population (e.g., major airlines, general aviation).

A *metric* is a specific mechanism used to quantify an indicator. This mechanism includes the method of calculation and the data source used.

One indicator may have more than one metric. Given the differences between operations in the various phases of flight, there may be more than one way to quantify a given concept. For example, predictability has distinct metrics for ground movement and airborne movement.

Measurement baseline. The baseline that is being measured in current analyses is one view of current system performance. Where possible, this view is shown over periods of recent time (e.g., average taxi time for the last ten months). However, the baseline should not be viewed as a measure of the performance goals for system operations. The baseline provides a reference to assess the impact of any potential change in the system.

One early insight gained from the early measurement baseline work is that the system is highly adaptive. When an improvement is made to the system, users soon adjust their schedules to make use of the new capacity and (to choose a familiar example) delay returns to its previous level. Therefore, a snapshot of the system at a given time may not convey a complete story. Therefore, trend analyses using a consistent set of metrics are essential to measure system performance over time.

Performance goals will be established by FAA management—the role of the system performance baseline may be to provide information for their decision process and to provide historical context.

1.3 Flexibility, Predictability, Access, and Delay

The FAA has identified four categories of system performance indicators to measure capacity and efficiency: flexibility, predictability, access, and delay. Each of the categories is described briefly below and in more detail later in this document. It must be emphasized again that safety will not be compromised as the FAA strives to make the ATM system more flexible, more predictable, and more accessible, and to reduce delays. The FAA also recognizes that it may not always be possible to improve flexibility, predictability, access, and delay simultaneously. There may be times when it will be necessary to balance one against another. Moreover, the appropriate balance may change as the FAA gains experience with these indicators and the needs of the user community change.

The *flexibility* indicators are measures of the ability of the system to meet users' changing needs and to permit users to adapt their operations to changing conditions. In the past, the FAA has assumed that it could estimate what values to optimize with respect to user operations. Delays on the ground were assumed to be preferred to airborne delays, to minimize fuel costs. In reality, users may desire to optimize their operations based on time-efficiency, fuel-efficiency, or distance-efficiency (flight trajectory), depending on their needs. For example, commercial air carriers may desire a flight to delay arrival in order to ensure the

availability of a gate, or to allow another flight to arrive first. From the *ATS Performance Plan* [7]:

“National airspace users expect more from the ATM system than strictly the minimization of delay. They desire the capability to optimize their operations based on their own objectives and constraints, which vary flight-by-flight and user-by-user. Measuring the flexibility of the ATM system allows ATS to evaluate its ability to permit users to adapt their operations to changing conditions. ATS has developed a new family of measures that quantify the flexibility of the ATM system from the perspective of the user.”

The *predictability* indicators are measures of the variation in the ATM system as experienced by the user. This variability is a result of the inherent uncertainty that accompanies the operation of the complex aviation system caused by such factors as changes in meteorological conditions and traffic demand (especially visual meteorological conditions). For users that rely on schedules, this variability is a key parameter used in planning and operating flights. Statistics are used to describe the variation of system performance from day to day. Again from the *ATS Performance Plan* [7]:

“...schedules are central to the operations of almost all commercial flights, driving crew scheduling, ground service operations, and other operational components. Even the smallest deviation from the planned schedule can cause drastic impacts. This unpredictability often manifests itself as “pads” in flight schedules. These time inefficiencies in schedules can result in high operating and opportunity costs to users.”

The *access* indicators are measures of the ability of users to enter the ATC system and obtain services on demand. The indicators reflect the quality and level of service, as well as the availability of system resources. These resources include physical resources such as airspace, airports, and other facilities, and information resources, such as Flight Services. According to the *ATS Performance Plan* [7]:

“Access to the ATM system and the services provided by ATS is the basic need of all airspace users. Users need to access airports, airspace, and services.”

The *ATS Performance Plan* goes on to describe several key aspects of system access:

- Most users gain access to the ATM system through airports.
- A critical component is general aviation access to ATC services.
- Another critical issue is utilization of SUA by both military users and civilian users.
- Access that users have to crucial information needed to operate flights is another aspect, in particular briefings on weather conditions and other factors associated with the flight environment which are primarily provided directly to pilots before and during flights.

The *delay* indicators are measures of the amount of time beyond expectations that it takes to complete an operation. Delay has traditionally been one of the key indicators for ATM system performance. Although a wide variety of delay metrics have been used, the delay metric currently used for reporting performance is a count of the number of flights delayed more than 15 minutes. Delay continues to be an important indicator. However, more sophisticated delay metrics will provide a more complete picture of ATC system performance.

The new delay indicators are not meant to necessarily be consistent with current delay metrics. In fact, because user expectations are a part of the definition of delay as presented above, one should not expect perfect consistency. In addition, note that although in general one can say that more flexibility, predictability, and access can probably be assumed to be indicators of better performance, less delay (in the traditional sense) may not always be better. Delay does not necessarily have a negative impact. If a small increase in delay should be the price of free routing through airspace, for example, many users might be willing to accept an increase in delay, especially if taken on the ground. Therefore, delay in this baseline is compared to user expectations, not to some absolute “minimum time” standard.

1.4 ATM System Performance Measurement Focus

One of the FAA’s primary goals is to provide a highly efficient service to ATM system users. This is a complex task, as users may include any person or organization who uses the services of the ATM system: commercial air carriers (majors and regional airlines and air taxis; passenger and cargo carriers), military flights, and general aviation pilots (commercial and private), plus manufacturers of airframes [6]. From many users’ perspectives, a high level of service means:

- approval for their desired flight plans which get them to their destinations on time, cost effectively, and safely;
- approval for flight plan changes; and
- easy access to timely and accurate information, reliable equipment, and helpful and informative ATM services.

The FAA’s measurement system is focused on the parts of the ATM system which directly affect user services and which the FAA can control. This focus is intended to help identify problems and opportunities and, when they are identified, to enable the FAA to make worthwhile, targeted, performance improvement changes in the ATM system. The measurement approach and metrics used to calculate performance have been, and will continue to be, developed in consultation with the user community. It is essential that the metrics represent the operational outcomes desired by the users and that they reflect the many different aspects of ATM system performance of interest to them. In addition, it is essential that the metrics evolve as the ATM system, user needs, and financial, technical, and political environments change.

The measurement process is focused primarily on the system-wide, high-level efficiencies and inefficiencies that are dictated by ATM system capacity. The basic ideas can be applied more broadly, as well as to the measurement of performance on a localized basis. In addition, it is important to note that there are many other aspects of ATM system performance that could be measured, but which are not within the scope of this project. In a sense, maintaining or improving safety acts as a constraint that leads to tradeoffs among performance indicators, since the FAA has repeatedly and explicitly stated that ATM system efficiency performance will not be improved at the expense of safety. Other important performance aspects such as project management, cost control, and innovation are also not addressed by this initiative because they are not part of operational ATM system performance. In addition, the impacts of ATM system performance on factors such as the United States' gross domestic product and the investment infrastructure are not addressed because they are not directly related to the level of services users receive.

Section 2

PERFORMANCE MEASUREMENT APPROACH

Throughout the development and implementation of the ATM system operational performance measurement project, emphasis has been placed on developing an approach that is theoretically sound, consistent with established approaches such as Total Quality Management and GPRA, and easy to understand and implement. This section contains descriptions of the ATM system measurement philosophy tenets, factors to consider when developing metrics, and the planned measurement process.

2.1 Philosophy and Tenets

The ATM system performance measurement philosophy is based on the following tenets:

1. ATM system performance will be measured using standards of service provided to users, from the perspective of the users.
2. No value judgments concerning performance indicators and metrics will be made. Given the system-level nature of the metrics and the global sources of the data, what is important and useful is not the absolute values provided by the metrics, but rather the historical direction and relative change in the metrics.
3. ATM system performance must be continually monitored, evolved, and improved.
4. The indicators and metrics which compose the ATM system performance measurement set will evolve over time.
5. The emphasis placed on the individual indicators and metrics included in the ATM system performance measurement set may change as the business, technical, and political environments change.
6. It will usually not be possible to optimize all indicators and metrics simultaneously; rather, tradeoffs will have to be made between performance, cost, technical concerns, and other factors, such as safety.
7. Measurement must include both the current values of metrics and historical trend information. After an initial baseline is established, the metrics should be regularly and routinely calculated and the resulting values compared to historical performance.
8. When current performance is found to deviate significantly from historical performance, the causes of the deviation should be identified. If the deviation is an improvement in performance, the cause of the improvement should be analyzed to see if it can be implemented throughout the system. If the deviation is a degradation in performance, lessons for improved operation of the ATM system should be derived from it.

9. When opportunities for improving performance (technological advances, procedural changes, or other means) are identified, prior to their implementation, analyses and simulations should be performed to generate the metrics to predict the impact the proposed change might have on ATM system performance.
10. When a change is implemented in the ATM system, the metrics should be routinely recalculated and monitored to identify the intended and unintended effects which result.
11. When metrics are calculated, any ambient conditions which may have impacted ATM system performance should also be measured. For example, if weather was unusually poor last month, an effort should be made to calculate the weather degradation relative to previous months'. The value of any new ambient condition metrics should then be reported along with the standard ATM system metrics.

2.2 Factors to Consider When Developing Metrics

Many factors must be taken into account during performance measurement efforts. Often, the performance of the system and the way users interact with it reflect factors such as user perceptions, expectations, and system constraints. In addition, besides producing performance information, the act of measuring performance almost always has residual and often unintended results. The following considerations were developed to help guide development of the ATM system performance measurement approach.

Users' current demands on the ATM system do not necessarily reflect their true needs. ATM system users often base their requests and expectations on what they know they will be allowed to do, rather than on what they really need. For example, a pilot might intentionally file an ATC preferred route because he believes that a direct flight would not be approved. Then, once the pilot is airborne, he might request a route more like the one he really wants since he believes that there may be a higher probability that the desired route will be approved while he is airborne. Because such restrictions currently exist, as the ATM system evolves and restrictions are lifted, user demands on the system will change. As a result, comparing user request-based performance metrics from year to year may not yield a consistent basis for comparison. These aspects of performance should still be measured, but additional metrics may be needed to assess user request changes and cause-and-effect of operational changes.

You get what you measure. When their performance is being measured, people often focus on those aspects of their performance which are being measured. When this happens, this results in increases in those performance aspects being measured and decreases in those performance aspects not being measured. It can also distort the effects of any improvement interventions which have been made. A reasonable set of metrics should provide consistent measures of the impact of proposed changes to help the FAA in making decisions and setting cost effective priorities.

If you measure something, people will expect that you have the authority and resources to improve it. Therefore, the majority of metrics should be focused on aspects of performance

that can actually be improved. Measuring things that cannot be improved is often non-beneficial and can cause disillusionment among the people who feel they are being held accountable for those aspects of performance. One notable exception is the measurement of events which cannot be controlled, but which significantly impact system performance. For example, measuring the weather and identifying its impact is essential in understanding ATM system performance and assessing whether or not system performance is improving. (Note that the FAA can take actions which can mitigate the effects of weather in the long run, such as by improving airport capacity during Instrument Flight Rules (IFR) weather conditions relative to VFR airport capacity.)

Regardless of their initial intent, metrics often inadvertently turn into goals. Two of the primary purposes of measurement are to identify areas for improvement and to ensure that performance trends improve over time. However, when performance metrics are identified, a common response is, “What percent improvement in this aspect of performance are you planning to achieve in the next year?” When goals are identified, it is important to have metrics for monitoring progress toward those goals; however, the reverse is not necessarily true or desirable. Numeric goals can be very useful if the metrics and data sources upon which they are based are accurate and consistent. However, the appropriate balance must be maintained between achieving goals and continuously improving system-wide performance. Improvement of the system as a whole should not be delayed by excessive attention to the attainment of individual goals.

The metrics set conveys information about what is important to the organization. Including a metric in an organization’s measurement set implies that that aspect of performance is very important to the organization. As a result, metrics sets must be complete, well balanced (measure a broad cross-section of performance), and politically sensitive if they are to endure. If a measurement set is incomplete or not well balanced, people will assume that the measurement team does not truly understand what they are trying to measure, and those people may go so far as to try to discredit the measurement system. If the set is not politically sensitive, people may try to impede the implementation of the measurement process.

Measurement teams must solicit and receive inputs to, and feedback on, metrics sets prior to their implementation. No one knows system performance better than the people out in the field who work with the system every day. Therefore, it is essential that those individuals have an opportunity to review and comment on the metrics before the metrics set is implemented. This step is even more important if human performance metrics (such as those developed to assess controller workload) are to be included in the set. In such cases, the individuals whose performance will be measured, as well as their organizations and representatives, such as unions, should be given ample opportunities to review and comment on the metrics set prior to its implementation.

2.3 Planned Measurement Process

The process planned for identifying, developing, refining, and using ATM system indicators and metrics is depicted in Figure 1 below. As metrics are applied and embedded throughout the FAA's business processes, they are used to evaluate systems and procedures, make decisions and budget allocations, and provide historical and current pictures of system performance. As new measurement needs arise, the existing indicators and metrics are refined, new ones are identified, and the identify-develop-refine-use process is repeated.

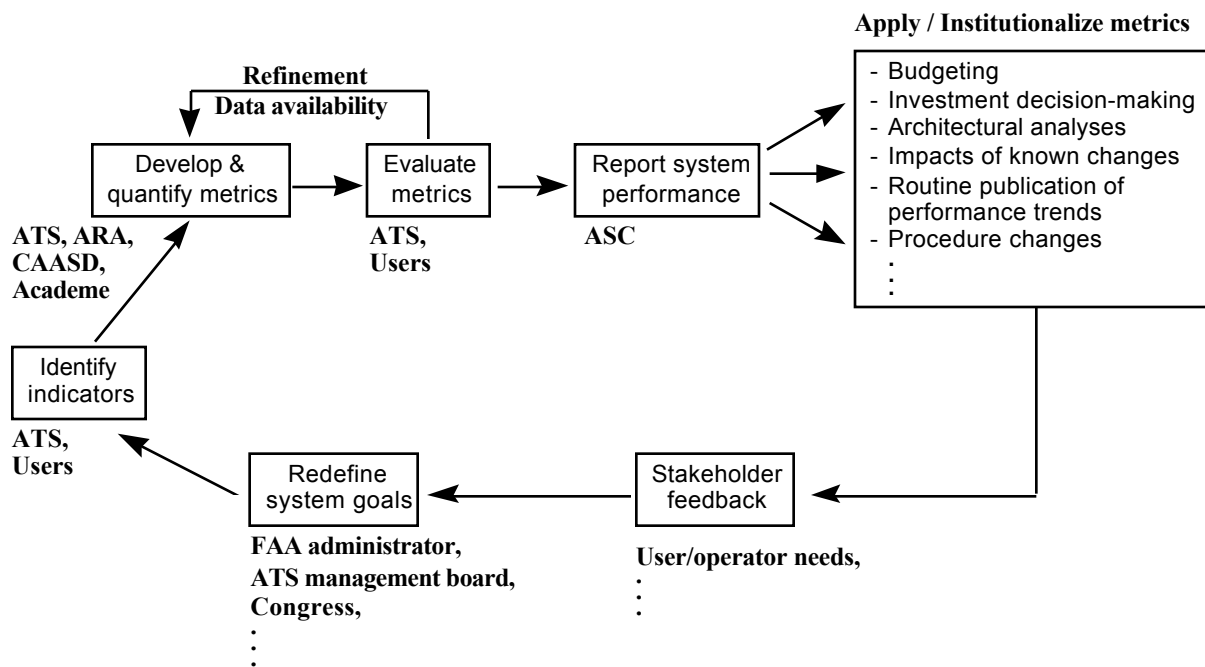


Figure 1. Metrics Identification, Refinement and Usage Process

Section 3

PROGRESS AND PLANS

This section contains descriptions of the completed, on-going, and planned ATM system performance measurement tasks, which are depicted in Figure 2 below.

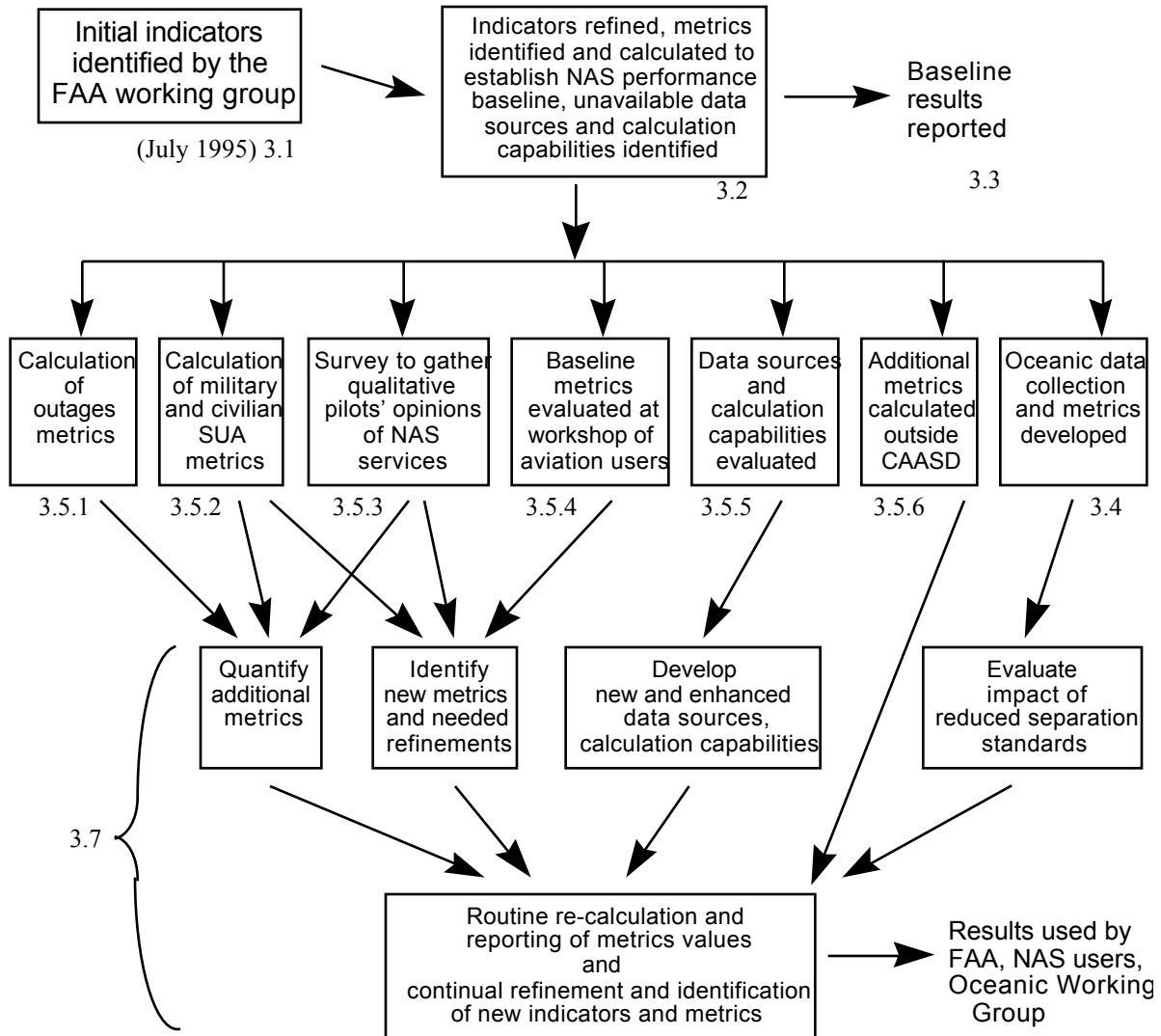


Figure 2. Project Flow and Tasks

3.1 Metrics Project History

Metrics work has been discussed and applied to U.S. aviation for years. Thus, in one sense, the general groundwork for this effort has been laid down over many years. However, this particular effort is the result of a conscious effort to break with much of the past work, much of which was based on some commonly-held beliefs:

- that the absence of delay represents optimum system performance to the user;
- other aspects of ATM system performance are not quantifiable, and are best managed through experience and intuition; and
- the quality and quantity of the data available from the operational system are inadequate to support effective performance measurement.

This paper presents ideas in support of analyses that may eventually cause these beliefs to be abandoned.

In parallel with the work which preceded this project (which focuses on performance from the perspective of user outcomes in domestic airspace), metrics were developed for oceanic air traffic services. These metrics for oceanic aviation cover a much broader range of concern, including internal FAA issues such as controller workload and productivity, in addition to user-oriented outcomes. These two projects have been closely coordinated; this document includes pertinent summaries of the oceanic metrics efforts where appropriate, particularly with the inclusion of Section 3.4, “Metrics for Oceanic Operations.”

The events which initiated this ATM system performance measurement project began in July 1995, when ASC sponsored an agency-wide Cost/Performance Measurement Working Group (CPMWG). The group’s discussions included an examination of GPRA and a review of the methods and definitions then in use for measuring the capacity and efficiency in the ATM system. These reviews were essential because GPRA mandates that, by fiscal year 1999, federal organizations must tie their budget allocations to their strategic performance improvement goals and they must have quantitative metrics in place for assessing their progress toward those goals.

The participants of the CPMWG agreed that a common definition for system performance needed to be developed and that an initial set of high-level system performance indicators should be proposed. The CPMWG then formed a system performance sub-group whose charter was to define and develop customer outcome-based system performance indicators. This sub-group, comprised of representatives from all parts of the FAA, explored and investigated numerous indicators, including safety. The group decided not to include safety in their final list of indicators, as it was already being fully addressed by the FAA Safety Office.

To continue with the development of these indicators, in October 1995, ASC tasked MITRE/CAASD to refine the initial set of indicators, identify calculable metrics for each indicator, and begin calculating the metrics. The initial computation was expedited in order to have values of current ATM system performance for possible use in that spring’s budget

allocation processes, and to move quickly from theory to practice so that a broader discussion of the utility of the new indicators and metrics could begin. Calculation of the metrics for which data was available was completed in January 1996 [3]. The resulting set of values was called the January 1996 Baseline.

Since the January 1996 Baseline was established, efforts have been underway to further enhance FAA's ATM system performance measurement capabilities. These efforts include:

- gathering data required for calculating new metrics;
- developing additional automated data feeds and analysis capabilities to facilitate on-going metrics calculations and tracking;
- holding meetings with user representatives in order to gain feedback regarding the indicators and metrics;
- coordinating with organizations that have access to data sources that permit expansion of the set of calculable metrics; and
- holding meetings with other FAA offices to increase their awareness of the ATM system performance measurement initiative, gain buy-in, share lessons learned, and promote use of the standardized terms and definitions developed as a result of this initiative.

In order to develop an initial set of metrics quickly, CAASD began by conducting a “triage” of the proposed indicators, and grouped them into three categories. The first contained those indicators that were quantifiable using CAASD in-house data sources—mainly composed of data on aircraft movement and airspace structure. The second group contained those indicators that could be measured using data available from aviation resources outside of CAASD (e.g., FAA facility operating parameters). The last group consisted mainly of the more qualitatively-oriented indicators that would require information from surveys of users. To meet the deadline for establishing a January 1996 baseline, CAASD developed specific metrics for quantifying each indicator only where in-house data was readily available. In the cases where data was not readily available, CAASD outlined an approach for measuring the indicator.

Since January 1996, efforts at obtaining metrics for the last two groups of indicators have been under way. Coordination with other parts of ATS has yielded sources of data previously unavailable to CAASD. Ongoing performance measurement activities have been incorporated wherever they fit the user-oriented requirements of this work. Companies working under contract to the FAA have adapted existing software to produce results that can expand the domains of several metrics and quantify additional indicators. A user survey has been developed to address the third class of indicators mentioned above.

In a parallel effort, the FAA's Oceanic and Offshore IPT selected a set of metrics to monitor and assess system efficiency and quality, user benefits, operator workload, and system development cost and schedule. The oceanic service set of measures are directed towards measuring the benefits provided to all stakeholders, including airlines, controllers, system

maintainers, and automation staff. In February 1995, the oceanic set of metrics were approved by the GSA. Even through the operations over the ocean have certain unique characteristics, Section 3.4 presents a subset of the oceanic metrics relative to the four categories of system performance indicators discussed in this paper, outcome measures that impact capacity and efficiency of the oceanic operations. With a strong global emphasis on reducing oceanic separation standards, which may have an effect on measures of safety, efforts are underway to define oceanic safety metrics independently.

3.2 Initial Performance Categories and Indicators

The four categories of system performance indicators developed by the FAA CPMWG working group (described in the previous section) were selected specifically for use in measuring ATM system capacity and efficiency. As previously stated, the working group noted that safety is essential and will not be compromised as the FAA strives to make the ATC system more flexible, more predictable, and more accessible, and to reduce delays. However, they did not include it as a performance category because it is already being address by the FAA Office of System Safety.

The working group also identified indicators for the four performance categories. They are listed in Table 1 below. These indicators have been revised and new ones have been added, so this table is provided for historical purposes only. However, metrics were developed for nine of these indicators and calculated in time to be included in the January 1996 Baseline, which is described in the next subsection.

Table 1. FAA’s Initial Performance Categories and Indicators

Performance Category	Nature of the Indicators	Initial Indicators
Flexibility	The ability of the ATM system to meet users' changing needs. Includes removal of restrictions and constraints to user preferences.	<ul style="list-style-type: none"> • Reduce the number of procedural restrictions in the system • Increase the number of user preferred route requests granted • Reduce the deviation between the route requested and the route flown • Reduce the difference between published preferred route distance and direct routing distance between city pairs at low altitude • Increase the peak acceptance rate of airports and/or airspace, and • Increase the number of decisions involving pilot-controller collaboration
Predictability	Variance in the system as experienced by the user. Includes removal of uncertainty that leads to padding, and information sharing.	<ul style="list-style-type: none"> • Reduce variation in system performance associated with changes in weather • Reduce the impact of system outages • Increase the timeliness and quality of data provided to the user on weather, traffic, and system status • Increase the number of delay allocation decisions made with direct user input
Access	The ability of users to enter the system and obtain services on demand. Includes physical resources such as airports, airspace, and SUA, as well as information resources.	<ul style="list-style-type: none"> • Increase the number of airports with approved approaches • Increase the number of airports with precision approach capability • Increase civilian utilization of Special Use Airspace (as measured by either hours available or number of civil flights using Special Use Airspace) • Increase the availability and quality of VFR in-flight services • Increase the availability of Flight Services to the system user • Increase the number of operations provided with VFR tower services • Increase the coverage of air traffic control surveillance and communication
Delay	The amount of time over the optimum that it takes to complete an operation. Includes variation from expected times, and national ground delays.	<ul style="list-style-type: none"> • Reduce ground movement times at key airports during peak operations • Reduce the difference between estimated and average en route time, and • Reduce the number, duration, and impact of ground delays imposed by the Air Traffic Command Center

This four-dimensional framework for measuring ATM system performance was deliberately left without an overarching optimization scheme. Tradeoffs among the system

performance categories will always be necessary, and the priority given to the individual indicators will change over time. Decision makers will have to understand their priorities and make judgments about what constitutes an acceptable level of performance in each category, so that enhancements in one category can be considered even if they cause a slight decrease in performance in another category.

3.3 January 1996 Baseline Metrics

Nine of the performance indicators identified by the FAA CPMWG working group had metrics identified and calculated in time for them to be included in the January 1996 Baseline. The details of this task are included in [3]. These initial calculations and resulting values were the first step toward understanding performance of the ATM system in a multidimensional sense. New views of the traditional measure of ATM system performance, delays, were calculated, but they were also balanced against the flexibility and predictability of operations and access to air traffic management services. All of the metrics have flaws; none is a perfect quantitative representation of aviation operations. This task showed that it is possible to calculate many metrics with existing data to a degree where the results are both practical and credible. Thus, the effort succeeded as a proof of concept demonstration for metrics.

The January 1996 Baseline analysis produced estimates of parts of the current performance of the ATM system. It did not create performance goals, or attempt to state whether the measured performance was good or bad. The question of whether the baseline values of each metric are acceptable performance must be addressed by aircraft operators, controllers, and passengers together. This analysis does begin to show measurement that is representative of true system-level performance and that does not focus in on any one function. The preliminary results highlight visible changes in performance by looking at trend data. This also helps to represent the non-linearity of ATM system performance.

A summary of the January 1996 Baseline indicators and metrics is presented in Table 2 below. Note once again that these indicators and metrics have since been revised and new ones have been added, so this table is provided for historical purposes only.

Table 2. January 1996 Baseline Indicators and Metrics

Performance Category	Indicators Assessed in January Baseline	Metrics Calculated for the January Baseline
Flexibility	<ul style="list-style-type: none"> • User-preferred route requests granted • Routes requested vs. routes flown • Deviation of ATC preferred routes vs. low-altitude direct routes 	<ul style="list-style-type: none"> • Distance deviation between direct & ATC preferred routes • Deviation between flight planned & actual altitudes • Distance deviation between flight planned & actual routes • Distance deviation between flight planned & ATC preferred routes • % flights off/on ATC preferred routes
Predictability	<ul style="list-style-type: none"> • Variation in system performance due to weather • Variation in system performance due to system outages 	<ul style="list-style-type: none"> • % difference between VFR & Category I EPS capacity • Deviation of ground movement times • Statistical spread of ground movement times • Statistical deviation of en route times • # cancellations & diversions • Arrival, departure, overflight delay
Access	<ul style="list-style-type: none"> • Airport approaches • Airport precision approaches 	<ul style="list-style-type: none"> • # airports w/ published approaches • # airports w/ precision approaches
Delay	<ul style="list-style-type: none"> • Ground movement times • En route times 	<ul style="list-style-type: none"> • Mean taxi-in & taxi-out time • Difference between estimated time of arrival & actual time of arrival

3.4 Metrics for Oceanic Operations

CAASD initiated an effort to routinely collect oceanic operations data from the Oakland Center in May 1996 and from New York Center in August 1996. These data are intended to provide a measure of the baseline performance of the current oceanic system. A capability has been developed to receive, clean, and store oceanic operational data, and to conduct daily and trend analyses of current operations. CAASD has also developed a simulation model to postulate the performance of future operations based on growth in demand, reduced separations standards, and airspace/track redesign. The simulation capability helps conduct “what if” analysis to understand the impact of future procedure and system enhancements and help the FAA in decision support and prioritizing system improvements cost effectively. The baseline data is also being used to help define Flight 2000. Future system assessment capability was used to analyze the benefits of reduced separations to 50/50 nmi for the FAA’s Major

Acquisition Review / Joint Resources Council (MAR/JRC), and to address future oceanic automation system requirements. Time and fuel savings metrics results for future system enhancements helped the airlines understand the impact on direct operating costs to assess return on investment with FANS 1 (Future Air Navigation System).

Table 3. Oceanic Baseline Indicators and Metrics

Category	Indicator	Metric
Flexibility	<ul style="list-style-type: none"> • User preferred track requests granted <ul style="list-style-type: none"> – Altitude requests granted • Tracks requested vs. tracks flown • First flight plan preference granted at oceanic gateway 	<ul style="list-style-type: none"> • Percentage of route/altitude/speed requests granted • Fuel savings • Percentage of preferred flight plans accepted
Predictability	<ul style="list-style-type: none"> • Rerouting and altitude changes due to weather • Estimation of aircraft flying times • Clearance delivery time 	<ul style="list-style-type: none"> • Number of reroutes and altitude changes due to weather • Variance or difference between predicted and actual flying times • Deviations of actual flight paths from desired • Clearance processing time • Percentage of responses granted within a specified time
Access	<ul style="list-style-type: none"> • Surveillance coverage • Communications coverage 	<ul style="list-style-type: none"> • Percentage of U.S. oceanic airspace with surveillance coverage • Percentage of responses made within a specified time (e.g., 1 minute)
Delay	<ul style="list-style-type: none"> • Departure windows met • Increase in flight time due to separation or track assignment 	<ul style="list-style-type: none"> • Ground delays (number of flights and mean delay) • En route delays (number of flights and mean delay) • Flight time savings

3.5 Additions and Refinements to the Baseline Metrics

Since the January 1996 Baseline results were published, several tasks have been underway to refine and enhance the metrics which are currently being calculated and reported. These tasks are described in the following subsections.

3.5.1 ARTCC Outages

One of the initial predictability indicators for which metrics were recently calculated is *to reduce the impact of major system outages*. One important example of this is outages at Air Route Traffic Control Centers (ARTCCs), for which an analysis has been completed. That effort is documented in [8]. As disruptions of service due to system outages can have detrimental effects on users of the system, the purpose of that study was to measure the degree to which the predictability of operations in the ATM system is reduced by analyzing changes in arrival, departure, and en route times near the affected areas.

This initial effort to develop and baseline the outages metrics began with a particularly visible set of outages in 1995—loss of en route ATC services due to failures in the Display Channel Complex (DCC), the Host computer, the power supplies, or other hardware at the ARTCCs. An outage of this magnitude has several levels of impact on the ATM system. These include, in order of proximity to the outage:

1. flights may be canceled or diverted to different airports;
2. departures and arrivals at major airports may be delayed;
3. flights passing through the ARTCC-controlled airspace may be required to hold or to change course;
4. an aircraft that was delayed may be unable to make later trips that day on schedule, leading to rippling delay through the nation; and
5. departures from smaller airports in the ARTCC-controlled area may be shut down entirely.

The above lists possible effects of ARTCC outages. Identifying which of these are actually observed to occur was the first task in the definition of outage-oriented metrics.

As a result of this task, the primary impact of an ARTCC outage was shown to be aircraft ground delays. Primary outage impacts are not nationwide, but are concentrated at the airports in the affected ARTCC. At least among the large-air-carrier flights that formed the data analyzed in this task, airborne delays were very rare, both for overflights and for arrivals. Departing flights are delayed during outages. These delays, and similar delays at the departure airports of inbound flights, propagate throughout the day such that the impacts of outages frequently continue to be felt for the rest of the day.

The metrics identified and baselined as a result of this task have been added to the metrics set and will continue to be re-calculated and refined.

3.5.2 Military Aviation and Military Airspace

Two Special Use Airspace (SUA) tasks have been performed since the January 1996 Baseline was calculated: one on metrics for military users of civilian and military airspace [4]

and one on civilian SUA use [2]. The metrics which were developed as a result of these tasks have now been included in the metrics set.

3.5.2.1 Military Metrics

The FAA and the Department of Defense (DoD) share a unique user/provider relationship. In many ways, the DoD is like any other aviation user—requiring air traffic control services and certification and inspection services for transport operations. In other ways, the military requires access to airspace for operations that are very dissimilar to conventional civilian flights, such as specialized performance aircraft training, combat training missions and non-aviation oriented missions, such as artillery testing. As a provider, the DoD also possesses capabilities to control both military and civilian flights. It is clear that defining a set of metrics to describe the level and quality of service provided by the FAA to the DoD user is a challenging activity.

To facilitate development of a set of SUA metrics, a joint DoD and FAA IPT was formed for the development of Air Traffic Services measures of effectiveness. This IPT was sanctioned by the Executive Director of the DoD Policy Board on Federal Aviation, and was chartered to identify metrics that the FAA could use to assess the adequacy, quality, and timeliness of air traffic services provided to the military. Given a six month period, the IPT was asked to propose an initial set of metrics and conduct preliminary analyses to test the validity of those metrics. As a member of this IPT, CAASD was charged with doing the technical analyses needed to determine the potential usefulness of several proposed indicators. The four areas of interest listed below were identified for review by the IPT. Metrics for each of the areas were developed and calculated, and they were added to the current metrics set and the performance baseline.

1. Special Use Airspace Utility: Defined to evaluate the efficiency of the management of special use airspace (SUA). Focused on answering questions such as: How effectively is the airspace being used? Does everyone (civilian and DoD) get access when they need it? If DoD is not utilizing airspace, is it returned to, and accepted by, the FAA for civilian use?
2. Administration: Defined to evaluate the timeliness and quality of airspace establishment, modification, or removal. This indicator area would include aeronautical and environmental concerns with rulemaking and non-rulemaking actions.
3. Certification: Defined to evaluate the timeliness of periodic flight check inspections. This area would cover FAA responsiveness to requests for certification of hardware (e.g., navigation aids and landing systems) and procedures.
4. Operations: Defined to evaluate whether DoD receives the air traffic management services needed to support their mission. This indicator area is most similar to the civilian measures proposed for civilian operations. Particular indicators include reducing delays to military operations and increasing planning and in-flight flexibility of military operations.

3.5.2.2 Civilian SUA Metrics

After the January 1996 Baseline was calculated, a metric was developed to measure civilian utilization of Special Use Airspace, another of the original indicators developed by the FAA CPMWG working group. Analysis showed that the indicator would be best assessed with a metric based on either SUA schedules or actual civilian access to SUA. A metric based on the latter was developed and calculated. The calculations involved counting the number of flights that penetrated SUA, for those origin–destination pairs that had at least one flight going through SUA. The measurement was carried out for four analysis days, using NASPAC (the National Airspace System Performance Analysis Capability) to analyze the flights. The metric includes only air carrier flights. This analysis is currently being documented and is being published as part of the Baseline documentation series [1].

3.5.3 Pilot’s Opinion Survey

To gather qualitative ATM system performance data that is not available through the existing operationally-oriented databases, a pilot’s opinion user survey has been developed and will soon be administered. This survey is intended primarily to capture VFR pilots’ opinions of the services they receive. Survey questions are focused primarily on the predictability and access indicators, as shown in Table 4 below.

Table 4. Indicators to be Supported by the Pilot’s Opinion User Survey

Performance Category	Indicators
Predictability	<ul style="list-style-type: none">• Data provided to the user on weather, traffic, and system status• Delay allocation decisions made with direct user input
Access	<ul style="list-style-type: none">• VFR in-flight services• Availability of flight services to the system user• VFR tower services

The survey is expected to be conducted annually. FAA has requested that CAASD develop and conduct the 1997 version of the survey; subsequent years’ surveys may be conducted by different organizations. This first survey will be used to establish a performance baseline; subsequent surveys will be used to track performance trends. Each year the survey will be revised to ensure that it addresses current issues.

The questions included in the 1997 survey were developed as a result of six focus groups held at the Experimental Aircraft Association’s (EAA’s) Oshkosh 1996 Fly-In Convention. The Aircraft Owners and Pilots Association (AOPA) and EAA were both informed that the focus groups would be taking place. Survey development has been coordinated with FAA’s Air Traffic Services (ATS) and the General Aviation division of the Flight Standards Service

(AFS-800) through the Office of System Capacity (ASC). Inquiries to other FAA divisions indicate that this is the only service assessment survey currently being planned by FAA.

Survey results will be reported to FAA by question for all respondents as a group, and broken down by region. Individual survey participants' responses will not be made available, and complete anonymity for individuals is assured. General aviation trade organizations and magazines may request copies of the general survey results from FAA.

3.5.4 Workshop on FAA System Performance Measures

In September 1996, a ATM system performance measurement workshop was hosted by Northwest Airlines. The purpose of this workshop was to review in detail each of the metrics included in the January 1996 Baseline and to receive feedback from the ATM system users regarding the metrics' appropriateness, accuracy, and usefulness. The workshop attendees primarily represented ATM system users from the commercial airlines.

The primary conclusion of the conference was that a user-based performance measurement system is essential and that one should be implemented directly, initially using imperfect data and methods if necessary. Refinement of the metrics should take place in parallel with efforts to embed the metrics throughout the FAA's business processes, not as a prerequisite for those efforts.

Following the recommendations of the workshop attendees, a number of initiatives were begun to enhance existing ATM system performance measurement capabilities. Generally, it was recommended that the simplest possible metric should be used at this time. For example, comparing routes to great circles would be sufficient—it is not necessary to find a pseudo-optimum wind route to accurately reflect user preferences. In addition, as a result of the workshop, one flexibility indicator, the *distance difference between ATC Preferred Routes and direct routes*, was extended to include High Altitude Preferred Routes (using great circle routes for comparison for now). The original indicator specified only low-altitude and Tower En route Control (TEC) Preferred Routes.

3.5.5 Metrics Calculation Enhancements

To improve the timeliness of ATM system metrics calculations and reduce the number of staff-hours required to calculate the baseline metrics, CAASD has enhanced its automated, in-house metrics calculation capabilities. New raw data sources have been identified and automated data feeds are being established wherever feasible, batch programs are being written to facilitate processing the large volumes of data into metrics values, and automated results formatting and reporting capabilities are being developed.

In addition, a study was performed to evaluate metrics prediction needs against available in-house capabilities [5]. CAASD's computer-based ATM system traffic models were evaluated, existing metrics prediction capabilities were identified, and recommendations were made for

enhancing the models in support of the metrics requirements. The results of the study are now being evaluated and a prioritized list of needed model enhancements is being developed.

CAASD has developed a prototype of oceanic operational data analysis capability to generate daily performance summaries that the FAA's oceanic control facilities could use on site.

3.5.6 Indicators and Metrics Addressed Outside CAASD

There are many other ATM system performance measurement efforts being performed within the FAA or under contract to the FAA. Some are geared toward benefits analysis of specific systems, some are special projects undertaken by an airport, ARTCC, or Region, and others are initiatives from the FAA headquarters that pre-date this project. Where the scope and methods of these projects permit, it is desirable to incorporate their results into the set of indicators and metrics resulting from this project.

Coordination is under way in a number of areas, including:

- Under contract to ASC-200, ATAC Corporation has developed the Graphical Airspace Design Environment (GRADE) tool, which, among other things, contains data and algorithms suitable for calculating flexibility and predictability in terminal (class B) airspace.
- An audit of procedural restrictions has been conducted by the FAA's Air Traffic Operations Service (ATO) since 1995, which applies directly to one of the flexibility indicators.
- As part of the development of the Collaborative Decision-Making (CDM) project, METRON Corporation, under contract to the FAA, has been analyzing ground delay programs in a way that can calculate metrics associated with one of the delay indicators, and with further development of the CDM system, a flexibility metric will be calculable. In addition, the project furnishes information about acceptance rates that may contribute to other flexibility indicators and metrics.

3.6 Current Set of Performance Indicators and Metrics

The indicators and metrics currently included in the ATM system performance measurement approach are summarized in the following tables. These indicators and metrics must continue to evolve as the ATM system evolves, the FAA gains performance measurement expertise, and the financial, technical, and political environments change. Table 5 lists the current descriptions of indicators and their associated metrics.

Table 5. Current Indicators and Metrics

Flexibility Indicators	Flexibility Metrics
Reduce the number of procedural restrictions in the system	
Increase the number of user preferred route requests granted	<ul style="list-style-type: none"> • The percentage of flights subject to an Air Traffic Control (ATC) preferred route, separately above and below flight level 180. • Among those flights subject to an ATC preferred route, the amount of aviation activity not on the ATC preferred route, by Center. • Among those flights subject to an ATC preferred route, the mean lateral deviation between the route in the flight plan and the ATC preferred route.
Reduce the deviation between the route requested and route flown	<ul style="list-style-type: none"> • Number of flights whose maximum altitude equaled the requested altitude in their flight plans. • Excess time from top of descent to wheels down. • Mean lateral deviation between the flight planned and the actually flown route.
Reduce the deviation between published preferred route distance and direct routing distance between city pairs	<ul style="list-style-type: none"> • The distance difference between the preferred routes Tower En Route Control (TEC), Low, and the high altitude preferred routes, contained in the National Flight Data Center (NFDC) preferred routes data base, and the great circle distance between the corresponding origin destination pairs, weighted by level of traffic.
Increase the peak acceptance rate of airports and/or airspace	
Increase the number of decisions involving pilot controller collaborations	

Table 5. Current Indicators and Metrics (Continued)

Predictability Indicators	Predictability Metrics
Reduce the variation in system performance associated with changes in weather	<ul style="list-style-type: none"> • The difference between the highest Engineered Performance Standard (EPS) Visual Flight Rules (VFR) capacity and the smallest Category 1 Instrument Flight Rules (IFR) capacity at the top 25 airports. • For each of the top 25 airports, treated individually, and for each day of a month, compute the average taxi-in time and the average taxi-out time, resulting in a series of average taxi-in and taxi-out times for each airport. For each series, the metric is the difference between the 75-th percentile of its values and average of its values. • For each destination airport, treated individually, and for each day of a month, compute the average of the differences between ETA and ETMS AZ time, resulting in a series of values, for each airport. The metric is the difference between the 75-th percentile of the values in the series and the average of the values in the series.
Reduce the impact of system outages	<ul style="list-style-type: none"> • Total number of cancellations and diversions at major airports within the affected area. Total delay of departures, and total difference between scheduled and actual arrival times.
Increase the timeliness and quality of data provided to the user on weather, traffic, and system status	
Increase the number of delay allocation decisions made with direct user input	

Access Indicators	Access Metrics
Increase the number of airports with approved approaches	<ul style="list-style-type: none"> • The number of airports that have at least one approach record in the Jeppeson Navigation Database.
Increase the number of airports with precision approach capability	<ul style="list-style-type: none"> • The number of airports that have at least one precision approach record in the Jeppeson Navigation Database.
Increase civilian utilization of special use airspace	<ul style="list-style-type: none"> • Percentage of air carrier flights that penetrate Special Use Airspace, for those origin destination pairs in which at least one flight penetrated Special Use Airspace.
Increase the availability and quality of VFR in flight services	
Increase the availability of flight services to the system user	
Increase the number of operations provided with VFR tower services	
Increase the coverage of air traffic control, surveillance, and communications	<ul style="list-style-type: none"> • Coverage of en route and terminal radar at 1500 ft. above ground level. • Histograms of coverage by various navigation, surveillance, and communication systems at 1500 ft. AGL, 10,000 ft. MSL, and 18,000 ft. MSL.

Table 5. Current Indicators and Metrics (Concluded)

Delay Indicators	Delay Metrics
Reduce ground movement times at key airports during peak operations	<ul style="list-style-type: none"> • Monthly average taxi in and taxi out times at 25 major airports.
Reduce the difference between estimated and actual en route times	<ul style="list-style-type: none"> • For arrivals at each of 25 major airports, treated individually, and for each day of a month, compute the average of the differences between ETA (calculated as estimated time en route from the flight plan added to the actual departure time) and ETMS AZ time, resulting in a series of values. The metric is the monthly average of the values in the series.
Reduce the number, deviation, and impact of ground delays imposed by the air traffic control system command center (ATCSCC)	

3.7 Next Steps

Future plans for domestic and oceanic ATM system performance measurement enhancement are focused on three areas: continual evolution of the indicators and metrics, sharing the performance measurement results and lessons learned, and actively using the performance information to improve planning and decision making. Planned tasks include:

- continuing to investigate and develop indicators and metrics by building on the current results, and identifying and analyzing new needs and capabilities;
- improving the efficiency with which metrics can be calculated and prediction models can be executed;
- continuing to communicate the current metrics results within the FAA and external user communities;
- promoting use of the ATM system performance measurement approach and terminology throughout FAA;
- encouraging ATM system performance (current values and historical trends) to be considered in budgeting and decision-making processes;
- predicting and analyzing the impacts proposed architecture and technology changes would have on overall ATM system performance;
- predicting and evaluating the impacts of new tools and field trials such as the National Route Program (NRP), the conflict probe tool, and reduced separation standards over oceanic airspace;
- calculating the actual impacts of known changes in the ATM system on the user communities; and

- routinely re-calculating the metrics and publishing their current values and graphs of their historical performance trends.

As these tasks are completed, progress reports will be published detailing what was accomplished and what changes in the ATM system performance measurement approach, indicators, and metrics resulted.

List of References

1. Bolczak, Catherine N., May 1997, *National Airspace System Performance Baseline II: Civilian Access to Special Use Airspace*, MTR97W0000041, The MITRE Corporation (DRAFT).
2. Bolczak, Catherine N., April 1997, *National Airspace System Performance Baseline I: ATC-Preferred vs. Great Circle Routes*, MTR97W0000019, The MITRE Corporation.
3. Bolczak, Catherine N., Lee M. Brown, Jonathan H. Hoffman, and Elizabeth S. Lacher, June 1996, *System Performance Baseline: Initial Indicators, Analyses, and Results*, WN 96W0000015, The MITRE Corporation.
4. Bowden, Samuel J., Lee M. Brown, David P. Burke, Elza P. Hisel, Jonathan H. Hoffman, June 1996, *System Performance Metrics: Initial Development and Assessment of Metrics for Military Aviation and Airspace*, MP 96W0000161, The MITRE Corporation.
5. Burke, David P., September 1996, *Availability of Models at CAASD for Analysis With User Oriented Metrics*, WN 96W0000072, The MITRE Corporation.
6. Federal Aviation Administration, Office of Air Traffic, *Air Traffic Service Plan 1995-2000*, August 1995.
7. Federal Aviation Administration, Air Traffic Services, *Air Traffic Services Performance Plan for Fiscal Years 1997-1999*, May 1997.
8. Hoffman, Jonathan H., March 1996, *A Survey of User Impacts of ARTCC Outages in 1995*, WN 96W0000141, The MITRE Corporation.

Glossary

AFS	FAA's Flight Standards Service
AIT	FAA's Office of Information Technology
AOPA	Aircraft Owners and Pilots Association
ARA	FAA's Administrator for Research and Acquisition
ARTCC	Air Route Traffic Control Center
ASC	FAA's Office of System Capacity
ATC	Air Traffic Control
ATCSCC	Air Traffic Control System Command Center
ATM	Air Traffic Management
ATO	FAA's Air Traffic Operations Service
ATS	FAA's Air Traffic Services
CAASD	Center for Advanced Aviation System Development of the MITRE Corporation
CDM	Collaborative Decision-Making
CPMWG	Cost/Performance Measurement Working Group
DoD	Department of Defense
DOT	Department of Transportation
EAA	Experimental Aircraft Association
ETA	Estimated Time of Arrival
ETMS	Enhanced Traffic Management System
EPS	Engineering Performance Standards
FAA	Federal Aviation Administration
FANS	Future Air Navigation System
GPRA	Government Performance and Results Act of 1993 [Public Law 103-62]
GRADE	Graphical Airspace Design Environment
GSA	General Services Administration
IFR	Instrument Flight Rules
IPT	Integrated Product Team
JRC	Joint Resources Council

MAR	Major Acquisition Review
MSL	Mean Sea Level
NAS	National Airspace System
NASPAC	National Airspace System Performance Analysis Capability
NFDC	National Flight Data Center
NRP	National Route Program
SUA	Special Use Airspace
TEC	Tower En Route Control
VFR	Visual Flight Rules

