COMMON METRICS FRAMEWORK FOR ATM PERFORMANCE ASSESSMENT AND MONITORING

Almira Williams, CSSI, Inc., Washington, DC, USA
Steve Bradford, Diana Liang, FAA, Washington, DC, USA
Wim Post, Ulrich Borkenhagen, EUROCONTROL, Brussels, Belgium
Jean-Marc Pomeret, CENA, Toulouse, France

Abstract

Across the Air Traffic Management (ATM) discipline there is a long-standing problem of inconsistencies in assessing and monitoring system performance. Research has offered limited solutions so far, and has failed to offer an internationally acceptable performance definition methodology that provides the ability to: (1) compare the analogous performance metrics and evaluations obtained by various organizations; (2) roll-up measures to the organizational goals; and (3) comprehend the impact of a change in metric-value onto the overall performance.

This paper describes the initial results of an on-going work proposed by FAA, Eurocontrol and their partners, that addresses the ATM performance measurement discrepancies. The problem is approached in a systematic manner by dissecting the ATM system into the smallest elements and processes, and by studying the International Standardization Organization (ISO) definition of Quality of Service (QoS) in Information Technology (IT). A common metric taxonomy is developed to provide for unambiguous translation of the terminology and metric specifications used by different organizations, and to facilitate developing a collection of the existing metrics to enable maintaining the connection between the identical metrics, metrics with identical definitions but different names, and metrics with identical names but different definitions. A common ATM performance framework is built upon the International Civil Aviation Organization (ICAO) concepts of Required ATM System Performance (RASP) and Required Total System Performance (RTSP). Illustrated through cause-effect diagrams, the framework maintains the connectivity between different levels of performance evaluation, performance aspects, and performance measures, and enables rolling-up measures to the performance requirements and organizational goals. A common 'dashboard' for presenting the achieved or expected performance is developed using radar diagrams to clearly and effectively accommodate simultaneous display of large amount of information. As such, this methodology enables common ATM performance monitoring, operational concept validation and comparison, and target setting.

The proposed framework is briefly illustrated on the Safety-evaluation example, and the further research guidelines are briefly described.

Introduction

Performance measurement is currently applied across a multitude of industries to quantify and compare the performance of systems, organizations and products. In the ATM discipline, there are numerous metrics used to evaluate the performance of ATM systems, concepts, operations, procedures and strategic goals. However, different organizations involved in evaluating ATM system performance at various levels are not provided with a unified methodology for the selection and definition of relevant metrics. As a result, metrics obtained by one organization are often not comparable to those obtained by another, limiting the ability to comprehend differences in system performance resulting from different technical and/or operational concepts. Also, the relationship between lower-level and higher-level metrics are poorly understood, limiting the ability to roll-up measures to organizational goals and objectives, which is a critical aspect of a performance-based organization. Finally, metrics are not always precisely defined, limiting the reproducibility of results, and lacking a directional quality (e.g., the value of a metric decreased; is that good or bad?).

Given the above situation, comparing operational and/or future systems and concepts, deciding upon R&D results, and acting to meet the expectations regarding system performance has been a long standing problem in ATM. Many research efforts have already tried resolving these discrepancies by developing different frameworks for ATM performance assessment focusing on more or less specific issues. However, they are often limited in scope, and address only the needs of specific ATM stakeholders. More importantly, they do not provide a connection between ATM performance metrics developed by different organizations with diverse objectives.

In 1995, FAA and Eurocontrol initiated a co-operative research and development effort to address these and similar common concerns through several R&D Action Plans (AP). The goal was to establish more common approaches through dedicated workshops or Technical Interchange Meetings, e.g. AP1 (Airborne Separation Concepts) and AP6 (ATM Decision Support Tools). These actions did improve (mostly within the observed areas), but did not resolve the concerns regarding the ATM performance assessment and monitoring.

Starting in 2002, AP2 (Air Traffic Operational Concepts) and AP5 (Validation and Verification Strategy) combined their activities searching for a more general and systematic ATM performance assessment approach. The ultimate objective was to define, through collaboration between the experts from Europe and United States, an ATM performance framework usable in operational performance measurement, concept validation and comparison, and target setting. The research began by analyzing the findings of the comparison between EACAC and AGIE metrics, and by creating an inventory of existing ATM performance metrics. This document describes the continuation of the joint research effort performed in 2003. The goal set for 2003 was to research frameworks applicable to the derivation of ATM performance measures, and select and apply a common framework to ATM performance measures. The research was to be performed through a systematic approach, starting from already agreed principles as described in the following two documents: (1) ISO, 1998, Information Technology — Quality of Service: Framework, ISO/IEC 13236:1998; and (2) ICAO, 2003, ATM Operational Concept Document, AN-Conf/11-WP/4.

Our Approach

In order to develop and apply a common framework to ATM performance measures, this research effort focused on two related tasks described below.

First, we developed common metrics taxonomy by adapting the classification of metrics as described in the ISO IT-QoS framework to the ATM arena [1]. This provides a structured classification of metrics consistent with that found through the metrics collection effort. ATM metric definitions were collected from differing organizations and research efforts to identify (1) identical metrics, (2) metrics with identical definitions but different names, and (3) metrics with identical names, but different definitions.

Second, we developed an ATM performance framework by extending the conclusions of the common metrics taxonomy onto the ICAO performance based ATM Operational Concept [2]. The proposed performance framework enables the ability: (1) to roll-up performance measures to the goals of disparate organizations and stakeholders; (2) to establish the relationships between different levels of performance evaluation (e.g. a single system versus the combined ATM system), between differing aspects of performance (e.g., safety, capacity, flight efficiency, etc.), and between different specific performance measures; and (3) to present performance information in a consistent manner at all levels of performance evaluation.

Both of the above tasks were necessary to resolve the previously identified problems. The application of common metrics taxonomy allows metrics to be categorized in a consistent manner across various organizations. This consistency enables cross-comparison of results. Furthermore, the derivation of metrics through the application of specializations and derivations establishes a clear relationship between metrics at different levels. The proposed framework applies these metrics to provide a consistent manner in which ATM performance can be evaluated along multiple aspects of performance. The framework facilitates the roll-up of measures, allows the development of cause-effect relationships and provides a common method for visualization of performance (Fig. 1).

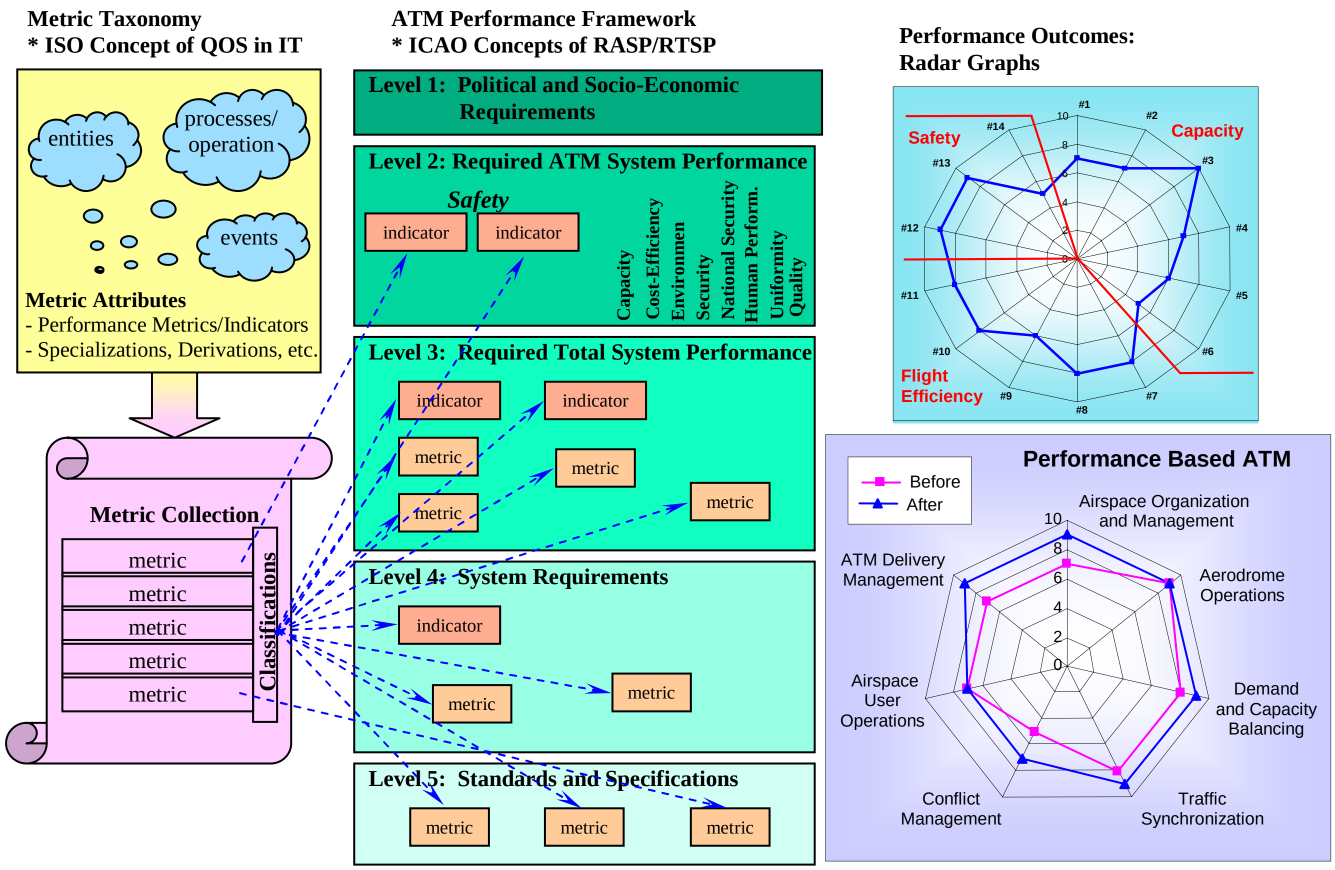

Figure 1. The Overall Framework for ATM Performance Assessment and Monitoring

Requirements

Prior to embarking on the development of the metrics taxonomy and framework, we identified certain considerations and requirements that would constrain our research effort. Specifically, the following three major considerations were addressed: (1) stakeholders –who will use the metrics? what issues are important?; (2) ATM system performance – what determines system performance? what determinants require simultaneous consideration?; and (3) available measurements – what can be measured? how do the measurements relate to specific assessments of performance?

Stakeholders. Categories of stakeholders include the user, the service provider and the regulator entities. ATM system stakeholders include air-carriers, private aircraft operators, military, and air cargo carriers (users); Air Traffic Control and Airport Authorities (service providers); and ICAO and FAA (regulators). The performance framework must be capable of accommodating the different, potentially contradictory, objectives of each of the stakeholders.

ATM system performance. The ATM system is a complex dynamical system with strong interactions between its numerous elements. We abstract this system by decomposing it into the following three components: (1) inputs to the system, comprising mostly of the demand; (2) elements of the system, comprised of the ATM system resources, procedures and policies; can be modified through a management/control process; and (3) output of the system, as described by the system performance.

Performance evaluation can relate to any combination of the three major components and their interaction (Fig. 2). The simplest interaction involves the demand stimulating the system to provide specific system performance. Other interactions will occur; e.g., the demand will also be influenced by the expected and actual system performance, and by the management/control decisions that change system resources, procedures or policies. The performance framework must enable the understanding of these complex interactions. In addition, performance evaluation can focus on only some of the aspects of performance, such as system safety or cost-efficiency. On the other hand, it can focus on overall system performance that encompasses all aspects and their trade-off mechanisms.

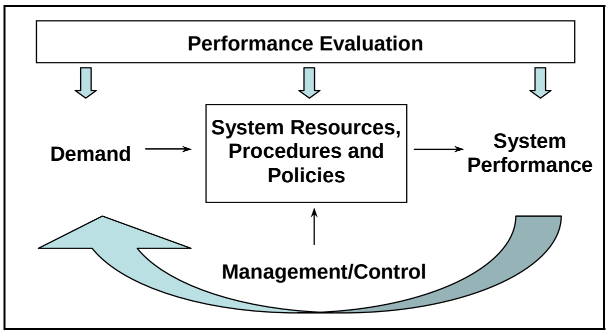

Figure 2. ATM System Performance Determinants

Available Measurements. Other factors can have an effect on measurements obtained for performance evaluation. These include data collection methodology, measurement location, data availability and environment (simulated and real). The performance framework and metrics definitions must be robust to these measurement differences.

Given the above requirements, the common performance framework must facilitate the following: (1) evaluation of overall system performance as well as individual aspects of performance; (2) evaluation of system-wide and regional performance; (3) evaluation within the real-world or a modeling environment; and (4) comparison of system performance before and after observed changes are implemented.

Metrics Taxonomy

The proposed metrics taxonomy represents a standardized methodology for the classification of ATM performance metrics. As such, it establishes a foundation for maintaining the connection between the analogous performance metrics that were specified using different terminology and/or defined by various organizations for various purposes.

The proposed metrics taxonomy is established upon the principles of the ISO IT QoS approach and is consistent with the ICAO performance based ATM operational concept. It also meets the requirements that the common performance framework must accommodate discussed in the prior section.

To ensure consistency among the existing ATM metrics obtained from a literature search, the classification was applied by providing each metric with its unique signature. Metric signature consists of metric name, definition, origin, and an array of relevant specifiers, defined as the attributes indicating the applicability and the scope of the metric.

The proposed taxonomy is developed through a bottom-up approach, by investigating the explicit information each metric conveys and by identifying all of the relevant specifiers. It is founded on the decomposition of the ATM system into the following elements: (1) entities, which define objects that a measurement is applicable to; (2) events, which define the scope of the measurement; and (3) processes and operations, which define the settings within which the performance is realized.

An entity has distinct and independent existence that indicates the object the observed metrics pertains to; e.g. air traffic controller, aircraft operator, flight, departure, arrival, flight segment, aircraft, airport, runway, etc.

An event is something that takes place and indicates a single occurrence of a process, or the beginning and the end of processes performed by or performed on an entity; e.g., engines-on, receiving a taxiing clearance, entering a departure queue, hand-off initiation, reaching a top of climb, etc.

A process is a series of actions or steps towards achieving a particular objective, and an operation is an instance of a process that can be singled out and observed independently. Both processes and operations are triggered and ended by the appropriate events.

A process is often associated with a development of a particular flight segment or a provision of a particular service during which simultaneous events and/or operations are taking place. The operation specifier may indicate the relevant performance aspect(s). For example, cruising is a process (relating to the cruise-flight segment and/or the current sector) during which many operations are taking place, some of which happen simultaneously. Among them are controlling the flight parameters (typically relating to the flight efficiency or pilot workload), traffic monitoring (controller or pilot workload), conflict detection and resolution (safety or workload), etc.

In short, entities and events define how a measurement is conducted, whereas processes and operations indicate what is being measured. Note that, depending on the perspective, the same object can be both an entity and a process, e.g. a flight.

The above generic structure allows metrics to be defined in a consistent manner by indicating specifiers referring to common entities and events.

Primary Classifications

Primary classifications are directly related to the essence of information communicated by each metric. The proposed classification enables the aggregation of metrics into categories representing the most important characteristics of the ATM system and its performance. This classification scheme was derived through discussions with researchers and a review of the ATM literature. As such, it is likely to be modified as new metrics are introduced into the metric collection. However, that will not affect the structure of metrics taxonomy, but will only allow for more precise metric classifications. There are 11 primary classifications at this time, summarized as follows: (1) Operational environment – relating to the ATM system resources, procedures, policies, and weather; (2) Demand – relating to the amount of service being requested of the system; (3) Capacity – relating to the amount of service that can be provided; (4) Flight efficiency – relating to the ability of the users to accomplish efficiency of gate-to-gate operations; (5) Cost efficiency – relating to the costs of all aspects of the ATM system (monetary value); (6) Safety – relating to the ability to prevent accidents and safety-related incidents at all levels in the ATM system; (7) Security – relating to the ability to prevent unauthorized access or utilization of all aspects of the ATM system; (8) CNS – relating to the ability of Communications, Navigation and Surveillance (CNS) elements to fulfill their purpose; (9) CDM and equity – relating to the ability of the ATM system users to participate in collaborative decision making (CDM) and receive equitable service by the service provider; (10) Environment – relating to the impact on the environment by the ATM system; and (11) Human performance – relating to the ability of human participants in the ATM system to perform their tasks satisfactorily.

Even though a metric can be relevant to one or more different characteristics of the system and its performance, it essentially conveys information directly relevant to only one of the characteristics. Therefore, metrics can only be assigned to a single primary classification.

For example, “Average Flight Delay” typically indicates flight efficiency as well as whether capacity is exceeded by demand within the observed resource and time frame. However, this metric does not provide any other specifics about capacity or demand. In fact, theoretically at least, delay can be accumulated for reasons that are not capacity related, such as a flight simply veered off its intended trajectory. This example might be far fetched, but it clearly illustrates that the essence of the information conveyed by the metric *precisely* relates to only one aspect of performance: flight efficiency. In addition, calculating the value of this metric involves observation of flight times, which are also flight-efficiency related. Consequently, the metric should be uniquely classified using flight efficiency primary classification and appropriate specifiers including flight entity, flight process, efficiency measurement aspect, and appropriate location-specifier.

Specifiers

The above primary classifications represent the starting point in the derivation of metrics. By applying a series of specifiers to the above classifications, categories of metrics can be obtained. The exact order in which specifiers are applied is irrelevant for the proper taxonomy. However, determining all relevant specifiers and including each one into the unique metric signature is essential for accomplishing consistency. The following generic specifiers have been identified:

- Entity – used to classify metrics corresponding to a particular entity;
- Process – used to categorize metrics according to the applicable process

(typically indicates the relevant flight segment);

- Operation – used to classify metrics according to the applicable operation;
- Location – used to classify metrics according to the location to which the metric applies;
- Measurement aspect – used to classify metrics according to the type of data collected (e.g., reliability, accuracy, predictability);
- Generic group (for all other specifiers) – Useful for aggregating specific ATM-type system states such as VFR/IFR or IMC/VMC.

Metric Categories

For each primary classification, specifiers can be applied in a unique order to develop a "tree" of metric categories. This tree is referred to as a base design unit. The base design unit can be applied to all primary classifications leading to a large collection of categories into which metrics can be placed. For example, metrics related to the runway departure queue can be applied to such classifications as the environment or flight efficiency.

Since the exact order in which specifiers are applied is irrelevant to the proper taxonomy, there are many different ways of grouping metrics with similar characteristics, and consequently many, more or less efficient, base design units. For example, a base design unit can be developed around distinct flight segments and repeated for each of the 11 primary classifications (Fig. 3.). This base design unit facilitates easy visualization of the smallest segments within which data collection takes place.

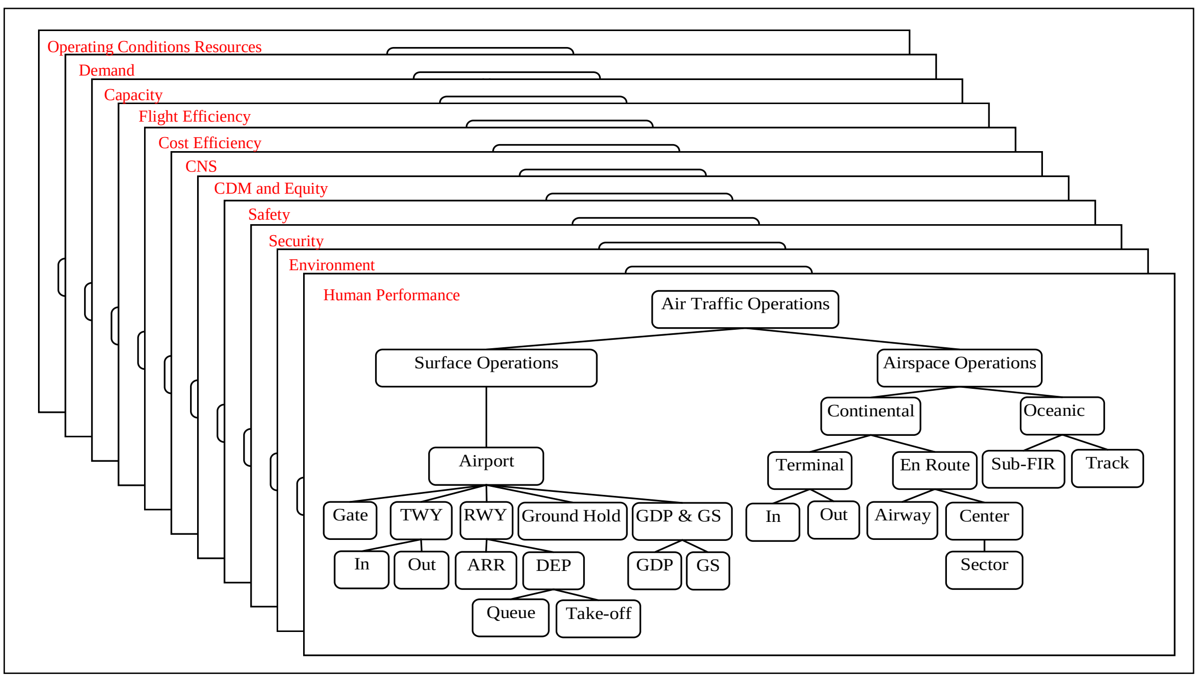

Figure 3. Base Design Unit Applied Onto Each of the Primary Classifications

Since metric categories are all derived through the addition of specifiers on higher-level categories, the ability to establish relationships between the lower levels and the higher levels is simplified. The relationship between two metric categories can be understood by examining the chain of links between them. The chain contains all possible intermediate partitions, each revealing the details of the metric signature, including metric's exact relevance, scope, and position relative to other metrics it affects or is affected by. The metrics placed within each category can further be filtered and grouped by the relevant entity, measurement aspect, or in some other way by applying the proper specifier. Such partitioning assists in locating appropriate metrics by breaking potentially long lists of metrics into groups that communicate similar information. Consequently, this structure facilitates an aggregation of metrics according to any relevant "dimension" (e.g., locations, operations, entities, etc.).

ATM Performance Framework

The proposed ATM performance framework is based upon the ICAO hierarchical structure of the performance-based ATM operational concept. The ICAO framework establishes a series of levels with political and socio-economic requirements at the top level, and standards and specifications at the bottom level (Fig. 4). However, it does not identify relationships between metrics across levels. This paper applies the previously described metrics taxonomy to help establish the inter-level metric relationships.

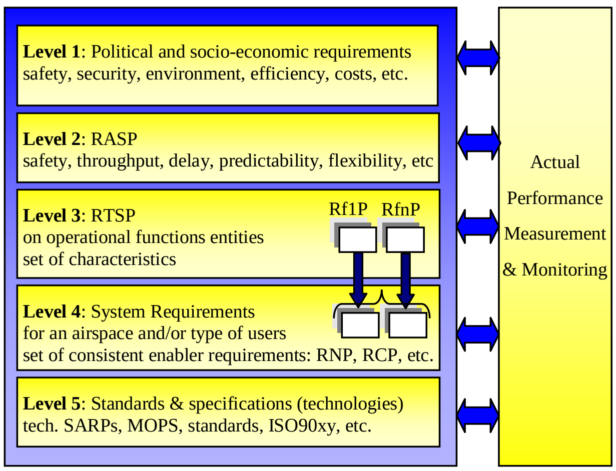

Figure 4. Hierarchical Structure of the ICAO ATM Performance Based Operational Concept (Source: ICAO, AN-Conf/11-WP/4, 2003)

ATM performance framework is developed through a top-down approach by investigating how the expectations translate into the required total system performance. Relevant metrics are identified along with their implication for ATM system performance. The proposed framework has the same five layers as the ICAO structure, and implements nine performance objectives as the "primary characteristics" of ATM system performance that correspond to the Eurocontrol ATM 2000+ Strategic Objectives [3].

Framework Hierarchy

Level 1. The highest-level manages political and socio-economic considerations determined by the societal role that the ATM system is expected to fulfill. These are contemplated by national governments, civil aviation and other regulatory authorities, and ICAO. While the expectations and requirements at this level were considered, their precise specification is out of scope for this effort.

Level 2. The RASP level manages the external perception of the quality of air transportation service provided by the ATM system, which streams from the expectations and the perceptions of ATM community and the traveling public. These expectations apply to the entire ATM system and correspond to the complex correlation of the distinct requirements with respect to safety, security, environmental impact, cost-effectiveness, capacity, equity of treatment, general efficiency and quality, flexibility, predictability, uniformity, and participation by the entire ATM community.

Level 3. The RTSP level manages the internal perception of the quality of services provided by the ATM system, which streams from the functionality of the ATM components as they collaboratively deliver the required levels of RASP. RTSP incorporate all system capability aspects, including the required services and information management, and the interoperability of systems and procedures. The RTSP requirements are designed to assure the delivery of RASP to the entire ATM community, and that the community implements and monitors them properly. Performance metrics at the RTSP level are defined through the specialization of the RASP performance metrics with respect to the relevant ATM service.

Level 4. The System Requirements level manages the internal perception of the quality of services provided by each of the ATM system components, which streams from the performance of the operational and technical functions performed within service. Therefore, performance metrics at this level are defined through a specialization of RTSP performance metrics with respect to the specific operational and technical functions performed within the relevant ATM service. Due to the change in the level of abstraction, the direct meaning of a performance metric at this level can be quite different from the RTSP level metric it directly affects.

Level 5. The Standards and Specifications level manages the internal perception of the performance of the operational and technical functions providing for ATM services, which streams from the performance of the implemented technologies and methodologies. Therefore, performance metrics at this level are defined through a specialization of the System Requirements, i.e., with respect to the specific technologies and methodologies used to implement the corresponding operational or technical function performed within the relevant ATM service. Again, the direct meaning of a performance metric at this level can be quite different from the corresponding higher-level metric it affects, and the connection between the two indicates the existence of a cause-effect relationship.

Primary Characteristics

Primary characteristics represent the distinct and quantifiable aspects of ATM system performance within each of the previously discussed levels. These primary characteristics are defined independently from the means used to represent or control it. They facilitate understanding the individual contribution of the distinct ATM performance aspects towards the total system performance, the complex relationships that these distinct aspects have with each other, and the trade-off mechanism or potential substitutions of one aspect for another that would trigger desired change in the total system performance. In addition, they correspond to the highest-level objectives that the ATM system pursues with the ultimate goal of meeting the stakeholders' expectations.

There are nine primary characteristics proposed at this time, briefly described as follows: (1) Safety, defined as the ability to ensure the safe separation of aircraft, both in the air and on the ground, expressed by the number of ATM-induced accidents and serious or risk bearing incidents; (2) ATM security, defined as the ability to enhance the response to security threats and events affecting flights (aircraft and passengers) or the ATM system; (3) Cost effectiveness, defined as the ability to reduce the direct and indirect ATM-related costs per unit of aircraft operations; (4) Capacity, defined as the ability to provide sufficient capacity to accommodate the demand of all users in an effective and efficient manner at all times, and during typical busy hour periods without imposing significant operational, economic or environmental penalties under normal circumstances; (5) National security and defense, defined as the ability to improve the effectiveness of existing, and determine new, mechanisms, criteria and structures to enhance civil-military co-operation and co-ordination, and to ensure access to, and availability of, airspace for military purposes through the implementation of special procedures where necessary; (6) Environment, defined as the ability to obtain improvements in ATM through work with ICAO and its member states, in particular the accelerated implementation of those CNS/ATM concepts, procedures and systems that help to mitigate the impact of aviation on the environment; (7) Uniformity, defined as the ability to ensure that ATM operations are compliant with ICAO CNS/ATM plans and global interoperability requirements, provide a seamless service to the user at all times, and operate on the basis of uniformity; (8) Quality, defined as the ability to foster, promote and enhance the use of recognized quality management standards, systems and continuous improvement methodologies in the provision of gate-to-gate ATM services, products and processes; and (9) Human involvement and commitment, defined as the ability to ensure human involvement and commitment in making possible the changes in future ATM so that operational, technical and support stuff can operate effectively, efficiently and safely within their capabilities, and obtain challenge and job satisfaction.

Note that the relationship between the ATM framework and the metrics taxonomy is not one-to-one. Primary characteristics are not equivalent to the primary classifications, even though some categories overlap. This distinction can be understood by realizing that the metrics taxonomy is used to classify measurable elements of the ATM system, whereas the ATM performance framework is used to understand the performance of the ATM system. In fact, metrics taxonomy is concerned only with where, how, and what is being measured and what entities, events, processes and operations the metric pertains to. ATM performance framework, on the other hand, is concerned with a potential value that the metric may have onto the overall system performance outcomes.

For example, in order to understand the performance of the system from a capacity point of view, it is typically necessary to seek metrics related to both capacity and demand. In fact, metrics may be pulled in from various *classifications* to represent performance at a particular level for any particular *characteristic*. If the specification has been conducted in a manner consistent with the definition of the levels, then rollup to the higher levels in the ATM framework will require consideration of parent metrics in the taxonomy (Fig. 5).

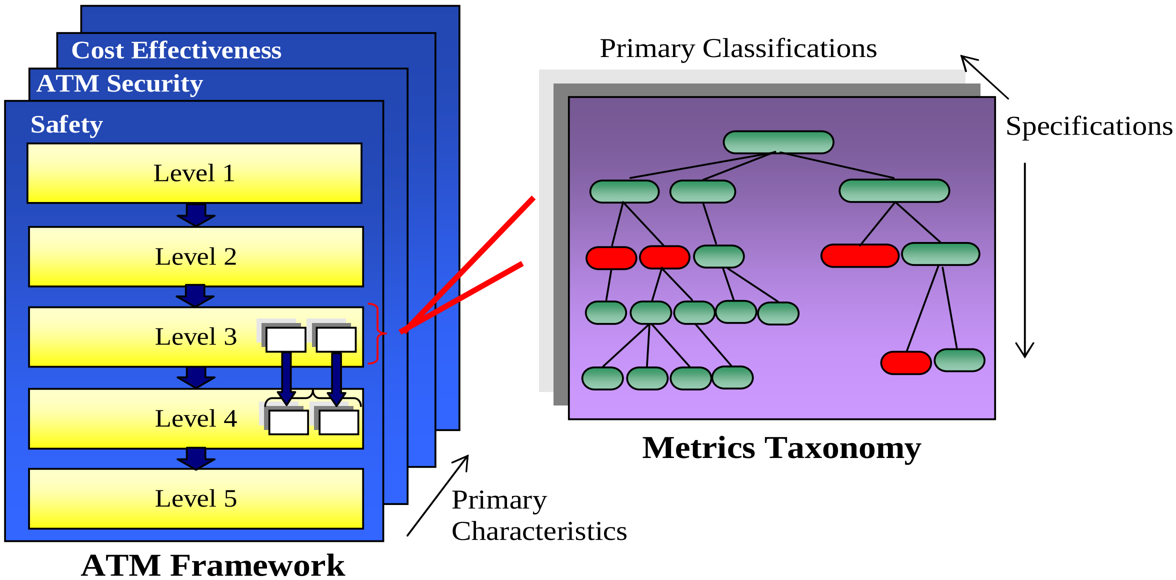

Figure 5. Relationship between Metrics Taxonomy and ATM Framework

Specifiers

In addition to the specifiers defined previously, an additional specifier indicating the relevant "ATM Service" is incorporated into the metrics taxonomy to support the ATM framework. ICAO defines seven principal ATM services: (1) Airspace Organization and Airspace Management – organization establishes a global airspace structure, and management applies strategies, rules and procedures; (2) Aerodrome Operations – enables safe and efficient use of the aerodrome infrastructure; (3) Demand and Capacity Balancing – enables safe and efficient traffic flow through strategic actions; (4) Traffic Synchronization – enables safe, orderly and efficient traffic flow through tactical supervision and actions; (5) Airspace User Operations – enables safe and efficient flight operations of variously equipped airspace users; (6) Conflict Management – assures separation and collision avoidance through co-operation with other services; and (7) ATM Service Delivery Management – enables seamless gate-to-gate services for all phases of flight and all service providers.

Examples: ATM System Safety Assessment and Monitoring

Due to space limitations, the following example summarizes only a segment of an ATM performance framework application that is developed for the safety primary characteristic.

RASP Level Safety Objective

To improve safety levels by ensuring that the numbers of ATM induced accidents and serious or risk bearing incidents do not increase and, where possible, decrease.

RASP Level Metric Example

Number of ATM induced accidents. This metric is defined as the total number of ATM induced accidents, within a specified period of time and specified geographical area. It can be quantified as accidents/year, accidents/month, etc., and can have numerous specializations including the particular ATM Services, type of accidents, part of the ATM system, period of time, etc.

RTSP Level Metric Example

Number of accidents due to among other things an error in the Aerodrome Operations Service. This metric is defined as the total number of cases where this service contributed to an ATM induced accident, within a specified period of time and geographical area. It can be quantified as accidents/year, accidents/month, etc., and can have numerous specializations including particular type of accidents, location or part of the ATM system, period of time, etc.

Other examples include the analogous metrics that correspond to the remaining six ATM Services.

System Requirements Level Metric Example

Runway Incursion Causes: Surface Movement Management. This metric is defined as the number of cases where a runway incursion occurred due to an error in surface movement management, within a specified period of time and at the specified airport of interest, divided by the total number of runway incursions that occurred in the same period of time and at the same airport. It can be quantified as percentage (%), and can have numerous specializations including particular type of incursion (aircraft with another aircraft, aircraft and vehicle), particular location (intersection of two runways, or a runway and taxiway, etc.), particular meteorological conditions, type of airport (with or without tower-control), etc.

Standards and Specifications Level Metric Example

Aeronautical Information System (AIS) technology: NOTAM deliver errors. This metric is defined as the number of times that a NOTAM did not reach one or more of its destinations, in a specified period of time. It can be quantified as per year, per month, per day, and can have numerous specializations, including particular type of NOTAM messages, particular means of delivering NOTAM, particular part of the ATM system.

Flight Data Processing System (FDPS) technology: Flight data processing errors: fix name errors. This metric is defined as the number of cases when a fix name in a flight plan entry was misinterpreted, per specified period of time. It can be quantified as per year, per month, per day, and can have specializations such as particular type of fix.

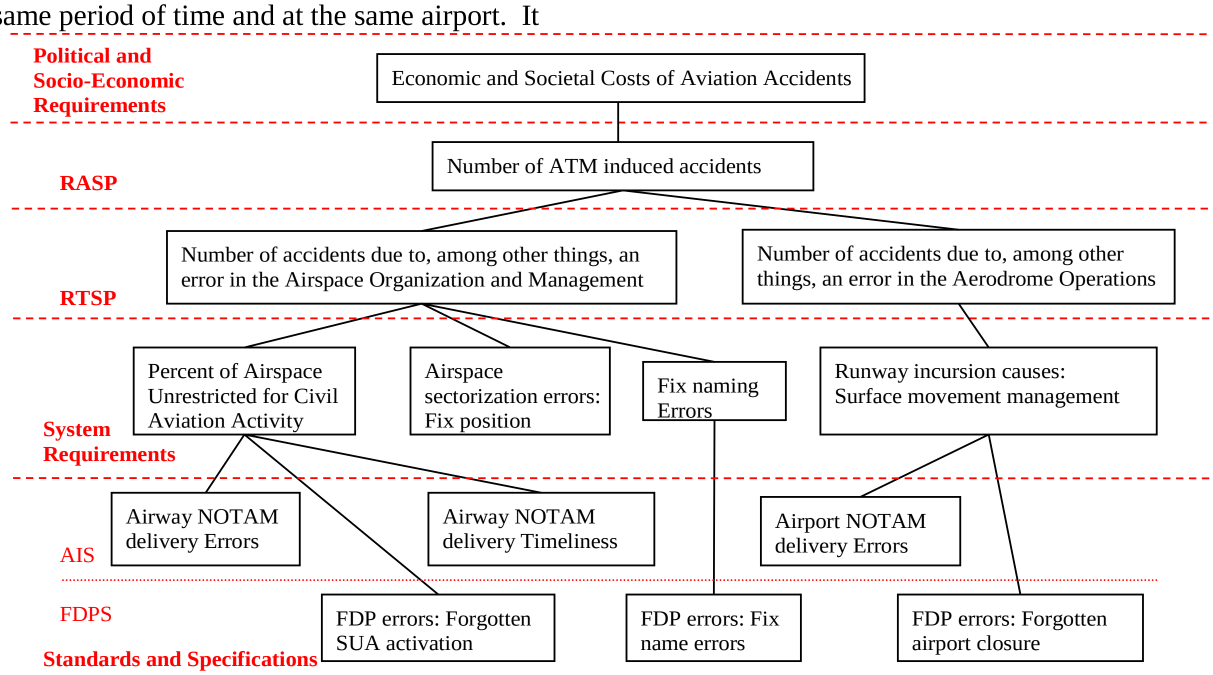

Figure 6. A Segment of the Cause-Effect diagram for the Safety Primary Characteristic

Cause-Effect Diagrams

ATM performance analysis involves applying the relevant metrics at each of the relevant levels to obtain quantitative performance measurements of the aspect(s) being investigated. Establishing the relationships between the different levels can be facilitated through the relationships identified in the metrics taxonomy, and can also be visualized through cause-effect diagrams. These diagrams establish a connection between the relevant metrics based on their influence on each other, which is represented by a link between the metrics. Each link signifies only the existence of the dependency and bears no meaning to its type. In fact, it is often difficult, if not impossible, to develop an exact mathematical expression of the relationship between the metrics in many cases. Quite often, the relationship is stochastic rather than deterministic, or may simply be based on empirical knowledge.

Figure 6 illustrates the segment of the cause-effect diagram that includes the metric examples presented in the previous section.

Performance Presentation

The final step in designing the common performance framework is designing common methodology for presenting the findings of a potential ATM performance assessment or monitoring, i.e., designing the “performance dashboard”. Since performance assessment is a goal driven process, the obtained results and the best way to present them are typically application specific. Therefore, one of the essential requirements governing the dashboard design is its flexibility in scope, type and amount of information it needs to accommodate. Radar graphs offer one of the most expressive ways to successfully meet all of these requirements. They are used to display many different variables simultaneously, and can utilize any number of axes and data sets, limited only by graph’s readability.

A radar graph is a two dimensional polar graph primarily used as a means of comparison between multiple data sets. It has a separate, uniformly-spaced radial axis for each included variable, and each axis can have different units as appropriate for evaluating the corresponding variable. Included variables emanate from the origin, or center of the polar plot and their values increase towards the outer border. By connecting the variable values on each radial axis with a straight line, a distorted polygon is created. Multiple polygons can be simultaneously plotted and visually compared, representing different outcomes resulting from the same experiment repeated under different conditions.

Radar graphs can be interpreted either by reading the actual values on the axes, or by comparing the areas enclosed by the polygons. Whether one data set is "better" depends upon whether the more favorable values lie in the center of the graph, or at the circumference.

At a given level of the performance framework, one can plot the performance for each of the nine primary characteristics as measured by an aggregate indicator for each characteristic (Fig. 7). Such radar graph, designed for the RASP assessment, has nine radial axes used to display the total system performance expressed by an appropriate indicator for each of the nine primary characteristics. The values chosen to display at this highest level of performance evaluation can correspond to a single best representative metric or to an index formulated as an appropriate combination of several representative metrics Each characteristic may not have the same units of measure, but the direction of change has to be consistent for each characteristic (e.g., more is better). Thus, in Figure 7, we see that all characteristics have improved with the exception of environment.

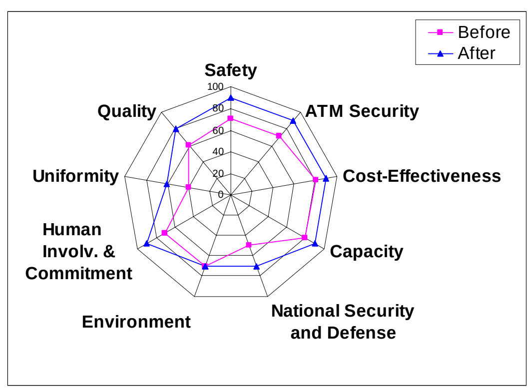

Figure 7. Radar Plot of RASP Assessment

This type of plot is useful in visualizing the areas of improvement and the high-level tradeoffs between different characteristics. When denominated in common units to account for the relative value placed on each characteristic, an individual may use such a plot to easily identify specific tradeoffs between each of the primary characteristics. An example of such would be the translation of all measures into the financial value of each (e.g. Dollars or Euros).

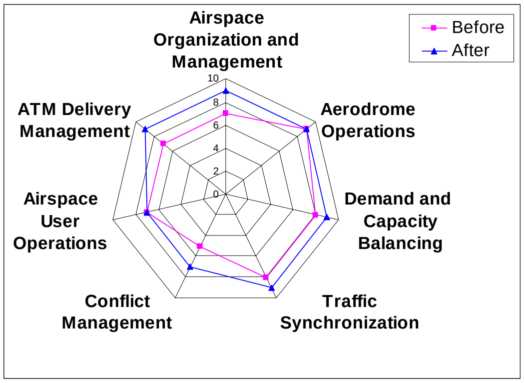

Figure 8. Presentation of the Safety Assessment Outcomes at the RTSP Level

The same type of plot can be used to drill-down into a specific primary characteristic for the purposes of identifying the relative contribution of each ATM service to the touted change (Figure 8). The radar plot can also illustrate different outcomes of a particular performance evaluation using a collection of differing metrics. In short, this method is flexible and may display changes in outcomes in such a way that facilitates the visualization of changes.

Conclusions and Further Research

This paper describes the initial results of an on-going joint research effort, involving FAA, Eurocontrol and their partners, that addresses the long standing problem of discrepancies in ATM performance measurement and monitoring. By analyzing a connection between the analogous terminologies, measures and generic performance considerations used across the international ATM community, we established a basis for the common classification of metrics. This common classification methodology provides the potential to greatly reduce, in not eliminate, issues associated with inconsistent and poor metrics definition. Moreover, we incorporated the common classification principles into the ICAO concepts of RASP and RTSP, thus creating a structure that established the connection between stakeholder expectations and system performance requirements to meet those expectations as propagating through several levels. The proposed methodology identified principles of drilling down from the highest-level objectives to the very specialized and detailed low-level requirements and metrics, and reveled distinct performance characteristics and considerations at each of the levels.

However, it is important to point out that this paper describes an intermediary product of a work in progress that also needs to be translated into practical reference material. As such, it is not fully adopted by the FAA and Eurocontrol. In order to proceed in this direction, the following steps are determined as the most important for continuing this research effort in 2004: (1) Strengthen the methodology for ATM performance assessment and monitoring, including the scope and definition of the primary classifications and primary characteristics, (2)Fully develop the safety objective example, and (3) Further research the aggregation of lower level measures into higher-level requirements and expectations by applying the cause-effect approach to other highest-level primary characteristics.

References

[1] International Standardization Organization, 1998, Information Technology — Quality of Service: Framework, ISO/IEC 13236:1998

[2] ICAO, 2003, ATM Operational Concept Document, AN-Conf/11-WP/4, www.icao.int/icao/en/anb/meetings/anconf11/

[3] Eurocontrol, 2003, *Air Traffic Management Strategy for the Years 2000+*, www.eurocontrol.int/eatmp/library/documents/