Final Project Results Report

Deliverable ID:	**D1.3**
Dissemination Level:	**PU**
Project Acronym:	**INTUIT**
Grant:	**699303**
Call:	**H2020-SESAR-2015-1**
Topic:	**Sesar-11-2015: ATM Performance**
Consortium Coordinator:	**Nommon Solutions and Technologies**
Edition date:	**01 June 2018**
Edition:	**01.00.00**
Template Edition:	**02.00.01**

Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
Rodrigo Marcos/Nommon	Researcher	30/04/2018
Ricardo Herranz/Nommon	Project Coordinator	30/04/2018

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
Rodrigo Marcos/Nommon	Researcher	04/05/2018
Núria Alsina/ALG	Researcher	04/05/2018
Gennady Andrienko/Fraunhofer	Researcher	04/05/2018
Eef Delhaye/TML	Researcher	04/05/2018
Luca Piovano/UPM-CeDInt	Researcher	04/05/2018

Approved for submission to the SJU By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
Ricardo Herranz/Nommon	Project Coordinator	01/06/2018

Rejected By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date

Document History

Edition	Date	Status	Author	Justification
00.00.01	23/04/2018	Draft	Rodrigo Marcos	First Draft, used as input to the Final Review & Close-Out Meeting
00.01.00	30/04/2018	Approved for submission to the SJU	Rodrigo Marcos	Approved by INTUIT Consortium
01.00.00	01/06/2018	Approved	Rodrigo Marcos	Approved by SJU

INTUIT

Interactive Toolset for Understanding Trade-offs in ATM Performance

This document is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 699303 under European Union's Horizon 2020 research and innovation programme.

Abstract

Performance orientation in Air Traffic Management (ATM) is one of the key pillars of the Single European Sky (SES) and of its technological pillar, the SES ATM Research programme (SESAR). Setting down performance targets and anticipating the performance impact of new solutions is a complex process. A targeted indicator may limit the maximum achievable value for other indicators due to interdependencies. Most indicators are correlated with more than one indicator, which makes trade-off evaluation a challenging task. Finally, trade-offs may also arise between stakeholders, as well as between short-term and long-term objectives. Data science provides a way to discover and model unexpected patterns and relationships among big and often heterogeneous data. INTUIT is a SESAR 2020 Exploratory Research project that has explored the potential of visual analytics and machine learning to improve our understanding of the trade-offs between ATM Key Performance Areas (KPAs), identify cause-effect relationships between indicators, and develop new decision support tools for ATM performance monitoring and management. This report summarises the main project outcomes and describes the potential uptake of results by the SESAR programme. The report describes the applications explored within the project, discusses the way they could contribute to improving European ATM performance management process, and outlines the future research and innovation activities required to realise this potential.

Table of Contents

List of Tables

List of Figures

1 Executive Summary

Performance orientation is one of the key pillars of the Single European Sky (SES) and of its technological pillar, the SES ATM Research programme (SESAR). However, the complex interrelationships between different Key Performance Areas (KPAs) make ATM performance analysis a challenging task. The goal of INTUIT was to explore the potential of visual analytics and machine learning to improve our understanding of the trade-offs between ATM KPAs, identify cause-effect relationships between indicators at different scales, and develop new decision support tools for ATM performance monitoring and management. The **specific objectives** of the project were:

- to conduct a systematic characterisation of the ATM performance datasets available at different spatial and temporal scales and evaluate their potential to inform the development of new indicators and modelling approaches;
- to propose new metrics and indicators providing new angles of analysis of ATM performance;
- to develop a set of visual analytics and machine learning algorithms for the extraction of relevant and understandable patterns from ATM performance data;
- to investigate new data-driven modelling techniques and evaluate their potential to obtain insights about cause-effect relationships between performance drivers and performance indicators;
- to integrate the newly developed analytical and visualisation functionalities into an interactive dashboard supporting multidimensional performance assessment and decision making for both monitoring and management purposes.

The project started by identifying the main data sources on ATM performance and conducting a qualitative assessment on quantity, validity, quality, and geographical and temporal resolution of each dataset. This work produced three main outcomes: a set of **Performance Data Factsheets** characterising each data source, a **Performance Data Guide** (Figure 1) which links ATM performance data with the sources where such data can be found, and the **INTUIT Data Repository**, which allowed the project partners to share the datasets used for the INTUIT data analysis work. These outcomes are documented in deliverable D2.1.

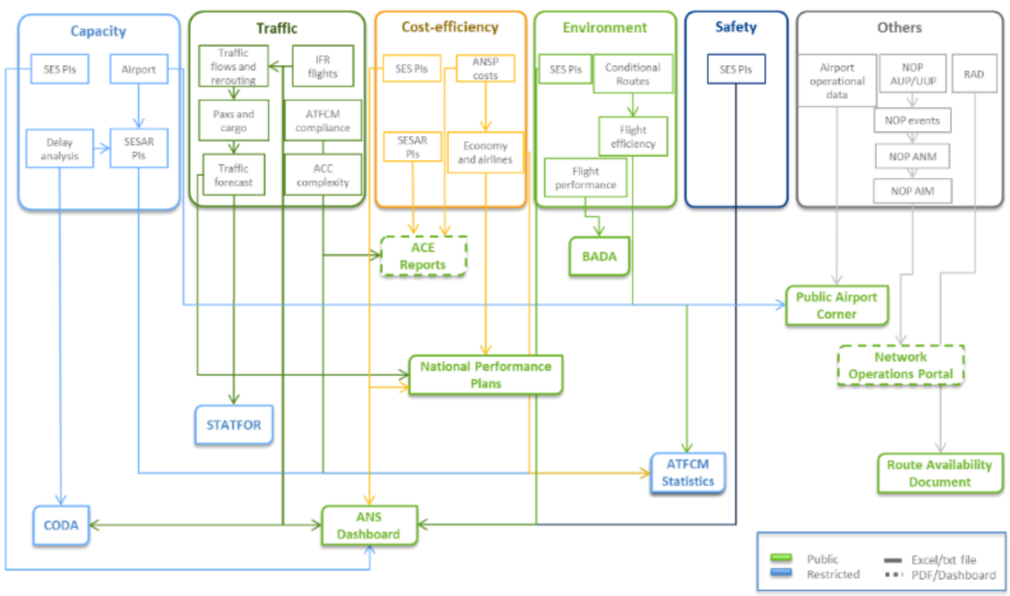


Figure 1. INTUIT Performance Data Guide.

Taking this work as a starting point, a combination of literature review and stakeholder consultation allowed the identification of a list of relevant research questions at the intersection of ATM performance modelling and data science, documented in deliverable D2.2. Based on a combination of factors, including the relevance of the research question, the expected impact of the results, the availability of sufficient data and the potential of data science to advance the state-of-the-art in that particular field, a subset of these research questions was selected to be investigated within the project in the form of three Case Studies (CS):

- CS-1: Effect of unit rates on airline route choices and impact on ATM performance.
- CS-2: Sources of en-route flight inefficiency.
- CS-3: Multi-scale representation of ATM performance indicators.

CS-1 analysed the effect of unit rates on en-route performance, and more generally the modelling of airline route choice decisions and their impact on ATM performance. Airline route choice decisions are an important factor influencing ATM performance (e.g., environmental performance) that is not sufficiently well understood. During the Air Traffic Flow and Capacity Management (ATFCM) strategic and pre-tactical planning phases, when flight plans are not available yet, traditional approaches to demand prediction need to rely on certain hypotheses about airline preferences and the permanence of airlines' behaviour over time, due to the sensitive nature of information such as airline cost of delay or aircraft take-off weight. To avoid this problem, INTUIT has proposed the development of data-driven models that are trained on the basis of the observed airline behaviour, by relying on historical data such as that available from EUROCONTROL's Demand Data Repository (DDR2). First, historical airline route choices for a number of Origin-Destination (OD) pairs within the European Civil Aviation Conference (ECAC) area were analysed by means of different visual analytics techniques, which led to the identification of a set of relevant variables to be considered in the modelling of airline route choices. Based on this initial data exploration, a variety of machine learning techniques, such as logistic regression, decision trees and neural networks, were used to develop models for the prediction of airline route choices. For a given OD pair, the proposed models predict the routes chosen by the different airlines serving such OD as a function of the characteristics of the available route options (route length, air navigation charges, expected congestion, etc.). The models were trained and evaluated using historical data from DDR2, showing significant potential to outperform traditional approaches such as those currently employed for pre-tactical traffic forecast. The newly developed models provide an enhanced understanding of Airspace Users (AUs) behaviour, which can in turn contribute to improving performance consolidation methodologies and Cost Benefit Analysis (CBA). Additionally, the models have a potential for traffic prediction during the pre-tactical planning phase and could thus contribute to enhancing pre-tactical ATFCM through the provision of better demand forecasts. Future development steps include working with more granular trajectory data, refining the prediction algorithms, conducting a more comprehensive validation and upscaling the proposed models to the whole network.

In the scope of CS-1, a decision support dashboard was developed to help in the decision of setting unit rates to optimise performance (see Figure 2). The main objective of this dashboard is to assess the performance impact of tuning unit rates on the flights covering a certain OD pair. A multi-objective optimisation tool was developed to find the setting of unit rates that optimises a series of Key Performance Indicators (KPIs) according to different weights for each KPI. The dashboard enables the evaluation of the trade-offs of a given optimal setting of unit rates in terms of flight efficiency, cost efficiency and capacity by means of different interactive visualisations and allows the assessment of the effect of unit rates and route choices on ATM performance. The dashboard provides additional functionalities to analyse the impact of a certain unit rate setting on different stakeholders, namely Air Navigation Service Providers (ANSPs) and AUs.

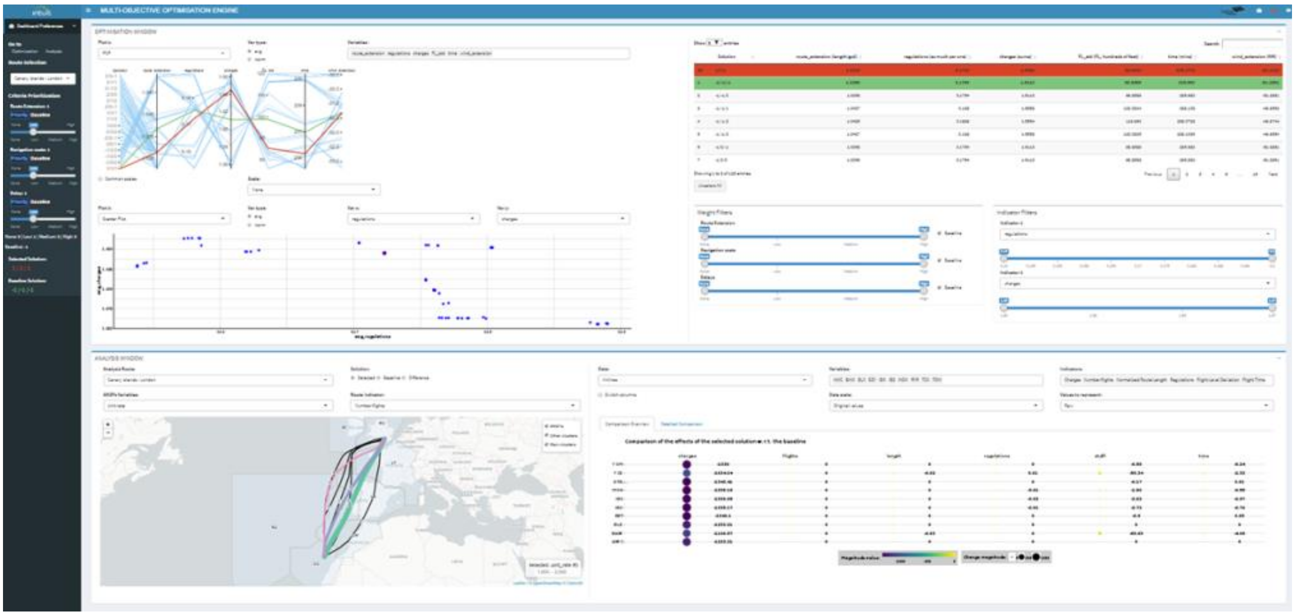


Figure 2. INTUIT Interactive Dashboard for Airline Route Choice Prediction and Optimisation of Unit Rates.

CS-2 aimed to develop new methods for the identification of the main sources of en-route flight inefficiency. The case study focused on analysing the influence factors that drive the horizontal flight efficiency of the flights crossing a certain airspace area, in order to isolate the contribution of different factors and stakeholders. We were interested, for example, in knowing how much of the inefficiencies measured in a certain Area Control Centre (ACC) are actually due to the closure of a certain airspace area in adjacent ACCs. The case study was conducted for the Bordeaux ACC (LFBBCTA). We started by a visual exploration of the correlation between horizontal flight efficiency and other variables such as take-off time, number of flights per Air Traffic Controller (ATCO), the distance between the ideal and the planned entry/exit points, the heading of the orthodromic trajectory, and the ACCs crossed by such orthodromic trajectory immediately before and after the ACC under study. The factors identified as potentially relevant through this visual exploration were used as model inputs for training a random forest model aimed to predict horizontal flight efficiency. The interrelationships revealed by the model serve to gain insights on the causes and effects of low performing flights in a certain airspace area, and to quantify the impact of the different influence factors. As an example application, the proposed approach can be used to investigate how much of the flight inefficiency observed in a certain ACC is due to suboptimal airspace design and to propose mitigation measures (e.g., improved design of the interfaces between adjacent airspace areas).

The results of the case study were consolidated into an interactive performance monitoring dashboard (see Figure 3). The dashboard provides a tool to identify and evaluate the causes of flight efficiency in a particular airspace (ACC or charging zone) of the ECAC area. Different types of visualisations and visual analytics techniques are used to support the exploration of the relationship between flight efficiency and other variables related to airspace structure. The tool allows the user to explore the influence of these factors on flight efficiency, with a view to support the identification of the main causes of flight inefficiency.

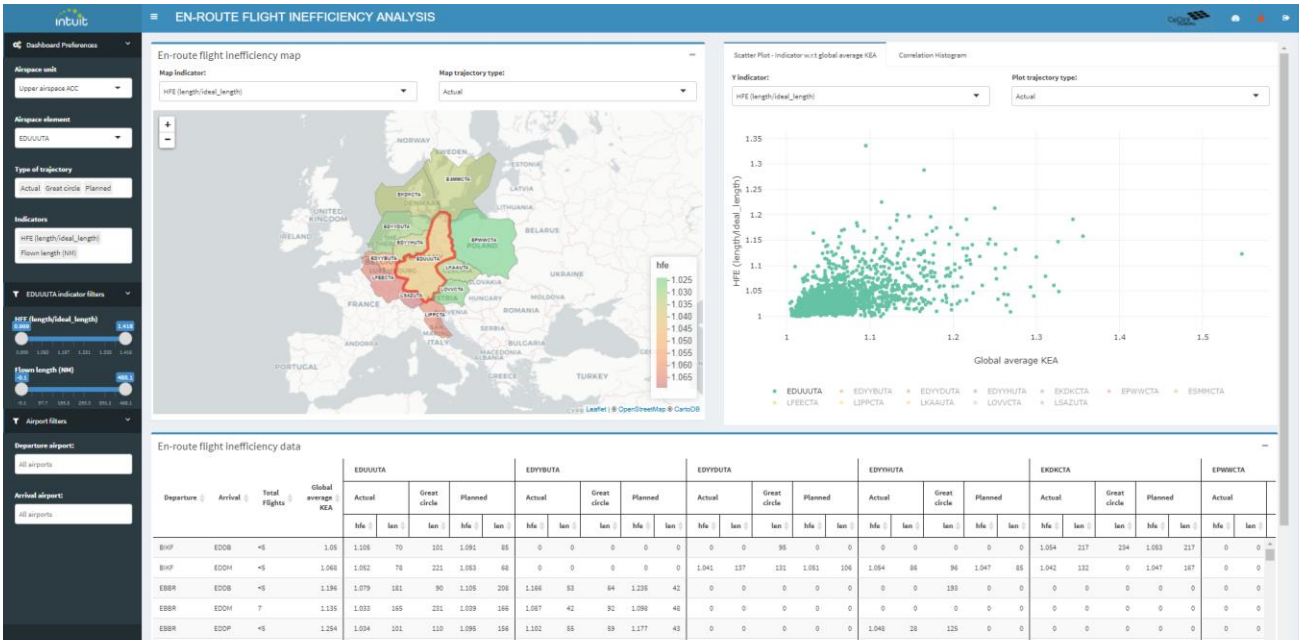


Figure 3. INTUIT Interactive Dashboard for the Analysis of En-Route Flight Inefficiencies.

CS-3 investigated new multi-scale representations of ATM performance indicators. In particular, Air Traffic Flow Management (ATFM) delay was studied at ACC level for different sector configurations. In order to select a representative set of ANSPs for the case study, European ANSPs were classified according to four characteristics: seasonal traffic variability, number of flight-km controlled, traffic complexity, and number of sectors. This led to four, well-differentiated clusters: big European countries, medium and small countries in Central and Northern Europe, Eastern European countries, and relatively small peripheral countries. A specific visual analytics tool was developed to support the clustering process and verify the robustness and soundness of the resulting clusters (see Figure 4).

The rest of the case study focused on two representative countries of the two clusters with the highest number of regulations: Spain and Poland. The correlation between the issued sector configurations along one year in the ACCs of Warsaw and Madrid and the associated number of regulations was investigated, with the purpose of identifying patterns emerging under different levels of congestion and determine whether certain configurations are more prone to generating ATFM delay. In future developments, the approach used in this case study could be extended to other KPIs, in order to identify different regimes of KPI behaviour and spot situations of suboptimal performance.

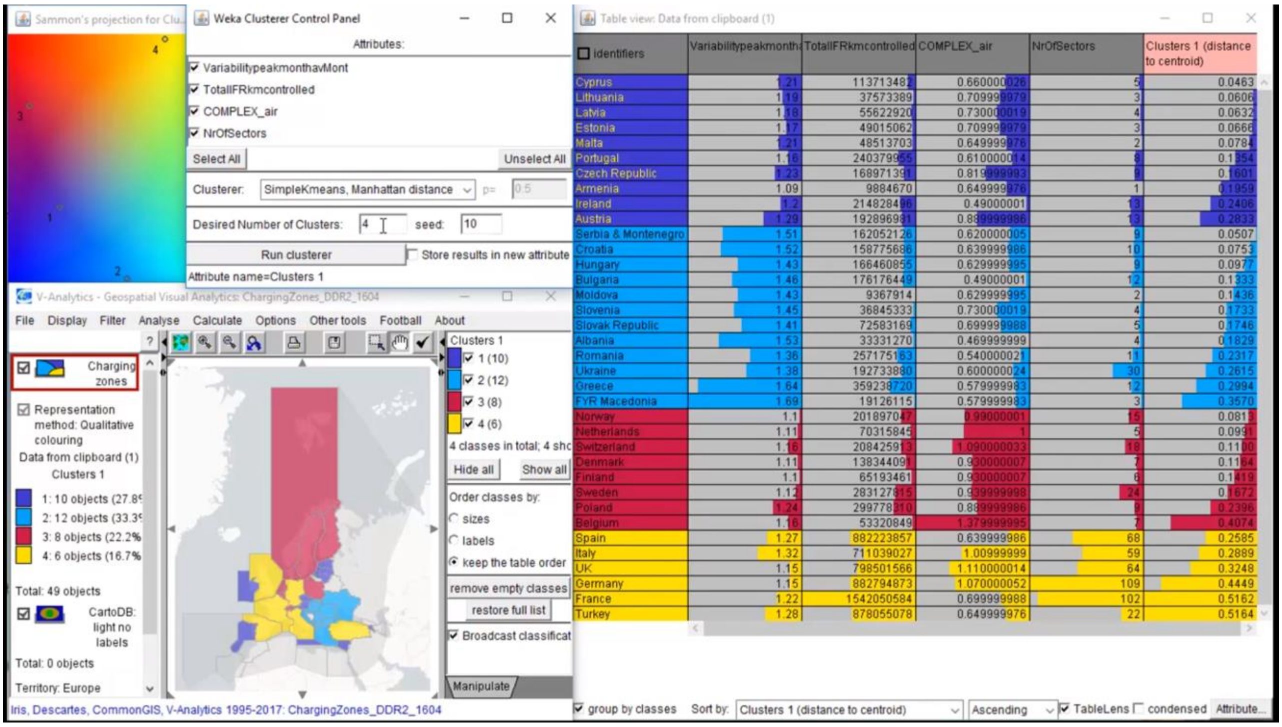


Figure 4. INTUIT Visual Analytics Tool for ANSP Clustering Analysis.

The INTUIT project prototyped a **performance monitoring and decision dashboard,** available at: https://viz2know.cedint.upm.es. Additionally, videos showing the main functionalities of the different visualisation tools developed by the project are available at the following links:

- https://drive.google.com/file/d/11e1QFjlF3z3MM2f17hHLkW6psQWF2EcI/view?usp=sharing
- https://drive.google.com/file/d/1_qiC9oYDtIFIjl-OT5ZKnmsAxJf-dabQ/view?usp=sharing
- https://drive.google.com/file/d/0ByvPezzMX-09RWJXM1FpSTF4Rm9nLXMwMkFaR25xNUt1T3d3/view

The models and visualisation tools developed by INTUIT are of exploratory nature. Nevertheless, they open the door to **future research and innovation activities**.

Regarding the **modelling of airline route choices**, a prospective application is the aggregation of route predictions into traffic demand volumes in order to predict the appearance of hotspots. To do so, the approach developed in INTUIT CS-1 should be applied to all OD pairs for which one or more possible routes cross the hotspot. Then, predictions would be aggregated in a probabilistic manner to obtain the predicted traffic volume in the hotspot. This application would be of use for Demand and Capacity Balancing (DCB) during the pre-tactical planning phase. On a more strategic level, the proposed modelling approach could also be used to investigate questions related to the interrelationship between ATM KPAs, e.g. the trade-offs between environment (flight efficiency), capacity (delay) and cost-efficiency. This approach would be of interest for SESAR PJ19.04 and the Performance Review Unit. Regarding the decision-support tool, further developments would integrate the predictions of several origin-destinations to provide a comprehensive analysis of the effects of unit rate tuning on all the routes affected, which would constitute a useful tool for the Network Manager or ANSPs. Finally, the work done in INTUIT opens the door to other exploratory research activities, such as the combination of the data-driven approach proposed by INTUIT with the model-driven approaches developed by other SESAR Exploratory Research (ER) projects, like APACHE and

AURORA, to enable better predictions of AUs' behaviour, for instance when a Free Route Airspace concept is applied to a certain airspace.

Regarding the **identification of flight efficiency influence factors**, further developments would be needed to generalise the proposed approach. The modelling should be extended to any ACC, and a systematic comparison of these models across ACCs and study periods should be carried out to find commonalities (global influence factors) and specificities. The application would then enable an accurate estimation of influence factors of flight efficiency, which would be a valuable input for the influence diagrams developed in SESAR PJ19.04. In addition, the model could be further developed and enhanced with more granular data to predict more sophisticated flight efficiency metrics such as vertical efficiency indicators and fuel consumption and/or performance metrics in other KPAs. This result would enable the measurement and reconstruction of new performance indicators, which otherwise would be very costly to measure. Regarding the decision-support tool, future developments would allow the user to further analyse and characterise low performing flows. The study could also be extended to other KPAs such as cost-efficiency or capacity. Finally, the tool could be enhanced to quantify the influence of the different factors and to provide prediction capabilities for the early detection of low performance episodes.

Regarding the case study on **multiscale KPIs representation**, the study could be extended to other representative countries and to other KPAs. Further developments of the proposed decision support tool would include the automation of the clustering process and the generalisation to different types of datasets, including high-dimensionality datasets such as trajectories. These improvements altogether would create a powerful tool to assess performance, automating the process of ANSP benchmarking.

In summary, INTUIT has developed a structured approach and a set of prototype performance analysis tools that have demonstrated the **potential of visual analytics and machine learning to improve the state-of-the-art in ATM performance analysis**. To realise this potential, different development paths are envisaged in terms of both exploratory research and industrial research.

2 Project Overview

2.1 Operational/Technical Context

The ongoing ATM modernisation programmes, including SESAR, build on ICAO Global ATM Operational Concept, one of whose cornerstones is performance orientation. A performance-based approach is defined by ICAO as one based on: (i) strong focus on desired/required results; (ii) informed decision making, driven by the desired/required results; and (iii) reliance on facts and data. A performance framework is intended to translate stakeholders' expectations into a shared set of values and priorities and be the basis for impact assessment, trade-off analysis and decision making.

ATM performance results from the complex interaction of interdependent policies and regulations, stakeholders, technologies and market conditions. Trade-offs arise not only between KPAs, but also between stakeholders, as well as between short-term and long-term objectives. To effectively steer the performance of ATM operations, metrics and indicators shall therefore be capable of capturing the full range of economic, social and environmental impacts of the ATM system on the different stakeholders and society at large, at different temporal and geographical scales. Performance modelling techniques shall be able to grasp the interdependencies between different KPAs and KPIs and allow the assessment of the possible future impacts of a range of policies and trends.

The need for improved indicators and modelling methodologies meeting these conditions has been acknowledged by the ATM stakeholders and the research community. While a lot of effort has traditionally been devoted to the development of microscopic performance models, there is a lack of useful macro approaches able to translate local improvements or specific regulations into their impact on high-level, system-wide KPIs. On the other hand, the increasing availability of data at different scales, together with recent advances in the fields of data analysis and visualisation, open new opportunities to develop new ATM performance metrics and modelling techniques.

2.2 Project Scope and Objectives

The goal of INTUIT is to explore the potential of visual analytics, machine learning and systems modelling techniques to improve our understanding of the trade-offs between ATM KPAs, identify cause-effect relationships between indicators at different scales, and develop new decision support tools for ATM performance monitoring and management. A detailed schema of the methodology can be found in Figure 5. The specific objectives of the project are:

- to conduct a systematic characterisation of the ATM performance datasets available at different spatial and temporal scales and evaluate their potential to inform the development of new indicators and modelling approaches;
- to propose new metrics and indicators providing new angles of analysis of ATM performance;
- to develop a set of visual analytics and machine learning algorithms for the extraction of relevant and understandable patterns from ATM performance data;
- to investigate new data-driven modelling techniques and evaluate their potential to provide new insights about cause-effect relationships between performance drivers and indicators;

- to integrate the newly developed analytical and visualisation functionalities into an interactive dashboard supporting multidimensional performance assessment and decision making for both monitoring and management purposes.

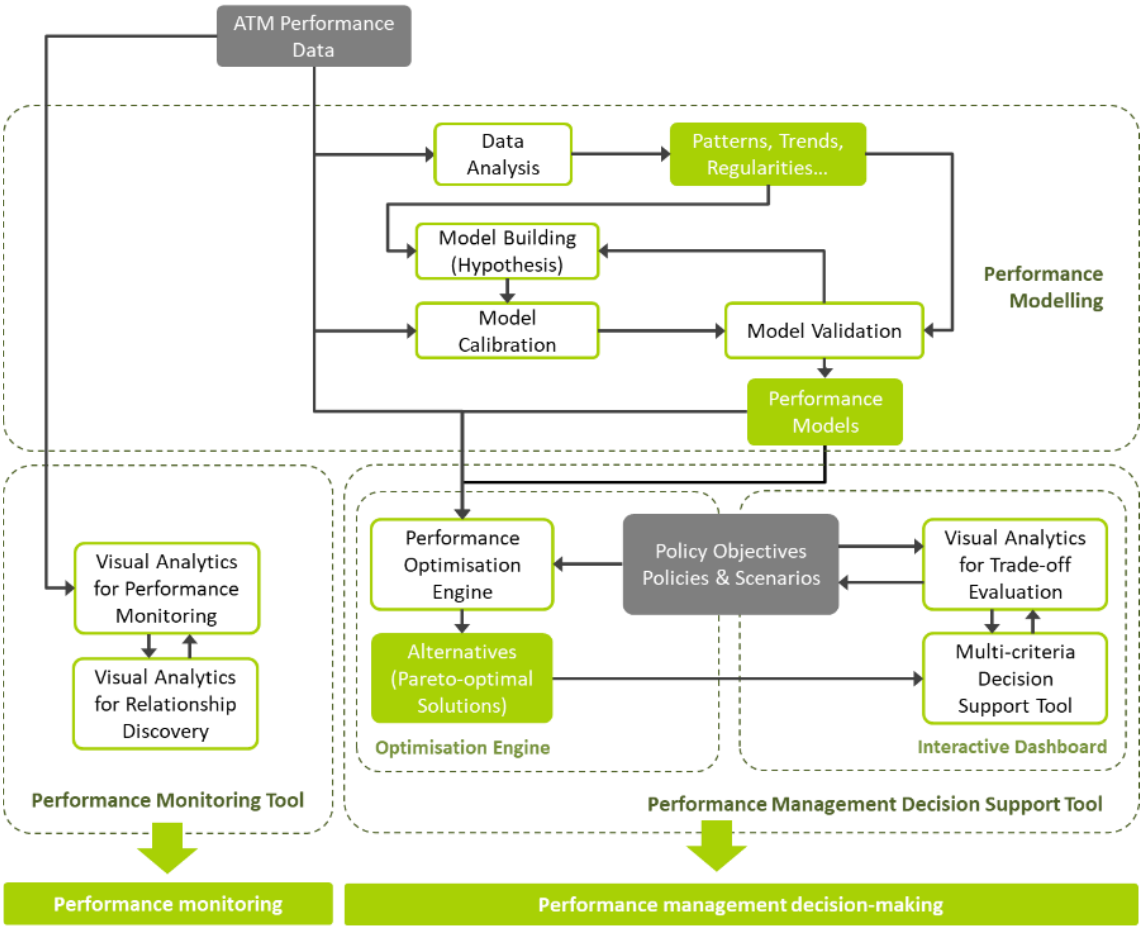


Figure 5. INTUIT: overall project concept.

To achieve these goals, the project started by identifying a set of relevant research questions at the intersection of ATM performance analysis and data science. The research work was structured around a set of case studies, each case study corresponding to one or several of the research questions previously identified:

- CS-1. Study of the effect of unit rates on en-route performance, and more generally the modelling of airline route choice decisions and their impact on ATM performance.
- CS-2. Development of new multi-scale representations of ATM performance indicators.
- CS-3. Identification of sources of en-route flight inefficiency.

The main results of these case studies were consolidated into a prototype performance monitoring and management dashboard providing interactive visualisation of the modelling results.

2.3 Work Performed

2.3.1 Characterisation of ATM performance datasets and evaluation of their potential for the development of new indicators and modelling approaches

The first task conducted by the INTUIT team was the identification of the available data sources on ATM performance, the analysis of their characteristics, and the evaluation of their usefulness for the project. This task consisted of several sub-tasks:

- Identification of the available data sources and organisation of these into different categories according to the type of data available from each source.
- Design of an analysis methodology for data quality assessment.
- Assessment of data quality and identification of the main issues for each data source.
- Creation of a visual guide of the information available from the different data sources (see Figure 6 below).
- Implementation of the INTUIT Data Repository, which was used to share different datasets among the project partners.

The results of this work are summarised in **D2.1 Performance Data Inventory and Quality Assessment**, which provides a detailed qualitative assessment of the quantity, validity, quality, and geographical and temporal resolution of each dataset.

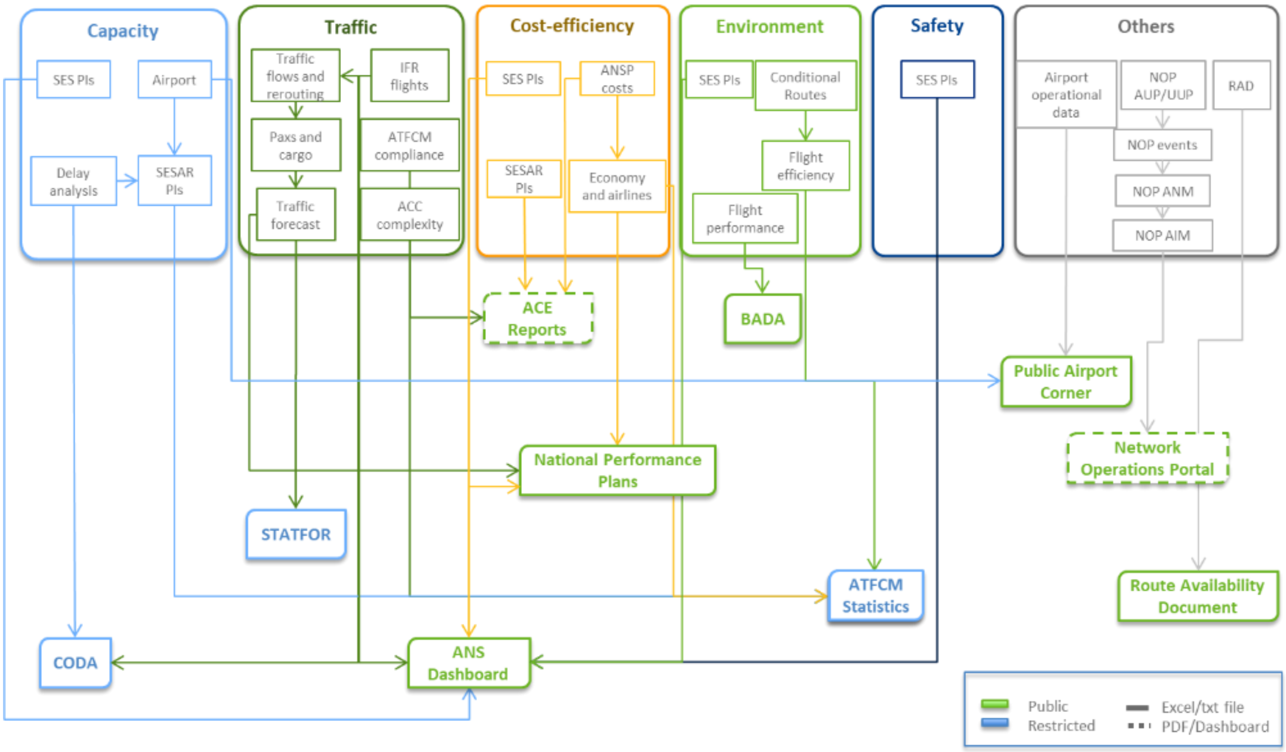


Figure 6. Visual guide of the identified performance databases.

2.3.2 Definition of a research agenda at the intersection of ATM performance and data science

In parallel to the analysis of the available data sources, a review of research papers and policy studies, together with a consultation with different ATM stakeholders represented in the project's Advisory Board, led to the definition of a set of research challenges related to ATM performance and data science. The main research conclusions of this analysis are depicted below and summarised in Figure 7:

- Regarding ATM cost-efficiency, existing indicators may be suitable means to measure performance, but in certain cases and under new operational paradigms, they may be improvable.
- Further development of the estimation of the cost of delay for an airline, which is highly nonlinear and depends on contextual and flight characteristics, would improve the delay indicators used for target setting.
- The trade-off between ANSP cost-efficiency (ATM navigation costs) and environmental cost-efficiency (fuel costs) in the presence of unit rate variability between ANSPs could be further explored.
- There is also room for improvement in the study of ATCO workload interdependencies with other areas such as: (i) capacity, by studying how traffic demand-capacity imbalance leads to an increase of sector complexity; (ii) uncertainty, by uncovering systematic relationships between volatility of demand and complexity patterns; and (iii) airspace complexity, by exploring the ATCO workload interdependencies between lower and upper airspace, for instance.
- Another research area is the development of interactive visualisations, able to disaggregate data at different scales. An example would be the enhancement of the Performance Review Unit (PRU) dashboard developed by EUROCONTROL. Visual analytics could be used to present disaggregated data with finer spatial and temporal granularity, with the aim to provide insights into the relationship between local decisions made in real-time operations and KPIs and identify the main influence factors of these KPIs.
- The implementation of new KPIs was also proposed. A new predictability KPI could compare forecasted and actual demand to improve ATM predictability and evaluate the benefits of such improvement. A KPI reflecting the effect of ATM system in constraining air traffic would also be interesting. In addition, KPIs could be modified to include airport operations performance. Finally, new metrics to measure safety and equity in a more significant way should be considered.
- The trade-off between safety and other KPAs should be further developed, in order to analyse how the targets imposed by the Performance Scheme to other KPAs affect safety performance.
- There is also a strong interest in the development of research questions related with uncertainty, such as studying the impact of on ground processes on gate-to-gate predictability.

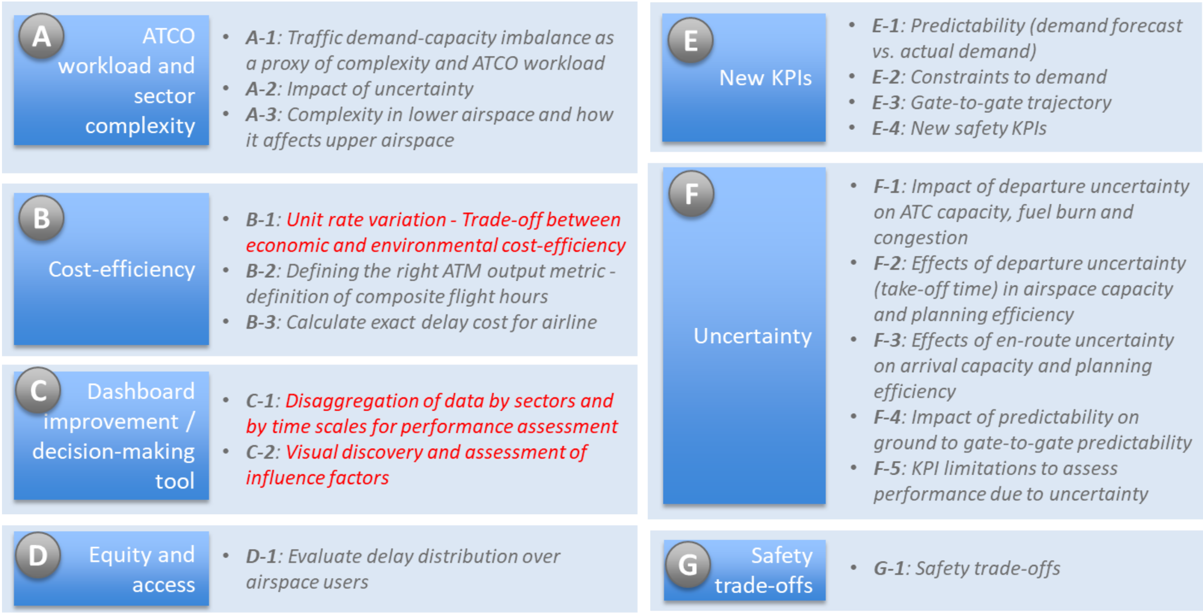


Figure 7. Overview of the proposed research areas at the intersection of ATM performance and data science. Red coloured threads are the ones studied in INTUIT.

The results of the literature review and the stakeholder consultation, and the research questions identified as a result of this process, are documented in **D2.2 Qualitative Analysis of Performance Drivers and Trade-offs**.

2.3.3 Visual analytics for the extraction of patterns from performance data

Visual analytics techniques were used to explore ATM performance datasets and extract patterns, regularities and trends that were used as an input to develop different performance models. The visualisation exercises conducted addressed the following topics:

- Route choice analysis. Airline route choice decisions were identified as an important factor influencing ATM performance (e.g., environmental performance) that is not sufficiently well understood. For this reason, a visualisation exercise was defined to explore such decisions. By using DDR2 flight trajectories, airline route choices were studied for different ODs to extract temporal and spatial patterns and identify relevant variables to be considered in the modelling of airline route choices and their impact on ATM performance.
- Flight efficiency analysis. Flight trajectories were studied to analyse visually the influence factors determining the horizontal flight efficiency of flights crossing a certain airspace. This analysis served to identify explanatory variables to be then used in the performance modelling tasks.
- Delay analysis. The visual analytics work here focused on the exploration of DDR2 data and ATFM daily delay summaries in order to identify temporal and spatial patterns of delays between OD pairs and detect bottlenecks in the European network. Different visualisations were developed to explore the interrelationships between different types of delays (e.g., arrival delay vs departure delay) under different network conditions, as well as to investigate the possible causes of delays affecting certain airports (e.g., during the French ATCOs' strike on 31st March 2016).

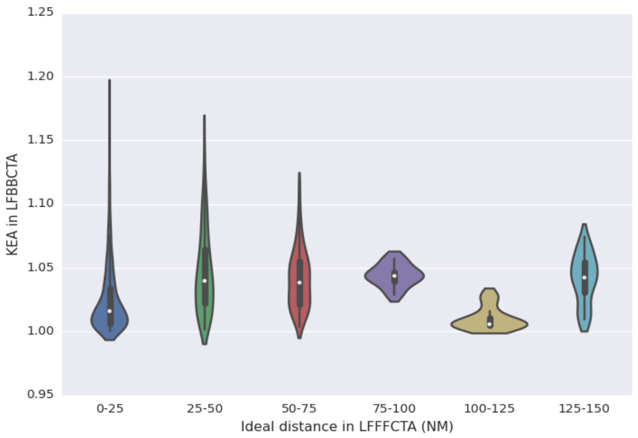
This visualisation work is reported in **D3.1 Visual Analytics Exploration of Performance Data**.

2.3.4 New data-driven modelling techniques to provide new insights about cause-effect relationships between performance drivers and performance indicators

In a first stage, several machine learning techniques were identified as potentially suitable to address the different research questions identified at the beginning of the project. After this initial review, the application of these techniques was articulated around three case studies:

- **CS-1. Modelling of airline route choices and the influence of unit rates on performance**. The goal was to develop new models able to predict airline route choices between different ODs in order to evaluate the performance trade-offs arising from these decisions (e.g., cost efficiency vs environment). The proposed approach has shown significant potential to improve the understanding of route choices, and it is of potential application to the problem of pre-tactical traffic forecast.
- **CS-2. Identification of sources of en-route flight inefficiency**. This case study, conducted in collaboration with the SESAR ER projects AURORA and APACHE, investigated the causes of inefficient routes in the European Network and their effects on performance, in order to isolate the contribution of different factors and stakeholders. An example of the visual exploration of influence factors can be found in Figure 8, where it is observed that nearby sectors can impact negatively the efficiency of a sector by not allowing flights to cross them in a certain direction, as LFFFCTA does with flights crossing LFBBCTA.
- **CS-3. Multi-scale representation of performance data**. This case study aimed to disaggregate traffic data and performance indicators at ACC and sector level, with different levels of temporal disaggregation, and later on model the relationship between these variables at different scales (e.g., what is the influence of sector configuration on the aggregated performance of a certain ANSP?).

The results of the performance modelling work are documented in **D4.1 Performance Metrics and Predictive Models**.



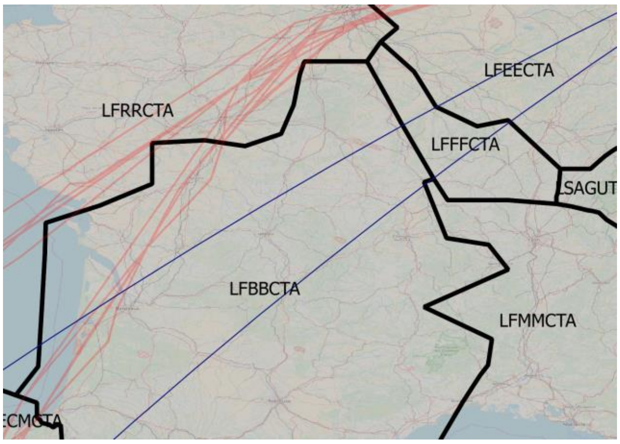


Figure 8. Left: Probability Density Function of flights crossing LFBBCTA and LFFFCTA grouped per ideal distance flown in LFFFCTA, showing how flights crossing transversally LFBBCTA (25-100 NM) have lower flight efficiency; Right: example of flights deviated that would ideally cross LFFFCTA transversally.

2.3.5 Integration of the developed functionalities into an interactive tool for performance assessment and decision making

The selected case studies and the new visualisation and modelling tools were integrated into an interactive performance monitoring and management dashboard comprising different components:

- A multi-objective optimisation engine to find the Pareto optimal solution for a set of KPIs was developed. The route choice predictor developed in the context of CS-1 was integrated into an optimisation tool to predict the effects on performance of a particular setting of unit rates and help in the selection of the best setting of unit rates according to these KPIs. The route choice predictor is used to calculate aggregated efficiency metrics for a given OD as a function of the unit rates setting, which determines the charges paid to fly each route. The dashboard allows the evaluation of the trade-off between flight efficiency, cost efficiency and capacity by means of different interactive visualisations and allows the user to assess the effect of unit rates and route choices on ATM performance.
- A flight efficiency monitoring dashboard. This dashboard provides a tool to identify and evaluate the causes of flight efficiency in a particular ACC. Flight efficiency indicators are presented with different type of visualisations versus other flight properties derived from both the flight plan and the ideal route, such as heading, altitude and airspace crossed. The tool allows the evaluation of the influence of these factors on flight efficiency. The extracted interrelationships may serve as a basis to perform an assessment of the causes and effects of low performing flights.
- A decision support framework for optimising the process of data clustering. Clustering is often used in model building for grouping objects (e.g., ANSPs) with similar characteristics and building models for these groups. Clustering results may vary depending on the selected method, its parameters, and initial settings such as randomisation seeds. Respectively, it is necessary to ensure that results of the clustering process represent meaningful groups of objects with similar characteristics, are stable in respect to the clustering settings and are easy to reproduce. The tool provides a suite of visual analytics tools that combines different clustering algorithms, distance measures, and clustering projections for representing multidimensional visual summaries and comparing cluster profiles.

The results of this work are documented in deliverable **D5.1 Performance Monitoring and Management Toolset** and **D5.2 Performance Monitoring and Management Toolset - Evaluation Report.** The developed performance dashboards can be accessed through the following link: https://viz2know.cedint.upm.es/. Figure 9 shows a snapshot of the dashboard for the first two case studies. The third case study was not incorporated into the dashboard.

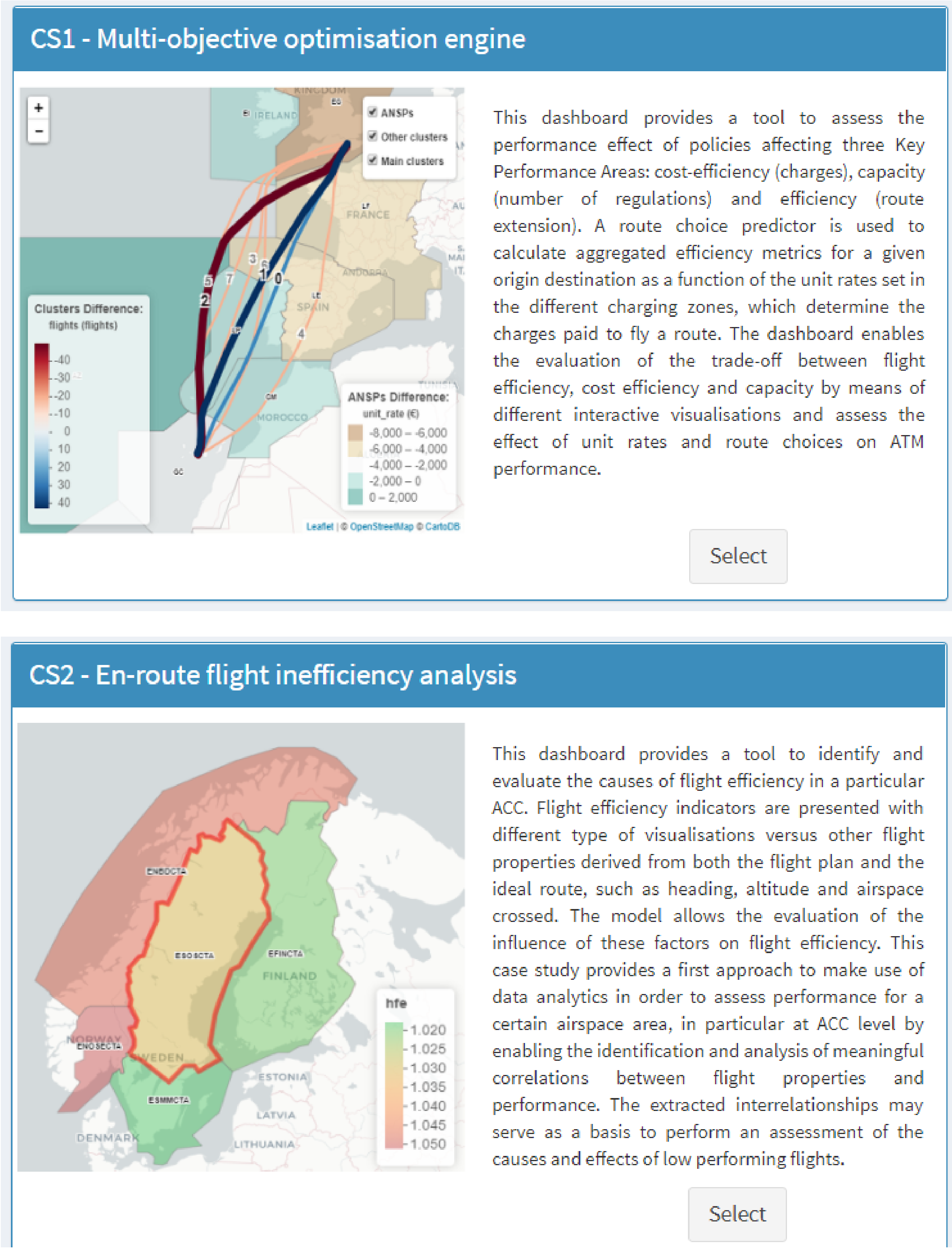


Figure 9. Overview of the introduction to the dashboards available in the INTUIT platform.

2.4 Key Project Results

This chapter discusses the main results of the work performed in the INTUIT project. The section is organised in: (i) results of the data source exploration, (ii) results of the analysis of ATM performance research challenges and (iii) the results from the 3 case studies developed in the project.

2.4.1 Performance Data Inventory

In this document, we have identified a set of data sources that provide the required level of detail and data quality for the achievement of the project objectives (see a complete list in Figure 10). These data sources can be coarsely classified into two groups:

- High granularity data sources, which provide large amounts of low-processed data. They are useful to compute new metrics. However, they typically require a high computing effort.
- Low granularity data sources, which provide highly processed and aggregated data. They contain relevant metrics that can be used without any processing.

The most useful high granularity databases identified are depicted below:

- DDR2: trajectories of individual flights in the European airspace and airspace design. This database has the highest potential for metrics computation. The geographical granularity is not the highest obtainable, but it is sufficient to compute a wide range of metrics.
- Daily summaries: aggregated statistics during the day of operation of the delays caused by ATFM regulations. This database is useful for research regarding delay-capacity and its interdependencies.
- ANM: information of regulations prior to the operation. This database is useful for research on delay-capacity and predictability.
- European AUP/UUP: airspace restrictions and alternative routings due to military activity. This database is essential for any study about cooperation between civil and military aircraft.
- CODA: reported delay statistics by airlines. This database is useful for any research related to delay-capacity and its interdependencies.
- STATFOR: actual and forecasted capacity, useful for research about delay-capacity.

The most useful low granularity databases identified are the following:

- ACE reports: statistics and balance of ANSP costs and staffing. This database is essential for any cost-efficiency analysis. Data is highly aggregated and could be improved with the provision of raw data.
- PRR: statistics regarding KPIs with broad temporal scope. This database is necessary for the study of interdependencies between metrics. Data is highly aggregated and could be improved with the provision of raw data.
- ANS dashboard: it provides KPIs and Performance Indicators (PIs) metrics during Reference Periods RP1 and RP2. These data are provided by PRR with a wider temporal basis and higher number of statistics. The advantage with respect to PRR is that data shown in the Air Navigation Service (ANS) dashboard is more easily exportable in table format, especially for RP2. Therefore, data downloadable from ANS dashboard is preferable when available, but the type analysis that can be performed is limited by the temporal scope of the date, which is narrower than in the case of the PRR.
- RAD: route restrictions during an Aeronautical Information Regulation and Control (AIRAC) cycle. This document can be used to study the interdependencies between airspace constraints and capacity/demand adjustment.

- NPP: ANSP plans to comply with KPI objectives. They are useful for computing statistics regarding the accomplishment of the SES objectives. Complete data is only available for RP1.
- Public Airport Corner: static operational data limited to some European airports. Forecasted runway capacity could be used to analyse airport demand-capacity balance.

Two other reviewed databases, which are not listed above, provide duplicated or less usable information. These databases are:

- NOP Events: it provides information of events that affect air traffic. The format, although standardised, is filled in different ways depending on the source. In addition, interpretation of description in text would require excessive computing effort.
- AIM: information about events or possible disruptions in the network. The format of messages is in text. The interpretation of text would require excessive computing effort.

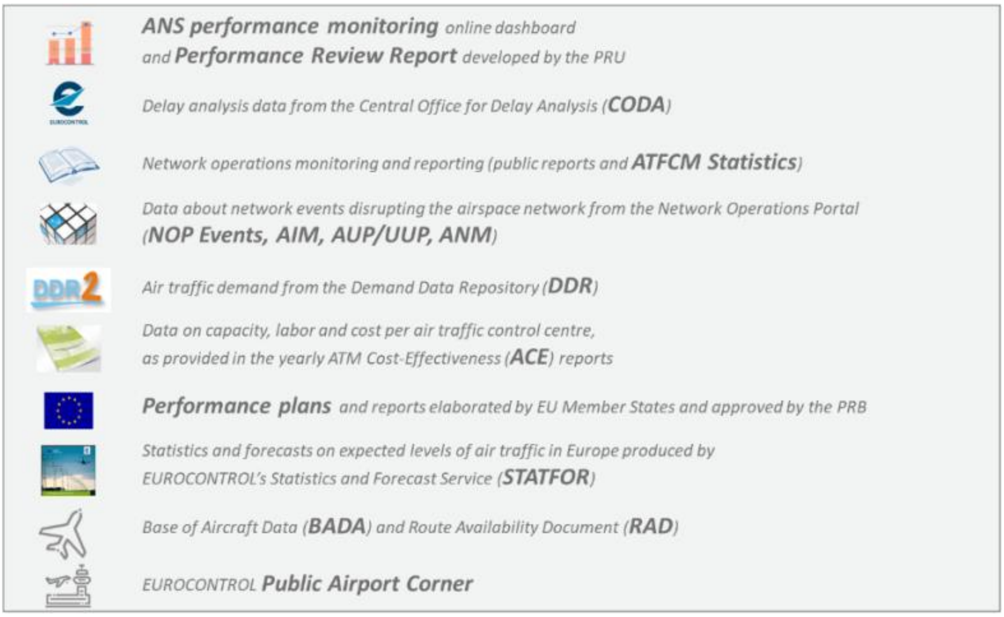


Figure 10. Overview of the performance databases analysed in INTUIT.

2.4.2 Qualitative Analysis of Performance Drivers and Trade-offs

The desk research performed by the INTUIT team, together with the consultation with ATM stakeholders and industry practitioners, allowed us to identify a set of research questions of interest for the upcoming INTUIT work packages. The assessment of ATM performance data sources conducted in parallel and documented in D2.1 was used to identify the relevant data sources that would be required to address the proposed research questions. The main conclusions regarding the relevance and feasibility of the different research questions for INTUIT are summarised below.

- **Cost-efficiency** is the performance area that has received more attention from researchers. However, there is room for improvement by focusing deeper on specific areas, such as the proper definition of composite flight hours, the evaluation of the cost of delay for airlines and the interdependencies between economic and environmental cost-efficiency, such the one arising from route choices avoiding airspace with high unit rates. One advantage of cost-efficiency studies is that required data seems easier to obtain with the adequate temporal and geographical scope.

- Another research topic is the development of **visualisation and decision-support tools** for identifying relationships in ATM operational performance, which could be implemented as an enhancement of the PRU Dashboard. Disaggregated KPI data and influence factors could be represented interactively to detect spatio-temporal patterns and assess such factors. This tool would provide insights into the relationship between local operational decisions (e.g., sector configurations) and KPIs, which are calculated ex-post. Due to the wide scope of this research question, several additional databases would be required for its development. Sector segregated data may only be available for certain collaborating ANSPs, which could limit the scope of the work.
- The **implementation of new KPIs** was also proposed. A new predictability KPI could compare the forecasted demand with the actual demand, using analytics and big data methodologies to improve ATM predictability. Another potential KPI could be developed in order to reflect the effect of the ATM system in constraining air traffic, assessing the potential cost/time savings for ANSPs. Modified KPIs to include airport operations performance or new metrics to measure safety in a significant way could also be considered. Linked with the KPIs definition, the **safety trade-offs** with other performance areas should also be further developed.
- There is also room for improvement regarding the study of **ATCO workload**, in order to better understand its dependencies with performance areas such as Capacity as well as with Uncertainty. Nevertheless, data on this subject is not easily accessible, as it may be considered as sensitive information by the ANSPs. Furthermore, it is likely that only limited subsets of historical data can be gathered. This means that the temporal and geographical scope of the research would be limited.
- A strong interest was detected in the development of research questions related with **uncertainty**. Up to five research questions were proposed within this field (e.g., regarding the effect of departure and en-route uncertainties, or the impact of on ground processes on gate-to-gate predictability). These research questions interact with other areas, requiring data availability from various data sources. For instance, flight planning data in several temporal-scopes would be needed.
- **Sector complexity** is significantly linked to other threads of research questions, such as ATCO workload and Uncertainty. However, its main drawback would be the required high granularity, which would reduce the scope of the project as it would not be accessible for all ANSPs.
- Finally, **equity and access** could be assessed in terms of ATM delay distributions over the different airspace users.

2.4.3 Case studies: key results

A subset of the previous list of research questions was selected to be investigated in the context of three case studies.

2.4.3.1 CS-1: Effect of unit rate variance on en-route performance

2.4.3.1.1 Visual exploration

The visualisation exercise developed for CS-1 extracted relevant insights regarding the route choice criteria of airlines. These findings are documented in D3.1 Visual Analytics Exploration of Performance Data. The main parameters identified affecting the route choice are:

- Horizontal length of the route. This variable, measured by the KPI KEA (horizontal en-route efficiency), provides the most significant parameter explaining fuel consumption.

- En-route charges. This variable explains the costs associated to flying through the different regions included in the route. Longer routes can avoid several expensive charging zones thus reducing significantly air navigation charges.
- Congestion. From the visualisation exercises it is clear that the abovementioned variables do not explain all the factors affecting route choice. Some routes may provide a more stable flight time and/or less delays or probability of getting a regulation than others. In addition, less congested routes allow airlines to fly the desired Flight Level (FL) and, thus, reduce fuel consumption. Therefore, some flights will choose those routes to ensure their schedule and reduce costs. Moreover, it is important to note that congestion is not constant as it varies during the day and its effect can be more evident during peak traffic days such as weekends. Thus, using a constant metric of congestion to explain route choice may lead to unreasonable results as it may happen that less-congested routes are mainly used during peak hours.
- Flight time. This variable is highly correlated with the horizontal length of flight. This variable explains the suitability of a route to recover delay and might indicate low congestion. For instance, the behaviour of cargo airlines, which tend to fly at non-congested hours, should be well explained by this variable. However, this variable presents high dispersion inside a cluster because it is also linked with wind and the assigned FL, which is in turn related to congestion.
- Airline. The explained factors have different impact depending on the structure of costs of an airline. Low cost carriers are much more affected by variable costs (such as fuel or charges) and are more prone to cost-optimise their routes. On the other hand, hub-and-spoke airlines might have a higher tendency to choose time-stable routes to ensure that their connections are not affected. Moreover, small airlines might not take into account all the factors because of lack of resources for route planning thus choosing sub-optimal routes.
- Weather. Another relevant factor affecting the choice of a route is weather. Weather can affect in two ways: weather events as Cumulo_Nimbus[1] (CBs) may deviate a route from the intended trajectory and tail winds may make one route choice better than other by reducing fuel consumption. The difference with the previous factors is that this effect is unknown a priori except for weather forecasts that may evolve differently in reality. For instance, the trajectory of a CB may change making a route unavailable.
- Other. This brief list does not include all the factors affecting route choice. There might be other factors that could not be observed in this approach, such as reactionary delay due to previous flights, military activity or seasonal effects. For instance, AIRAC 1603 showed much different airline schedules than AIRACs 1601 and 1602 and they should not be modelled together. These effects would be transformed in a variability or error when modelling route choice.

The abovementioned variables provide a large list of factors affecting route choice. Some of these factors are intrinsic properties of the routes (e.g., average horizontal length), others depend on the airline, and others are evolving variables (e.g., wind). Each variable has different levels of uncertainty, e.g. a flight flying one route may fly longer or shorter due to a variety of factors such as Air Traffic Control (ATC) instructions or to avoid some storm. In addition, the decision of flying one route might depend on other variables not stated in this approach, thus creating a variability that cannot be explained by the means used in the exercise. This variability depends on the particular airline, while some airlines usually fly the same route no matter external variables, others decide from the whole set of routes. This means that two or more routes

1 Cumulo nimbus is a dense, towering vertical cloud, forming from water vapour carried by powerful upward air currents. These clouds are capable of producing lightning and other dangerous severe weather, such as tornadoes.

can result similarly cost-worthy and, in theory, have the same probability of being chosen, but in practice they can show considerably different traffic shares.

2.4.3.1.2 Route choice modelling

Two machine learning techniques were tested in CS-1: multinomial regression and decision tree regression. The best performance for two of the three OD pairs studied was obtained with the multinomial regression technique. This technique proved to improve the traditional methodology (null model, see Figure 11), which predict flows based on similarity criteria with previous flights.

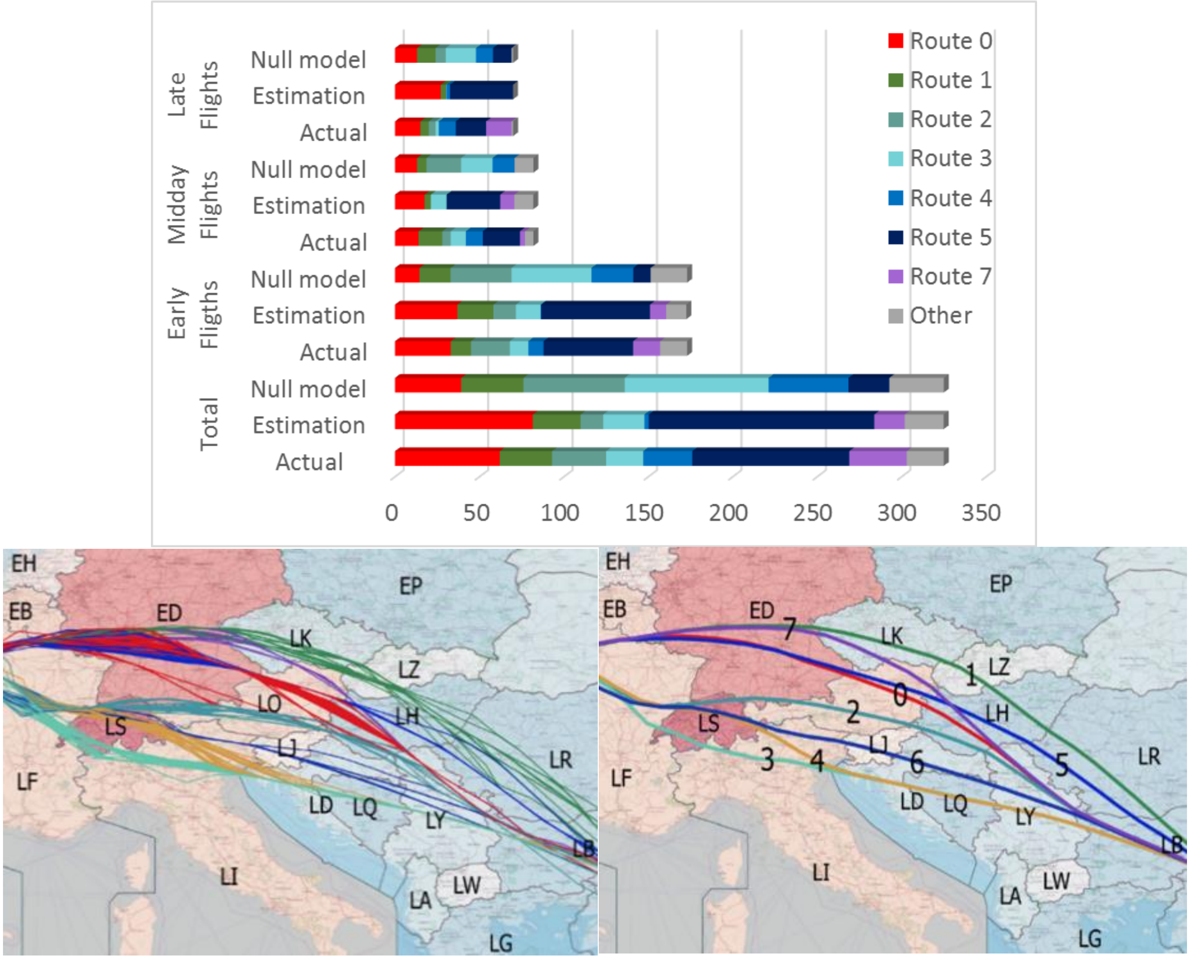


Figure 11. Top: Results of testing a multinomial regression model to predict election of routes from Istanbul to Paris airports compared to actual choices and the null model, which emulates current pre-tactical forecast methods. Bottom: simplification of the routes between Istanbul and Paris airports into 8 cluster trajectories.

A prospective application of the proposed modelling approach is the aggregation of route predictions into traffic demand volumes in order to predict the appearance of hotspots. To do so, the current approach should be applied to all OD pairs for which one or more possible routes cross the hotspot. Then, predictions should be aggregated in a probabilistic manner to obtain the predicted traffic volume in the hotspot. This application would be of use for demand-capacity balancing and planning during pre-tactical planning phase.

On a more strategic level, the modelling approach developed in this paper could also be used to investigate questions related to the interrelationship between ATM Key Performance Areas, e.g. the trade-offs between environment (flight efficiency), capacity (delay) and cost-efficiency.

To sum up, the presented models have a potential for traffic prediction during the pre-tactical planning phase, when no flight plan is available to know which route the airline will choose. This represents a step forward in enhancing ATFCM by the provision of better estimations of traffic evolution. However, the current approach requires further development and enhancement to produce more reliable traffic forecasts in terms of trajectory granularity, generalisation of the algorithm and prediction accuracy.

2.4.3.1.3 Optimisation dashboard

The route choice predictor developed in the context of CS-1 was fed into an optimisation tool to predict the effects on performance of a particular setting of unit rates and help in the selection of the best setting of unit rates according to certain KPIs. The main objective of this dashboard is to assess the performance impact of tuning unit rates on the flights covering a certain OD. A multi-objective optimisation was performed to find the settings of unit rates that optimise a series of KPIs according to different weights given to each KPI. The dashboard enables the evaluation of the trade-offs of a given setting of unit rates in terms of flight efficiency, cost efficiency and capacity by means of different interactive visualisations and assess the effect of unit rates and route choices on ATM performance. The dashboard provides additional functionalities to analyse a certain unit rate setting in detail to assess the impact on the different actors (ANSPs, airlines) affected by the unit rate tuning. Further developments would integrate the predictions of several ODs to provide a comprehensive analysis of the effects of unit rate tuning on all the routes affected.

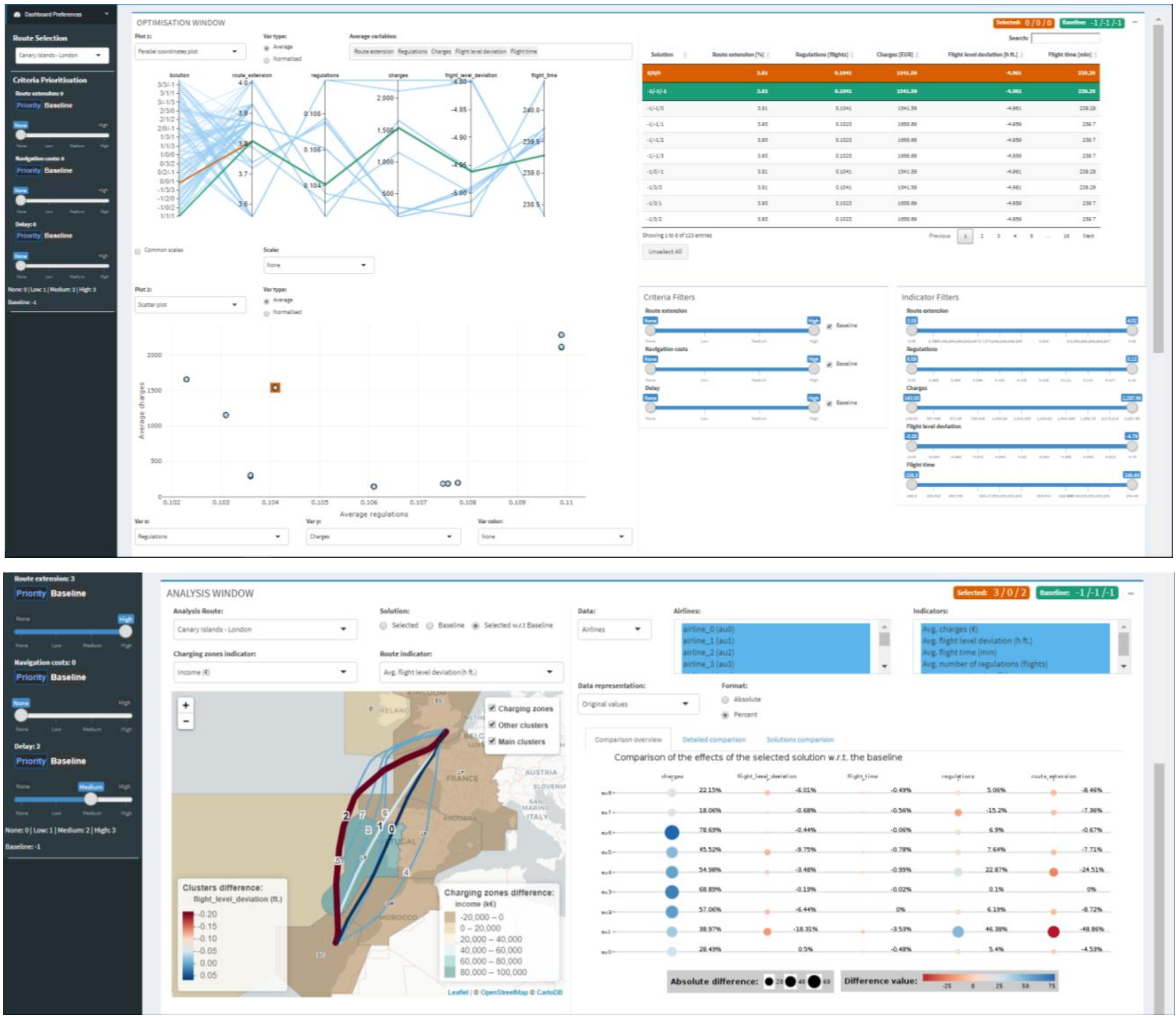


Figure 12. Overall view of the optimisation dashboard developed within CS-1. Top: overview of the optimisation window. Bottom: overview of the analysis window.

This dashboard is integrated in the INTUIT visualisation platform, currently hosted on UPM-CeDInt servers, and can be accessed through the following link: https://viz2know.cedint.upm.es[2]. A description of the type of visualisations developed in this dashboard is documented in D5.1 Performance monitoring and management toolset.

The optimisation dashboard was evaluated by a series of experts in the ATM domain to ensure that the objective of the tool is fulfilled and to receive comments on future developments of the tools. These comments are summarised in deliverable D5.2.

[2] Due to the reserved nature of the data used throughout the project and in compliance with EUROCONTROL DDR2 data policy, the access is restricted to authorized personnel only. Currently, the users allowed to enter the platform and explore the performance dashboards contained therein are people belonging to the INTUIT consortium.

2.4.3.2 CS-2: Identification of flight efficiency influence factors

2.4.3.2.1 Performance modelling

A machine learning model was trained in order to assess performance for a certain airspace area, in particular at ACC level. This case study was conducted in coordination with the other SESAR ER projects AURORA and APACHE by selecting a common scope: analyse flight efficiency in Bordeaux ACC (LFBBCTA). In the proposed approach, performance is modelled as a function of the flight properties by means of a random forest regressor. This model enables the analysis of meaningful correlations between input properties and output performance. The extracted interrelationships are used to perform an assessment of the causes and effects of low performing flights, which is highly useful for the evaluation of performance gains of a certain improvement. For instance, in LFBBCTA, it was discovered that flows from Northern Spain (e.g., Asturias, Bilbao) are less optimal than flows from Canary Islands and Portugal.

The proposed approach can be of use for regulatory authorities, ANSPs and other organisations such as the Performance Review Unit, in order to better recommend mitigation measures or focus areas for performance improvement. The additional value with respect to traditional business intelligence lies in the ability to: (i) isolate the magnitude of inefficiencies; (ii) perform a diagnosis of low performance episodes; and (iii) predict potential improvements due to new solutions.

To generalise the application of the tool, further development would be needed. The modelling should be extended to other ACCs and time ranges, and a systematic comparison of these models across ACCs and study periods should be carried out to find commonalities (global influence factors) and specificities. In addition, the minimum quality and granularity of data used for such assessment should be determined. Finally, the model could be further developed to assess more sophisticated flight efficiency metrics such as vertical efficiency indicators and fuel consumption.

It is important to recall that the efficiency predictor is not meant to be used as an accurate prediction tool but rather as a trend and correlations identification tool. In order to enhance current explanatory variables, future improvements include:

- addition of an absolute congestion variable (number of flights) to spot saturation of the airspace,
- calculation of congestion variables at the time when the flight enters the ACC (entry time and average flights per ATCO),
- use of the expected efficiency from the flight plan as an input, and
- addition of other interface explanatory variables, such as the turn angle at the interface, or the distance from the origin to the entry point.

Another limitation of the proposed approach is the low granularity of the input data (DDR2 actual trajectories), which may lead to different flight efficiency metrics from those actually achieved. In this regard, a further exploration of this approach is planned, consisting in fitting efficiency with ADS-B trajectory data, which has a higher granularity and reflects better the actual trajectories of flights.

2.4.3.2.2 Performance assessment dashboard

A performance monitoring dashboard was developed to help in the identification of sources of en-route flight inefficiency. This dashboard provides a tool to identify and evaluate the causes of flight efficiency in a particular airspace (ACC or charging zone) of the ECAC area by using DDR2 data. Different types of visualisations are used to explore the relationship between flight efficiency indicators and other flight properties derived from both the flight plan and the ideal route (the great circle route), such as average heading, altitude and airspace crossed. The model allows the evaluation of the influence of these factors on flight efficiency by computing different correlation metrics between them. The extracted interrelationships may serve as a basis to perform an assessment of the causes and effects of low performing flights.

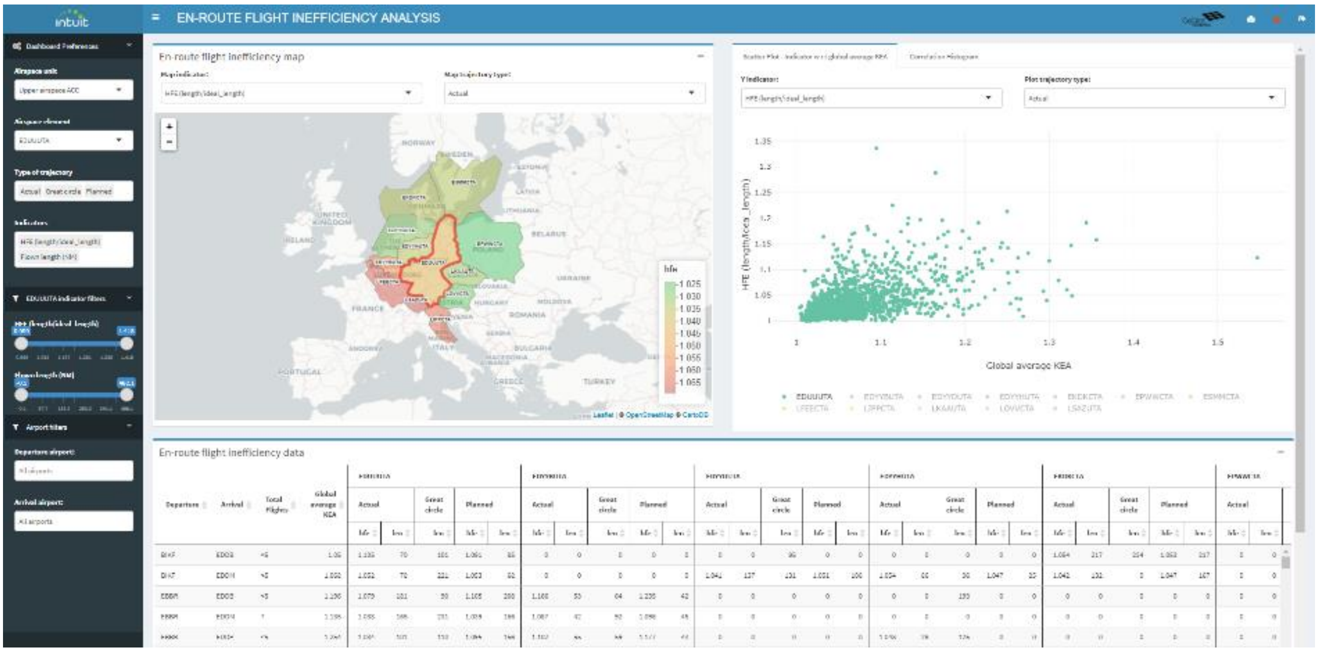


Figure 13. Overall view of the performance assessment dashboard developed within CS-2.

Future developments of the tool should allow the user to analyse in more detail low performing flows and the influence factors determining such low efficiency. Moreover, the study could be extended to other KPAs such as cost-efficiency or capacity. Finally, the tool could be enhanced not only to study correlations but also to measure the influence of such factors in quantity and sense and to provide prediction capabilities to early detect low performance episodes.

This dashboard is integrated in the INTUIT visualisation platform. A description of the type of visualisations developed in this dashboard is documented in D5.1 Performance monitoring and management toolset.

2.4.3.3 CS-3: Multiscale representation of KPI data

CS-3 aimed to study a certain KPA (capacity, in this case) and its related KPIs on a more disaggregated level. In this regard, the relationship between ATFM delays, delay reasons, temporal trends, delay propagation and issued configurations was investigated by means of visual analytics and data analysis at ACC level.

2.4.3.3.1 ATFM delay visualisation

The visualisation exercise developed for CS-3 explored the possibility of disaggregating performance metrics, namely ATFM delay indicators, and the kind of analyses that a disaggregation would enable. The approach computed statistics directly from the aggregates obtained in the ATFM daily summaries. Further conclusions would be expected by calculating aggregates based on complete lists of regulations and the corresponding flights.

2.4.3.3.2 Analysis of ATFM delay

A further analysis of ATFM delay metrics at a disaggregated level was performed and documented in D4.1 Performance Metrics and Predictive Models. The PI "En-route ATFM delay per flight attributable to ANS" was analysed during a whole year in a disaggregated manner as described below:

- ANSP selection. All the ANSPs in the ECAC were considered in a first step to evaluate the similarities and differences between them. Based on these, ANSPs were grouped by means of a hierarchical clustering according to their characteristics (km controlled, traffic seasonality and traffic complexity). From these groups, two ANSPs (Spain and Poland) were selected to be studied due to their prototype characteristics and number or regulations, ensuring the usefulness of the disaggregation.
- Individual regulations and imbalances, trends and seasonal effects in the issued ACCs were analysed taking into account delay reasons. For instance, seasonality is an issue on Spanish ACCs, with almost all regulations issued because of capacity during Summer (see Figure 14).
- Coupled analysis of relations between sector configurations and regulations. The issued configurations (number of open sectors) in each of the ACCs composing the ANSP during the period of study were analysed in terms of the generated ATFM delay under each configuration. This allowed us to assess how ANSPs deal with high demand. For instance, Spanish ACCs have regulations mostly for the configurations with highest number of sectors (see Figure 14).

Based on the analyses performed, the following conclusions were derived:

- ATC capacity related delays (indicated via the registered regulation plans) are the most occurring, followed by those related to 'other reasons'. This also holds true for the total accumulated delay taken over the whole year. The most-occurring ATC capacity related delays happen in the weekends, with Sunday counting the highest average duration due to 'other reasons'.
- When looking at the average delay, weather-related delays also cause a significant disturbance. The average delay remains more or less high throughout the week, with no exceptions on any day.
- The number of active sectors has no strong connection with the average ATFM delay of the issued regulations. Delays of various degrees occur, mostly for the configurations with higher number of sectors. In this regard, it is positive that regulations (and particularly ATC capacity related regulations) mainly occur at high number of active sectors, meaning that the ACCs under study aim to reach maximum capacity before issuing any regulation (see as an example Figure 14).

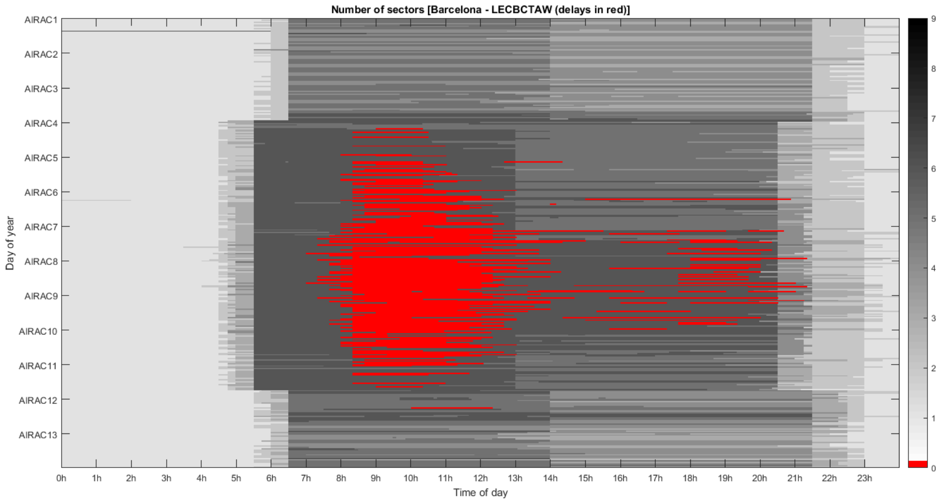


Figure 14. Mapping of regulation activation periods (red) and number of sectors (grey scale) for the regulations assigned to Madrid North (LECMCTAN) Air-traffic Control Centre during 2016. The number of active sectors is shown in grey scale, with lower numbers white to grey, and higher numbers with grey to black.

- There is a small distinction to be made between the delays based due to ATC capacity related issues and those due to 'other reasons': the latter typically occur in more proportion in configurations with higher number of sectors.
- The coincidence of high number of ATC capacity related regulations together with the fact that these occur mostly for configurations with high number of sectors entails that capacity issues in Spanish ACCs are mainly due to saturation of traffic, with Barcelona West suffering the most during the Summer period.

CS-3 assessed whether or not it is worthwhile to look at capacity (en-route ATFM delay attributable to ANS PI) on a more disaggregated level. To that end, the relationship between ATFM delays and the number of active sectors for a set of representative ACCs was investigated. The proposed approach helped to unveil relationships between ATFM delay and regulations occurrences, as well as between the number of active sectors and traffic peaks at a highly disaggregated level, allowing a deeper study on the causes of inefficiencies and low performance in terms of KPI values. In order to further develop the current approach, the approach could be extended to consider longer time horizons, i.e. multiple years. This would allow the study of both seasonal and global trends. Secondly, the approach could be extended to compare different ACCs. This would enable the evaluation of performance linked with operational measures at local level. Such analysis would be particularly useful for the assessment of the issued sector configurations, which are a key tactical decision for ANSPs, and the relationship with traffic demand. Finally, this approach can be extended to other KPAs, KPIs and operational decisions, such as cost-efficiency and ATCO rostering.

2.4.3.3.3 Decision support tool

The clustering tools used in CS-3 to evaluate similarities between ANSPs were fed into a decision support tool for optimising the process of clustering by combining different types of clustering algorithms, distance measures and clustering projections for representing multidimensional visual summaries of clusters and comparing cluster profiles (see Figure 15). This tool was not integrated into the INTUIT platform. Instead, Fraunhofer IAIS relied on its own visual analysis tool, V-Analytics.

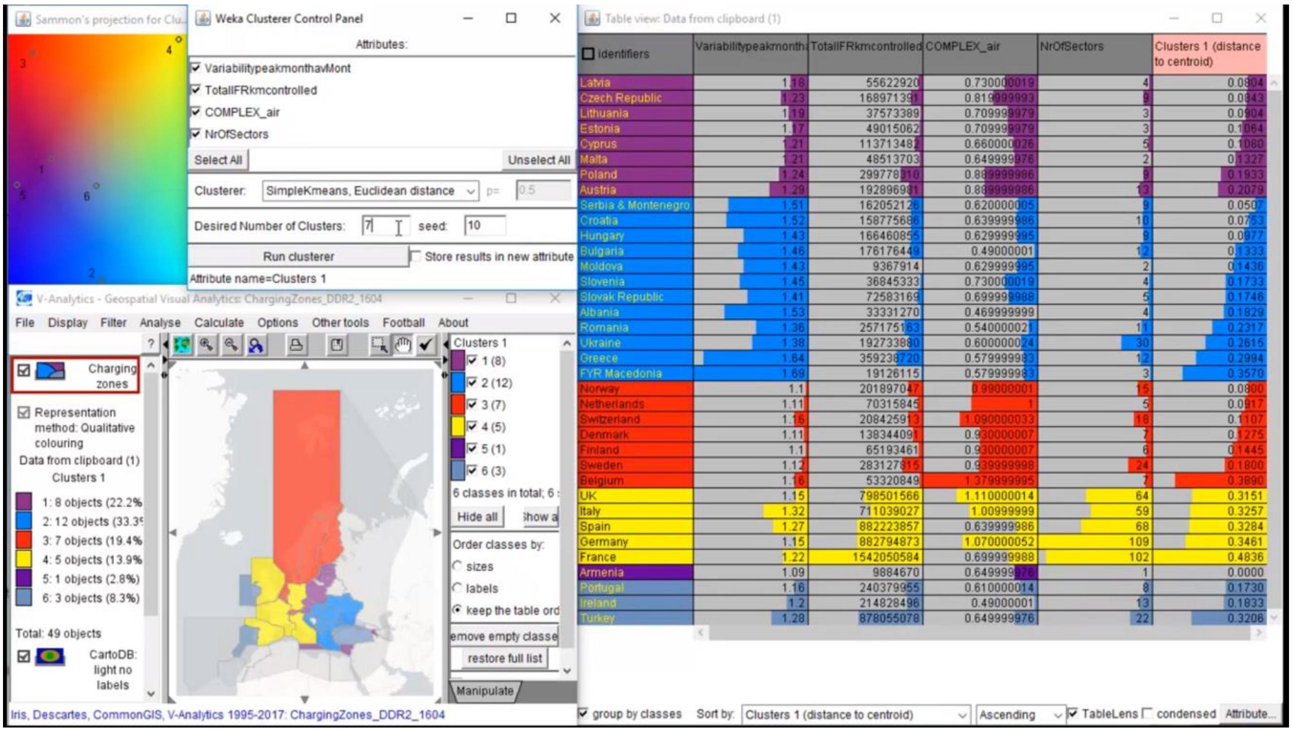


Figure 15. Overview of the decision support tool developed in CS-3.

2.5 Technical Deliverables

Ref.	Title	Delivery Date	Dissemination Level
Description			
D2.1	Performance Data Inventory and Quality Assessment	15/12/2016	Public
This document identifies the main data sources on ATM performance and provides a qualitative assessment on quantity, validity, quality, and geographical and temporal resolution of each dataset. The work documented in this deliverable has produced three main outcomes: a set of Performance Data Factsheets characterising each data source, a Performance Data Guide which links ATM performance data with the sources where such data can be found, and the INTUIT Data Repository, which allows the project partners to share the datasets used for the INTUIT data analysis work.			
D2.2	Qualitative analysis of Performance Drivers and Trade-offs	24/11/2016	Public
This deliverable provides a detailed qualitative analysis of ATM performance drivers and trade-offs between performance areas. This analysis has been conducted through the combination of the analysis and comparison of the different performance frameworks in ATM, desk research focused on studying research papers on ATM performance interdependencies, and a consultation workshop with a variety of ATM stakeholders. The outcome of this research is a list of research questions gathering challenges and opportunities in the ATM domain which are potentially of interest for the upcoming work packages of the INTUIT project.			
D3.1	Visual Analytics Exploration of Performance Data	06/10/2017	Public
This deliverable reports the use of state-of-the-art visualisation and visual analytics techniques to explore ATM performance datasets and extract patterns, regularities and trends, with a view to support the INTUIT performance modelling work. Three visualisation exercises are presented: (i) visualisation and analysis of ATFM delay; (ii) exploration of airline route choices; and (iii) visualisation of flight delays. The specific objectives of each of these exercises are to explore the available data, extract valuable information and unveil complex patterns, and ultimately suggest hypothesis and possible modelling approaches.			
D4.1	Performance Metrics and Predictive Models	12/02/2018	Public
This document reports the results of INTUIT WP4. The purpose of WP4 is to develop new ATM performance modelling approaches taking into account the interdependencies between KPIs at intra and inter KPAs level and the cause-effect relationships between performance drivers and KPIs. The document includes three case studies: (i) characterisation of the trade-offs between cost-efficiency, environment and capacity arising from airline route choices; (ii) identification and assessment of inefficient routes; and (iii) disaggregation of efficiency metrics at finer spatial and temporal granularity. The conclusions of each case study include recommendations for the development of performance monitoring and management tools that make use of the newly developed models.			
D5.1	Performance monitoring and management toolset	19/04/2018	Public
This document reports the results of INTUIT WP5. The purpose of WP5 is to design and develop a performance monitoring and management toolset organised around the concept of an interactive dashboard equipped with a set of visual analytics tools. The document describes the three visualisation tools developed in the context of INTUIT WP5: (i) optimisation of unit rates to optimise a series of KPIs; (ii) performance assessment dashboard for the identification of flight efficiency influence factors; and (iii) decision support tool for data clustering.			

Ref.	Title	Delivery Date	Dissemination Level
Description			
D5.2	Performance Monitoring and Management Toolset Evaluation Report	20/04/2018	Public
This document describes the evaluation of this dashboard. The evaluation has been conducted with a small set of participants with professional background and expertise in projects related to ATM performance, in order to assess the usefulness and applicability of the dashboard and derive recommendations for future improvement.			

Table 1: Project Deliverables

3 Links to SESAR Programme

3.1 Contribution to the ATM Master Plan

INTUIT aims to enhance current performance modelling and management in the ATM domain through the use of machine learning and visual analytics. This novel approach and the project outcomes are mainly related to SESAR transversal activities and do not focus on a particular Operational Improvement (OI) Step or Enabler (EN). Nevertheless, two related SESAR Solutions are identified where INTUIT can contribute advanced.

- The main related Solution is **Solution PJ19.04 — Performance Management**: "provide and maintain the performance framework (including support, guidance and training), define and execute the dedicated performance evaluation and consolidation process in order to integrate and consolidate the performances assessments done at SESAR Solution level". In this regard, INTUIT has advanced in the identification of needs for the Performance Framework and in the use of data-driven techniques to characterise influence factors of KPIs, which could be used to enhance feasibility assessment and Cost-Benefit Analysis.
- Other links: **Solution PJ.09-01 — Network Prediction and Performance**: "Network Prediction and Performance relies on shared situational awareness with respect to demand, capacity and performance and has an impact on regional, sub-regional and local DCB processes. It consists of improved traffic and demand forecast based on Shared Business Trajectory and the computation of confidence indexes. Prediction of DCB constraints and complexity issues will be based on the definition of metrics and algorithms for prediction, detection and assessment of traffic complexity, thus improving the accuracy and credibility of the diagnosis and awareness of hotspots. Network Operations will be monitored through Network Performance KPA/KPI while a Network impact assessment will analyse trade-offs and facilitate collaborative decision-making processes". In this regard, INTUIT has advanced in the characterisation of Airspace Users and in the development of new data-driven techniques to enhance pre-tactical traffic predictions and, therefore, DCB effectiveness.

Code	Name	Project contribution	Maturity at project start	Maturity at project end
PJ19-04	Performance Management	Analysis of performance influence factors	TRL-0	TRL-2
PJ19-04	Performance Management	Interactive performance assessment dashboards	TRL-0	TRL-2
PJ09-01	Network Prediction and Performance	Machine learning techniques for pre-tactical traffic forecast	TRL-1	TRL-2

Table 2: Project Maturity

3.2 Maturity Assessment

The assessment of the maturity of INTUIT results is presented in the following table.

Table 3: ER Application-Oriented Research Maturity Assessment

Thread	ID	Criteria	Satisfaction	Rationale - Link to deliverables - Comments
OPS	OPS.ER.1	Has a potential new idea or concept been identified that employs a new scientific fact/principle?	Achieved	- New performance modelling techniques - D4.1 - New visual analytics techniques for performance analysis - D3.1, D5.1
OPS	OPS.ER.2	Have the basic scientific principles underpinning the idea/concept been identified?	Achieved	D3.1, D4.1, D5.1
OPS	OPS.ER.3	Does the analysis of the "state of the art" show that the new concept / idea / technology fills a need?	Achieved	The proposed solutions address specific needs identified from the literature review in D2.2 and contrasted through consultation with ATM stakeholders.
OPS	OPS.ER.4	Has the new concept or technology been described with sufficient detail? Does it describe a potentially useful new capability for the ATM system?	Achieved	- New performance modelling techniques - D4.1 - New visual analytics techniques for performance analysis - D3.1, D5.1
OPS	OPS.ER.5	Are the relevant stakeholders and their expectations identified?	Achieved	Stakeholder needs and potential use of project results are discussed in D1.3
OPS	OPS.ER.6	Are there potential (sub)operating environments identified where, if deployed, the concept would bring performance benefits?	Partial - Non Blocking	(Sub)operating environments (e.g., integration of the newly developed techniques into SESAR Performance Framework) still require further definition
SYS	SYS.ER.1	Has the potential impact of the concept/idea on the target architecture been identified and described?	Partial - Non Blocking	Impact on architecture still requires further definition
SYS	SYS.ER.2	Have automation needs e.g. tools required to support the concept/idea been identified and described?	Not Applicable	
SYS	SYS.ER.3	Have initial functional requirements been documented?	Partial - Non Blocking	D4.1 and D5.1 provide the necessary elements, but a formal specification is still to be developed.

Thread	ID	Criteria	Satisfaction	Rationale - Link to deliverables - Comments
PER	PER.ER.1	Has a feasibility study been performed to confirm the potential feasibility and usefulness of the new concept / idea / Technology being identified?	Partial - Non Blocking	D1.3 includes a preliminary qualitative assessment that suggests the feasibility of implementing the proposed solutions and identifies the required enablers, but a more comprehensive assessment is needed.
PER	PER.ER.2	Is there a documented analysis and description of the benefit and costs mechanisms and associated Influence Factors?	Not Applicable	
PER	PER.ER.3	Has an initial cost / benefit assessment been produced?	Partial - Non Blocking	D1.3 provides a qualitative assessment of benefits, but a comprehensive assessment of implementation costs is still needed.
PER	PER.ER.4	Have the conceptual safety benefits and risks been identified?	Not Applicable	
PER	PER.ER.5	Have the conceptual security risks and benefits been identified?	Not Applicable	
PER	PER.ER.6	Have the conceptual environmental impacts been identified?	Not Applicable	
PER	PER.ER.7	Have the conceptual Human Performance aspects been identified?	Achieved	D5.2 provides an evaluation of human performance aspects of the proposed tools
VAL	VAL.ER.1	Are the relevant R&D needs identified and documented? Note: R&D needs state major questions and open issues to be addressed during the development, verification and validation of a SESAR Solution. They justify the need to continue research on a given SESAR Solution once Exploratory Research activities have been completed, and the definition of validation exercises and validation objectives in following maturity phases.	Achieved	D1.3, D4.1, D5.1 and D5.2 describe future research and innovation needs for each of modelling approaches and performance analysis tools developed by INTUIT.
TRA	TRA.ER.1	Are there recommendations proposed for completing V1 (TRL-2)?	Achieved	D1.3, D4.1, D5.1 and D5.2 describe future research and innovation needs for each of modelling approaches and performance analysis tools developed by INTUIT.

4 Conclusion and Lessons Learned

4.1 Conclusions

INTUIT has developed a performance modelling and management toolset for a series of case studies that represent a first approach to apply data science techniques, including machine-learning and visual analytics, to relevant problems related to ATM performance. The project has proven the usefulness of such approaches to improve the state-of-the-art in ATM performance analysis, opening the path to further apply this type of analytical tools to ATM performance management and monitoring. The outputs of the project include:

- a detailed review of available databases relevant for ATM performance research, including a traceability exercise between KPIs/KPAs and databases;
- a list of research questions at the intersection of ATM performance and data science;
- a set of new modelling approaches and interactive visualisation tools for ATM performance analysis focused on three specific applications: CS-1 - Study of the effect of unit rates on en-route performance, and more generally the modelling of airline route choice decisions and their impact on ATM performance; CS-2 - Identification of sources of en-route flight inefficiency; and CS-3 - Development of new multi-scale representations of ATM performance indicators.

For each of these applications, the following outputs have been produced:

- definition of the overall concept;
- detailed review of the state-of-the-art;
- development of a data-driven methodology to tackle the research question, including a prototype visualisation tool;
- identification of the potential interested stakeholders and expected outputs and impact of the new tools;
- description of the required developments to upscale the models and visualisation tools to meet the required expectations in an operational environment;
- documentation of the considerations and requisites to be taken into account to further develop such tools to ensure feasibility and usability, including evaluation by external experts;
- a list of future recommended research topics to achieve further maturity levels, built from the previously identified research needs.

4.2 Technical Lessons Learned

INTUIT has served to develop the expertise of the Consortium members in ATM performance:

- expertise in ATM performance management and monitoring and knowledge of the available databases, such as the Demand Data Repository, including their usefulness and limitations;
- state-of-the-art data analysis and performance modelling techniques;
- visual analytics techniques, focused on the visualisation of flight trajectories, clustering and simplification of air traffic flows, representation of performance metrics and pattern identification; and
- web-based tools useful for the implementation and prototyping of performance dashboards and decision-support tools.

Regarding the investigated case studies, several research needs can be extracted for further development of the project results:

- Access to better data and/or easier access to data, such as Correlated Position Reports.
- Upscaling: a key condition for delivering an operational version of the INTUIT prototype tools is to develop and document a stable version of the tools that can be scaled to the whole European network.
- Validation: once a stable and fully documented version is obtained, the tools should be verified and validated to ensure their applicability to the whole network. In addition, dashboards should be evaluated with the final users to ensure a stylised workflow.
- Detailed feasibility assessment and Cost-Benefit Analysis.

4.3 Recommendations for future R&D activities (Next steps)

INTUIT has identified a set of relevant research questions related to ATM performance modelling that can be tackled by means of machine learning and visual analytics techniques. A subset of these research questions were developed in the form of case studies and performance dashboards were prototyped for each case study. Several further research activities are recommended to make progress towards the implementation and industrialisation of such techniques. These recommendations have been enriched and validated through several stakeholder consultation mechanisms, including workshops organised by the project, participation in other performance workshops and collaboration with the other performance-related ER projects, APACHE and AURORA.

Two types of continuation paths have been identified by the INTUIT project: (i) exploration of new data-driven techniques through Exploratory Research funding, such as the one provided by the Knowledge Transfer Network Engage and the next SESAR Exploratory Research call; (ii) industrial development of the most mature results, through SESAR Industrial Research programme.

Regarding the research topics to address in the context of future SESAR ER activities, we can highlight the development of **enhanced performance assessment based on the integration of data-driven and model-driven approaches**. Model-driven approaches are usually based on assumptions of hidden or sensible data (like take-off weight). Data-driven techniques provide a tool to infer such variables. In this regard, we propose to extract such information directly with data-driven models. The tool could be used by other performance modelling tools, such as the ones developed by APACHE and AURORA to predict fuel consumption, as an input to their performance prediction tools. In addition, the list of research questions created at the beginning of the project and reported in D2.2 constitutes a thorough revision of the potential of data-driven techniques to enhance ATM performance assessment and monitoring and will be further developed in the form of a **white paper outlining a possible research agenda** in this domain.

Regarding the research topics with the potential to feed into the SESAR IR, the main opportunities are the following:

- **Data-driven characterisation of Airspace User behaviour.** Performance assessment is sometimes limited by the non-availability of sensitive information from AUs like take-off weight, cost of delay or AU preferences. Data-driven techniques provide a tool to infer such preferences. In this regard, we propose to further develop INTUIT CS-1 to characterise airspace users flying in the whole European airspace in terms of preferences. The tool could be used by other performance projects, e.g. as an input to the performance consolidation methodologies employed by PJ19.04. This would enhance Cost-Benefit Analyses through better characterisation of the performance trade-offs arising from decisions taken from airlines. Ultimately, this tool would impact positively the ATM domain by improving decision-making.

- **Identification and characterisation of performance influence factors by means of data-driven techniques.** The work developed in CS-2 could be generalised to other KPAs and airspaces to identify performance influence factors and measure their relative influence. This tool would serve to validate and/or refine the current influence diagrams developed within SESAR to validate the benefits of a solution, often obtained from expert consultation. This analysis could also be generalised to be applied to other, new performance indicators such as the ones developed in other performance projects such as APACHE and AURORA. This solution would impact positively the decision-making in SESAR by improving the capabilities to isolate the effects of a certain technology in a series of indicators and enabling to spot any trade-offs arising from the application of such technology. In addition, this kind of tools would serve to policy makers and other stakeholders such as ANSPs to better spot the causes of inefficiencies and evaluate mitigating solutions.
- **Operational assessment of ANSPs with disaggregated data**. Related to the work done in CS-3, further developments would focus on the enhancement of the assessment of ANSPs operations. The approach could be extended to consider: (i) longer time horizons to study both seasonal and global trends, (ii) all ACCs in the ECAC, to evaluate performance globally linked with operational measures (such as configurations) and traffic demand at local level, and (iii) other KPAs, KPIs and actions, such as cost-efficiency and ATCO rostering. This type of tools would be of interest for users such as the Performance Review Body and ANSPs to better assess performance, identify the causes of performance degradation, and recommend more effective countermeasures.
- **Performance dashboard coupling interactive visualisation with analytic functionalities.** The dashboards developed within INTUIT project have explored a series of classical and state-of-the art visualisation techniques that could be further developed to enhance current ATM performance dashboards. This type of tools would be of highly interest for users such as the Performance Review Body and ANSPs to better assess performance, identify the causes of performance degradation, and recommend more effective countermeasures.

Finally, it is worth highlighting that, during the execution of INTUIT, collaboration with other ongoing performance projects has been actively sought to ensure the complementarity of the approaches adopted by the different projects and identify potential synergies. This collaboration resulted in a commonly agreed definition of the performance modelling tasks to be developed in CS-2. After the development of the case study, the results were shared among the three projects and led to the identification of the following synergies, which could be developed in future research activities:

- **Synergies with APACHE: combination of model-based and data-driven approaches for performance impact assessment**. Model-based simulations of airspace behaviour (APACHE-model) can be enhanced with INTUIT predictions of contextual aspects such as Airspace User and ANSP behaviour inferred from historical data.
- **Synergies with AURORA**: improved KPIs. The new KPIs developed by the AURORA project have often to rely on hypotheses about Airspace User preferences or sensitive data such as aircraft weight. The approach developed by INTUIT could provide new tools to better characterise that behaviour and thus enhance these metrics, such as the new Equity metrics, which require to compare the Airspace User intentions to the actual flown trajectories.

5 References

5.1 Project Deliverables

- D1.1 Project Plan
- D6.1 Project Website
- D7.1 POPD - Requirement No 1
- D7.2 H - Requirement No 2
- D1.2 Data Management Plan
- D2.1 Performance Data Inventory and Quality Assessment
- D2.2 Qualitative Analysis of Performance Drivers and Trade-offs
- D3.1 Visual Analytics Exploration of Performance Data
- D4.1 Performance Metrics and Predictive Models
- D5.1 Performance Monitoring and Management Toolset
- D5.2 Performance Monitoring and Management Toolset Evaluation Report
- D1.3 Final Project Report

All public deliverables are available from: https://www.intuit-sesar.eu/publications

5.2 Project Publications

- R. Marcos, O.G. Cantú Ros and R. Herranz (2017) "Combining Visual Analytics and Machine Learning for Route Choice Prediction: Application to Pre-Tactical Traffic Forecast", in D. Schaefer (Ed.) Proceedings of the SESAR Innovation Days 2017, EUROCONTROL.
- G. Andrienko, N. Andrienko, W. Chen, R. Maciejewski and Y. Zhao (2017) “Visual Analytics of Mobility and Transportation: State of the Art and Further Research Directions”, IEEE Transactions on Intelligent Transportation Systems, vol. 18(8), pp.2232-2249.
- G. Andrienko, N. Andrienko, G. Fuchs and J. M. Cordero (2017) “Clustering Trajectories by Relevant Parts for Air Traffic Analysis”, IEEE Transactions on Visualization and Computer Graphics, vol. PP, no. 99, pp. 1-1
- R. Marcos, D. Toribio, N. Alsina, L. Garrigó, N. Andrienko, G. Andrienko, L. Piovano, T. Blondiau and R. Herranz (2016) "Visual Analytics and Machine Learning for ATM Performance Modelling: Preliminary Findings of the INTUIT Project and Prospects for Future Research", in D. Schaefer (Ed.) Proceedings of the SESAR Innovation Days 2016, EUROCONTROL.

Appendix A Acronyms and Terminology

Term	Definition
ACC	Area Control Centre
ACE	ATM Cost-Effectiveness Benchmarking Report
ADS-B	Automatic Dependent Surveillance-Broadcast
AIM	ATFM Notification Message
AIRAC	Aeronautical Information Regulation And Control
ANM	ATFM Notification Message
ANS	Air Navigation Services
ANSP	Air Navigation Service Provider
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATFCM	Air Traffic Flow and Capacity Management
ATFM	Air Traffic Flow Management
ATM	Air Traffic Management
AU	Airspace User
AUP/UUP	Airspace Use Plan / Updated Airspace Use Plan
CB	Cumulo-Nimbus
CBA	Cost-Benefit Analysis
CODA	Central Office for Delay Analysis
CS	Case Study
DCB	Demand and Capacity Balancing
DDR2	Demand Data Repository
ECAC	European Civil Aviation Conference
EN	Enabler
ER	Exploratory Research
FL	Flight Level
ICAO	International Civil Aviation Organisation

Term	Definition
KEA	Horizontal En-Route Flight Efficiency Key Performance Indicator
KPA	Key Performance Area
KPI	Key Performance Indicator
NM	Nautical Miles
NOP	Network Operations Portal
NPP	National Performance Plan
OI	Operational Improvement
OPS	Operations
PER	Performance
PI	Performance Indicator
PJ	Project
PRR	Performance Review Report
PRU	Performance Review Unit
RAD	Route Availability Document
RP	Reference Period
SBT	Shared Business Trajectory
SES	Single European Sky
SESAR	Single European Sky ATM Research Programme
SJU	SESAR Joint Undertaking (Agency of the European Commission)
STATFOR	Statistics and Forecasts
SYS	System
TRA	Transversal
TRL	Technology Readiness Level
VAL	Validation

Table 4: Acronyms and terminology