Abstract

This paper discusses the need for a Performance Framework (PF) with a methodology to support the assessment of a European-wide concept in ATM.

The article describes the work undertaken in Episode 3 (a European Commission Sixth Framework Programme research project) to develop a Performance Framework based on deliverables from the SESAR Definition Phase and associated Influence Models (IM) used to express the links between required performance and proposed improvements, including results from validation exercises.

The Performance Framework includes the SESAR target performance, while the Influence Models are an extension of work started in SESAR to trace the relationship between Target Performance and the SESAR proposed concept.

This article also covers the approach, illustrative assessments using the Performance Framework and Influence Models, and early lessons learned with an outline of future tasks.

Introduction

At the launch of the SESAR\textsuperscript{1} Definition Phase in 2005, European Commission Vice President Jacques Barrot set highly challenging performance targets for the future European Air Traffic Management (ATM) system of 2020 and beyond.

These targets challenged the designers of the future European ATM system to:

- enable a 10% reduction in the effects that each flight has on the environment; and
- reduce the cost of ATM services to airspace users by at least 50%.

These statements have been widely communicated among the ATM community and the general public. Although these targets are stated as broad goals, they convey a political vision for the performance of the future European ATM system. The vision sets out the main challenges for the SESAR programme.

In the SESAR Definition Phase, the ATM stakeholders elaborated on the above vision with a more detailed set of strategic performance objectives, indicators and targets, called the SESAR Performance Framework, which is structured around the 11 ICAO Key Performance Areas (KPAs) as defined in ICAO Doc 9883 [1].

The strategic performance objectives challenged the Definition Phase designers of the SESAR operational concept, architecture and technical enablers to show how these contribute to the overall performance of the future system. The performance experts were also challenged to define how the concepts, enablers and performance objectives could be framed in a coherent and integrated manner.

Now validation experts are working to prove the performance contributions brought by the concept and associated technical enablers while ensuring traceability from the multitude of validation projects of operational improvements to the top-level performance targets.

The SESAR Definition Phase provided the general direction by developing an initial performance framework in Deliverable 2 “The Performance Target” [2]. Further work was done in the Definition Phase when Influence Diagrams were used to trace the proposed improvements to performance objectives.

\textsuperscript{1} Single European Sky ATM Research programme.
Episode 3 (EP3) is a project funded under the European Commission’s Sixth Framework Programme (FP6). The objective of the project is to clarify the SESAR 2020 concept and to perform some early validation activities.

The EP3 project has developed a Performance Framework with the participation of several air navigation service providers (France DSNA, Germany DFS, Spain AENA, UK NATS, China ATMB, CAST), Engineering ISDEFE and EUROCONTROL research. This Performance Framework provides a structure (in a top-down fashion) for the target performance through KPI to low-level proposed improvements. When EP3 started, no such framework existed outside of the SESAR Definition Phase material. EP3 provides for the integration of the various performance contributions expected from the many proposed improvements at a network (European-wide) level.

For a deeper understanding of the links between performance objectives and the proposed improvements, the EP3 work revisited the SESAR Influence Diagrams. These were integrated in a common tool to ensure consistency between the diagrams and the model while rendering them more dynamic and open to quantifiable analysis.

The EP3 research noted that there are very few existing models or methods that provide for a consolidated European-wide assessment of an operational concept. EP3 identified the need adopt system engineering techniques: to ensure the concept and associated enablers remain coherent and aligned; and to apply these to the validation activities.

EUROCONTROL has further supported EP3 with the development of the Influence Model by HELIOS.

As a result of initial application of a performance-based validation approach within EP3, this paper provides a first overview of the challenges related to the further development of the Performance Framework and the Influence Models.

Some issues

For the EP3 research, the validation approach faced a number of issues:

- validation is usually done at the level of individual projects on low-level concept elements; however, EP3 is charged with providing a “system view” of the benefits expected of the SESAR concept on a European-wide scale;
- there were no existing tools identified for the integration of all the SESAR proposed improvements, KPIs and validation on this scale;
- the ability to integrate results (qualitative and quantitative) from past and ongoing projects to increase the available data is considered important;
- this data and EP3 output would have different levels of maturity, which needs to be considered when aggregating results;
- the ability to trace the maturity or uncertainty surrounding a proposed improvement and its validation as this progresses through concept lifecycle (e-OCVM [3]), should be possible; and
- it should be possible to efficiently undertake sensitivity analysis, while managing uncertainty and trading between Key Performance Areas (KPA)/Key Performance Indicators (KPI) and the proposed improvements with a view to future prioritisation of validation exercises.
The Performance Framework

Problem setting

How to develop a very large system using a performance-based approach?

The target performance of the new system is described in terms of Key Performance Areas, performance targets and metrics gathered into a Performance Framework.

So as to achieve the performance targets, the system will incorporate a number of research solutions. Each solution is represented by a proposed improvement.

Figure 1: Episode 3 Performance Framework layers

The proposed improvements are at different levels of maturity and only partially validated and mostly on an individual basis; at the deployment the interactions between them are complex and not always well understood.

The proposed improvements have been defined in the SESAR Definition Phase and are laid out in the SESAR Master Plan [4].

How to ensure that the proposed improvement (the target concept) will provide the target performance at the system-wide level?

The SESAR approach aims to develop a Performance Framework (PF). The PF is the structure for the performance log of the system that lives through its development, recording its performance level at each validation step, and allowing gaps in system performance to be identified. The PF helps to identify what enablers are able to bring in terms of performance to meet the performance target. Along the development life cycle, validation steps increase the level of certainty of the future performance.

How to combine the performance of individual proposed improvements at a local level into a European-wide performance?

With 11 KPA, 183 proposed improvements (and many validation results already known or planned for production) the complexity arises from the combinations between all of these elements, thus:

- each proposed improvement contributes to several KPAs with different impacts;
- a proposed improvement designed to have a positive impact on one KPA, may have a negative impact on another KPA; traditionally, for example, trade-off has been required between flight efficiency and capacity;
- each validation exercise includes a particular combination of proposed improvements;
- each validation exercise provides results about several KPA.

An understanding of the factors that drive performance is essential for managing this complexity. The ability to trace the contributions of a proposed improvement to KPAs through the validation results is also important. Therefore, it is necessary to record the conditions under which validation results are obtained, the step included, the results in respect of the different KPA, and any uncertainty about results.

EP3 Performance Framework [12]

The objective of the EP3 Performance Framework is to provide a methodology by which to develop a model for future assessment of the SESAR 2020 ATM concept [5] on a European1 wide basis since no such model currently exists.

[1] Example: Continuous Descent Approach (CDA) improves fuel efficiency, but tends to degrade capacity.
This methodology allows the aggregation of validation measurements of different levels of granularity, e.g., local versus regional and uncertainty, e.g., expert judgment versus simulation. The methodology is based on:

- an understanding of the elements that contribute and influence the performance (Influence diagrams) [13];
- a European model that represents the elements that are linked and the mechanism to combine their influences. (i.e., Influence Models) [14];
- The definition of a catalogue of common Performance Indicators (PIs) as references to ensure consistency and capture data about the influencing factors from exercises, expert groups, as well as current and past studies.

The EP3 Performance Framework has further developed the SESAR Definition Phase material in six of the Key Performance Areas (KPAs):

- Capacity
- Efficiency
- Flexibility
- Predictability
- Safety, and
- Environment.

The EP3 Performance Framework has defined different layers of the Performance Indicators (PIs) according to the Performance Areas (PAs). The catalogue of PIs is divided into three layers:

- European-wide
- European Airport, TMA, En Route
- Local Performance Indicators for Airport, TMA and En-Route.

The document provides a methodology by which to develop a model for future assessment of the SESAR concept on a 2020 European-wide basis; no such model currently exists. In view of this, the methodology explains how the validation measurement of different levels of granularity and uncertainty will be aggregated and also provides common reference on "WHAT" to measure.

The EP3 Performance Framework is based on the definition of SESAR KPAs, Focus Areas and PIs and their targets and how these fit with the requirements defined in the operational concept. The Performance Framework is composed of KPAs, Focus Areas and performance indicator layers. A common approach to integrate the results from the validation exercises is described in the scheme (see figure below).

![Figure 2: Performance Framework relationships](image-url)

The Influence Diagrams developed in SESAR D3 deliverable (ATM Target Concept [5]) and in deliverable D4 (ATM Deployment Sequence) were used as a starting-point to build the relationships among the different layers of Performance Indicators and Key Performance Areas. The EP3 Influence Diagrams were developed during 2008 and delivered in early 2009. The EP3 improvement to the SESAR Influence Diagram work includes a wider coverage of the KPA and the use of a decision support tool that ensures consistency between the Influence Diagrams and the Model and allows diagrams to be presented at various levels of detail.

The Influence Model is a performance log: an evolving repository of performance targets and achievements with assessment of confidence level. Validation assessment results are regularly stored in the Influence Model and this reduces the uncertainty on performance achievements.
Influence Models

*All models are wrong but some are useful.* - G. Box.

Influence Models and Episode 3

Influence Diagram techniques were initially developed in the 1970s and became more widely used following the work of Professor Ronald Howard and Dr James Matheson (Morgan2003 [7]). As described by Kjaerulff and Madsen in (Uffe2008 [8]), an Influence Model can be defined as a probabilistic network for reasoning about decision making under conditions of uncertainty. The representation of a decision problem involving a sequence of interleaved decisions and observations is both graphical and mathematical as follows:

- Graphical representation offers an intuitive way to identify and display essential decision elements, and reduces the apparent complexity of a problem.
- Mathematical representation enables uncertainties and objectives to be assessed.

An Influence Diagram is a graphical representation of a decision problem; whereas an Influence Model (IM) - that is, an Influence Diagram with an associated set of mathematical formulae and data - is an executable representation of a decision problem.

In the EP3 project, an Influence Model is being constructed of the SESAR Target ATM Concept for 2020 - using the results of the SESAR Definition Phase as a starting-point - to better understand and evaluate the links between the PF and the proposed improvements.

The proposed role of the Influence Model is three-fold:

- capture the links between the steps and PF with diagrammatic representation of the associated influences; for example, a step which reduces runway occupancy time could lead to increased runway capacity.
- structure the available validation evidence (e.g., experimental results, simulation outputs, expert judgements)
- enable an initial performance assessment of some steps within an integrated Influence Model.

Creating Influence Diagrams

In view of the complexity of the problem space in EP3, an integrated tool is essential to support the IM. A number of tools and notations were investigated (EP3-2008b [9]). The final choice was ‘Analytica’ - a COTS tool specially designed for the development of Influence Models. Analytica combines a graphical user interface with powerful mathematical modelling capability and an ability to link to external results using Microsoft OLE.

An integrated set of Influence Diagrams have been constructed within the tool to show the influences between the steps and the focus areas - Capacity, Efficiency, Predictability, Environment, Flexibility, and Safety - of the Performance Framework, plus Cost Effectiveness (EP3-2009a [10]).

The EP3 project made it clear that modelling rigor must be maintained if Influence Models are to be used as a tool for stakeholder consultation. This includes: strict adherence to the methodology; and deployment of similar modelling techniques across the whole model; for example, the level of abstraction, the use of absolute or relative values, and the use of “phase of flight” as a criterion for structuring the model.

![Figure 3 Fuel efficiency Influence Diagram](Source: EP3)

This figure above presents a top-level view of the fuel efficiency in which Airport, TMA and En-route phases of flight are treated separately.
The fuel efficiency diagram has been developed to compare the actual amount of fuel burnt with a reference target.

This illustration is a representation of a first layer. As an example, we will now examine in detail the content of the En-Route module.

**Figure 4 En-Route fuel burnt trajectory influences**

The main driver for the fuel burnt en-route is the trajectory that the aircraft flies. This trajectory is influenced by conflict management actions, such as vectoring and speed control, the design, availability and utilisation of the airspace, and the structure of routes. The trajectory flown is also influenced by such factors as designated military areas, meteorological conditions, and route charges. The Influence Diagram below represents a transcription of these main influence factors.

**From Influence Diagrams to Influence Model**

This first step has demonstrated the value of Influence Diagrams as a means of capturing influences. Integration of a series of influence diagrams enables the linkages between various KPAs - e.g., Efficiency (fuel) to Environment (global emissions, etc) - to be identified and the need for trade-offs established.

The next stage is *quantification* and is crucial in supporting an understanding of the potential performance and trade-off between competing concepts.

The process of quantification requires data that derive from sources such as expert judgement, experimental observations, or - where possible - observed performance (in particular, for the 2006 baseline). Indeed, quantification is the process of enabling this data to be used for an estimate of the performance gain from implementing a coherent set of proposed improvements and to assess the trade-offs between them.

Now, to continue with the En-route fuel efficiency example. The Performance Review Report (PRR) [11] data published have been used to provide an initial 2006 baseline for each of these three variables represented in the Influence Diagram (see Figure 5 above):

- amount of fuel burnt per flight for a flight following a great circle route, - presented in the Influence Model as ‘Average direct fuel burnt en-route’.
- additional fuel burnt as a result of additional kilometres flown compared with the great circle distance - presented in the Influence Model as ‘Fuel burnt en-route due to additional distance’.
- an estimate of the additional fuel burnt as a result of vertical deviation from the optimum flight level - presented in the Influence Model as “Additional fuel burnt en-route due to vertical deviations”.

The data used to populate these three variables in the model shall include the influence factors represented farthest from the left-hand side up to the proposed improvements. Where possible, the Influence Model will contain validation results from existing simulation
tools. In this case, the model will either capture the results as an array of possible input values or an approximation of the results (e.g., using curve-fitting) where a greater range of values is required.

These variables then flow through the influence model to provide a figure for the total amount of fuel burnt in en-route airspace that contribute to calculate the total European wide fuel burnt.

Validation experiments are required to verify the correctness of expert judgment in particular for the assessment of future horizon such as 2013 and 2020 were not many data are available.

A key benefit of using an integrated tool is the ability to ensure that all results are based on the same scenarios – that is using a consistent context.

Initially for this illustrative example data input to the model have been estimated using PRR report [11], expert judgment and some experiment results. In the future results from future project such as SESAR Development Phase will become available to populate this model.

**Initial Performance Assessment**

“It is better to be approximately right than precisely wrong.” - D. Pauly

In the project the Initial quantification work has focused specifically on the efficiency (fuel and time), Environment (global emissions), predictability (knock-on effects) and airport capacity as initial example to illustrate the potential feasibility and power of the methodology.

An initial performance assessment has been conducted with the primary aim of evaluating the Influence Model with illustrative data input.

Figure 6: Illustrative example of the EnRoute data input (Fuel and Time) (Source: EP3)

In the future, European-wide performance assessments could allow for:

- the combination of local results from a number of small-scale validation exercises effectively provides the expected performance at European-wide level
- sensitivity analysis which also provides precious information on uncertainty of results, allowing better prioritizing of the next validation exercises by focusing effort where results matter the most.

The Figure 7 below shows an example of Fuel Efficiency computed by the model. IP0 is the 2006 baseline and IP1, IP2 represent the clusters of potential improvements that will be deployed at different time horizon (2013, 2020).

Figure 7: Example of 2006, 2013 and 2020 fuel consumed per flight (Source: EP3)
There is the possibility for the user to obtain result for any module defined on the Influence diagrams. The Figure 8 is another example of the total fuel burnt for the top 150 European airport that carry 90% of the traffic.

Sensitivity analysis helps to identify which variables contribute the most to the overall results. The Figure 9 below shows the amount of fuel burnt in each phase of flight (Airport on stand, taxi in, taxi out, TMA departure, arrival and En-Route). As the inputs for the fuel consumption are currently single values, a tornado analysis is the appropriate representation for assessing the sensitivity of the results to the input data. In this example the tornado analysis is done by showing the impact of a 10% uncertainty on the inputs for the fuel burnt in each phase of flight that could be due to the simulators errors and expert opinion. In this illustrative example this figure shows that the en-route (+/-10% of 4250kg) and 100NM radius TMA departure (+/-10% of 1480kg) phases of flight contributes most to the overall amount of fuel (6400kg) that is burnt in Europe. A 10% uncertainty in this value can have large implications on the overall result and help to prioritise future validation exercise efforts and the potential improvements.

Figure 8: Example of 2020 European airport total fuel consumed (Source: EP3)

Figure 9: Example of sensitivity analysis of 2020 fuel consumed per flight (Source: EP3)

Lessons learned and further steps

A number of lessons have been learned (and captured) with the development of this initial Influence Model:

- the Performance Framework and Influence Model should have been developed before the system validation plan was defined. This would have allowed the validation effort to focus on producing results for use in performance assessment. In addition, defining validation exercises could benefit from description of the influences.
- Influence modelling has proved to be a powerful technique for representing (in a simple manner) the performance of the complex ATM system; and, in particular, the influences between KPAs, which are often not addressed in validation exercises.
- start the modelling simply, and keep it simple. The level of detail can be adjusted to the need when necessary.
- acceptable assessment results can be obtained with high-level data obtained through Fast-time simulations. Such initial high-level assessment can be refined further using more detailed tools focusing on critical performance elements in an iterative validation process. Each iteration can help identify:
missing data that will need to be produced by validation exercises
- uncertainties about the benefits, where more validation will be needed
- resulting prioritisation of exercises
- potentially, concept and targets refinement.

- in the project, we have been confronted with the representative-ness of validation data: often valid at a local-level (airport, portion of airspace), validity at the global European-level is uncertain or unknown. For an idea of how other areas could also benefit from the results to build a picture at a more global level, key representative characteristics of the required data should be part of the definition of validation exercises.
- developing an IM requires significant resources to produce a quantified result.
- the development team of the Influence Model needs to include the right balance of expertise for ATM operations, design, safety and environment, along with strong modelling expertise. The stability of the Core Team is the condition for the consistency of the model. This can become tricky in a multi-organisational project with a duration of more than one year.

The benefit of an IM derives from a trade-off between readability - for communication about results assumptions - and preciseness - for the production of quantified result-. The credibility of the IM results comes from both its technical completeness and from the stakeholders’ “buy-in”. A considerable effort will likely be required to achieve “buy-in” from those stakeholders who have not participated in the development of the Influence Model.

The Influence Model has potential as a more dynamic “what-if” scenario tool. This would require extension of the IM computational capabilities and user interface to support scenario execution in a flexible way. The Influence Model could then be deployed as a tool to support decision-making in respect of the overall system performance.

**Future plans**

The Performance Framework and associated Influence Models will be further elaborated and refined as part, and in support, of the ATM R&D Work Programme executed during the SESAR Development Phase in the 2009-2014 timeframe. (An overview of the Work Programme is beyond the scope of this paper).

The complete description can be found on the website for the SESAR Joint Undertaking (www.sesarju.eu). Relevant projects in this Work Programme that enhance and/or use the Performance Framework and Influence Models include:

- Development & Maintenance of the ATM Performance and Business aspects of the European ATM Enterprise Architecture: to deliver an ATM Enterprise view for SESAR that demonstrably connects the performance framework, business model, operational improvements, services and technical architecture through its evolutionary steps towards a 2020 target, answering key questions, and is responsive to changes in the ATM market by being able to support “what-if” analyses.
- Performance Analysis of ATM Target Concept: to provide evidence that the Target Concept (and associated ATM services) meets the performance targets.
- Content Coordination on Performance Requirements and Analysis (ICAO Performance Management Process): to provide coordination between the above-mentioned B4.1 and B5 projects to ensure proper conduct of the performance management process according to the ICAO Model as defined in Doc 9883 [1].
- Deployment/Performance planning and reporting: to perform performance planning from an ATM Master Plan perspective and support when relevant the development and maintenance of the SESAR Business Case(s).
- Maintain Business Case: to summarize relevant information (qualitative, quantitative, monetary) necessary to decide on an investment expenditure, addressing both problems raised by the decision-makers and issues of a nature to significantly change a resource allocation (*inter alia* values, risks, opportunities, relative priorities, timeliness
deployment). Business Case covers all aspects of business necessary to substantiate a decision to implement results from SESAR, such as: Deployment Plan, Cost-Benefit Analysis, Financing Plan, Safety Case, Capacity Case, Environmental Case, Security Case, and Human Factors Case. The overall Business Case consolidates all business bases at a pan-European level. One of the objectives of this project is to ensure that further performance planning and monitoring is made possible so as to compare \textit{ex ante} (forecast) to \textit{ex post} (actual) performance.

The underlying approach in SESAR will be to work both top-down and bottom-up:

**Top-down target setting for design (R&D) and deployment.**

A strategy for the evolution of ATM performance is (and will be) expressed in terms of a top-down hierarchy of objectives and targets (as schematically illustrated in Figure 10):

- At the highest level, there is the political vision for capacity, safety, environmental effects and cost. These statements drive the SESAR programme at political level, and are widely communicated among the general public.
- Below this are the SESAR Strategic Performance Objectives and Targets in the 11 ICAO Key Performance Areas. This is the set of high-level targets that have been elaborated during the SESAR Definition Phase, which is design guidance for the SESAR Development Phase.
- The third level is the home for the so-called “secondary” (i.e., cascaded-down, more specific) objectives, indicators and design targets, which underpin the strategic objectives and targets. Most of these are yet to be developed. On the one hand, they will incorporate the top-level (strategic) trade-off decisions to be made during the SESAR Development Phase. On the other hand, by being more specific, they will trigger and drive change at a level which is closer to the “solution level” described by the concept. In other words, the role of the proposed improvements will be to respond to these secondary objectives and targets, rather than provide direct traceability to the highest level of the Performance Framework (i.e., to the KPA level).
- The bottom level comprises the community-wide plus the local and regional deployment targets. This includes targets set within the legislative context of the Single European Sky second package (SES II). Only this level represents commitment to deployment. This information will be found in the European and Local Single Sky Implementation (ESSI/LSSI) Plans.

![ATM Performance hierarchy](image)

**Figure 10: ATM Performance hierarchy**

The Influence Models will play an essential role in making the link between the Strategic Performance Objectives and Targets and the “secondary” (i.e., cascaded-down, more specific) objectives, indicators and design targets.

**Bottom-up performance assessment during design (R&D) and deployment**

As part of executing the R&D life cycle phases, the expected ATM system performance will be evaluated based on validation results and Master Planning information.

At the same time, annual Performance Review Reports will be used to update the (rolling) baseline information and to develop a better understanding of performance. Assessments will be carried out both at local level (to prove concept performance) and at network-wide level (to prove respect to foreseen future traffic patterns). As part of executing the deployment life cycle phase, progress monitoring adequacy of potential deployment scenarios with of European and
Local Single Sky Implementation (ESSI/LSSI) Plans will be used to track the performance impact of ongoing deployment. All this information will be used to support the assessment of the risk of SESAR not delivering the planned and expected performance.

The data collected during the bottom-up performance assessment will be stored inside the Influence Models (repository function) and be used to prove the correctness of these models and adjust them where necessary.

Conclusions

ATM modernisation in Europe, as delivered through the SESAR programme, will be performance led.

The performance approach will be supported by the introduction of a performance scheme within the second Single European Sky package (SES II). This provides additional political weight to the need for an established link between the Performance Framework and the deployment framework. Regulators, service providers and airspace users will require evidence of the performance benefits of proposed investments to support target setting.

The EP3 Performance Framework and Influence Models provide a first tangible attempt at creating and maintaining a link between proposed improvements and performance targets supported by validation data.

This Performance Framework provides a mapping from political objectives, through the ICAO KPAs and KPIs to regional and local performance indicators, for assessment in validation tasks.

The fledgling Influence Model is providing a structure for:

- an analysis of step-contribution to performance improvement, and
- structuring of validation evidence to support the quantification of the Performance Framework.

The results reported at conference will provide an input for developing a European wide Performance Framework to help understand the contribution of operational improvements to a performance-based ATM system.

References

[3] EUROCONTROL: European Operational Concept Validation Methodology V2
[4] SESAR Definition Phase: Milestone Deliverable D5 - The ATM Master Plan
[5] SESAR Definition Phase: Milestone Deliverable D3 - The ATM Target Concept
[6] SESAR Definition Phase: Milestone Deliverable D4 - The ATM Deployment Sequence

28th Digital Avionics Systems Conference
October 25-29, 2009