Hindawi
Journal of Advanced Transportation
Volume 2018, Article ID 1601295, 17 pages
https://doi.org/10.1155/2018/1601295

Research Article

A Novel Performance Framework and Methodology to Analyze the Impact of 4D Trajectory Based Operations in the Future Air Traffic Management System

Sergio Ruiz,[1] **Javier Lopez Leones,**[2] **and Andrea Ranieri**[3]

[1] *Technical Innovation Cluster on Aeronautical Management, Universitat Autònoma de Barcelona, Sabadell, Barcelona, Spain*
[2] *Boeing Research & Technology Europe, Madrid, Spain*
[3] *Advanced Logistics Group, Indra, Barcelona, Spain*

Correspondence should be addressed to Sergio Ruiz; sergio.ruiz@uab.es

Received 10 July 2017; Revised 6 December 2017; Accepted 31 December 2017; Published 11 February 2018

Academic Editor: Andrea D'Ariano

The introduction of new Air Traffic Management (ATM) concepts such as Trajectory Based Operations (TBO) may produce a significant impact in all performance areas, that is, safety, capacity, flight efficiency, and others. The performance framework in use today has been tailored to the operational needs of the current ATM system and must evolve to fulfill the new needs and challenges brought by the TBO content. This paper presents a novel performance assessment framework and methodology adapted to the TBO concept. This framework can assess the key performance areas (KPAs) of safety, capacity, and flight efficiency; equity and fairness are also considered in this research, in line with recent ATM trends. A case study is presented to show the applicability of the framework and to illustrate how some of the complex interdependencies among KPAs can be captured with the proposed approach. This case study explores the TBO concept of "strategic 4D trajectory deconfliction," where the early separation tasks of 4D trajectories at multisector level are assessed. The framework presented in this paper could potentially support the target-setting and performance requirements identification that should be fulfilled in the future ATM system to ensure determined levels of performance.

1. Introduction

The Single European Sky (SES) High-Level Goals are ATM performance targets set by the European Commission with the support of the Single Sky Committee to steer the design and management of future ATM operations. In 2005, the commission set four high-level performance goals for the SES to be met by 2020 and beyond: (1) achieve a 3-fold increase in capacity; (2) improve the safety by a factor of 10; (3) enable a 10% reduction in the effects that flights have on the environment; and (4) provide Air Traffic Management (ATM) services to the airspace users at a cost of at least 50% less. These overarching goals set the initial foundation of the SES initiative and, despite the evolution of aviation since 2005, they remain a valid reference in 2017 for assessing the performance of the ATM system in Europe.

As a contribution to the SES High-Level Goals, the Single European Sky ATM Research (SESAR) Concept of Operations [1, 2] is structured around the new concept of Trajectory Based Operations (TBO), which consists of a coordination of four dimensional trajectory predictions and executions subject to generic ATM constraints imposed across all involved operational stakeholders during the last 24 hours before flight arrival [3]. The trajectory element is intended to act as the coordination instrument among the different ATM components identified in the ICAO's Global ATM Operational Concept [4], one of these being conflict management. This consists of three layers, depending on the timeframe on which it applies: (a) strategic conflict management achieved through airspace organization and management, demand-capacity balancing, and traffic synchronization; (b) separation provision; and (c) collision avoidance.

The steering of TBO at planning and execution phases will evolve towards the so-called Performance Based Operations (PBO), which will take into consideration a holistic view of the ATM and the different performance areas of interest. However, the performance assessment framework used today in pre- and postoperations has not been designed for the purpose of assessing and target-setting the performance of future ATM and therefore a novel framework is needed. In order to validate the operational feasibility of TBO concept and to anticipate its impact on the different ATM stakeholders, the new performance framework should be applied from R&D phase onwards. This is a necessary step in order to generate commitment from all ATM actors to proceed with implementation.

This paper presents a novel performance assessment framework that includes a set of advanced metrics and a methodology based on simulation, deconfliction, and optimization tools. This framework can be used to assess the performance of any traffic demand and airspace organization plan, including the airline operational efficiency and ATM performance in terms of safety, capacity, environmental impact, equity, and fairness. The added value of this new performance framework is its capability to reproduce through high fidelity simulations the dynamics and complex interactions of the flights and their corresponding 4D trajectories subject to ATM constraints. The allocation of 4D trajectories to flights and the management of trajectories are topics still under research in the context of SESAR and the ATM scientific community, but the proposed framework allows different implementation options, thus being flexible to be adopted by other ATM R&D programmes centered around TBO concept, such as NextGen, CARATS, or MEAP.

To show the applicability of the proposed performance framework, a specific TBO concept is used as a case study: "strategic 4D trajectory deconfliction." For this purpose, the trajectory deconfliction algorithms presented in [5, 6] will be used during the performance assessment. These algorithms can be applied for collaborative and optimized planning of precise 4D trajectories at ECAC level, which will be subject to the constraints imposed by strategic deconfliction mechanisms. In this paper, these are referred to as *STREAM algorithms* (i.e., *strategic trajectory deconfliction to enable seamless aircraft conflict management*), which are consistent with the SESAR ATM target concept of operations; that is, they aim at proposing a set of deconflicted trajectories that respect the active ATM network constraints as well as the airspace users (AU) preferences in a free-route environment that covers the entire European Civil Aviation Conference (ECAC) airspace [7]. Simulation and performance assessment results will be analyzed and discussed to show how the proposed performance framework may contribute to increasing the current knowledge of the ATM system dynamics and the complex interdependences between the different KPAs.

The remainder of this paper is structured as follows. Section 2 presents the review and limitations of the current performance framework. Section 3 presents the novel performance framework proposed. Section 4 introduces details on the methodology and the case study used to assess the performance of the STREAM algorithms. Section 5 presents the assessment and discussions of the simulation results, and Section 6 outlines the conclusions and future research.

2. State of the Art

2.1. Current Performance Scheme. The Single European Sky (SES) Performance Scheme [8] is an important management tool to steer and monitor the continuous improvement of European ATM performance as the SES is implemented. The scheme establishes EU-wide targets for 4 Key Performance Areas (KPAs): safety, cost efficiency, capacity, and environment. These targets are reviewed and updated periodically in reference periods (RP, being RP1 covering 2012–2015 and RP2 covering 2015–2019).

The SES Performance Scheme defines a set of Key Performance Indicators (KPIs) for each of the KPAs, providing common metrics to evaluate the performance of the system based on measurable operational data. The Performance Scheme sets targets for each of the KPIs, establishes mechanisms to monitor their values with respect to the targets, and contemplates incentives and corrective actions at both European and national/FAB levels to enforce compliance with the targets. The KPAs and KPIs and PIs adopted by the SES Performance Scheme in two reference periods are indicated in Table 1 (indicators that are KPI are identified with symbol (K)).

The SES Performance Scheme (Key) Performance Indicators measure the impact of ATM service provision on system-wide performance. The prefix *Key* means that the indicator is subject to a target, when it constitutes a measure of the achievement of a business objective for service providers, while the others are supportive of the *Key* ones or represent their measurable outcome (i.e., the metric) but are not targeted directly.

The practical experience so far is that before becoming a targetable KPI; the equivalent PI is first defined and monitored by collecting a representative set of historical observations by the regulator, to allow setting adequate and commensurate quantitative targets as soon as the PI is considered "mature." For the purpose of this paper no distinction is made however between KPIs and PIs, assuming that whatever variable that can be measured as a PI can undergo its evolution process, becoming a KPI at a certain point in time.

Further development and update of the SES Performance Scheme are undergoing at the time of writing this paper, to broaden its scope, enhance transparency of definitions and operational meaningfulness, and streamline the data collection and reporting processes in view of the RP3 (2020–2024). In particular the Safety KPA is receiving much attention to extend the performance indicators set, by including a combination of lagging indicators (outcome-based) and leading indicators (process-based) to monitor safety performance.

2.2. Advanced Performance Metrics and Tools Not Used in the Current Performance Framework. To steer and monitor the performance in a TBO environment, additional indicators and aggregation levels with respect to the current performance scheme metrics are needed. According to [4] one of the main pillars of the future ATM systems should be an efficient Performance Management System, to steer and monitor the performance in 11 KPAs, that is, safety, capacity,

Table 1: KPAs and KPIs/PIs adopted by the SES Performance Scheme.

KPA	KPI/PI	Targeted since	Monitored since
Capacity	Arrival ATFM delay (K)	RP2	RP1
	En-route ATFM delay (K)	RP1	
	ATC predeparture delay		RP2
	Adherence to ATFM slots		RP2
Cost efficiency	Determined unit cost for en-route ANS (K)	RP1	
	Determined unit cost for terminal ANS (K)	RP2	RP1
	Costs of Eurocontrol		RP2
Environment	KEP (horizontal en-route flight efficiency of last filed flight plan) (K)	RP1	
	KEA (horizontal en-route flight efficiency of actual trajectory) (K)	RP2	
	Additional time taxi-out phase		RP2
	Additional time in arrival sequencing and metering area (ASMA)		RP2
	Effective use of conditional routes		RP1
	Effectiveness of booking procedures for FUA		RP2
	Rate of planning of conditional routes		RP2
Safety	Effectiveness of safety management (K)	RP2	RP1
	Application of severity classification scheme (RAT methodology) (K)	RP2	RP1
	Application of just culture (K)	RP2	RP1
	Airspace infringements		RP2
	Application of automatic data recording for runway incursion monitoring		RP2
	Application of automatic data recording for separation minima infringement monitoring		RP2
	ATM-specific occurrences at ATS units		RP1
	Level of occurrence reporting		RP2
	Runway incursions		RP1
	Separation minima infringements (K)		RP1

cost-effectiveness, efficiency (operational), flexibility, predictability, security, environmental sustainability, access and equity, participation, and interoperability.

In order to be valuable and practical, metrics must (a) be measurable (or be calculable through other measures available); (b) have a clear definition (including boundaries of the measurements); (c) indicate progress toward a performance target; and (d) answer specific questions about performance. Due to that, some of the KPAs are not easy to measure and as a consequence not all of them have the same number of metrics available, even though some of them may not still have metrics [9]. In addition, there is no global consensus on a standardized set of performance indicators, where this is one of the biggest challenges for the ATM community. Large variety of indicators can be explained by different interpretations of the ICAO high-level guidelines, different needs, different perspectives, and/or different approaches, among others.

New metrics have been proposed in many recent researches that require the use of simulations to assess or approximate the performance of given traffic records. For instance, in [10] (Alligier et al. 2013), a method is described that typically requires some aircraft performance data to calculate fuel burn (such as aerodynamic coefficients or specific fuel consumption figures), since fuel data is not available from surveillance data, and it will most likely not be available in the near future due to privacy concerns of the AUs.

A very important requirement for a performance management system is the capability for balancing between various KPAs by including their interdependencies into the analysis. By now, all relevant organizations in charge of measuring and reporting past, current, or expected future ATM system performance observe the KPAs independently of one another. The use of simulation and optimization tools can be useful for recreating future operational and business model changes of the ATM main actors (airspace users, air navigation service providers, and ATFCM, in the case of en-route operations), and the new metrics to capture the impact of these new TBO concepts in ATM performance, along with the complex interdependences and trade-off among the different KPAs [5]. For instance, if all the AUs could fly their "User-Preferred 4D Trajectory" (UP4DT), that is, with minimum ATM constraints, the metrics of operational efficiency would be close to their maximum levels; however, too much freedom for flights may affect negatively the KPAs of safety and capacity; finding what is the optimal balance between freedom and constraints to flights is still a subject of scientific interest because the exact relationship among these KPAs is still not well understood today.

In line with the above, a KPI may be defined to measure the difference between the minimum possible cost of a certain flight trajectory, that is, the cost of the UP4DT without any ATM constraints or interventions, and the cost of the trajectory actually flown by the aircraft. Such a KPI (referred to in this paper as *user cost efficiency metric*) can measure the cost penalty caused by the ATM system, that is, the increment in operational cost that results from flying a trajectory that is compliant with ATM rules, procedures, and air traffic control (ATC) instructions to ultimately guarantee safety [11].

Fairness and equity are also important aspects to be considered during the evaluation of the performance of TBO concept, since one of the key objectives of the SES is to accommodate better the trajectories that are preferred by airspace users within a framework underpinned by equity and impartiality principles. Historically the first-come-first-served rule has been considered fair [12] but one of the objectives of SESAR is to move away from such principle in favor of performance oriented prioritizations. Metrics could also be defined to evaluate the fairness of the ATM service from the point of view of how any cost penalties are distributed among different users as a consequence of the ATM constraints imposed to their flights. Nevertheless, today there is still no general agreement about what metrics can best represent this aspect; therefore the definition of fairness as proposed by [13] is incorporated into the performance framework proposed in this paper.

3. Performance Framework Definition

3.1. TBO Compliant Performance Metrics. The following three subsections define the set of metrics (KPIs) proposed in this paper for the evaluation of performance before and after the application of the conflict detection and resolution (CD&R) algorithms used for strategic 4D trajectory deconfliction. Three categories are considered: (1) *ATM system performance metrics*, corresponding to the KPAs defined by the SES Performance Scheme and aimed at measuring the impact of the algorithms on system-wide ATM performance aspects, (2) *operational efficiency metrics*, measuring the impact of the CD&R algorithms on the operational costs of the individual AUs, and (3) *fairness and equity metrics*, which measure how well the distribution of deconfliction costs (i.e., costs caused by the deviations from UP4DT) is balanced among AUs, considering their revealed preferences. Note that the set of metrics proposed in this paper is not covering all the KPAs defined by ICAO, but the framework could be easily updated as soon as new metrics are available or defined. Also note that all of these metrics will be calculated with the support of traffic and ATM simulation tools as explained in Section 4.

3.1.1. ATM System Performance Metrics. *Number of conflicts* is a metric that has been calculated as the number of infringements of the specified separation minima within each sector over the time interval of the scenario (2 hours). The STREAM Conflict Detection (STREAM CD) algorithm has been used to evaluate this metric with the following aircraft separation minima: 5 NM for lateral separation and 1000 feet for vertical separation (in line with the Reduced Vertical Separation Minima definition). This metric can be interpreted as a partial indicator of the ATC officers workload [14] and may therefore provide some indication about the operational capacity at sectors; that is, for similar operational conditions and traffic patterns, a larger number of conflicts may require more operational capacity available. The number of conflicts can also be interpreted as an indicator of safety; that is, the lower the probability/risk of conflicts (ATC predictions that are precursors of potential separation infringements), the lower the probability/risk of actual aircraft separation infringements and consequently the lower the probability/risk of mid-air collisions. This metric is in line with the modern dynamic risk modelling approaches emerging in several fields, including the ATM field. See, for instance, [15].

Traffic density is defined in the literature as the average number of aircrafts present in a specific volume of airspace over a defined time interval. Given a sector S defined as a 3-dimensional volume in space, the traffic density metric is defined as follows:

$$\mathrm{KPI}_{\mathrm{td}}^{S} = \frac{N_T^S}{T}, \tag{1}$$

where N_T^S is the number of aircrafts flying through a sector S during a reference time interval T, for example, 1-hour interval or the duration of the scenario considered. This metric is also a partial indicator of capacity.

Peak load is defined in the literature as the maximum number of aircrafts that are simultaneously flying within a specific volume of airspace over a reference time interval [16]. Given a sector S, the peak load metric is defined as follows:

$$\mathrm{KPI}_{\mathrm{pl}}^{S} = \max_{t \in T} \left(N_t^S (t) \right), \tag{2}$$

where $N_t^S(t)$ is the number of aircraft present in sector S at time instant t and T is the reference time interval over which the peak load is measured, for example, 1 hour or the duration of the scenario. Peak loads have a direct impact on the controllability of a sector and therefore can be used as a partial indicator of operational capacity.

The *impact on capacity* is defined in this paper as a composed metric that is evaluated taking into consideration the traffic density, the peak load, and the number of conflicts in each sector:

$$\mathrm{KPI}_{\mathrm{ic}}^{S} = \left(\frac{\mathrm{KPI}_{\mathrm{td,up}}^{S}}{\mathrm{KPI}_{\mathrm{td,dc}}^{S}}, \frac{\mathrm{KPI}_{\mathrm{pl,up}}^{S}}{\mathrm{KPI}_{\mathrm{pl,dc}}^{S}}, N_{c,\mathrm{up}}^{S} - N_{c,\mathrm{dc}}^{S} \right), \tag{3}$$

where $\mathrm{KPI}_{\mathrm{td,up}}^{S}$, $\mathrm{KPI}_{\mathrm{pl,up}}^{S}$, and $N_{c,\mathrm{up}}^{S}$ are, respectively, the traffic density, peak load, and number of conflicts for sector S considering the trajectories before applying the CR algorithms and $\mathrm{KPI}_{\mathrm{td,dc}}^{S}$, $\mathrm{KPI}_{\mathrm{pl,dc}}^{S}$, and $N_{c,\mathrm{dc}}^{S}$ are, respectively, the traffic density, peak load, and number of conflicts for sector S after deconfliction.

Due to the difficulties to actually measure the impact on capacity [16–19], this metric will be interpreted in this paper according to the following categories:

(1) ($\mathrm{KPI}_{\mathrm{td,up}}^{S}/\mathrm{KPI}_{\mathrm{td,dc}}^{S} = 1$, $\mathrm{KPI}_{\mathrm{pl,up}}^{S}/\mathrm{KPI}_{\mathrm{pl,dc}}^{S} = 1$, $N_{c,\mathrm{up}}^{S} - N_{c,\mathrm{dc}}^{S} = 0$): no impact on sector capacity

(2) ($\mathrm{KPI}_{\mathrm{td,up}}^{S}/\mathrm{KPI}_{\mathrm{td,dc}}^{S} \geq 1$, $\mathrm{KPI}_{\mathrm{pl,up}}^{S}/\mathrm{KPI}_{\mathrm{pl,dc}}^{S} \geq 1$, $N_{c,\mathrm{up}}^{S} - N_{c,\mathrm{dc}}^{S} \geq 0$): reduction of ATC workload (more capacity available is assumed)

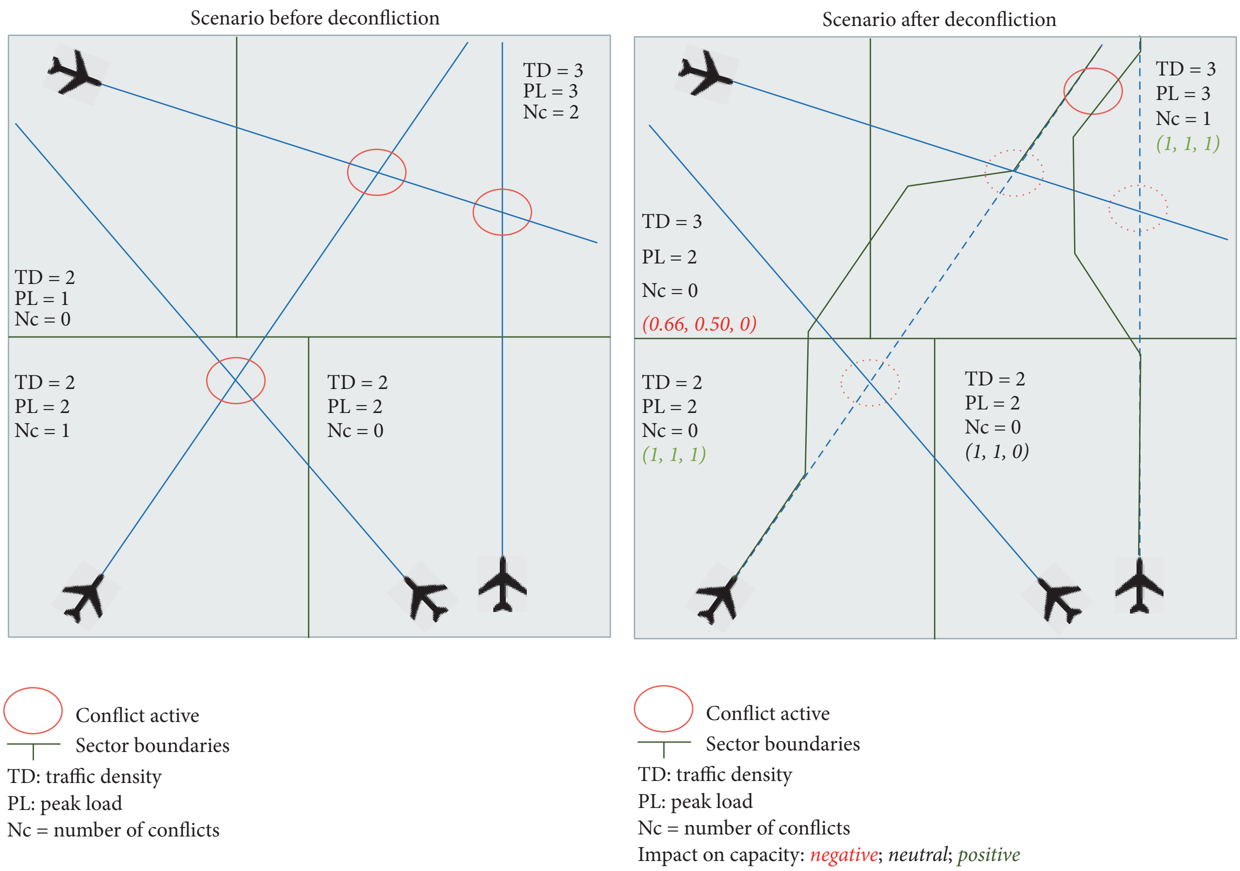


FIGURE 1: Notional illustration of the application of the impact on capacity metric. Dotted circle means "conflict solved."

(3) $(\mathrm{KPI}^{S}_{\mathrm{td,up}}/\mathrm{KPI}^{S}_{\mathrm{td,dc}} \leq 1,\ \mathrm{KPI}^{S}_{\mathrm{pl,up}}/\mathrm{KPI}^{S}_{\mathrm{pl,dc}} \leq 1,\ N^{S}_{c,\mathrm{up}} - N^{S}_{c,\mathrm{dc}} \leq 0)$: increase of ATC workload (less capacity available is assumed).

In principle, the higher the positive values of the elements of this metric are, the larger the potential capacity increment is as a result. Conversely, the smaller the values are, the more constrained the available sector capacity will be. Those other cases in which one of the elements of the impact of capacity metric increases while another element decreases are not conclusive regarding the impact on capacity and ATC workload. Note that this metric does not weigh which element, that is, whether traffic density, peak load, or number of conflicts, influences more the ATC workload. An example that shows how the impact of capacity works can be found in Figure 1: after the deconfliction two of the sectors present a potential increment on capacity, whereas one sector is affected negatively for the conflict resolution (a new aircraft that was not expected initially would appear).

Delay is defined in literature as the time difference between the scheduled time at a certain point, such as a runway threshold or the entry/exit to an airspace sector, and the actual time over that point. Given a set of N aircrafts and a reference point, a commonly used delay metric is the average delay per aircraft over that point:

$$\mathrm{KPI}_{\mathrm{delay}} = \frac{1}{N}\sum_{k\in A} t_{a,k} - t_{s,k}, \tag{4}$$

where A denotes a set of N aircrafts, $t_{s,k}$ is the scheduled time at a certain point for aircraft k, and $t_{a,k}$ is the actual time over that point for the same aircraft k.

Emissions are a set of three metrics to assess the potential impact of traffic on emissions, that is, to compare the hypothetical amount of CO_2, NO_x, and SO_2 that would be released if the 4D trajectories planned by the AUs and cleared/deconflicted by the air traffic services were actually flown, from departure to destination, against the amount of pollutants that would be released to the atmosphere if the aircraft flew their original user preferred 4D trajectories; that is, with no air traffic service constraints incorporated from departure to destination (hypothetical best case). The following metrics are used in line with ICAO Annex 16 [20]:

$$\begin{aligned}
\mathrm{KPI}^{\mathrm{CO}_2}_{\mathrm{emissions}} &= \frac{1}{N}\sum_{k\in A} c_{\mathrm{CO}_2}\left(\Delta m^{\mathrm{dc}}_{F,k} - \Delta m^{\mathrm{up}}_{F,k}\right), \\
&\quad \text{with } c_{\mathrm{CO}_2} = 3.149\ \text{Kgs/Kg of fuel burnt}, \\
\mathrm{KPI}^{\mathrm{NO}_X}_{\mathrm{emissions}} &= \frac{1}{N}\sum_{k\in A} c_{\mathrm{NO}_X}\left(\Delta m^{\mathrm{dc}}_{F,k} - \Delta m^{\mathrm{up}}_{F,k}\right), \\
&\quad \text{with } c_{\mathrm{NO}_X} = 1.230\ \text{Kgs/Kg of fuel burnt}, \\
\mathrm{KPI}^{\mathrm{SO}_2}_{\mathrm{emissions}} &= \frac{1}{N}\sum_{k\in A} c_{\mathrm{SO}_2}\left(\Delta m^{\mathrm{dc}}_{F,k} - \Delta m^{\mathrm{up}}_{F,k}\right), \\
&\quad \text{with } c_{\mathrm{SO}_2} = 0.00084\ \text{Kgs/Kg of fuel burnt},
\end{aligned} \tag{5}$$

where each KPI expresses the amount of emissions of a certain pollutant per flight for a given set of flights, A denotes a set of N aircrafts, $\Delta m_{F,k}$ is the amount of fuel consumed by aircraft k (which includes the evolution of the aircraft mass), "dc" refers to the trajectories after deconfliction and the superscript "up" refers to the trajectories before deconfliction (user-preferred), and c_{CO_2}, c_{NO_X}, and c_{SO_2} are constant values obtained from [21].

3.1.2. Operational Performance Metrics.

User cost efficiency is a novel metric defined in this paper to measure the cost penalty caused to airspace users by the route/trajectory changes proposed by the air traffic services, which are represented by the CR algorithms in the case study. To evaluate this metric, the cost associated with each flight has been defined according to the simplified cost model in the following expression:

$$C = C_F \Delta m_F + C_T \Delta t, \tag{6}$$

where Δm_F is the fuel consumed during the flight, Δt is the flight time, and C_F and C_T define the unit cost assigned by the airline to the fuel consumed and the flight time, respectively.

It is assumed that the operational objective of AUs is to minimize the total cost of each of their flights, given by the simplified cost function C. The values of the coefficients C_F and C_T defining this cost function depend, among other factors, on the airline's cost structure and business strategy. The value of this ratio, which determines the cost to be minimized, is the cost index (CI) defined in literature as

$$\mathrm{CI} = \frac{C_T}{C_F}. \tag{7}$$

Since C_F and C_T values are typically confidential to the airline, the total cost normalized with the unit cost of fuel for the flight, denoted as C^*, is used. C^* can be calculated using the following expression (obtained dividing by C_F both sides of the formula of the total cost C):

$$C^* = \frac{C}{C_F} = \Delta m_F + \mathrm{CI}\Delta t, \tag{8}$$

where C^* is the total cost normalized with the unit fuel cost for the flight, CI is the cost index for the flight, and Δm_F and Δt are obtained from the synthesized trajectory for that CI (original UP4DT or deconflicted). For simplicity, it is assumed that C_F is the same for all the flights in the scenario. The comparisons and metrics calculations that involve cost are done in terms of normalized cost C^*, whose units are kilograms of fuel, instead of total cost C, whose units are dollars or euros. It is also assumed that each airspace user in the scenario assigns a specific value of CI to each of the flights it operates. For the simulations the CI of each specific flight has been randomly selected from a predefined interval (see Section 4.2).

Considering the above, the cost efficiency metric for a flight is given by the difference between the normalized cost of the trajectory that results from implementing the trajectory amendments proposed by the CR algorithm, C^*_{dc}, and the normalized cost of the original UP4DT, C^*_{up}. The value of C^*_{dc} is calculated using the CI assigned to the flight and the values of Δm_F and Δt obtained from the corresponding synthesized trajectories. The paper defines the user cost efficiency metric as the average value of the differences between normalized costs for each airspace user over all the trajectories it operates in the scenario, using the following expression:

$$\mathrm{KPI}_{\mathrm{uce,AU}_j} = \frac{1}{N_j} \sum_{k \in A_j} \left(C^*_{\mathrm{dc},k} - C^*_{\mathrm{up},k} \right), \tag{9}$$

where A_j is the set of flights operated by the airspace user AU_j in the scenario, N_j is the number of flights in that set, $C^*_{\mathrm{dc},k}$ is the normalized cost of flight k after deconfliction, and $C^*_{\mathrm{up},k}$ is the normalized cost of the UP4DT of that flight (before deconfliction). If a flight trajectory is not modified by the TBO model, then $C^*_{\mathrm{dc},k} = C^*_{\mathrm{up},k}$.

User fuel efficiency measures the difference in fuel consumption between the UP4DT and the trajectory that results from implementing the deconfliction strategies proposed by the CR algorithms. The metric is calculated in this paper similarly to the user cost efficiency, but focusing on fuel consumption instead of focusing on the normalized costs:

$$\mathrm{KPI}_{\mathrm{ufe,AU}_j} = \frac{1}{N_j} \sum_{k \in A_j} \left(\Delta m_F^{\mathrm{dc},k} - \Delta m_F^{\mathrm{up},k} \right), \tag{10}$$

where A_j is the set of flights operated by airspace user AU_j in the scenario, N_j is the cardinality of that set, $\Delta m_F^{\mathrm{dc},k}$ is the fuel consumption of flight k after deconfliction, and $\Delta m_F^{\mathrm{up},k}$ is the fuel consumption of the last up-to-date trajectory before deconfliction for that flight k. If a flight's trajectory is not modified by TBO deconfliction algorithms, then $\Delta m_F^{\mathrm{dc}} = \Delta m_F^{\mathrm{up}}$.

3.1.3. Fairness and Equity Metrics.

Fairness can be defined as the quality of distributing something among a set of individuals in a manner such that each receives a share that fulfills its individual satisfaction threshold [13]. In order to measure fairness objectively, it is essential to agree on a common way to quantify such individual satisfaction thresholds. On the other hand, *equity* measures how uniformly the distribution of the good is performed, that is, without taking into account individual satisfaction thresholds.

Our interest is to measure the fairness of the allocation of cost penalties that result from implementing the trajectory changes proposed by the STREAM algorithms. Specifically, the purpose of the fairness metric is to measure how fairly cost penalties are distributed among a set of flights assuming that each flight has associated a cost penalty threshold defined by the airline, that is, a maximum cost penalty that the airline considers tolerable for that flight according to its business objectives.

The fairness metric Φ used in this paper is the one defined in [13]. It measures how balanced among the N affected flights is the distribution of the cost penalties implied by the amended trajectories proposed by the TBO model

(STREAM algorithms), considering as a reference the cost penalty threshold for each flight:

$$\Phi = \frac{\left(\prod_{i=1}^{N}\left(1-\wp_i\right)\right)^{1/N}}{\sum_{i=1}^{N}\left(1-\wp_i\right)}\cdot N, \tag{11}$$

where $\wp_i$ is the relative cost penalty for flight i, calculated as

$$\wp_i = \frac{C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i}}{C^*_{\mathrm{thr},i}}, \tag{12}$$

and $C^*_{\mathrm{thr},i}$ is the cost penalty threshold for flight i, that is, the maximum acceptable cost associated with the trajectories amendments required for deconfliction. It is assumed that this maximum additional cost to the UP4DT is considered acceptable by the airline when it is not excessively onerous and only in exchange of higher safety and efficiency. Since the cost penalty threshold is inherently dependent on each airline's business strategy, the following expression is used to approximate it:

$$C^*_{\mathrm{thr},i} = C^*_{\mathrm{up},i} + \Delta m_F^{\mathrm{dc,max}} + \mathrm{CI}\Delta t^{\mathrm{dc,max}}, \tag{13}$$

where $\Delta m_F^{\mathrm{dc,max}}$ and $\Delta t^{\mathrm{dc,max}}$, respectively, correspond to the maximum values of additional fuel burn and to the flight time that the airline considers acceptable for the flight in question. For simplicity, these values are assumed to be equal for all flights: 300 Kgs of total extra fuel and 600 seconds of arrival delay with respect to the UP4DT (note that 300 kgs is about 6% of the average European consumption and 600 seconds is the maximum delay allowed without losing an ATFM slot). Thus, the cost penalty threshold assigned to a flight is a function of the CI assigned to that flight.

Considering the above, since it is assumed that $C^*_{\mathrm{dc},i} \le C^*_{\mathrm{thr},i}$, then $C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i} < C^*_{\mathrm{thr},i}$ and, consequently, $\wp_i$ is always smaller than 1. The fairness metric is therefore defined in the interval $\Phi \in [0,1]$. Note that Φ is indeed a ratio of the geometric mean divided by the arithmetic mean of "user relative satisfaction" (i.e., $1-\wp_i$). The geometric mean penalizes the dispersion lowering the value of the mean and will be only equal to the arithmetic mean when all the values are the same, thus when the dispersion is minimal. In consequence, Φ will be maximum and equal to 1 (highest possible fairness) when the relative cost penalty is the same for all flights, whereas Φ will decrease as soon as the relative cost penalties spread apart from their arithmetic mean (the dispersion, δ, of the relative cost penalties associated with different flights gets larger); that is, $\lim_{\delta\to\infty}\Phi_{\min} = 0$.

The metric of fairness among N flights as defined in [13] is extended in this paper to the one of *fairness among K airspace users*, defined as follows:

$$\Phi_{\mathrm{AU}} = \frac{\left(\prod_{j=1}^{K}\left(1-\wp_{\mathrm{AU}_j}\right)\right)^{1/K}}{\sum_{j=1}^{K}\left(1-\wp_{\mathrm{AU}_j}\right)}\cdot K, \tag{14}$$

where $\wp_{\mathrm{AU}_j}$ is the aggregated relative cost penalty for airspace user AU_j, calculated as follows:

$$\wp_{\mathrm{AU}_j} = \frac{\sum_{i=1}^{N_j} C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i}}{\sum_{i=1}^{N_j} C^*_{\mathrm{thr},i}}, \tag{15}$$

where N_j is the number of flights operated by airspace user AU_j whose trajectories are modified by the algorithms.

This metric has the advantage of measuring the balance of the relative overall cost penalties among different airspace users operating in the scenario, considering only the set of flights interacting in such scenario.

Equity is the metric that evaluates how uniformly the costs of the trajectory changes required by the algorithms are distributed among the N flights whose trajectories are changed by the algorithms. The metric has been defined in [13] as follows:

$$E = \frac{\left(\prod_{i=1}^{N}\left(C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i} + e\right)\right)^{1/N}}{\sum_{i=1}^{N}\left(C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i} + e\right)}\cdot N, \tag{16}$$

where e is a positive real number such that $0 < e \ll \min_{i\in\{1,\dots,N\}} C^*_{\mathrm{up},i}$.

The use of the small positive constant e in the definition of the metric ensures that the equity metric is not zero when $C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i}$ is zero for all flights (in such situation the metric is equal to 1).

The equity metric is defined in the interval $E \in [0,1]$. As in the case of the fairness metric, E is a ratio of the geometric mean divided by the arithmetic mean. Thus, E is maximum and equal to 1 (highest possible equity) when the cost penalties are uniformly distributed across all flights, whereas the value of E decreases as the cost penalties spread apart from their arithmetic mean (i.e., the dispersion, δ, of the cost penalties associated to different flights gets larger); that is, $\lim_{\delta\to\infty} E_{\min} = 0$.

The metric of equity among N flights as defined in [13] is extended to the one of *equity among K airspace users* with at least one flight modified by the algorithms, defined as follows:

$$E_{\mathrm{AU}} = \frac{\left(\prod_{j=1}^{K}\left(\sum_{i=1}^{N_j} C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i} + e\right)\right)^{1/K}}{\sum_{j=1}^{K}\left(\sum_{i=1}^{N_j} C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i} + e\right)}\cdot K, \tag{17}$$

where N_j is the number of flights operated by airspace user AU_j whose trajectories are modified by the algorithms and e is a positive real number such that $0 < e \ll \min_{j\in\{1,\dots,K\}} \sum_{i=1}^{N_j} C^*_{\mathrm{dc},i} - C^*_{\mathrm{up},i}$.

3.2. Simulation and Optimization Tools.

The *performance framework* architecture presented in this paper is depicted in Figure 2. As input, a sample of real traffic data in the form of *flight plans* is used, in representation of the AUs demand characteristics for each particular scenario under analysis. The flight plans are introduced as input to the *Trajectory Synthesis* (TS) module together with airspace information

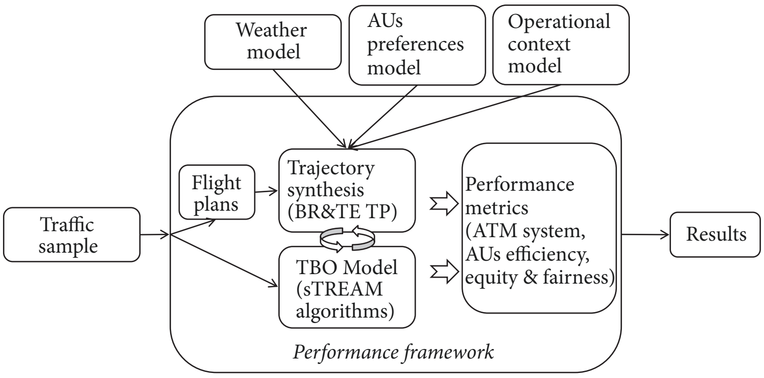


Figure 2: Simulation and optimization tools architecture.

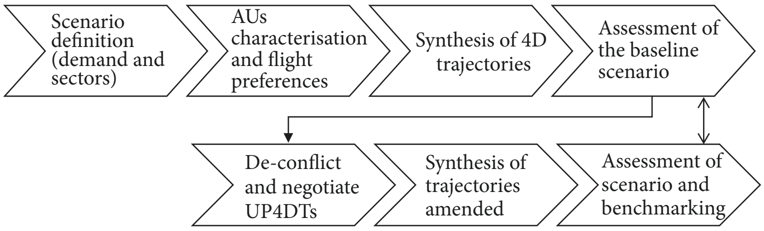


Figure 3: Methodological workflow.

(e.g., waypoints, routes), weather information (e.g., wind, pressure, and temperature), user preferences (e.g., cost index, cruise altitude, and speed), and performance data (e.g., thrust and drag, rate of climb), to generate the full aircraft state trajectories (i.e., a trajectory containing 4D points plus other relevant variables such as mass, fuel rate, thrust, and different speeds). The resulting synthetized trajectories are used by the TBO model, which in this paper is represented by the *STREAM algorithms* (SA) module. The SA module has been used to detect conflicts among those trajectories and to solve those conflicts according to different optimization strategies that emulate a futuristic collaborative flight planning among the AUs and the ANSPs. This process is iterative and requires a continuous communication between the SA module, which updates the flight plan to solve the conflict among a cluster of trajectories, and the TS module, which generates those trajectories according to the updated flight plans coming from the conflict resolution algorithm in the the SA module. Finally, the *performance metrics* module takes all the information coming from both the initial set of trajectories and the deconflicted ones and assesses the metrics dealing with ATM performance, efficiency, and others.

BR&T-Europe AIDL-based trajectory predictor [22, 23] has been used as the core of the TS module, but other alternative trajectory predictors could be used instead [24]. Weather information has been taken from the National Oceanic and Atmospheric Administration (NOAA) [25] and the performance data from BADA 4 [26]. A simplistic model of user preference has been considered for illustration purposes in which the cost index for each flight has been randomized and cruise level and speed extracted as given in the original flight plan.

3.3. Assessment Methodology. Performance assessment can be done through the comparison of metrics measured in a given historical/recorded traffic scenario against a reference baseline scenario measured with the same metric (benchmarking of scenarios) or against a target reference set for that metric (monitoring). In both cases simulations, trajectory and network optimizations, and an assessment framework are needed for the purpose of evaluating and target-setting the performance of the future ATM system. Simulation and optimization tools can recreate the application of future ATM operational concepts and enable the identification of the optimal reference values for each KPA under different assumptions. Benchmarking among different scenarios is also possible to extract knowledge about the complex interdependencies among the different KPAs. The assessment methodology used in this research consists of the seven steps described in Figure 3.

4. Case Study Description and Application of the Assessment Methodology

This section describes the case study that will be analyzed in this paper to exemplify how the metrics, tools, and methodology proposed in this paper can be used in practical applications. The case study is based on an ideal and simplified TBO concept of operations that is applied at ECAC level. UP4DTs are assumed to be direct routes from origin to destination, flying at flight level specified in the original flight plans and at a cruise speed obtained as a function of the CI considered. The introduction of ATM constraints will be analyzed in two different cases: requiring separation among flights at the planning phase (i.e., strategic 4D trajectory deconfliction) of 7 NM and 10 NM, respectively. These two cases will be compared against each other and against the nominal case (the case in which no separation is applied strategically), to show how the KPIs are affected differently in the three cases.

Due to the novelty of the performance framework proposed, the case study analyzed, and the lack of available data (e.g., AUs preferences), some important simplifications and assumptions are done and will be pointed out during the description of the case study. For instance, the TBO model considered (STREAM algorithms) focuses on introducing a dynamic and collaborative flight planning together with new strategic trajectory deconfliction strategies that enables an efficient allocation of 4D flight trajectories while air traffic flow and capacity management (ATFCM) and ATC separation provision constraints are assumed to be always fulfilled (under ideal conditions trajectory planning could be fully coordinated with ANSPs constraints). Therefore, although ATFCM and ATC will still be implemented in the future concept of operations, they are not fully modeled in this paper.

Delivering conflict-free traffic to the ATC services is expected to have a significant positive impact on the ATM performance, by reducing the traffic complexity, the tactical interventions on the trajectories, and in turn the air traffic controller workload. Thus, the strategic deconfliction of 4D trajectories can contribute to increase operational capacity at sectors, while operational efficiency can be enhanced through a collaborative optimization approach in which the airspace users (AUs) preferences and network constraints are adequately negotiated and prioritized based on their nature [1, 2]. In this sense, a more dynamic optimization and allocation of airspace to enable the airspace users to access required airspace with minimum constraints are foreseen. In addition, the anticipation of a number of aircraft separation tasks through strategic trajectory deconfliction may also contribute to improving safety, since a better organization of traffic and the reduction of controllers' separation tasks could potentially reduce the probability of separation infringements and consequently the probability of mid-air collisions.

The predictability of the system and robustness of the conflict-free scenarios to disturbances are key factors in determining the operational feasibility of strategic trajectory deconfliction. It is assumed however that the massive introduction and use of the technologies for precise navigation of aircraft (RNP) and for 4D Flight Management System (FMS), coupled with computationally efficient algorithms for collaborative flight planning and strategic trajectory deconfliction at large scale, will permit achieving the accuracy required to build increasingly reliable plans and to update them quickly when necessary, possibly in real-time.

4.1. Definition of the Scenario: Trajectories and Sectors. The case study is based on real air traffic demand data corresponding to a peak traffic day as provided by Eurocontrol and simulated with a high fidelity air traffic simulator developed by Boeing Research & Technology Europe [23], to obtain the direct routes corresponding to such structured trajectories.

To simplify the analysis only the en-route segments of the trajectories are considered, that is, from Top of Climb to Top of Descent. The resulting scenario included 4010 flights (representative of the two busiest hours of that day) within a spatial region covering most of the European airspace and defined within the latitude interval $[30, 70]$, longitude interval $[-20, 30]$, and all flight levels included between FL130 and FL430.

Figure 4 depicts the resulting traffic demand scenario, composed by the 4010 flights with routes planned according to a theoretical unique free-route airspace at ECAC regional level. (a) shows the original demand of trajectories including potential conflicts (i.e., the nominal scenario), whereas (b) shows the conflict-free scenario after the collaborative 4D trajectory planning process (after considering the AUs preferences and network constraints).

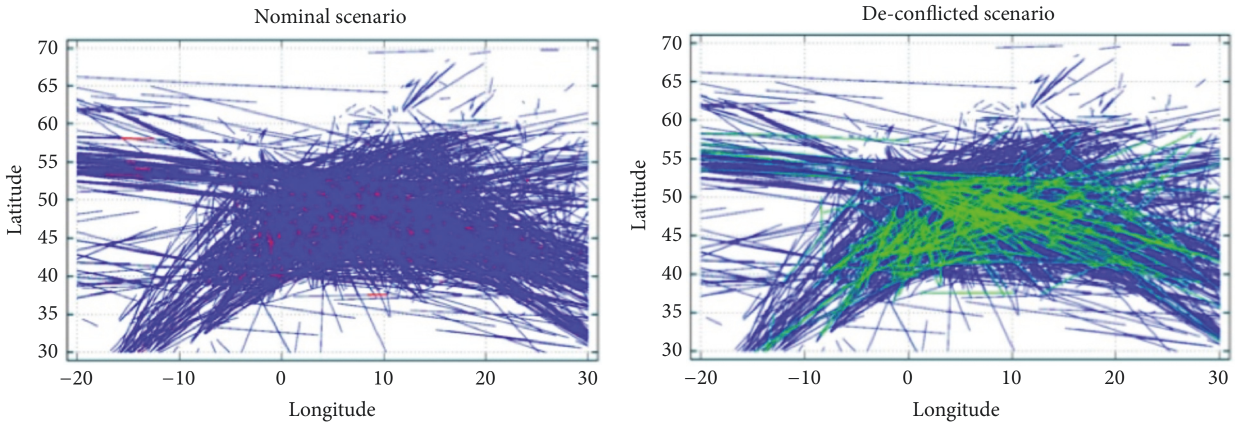
To compare different results of the ATM System Performance metrics defined, a simplified upper airspace sectorisation has been considered, with just 16 sectors covering the entire ECAC airspace, that is, macrosectors comparable to Functional Airspace Blocks. Every sector is assumed to cover the entire airspace in the vertical domain, that is, from 13000 ft to 43000 ft. For simplicity, the sectors are laid out in a grid pattern 4-by-4, with the sector borders following equally spaced geodesic lines of constant latitude and constant longitude. Constant latitude lines are spaced by 10 degrees and constant longitude lines are spaced by 12.5 degrees. Figure 5 shows the resulting geographic distribution of the 16 sectors. Note that the performance framework presented could also be applied to real sector shapes; however, such simplification has been considered useful for illustration purposes in this paper.

The sectors were numbered for their identification sequentially in longitude and latitude, first numbering sectors at the same latitude and then moving to higher latitudes, as shown in Figure 5.

4.2. Characterization of the Airspace Users and Flight Preferences. Each AU that has been considered in the final performance assessment scenario is characterized by a specific cost index and a cost penalty threshold (for all his flights). The cost index and cost penalty threshold were randomly generated in a specific interval, as a simplified way to illustrate heterogeneous AUs preferences, and assigned to each AU. The three intervals considered were $\mathrm{CI} \in [0]$ (all flights with $\mathrm{CI} = 0$, i.e., equivalent to minimizing the fuel consumption), $\mathrm{CI} \in [0, 400]$, and $\mathrm{CI} \in [0, 10000]$. This approach allows us to perform a sensitivity analysis for low, medium, and high values of the CI, to understand how different categories of AUs preferences may impact differently on the ATM performance.

In order to generate a meaningful set of AUs from the 4010 flights considered in the traffic sample, two theoretical cases were generated: (a) each flight belongs to a different AU (i.e., 4010 AUs) and all the airspace users have the same cost penalty threshold; (b) the flights belong to 100 different AUs (with random the distribution of the 4010 flights among the 100 AUs). All the flights belonging to the same AU were set with the same CI (i.e., the AU cost index) and with the same cost penalty threshold.

In total, six different subcases (three CI cases and two possible flight-distribution per case) will be evaluated in terms



(a)

(b)

Figure 4: Nominal scenario of a single free-route airspace in ECAC, including 4010 UP4DTs (in blue) and producing more than 300 potential conflicts colored in red (a) and the same scenario globally deconflicted with 283 trajectory amendments colored in green (b).

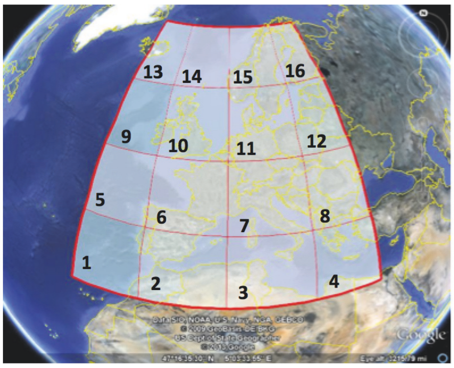


Figure 5: Geographic layout of the 16 sectors of the case study.

of operational efficiency, fairness, and equity assessment. In each simulation run for these subcases, a UP4DT has been defined for each of the 4010 flights in the scenario. The UP4DT is assumed to be the a trajectory planned freely by each AU, following purely its business strategy, and shared with the network manager approximately 3 hours prior to the takeoff, in line with the current practices for flight planning [27].

The definition of the UP4DT for the case study includes the following information: departure time, intended route, intended cruise altitude and cruise speed, initial weight, cost index, and cost penalty threshold. The information required to define the UP4DT was available from the flight plan included in the traffic data provided by Eurocontrol, except for the initial mass, CI, and cost penalty threshold. The initial mass was defined as an average takeoff mass for the aircraft type in question, while the CI and the cost penalty threshold for each flight were obtained as detailed in Section 3.1.3.

4.3. Flight Simulation and Generation of the Baseline Scenario. The trajectories were synthesized from the flight plan information and AUs preferences considered using a BR&TE's high fidelity trajectory simulation infrastructure; thus a realistic trajectory was synthesized for each UP4DT. These synthesized 4D trajectories represent the ideal trajectories that would be flown according to the initially intentions of the AUs and without any ATM intervention during the execution phase, serving as the reference baseline for the final assessment of the STREAM algorithms. The synthesized trajectories have been calculated emulating the trajectory planning process of a Flight Management System (FMS), that is, including detailed aircraft state information (Lat/long position, velocities, weight, etc.) at one-second intervals. The trajectories are synthetized from departure to destination, including the climb and descent phases, although in this case study only the cruise phase is considered for the application of the deconfliction algorithms and performance assessment, as explained in Section 4.1.

4.4. Assessment of the Baseline Scenario. For the airspace sectorisations considered, the metrics for traffic density, peak load, and number of conflicts have been evaluated in each sector based on the trajectories synthesized in previous step. The results determine the performance of the system assuming that no further ATM intervention or constraint imposition occurs. The evaluation of the number of conflicts has been carried out using the STREAM CD algorithms.

Based on the trajectories synthesized in the previous step, the fuel consumed, the flight time, and the total flight cost were calculated for each flight. Three different values of total flight cost were calculated for each simulated flight, one for each of its three possible CI intervals. The values found in the baseline scenario represent an upper limit of the operational efficiency that can be achieved under ideal conditions (minimal ATM constraints) for each of the selected CIs.

4.5. Generation of Conflict-Free Scenarios. The STREAM CD&R algorithms were used to deconflict strategically the 4D trajectories synthesized previously. The result of the CD&R process is a set of amendments to some of the UP4DTs in the sample. For simplification, the types of amendments considered for deconfliction were heading changes or altitude

Table 2: Main results obtained during the CR process with separations of 7 NM and 10 NM.

	7 NM	10 NM
Nominal trajectories	4010	4010
Nominal conflicts	211	211
Total modified trajectories	193	186
HAC maneuvers in solution	190	180
FL changes in solution	3	6

changes. Due to the fuel consumption requirements, heading change maneuvers were prioritized in the strategic deconfliction process, while flight level changes were only applied to some flights when no other solution in the horizontal plane was found. Speed changes, departure time changes, and cooperative maneuvers were not considered.

Due to the presence of inaccuracies in the trajectory execution and traffic synchronization with respect to the plan, the CR algorithm can introduce a buffer to the 5 NM horizontal separation standard, in order to achieve higher levels of robustness in the deconflicted traffic plans. In this paper, due to the large look ahead times of the trajectory amendments, two different buffer parameterization versions of the CR algorithm were applied, thus originating two different sets of experiments: (a) *7 NM version*: the CR is parametrized to produce deconflicted trajectories with a minimum separation of 7 NM (2 NM of buffer); and (b) *10 NM version*: the CR is parametrized to produce deconflicted trajectories with a minimum separation of 10 NM (5 NM of buffer).

The selection of these two buffers (i.e., 2 NM and 5 NM for the 7 NM and 10 NM versions, resp.) represents 40% and 100% of nominal safety distance and their main purpose in this research is to illustrate how different algorithm parameterizations can be assessed with the proposed methodology and how this parametrization could affect the ATM performance in terms of safety, capacity, and airline operational efficiency.

After the CD&R optimization process, in which a collaborative flight planning is emulated, a *globally preferred* conflict-free scenario was selected for each CR configuration, under some simplifying assumptions in which the airspace users have been assumed to express their preferences according to their business targets.

4.6. Trajectory Generation of the Amended UP4DTs. The trajectories of the flights whose UP4DTs were amended as a result of the CD&R computations (193 for the 7 NM scenario and 186 for the 10 NM scenario) were resynthesized using BR&TE's trajectory simulation infrastructure according to the modifications proposed by the CR algorithms (changes in the horizontal route and/or flight level). Note that the number of amended trajectories was lower in the case of 10 NM due to the higher presence of positive "domino effects"; that is, when a trajectory with several conflicts is amended to solve one of them it could lead to the resolution of more than one conflict [5, 28]. Those trajectories were tightly coupled due to peak hour operations. Thus, when some of them were amended with larger horizontal separation buffers, a larger delay was introduced in those trajectories, resulting in a natural decongestion and trajectory decoupling.

At the end of the process, two sets of 4010 deconflicted trajectories were generated, one for each version of the CR algorithms. The trajectory simulations indicated that all the trajectories proposed by the STREAM CD&R system were flyable (according to BADA models for the aircraft types under consideration). Table 2 reports the main statistics for each of the CR parametrization cases considered.

5. Simulation Assessment Results

The performance assessment results of the trajectories that were optimally deconflicted with the STREAM algorithms are structured in this section according to different analysis cases. An analysis case is defined by the version of the algorithms applied (7 NM or 10 NM) together with a set of scenario parameters that are relevant to the performance aspects that are measured. These parameters are the airspace sectorisation (16 Sectors), the number of AUs operating in the scenario (1 or 100), and the specific allocation of CI and cost penalty threshold to the flights in the scenario (CI = 0, CI randomly selected between 0 and 400, or CI randomly selected between 0 and 10000, and the corresponding values of cost penalty threshold per flight). The airspace sectorisation is required to measure the metrics of impact on capacity (given by the combination of traffic density, peak load, and number of conflicts) and the Conflict Resolution Success Rate. The number of AUs is required to measure delay, emissions, the operational efficiency metrics, and fairness and equity among AUs. The CI allocation is required to measure user cost efficiency and all the fairness and equity metrics.

The evaluation of the metrics in each analysis case produced different results, allowing the comparison of performance results between different cases. In summary, the cases considered for the performance assessment in this research are

(a) Two different separation distances required during deconfliction: 7 NM and 10 NM distances

(b) One sectorisation: 16 sectors

(c) Six different airspace users' subcases: three CI distributions allocated (and cost penalty thresholds) combined with two different sizes for the sets of AUs (1 or 100)

5.1. Analysis Cases: 7 NM/16 Sectors and 10 NM/16 Sectors. The performance assessment results show very small differences between the two versions of the algorithms for all the metrics calculated across sectors, as it can be observed in the charts in Figures 6–8.

In both versions of the algorithms, the differences in traffic density and peak load of deconflicted trajectories with respect to the nominal trajectories are small. These results could be expected considering the fact that the airspace considered was divided into a small number of relatively large sectors, and consequently the modified trajectories tend to cross the same sectors as the nominal ones (not many changes are captured across the sectors borders).

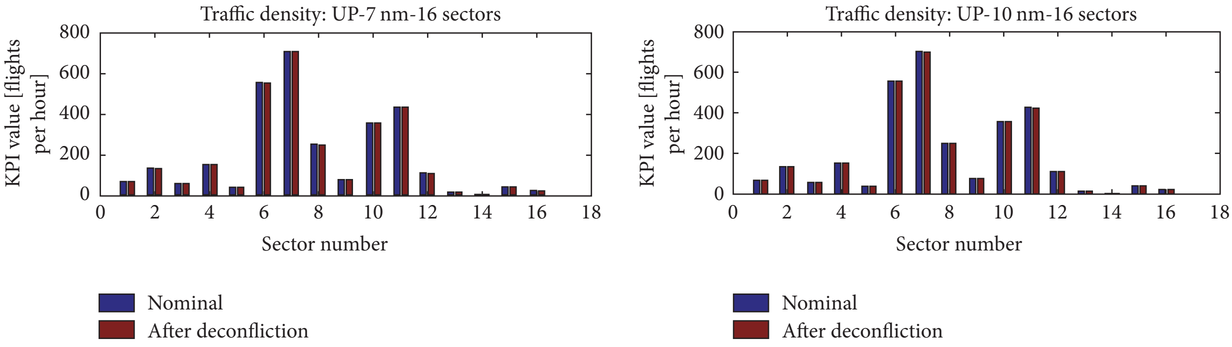


Figure 6: Traffic density per sector in the analysis cases of 7 NM/16 sectors and 10 NM/16 sectors.

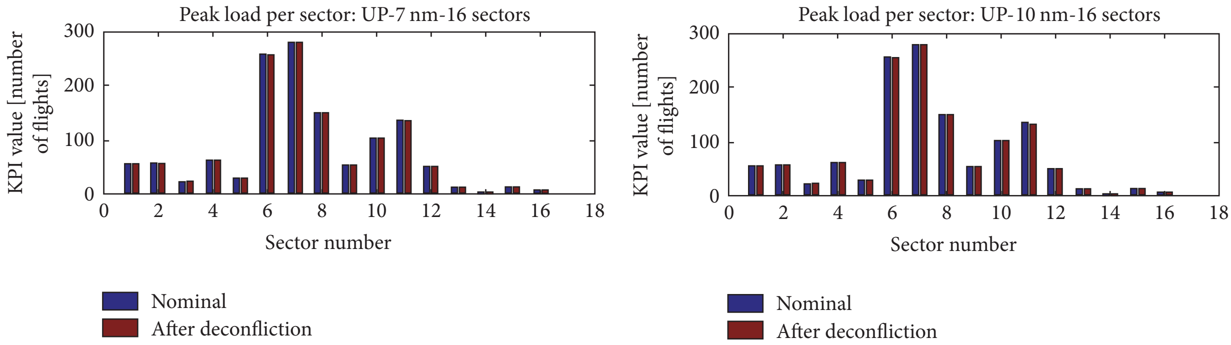


Figure 7: Peak load per sector in the analysis cases of 7 NM/16 sectors and 10 NM/16 sectors.

The impact on capacity suggests that the TBO concept of strategic trajectory deconfliction may potentially reduce the ATC workload and consequently increase the operational capacity in all the sectors, especially in those with higher levels of traffic demand. As explained in Section 3, the higher the values of the components of the metric impact on capacity are, the larger the potential capacity increment is, as a result of the strategic deconfliction process. It can be observed that both versions of the algorithms result in a positive impact on capacity in all the sectors, mainly due to the actual reduction of the number of conflicts (traffic load and peak load values in this metric remain "1" or "close to 1"). Safety is expected to improve as well, since the anticipation of traffic separation tasks, together with the reduction of the ATC officers' workload, may contribute to reduce significantly the probability of mid-air collisions.

Note that some unanticipated conflicts appeared after the high fidelity TP simulations, that is, conflicts among the "true" trajectories flown after the trajectory planning. From a total number of conflicts in the nominal baseline scenario of 211, the 7 NM version resulted in 30 unresolved conflicts and the 10 NM version resulted in 35 unresolved conflicts. If there were no model discrepancies in the planning tools as well as no navigational uncertainties in the trajectory executions, all the conflicts detected and solved at the strategic phase would not reappear during flight execution. Such conflicts appeared due to some degree of uncertainty/discrepancy in the models used by the strategic deconfliction algorithms (e.g., Earth models, flight dynamics, and others), in which some simplifications were implemented to reach a reasonable trade-off between computational efficiency and resolution accuracy. Model discrepancies and other sources of uncertainty are also present in real operational systems, resulting also in traffic desynchronization and conflicts as observed in this simplified case study.

Figure 8 shows that some sectors are more benefitted than others from the capacity gains enabled by the strategic deconfliction strategies (sector structure and traffic patterns have a great influence on the results of this metric). In this case it can be observed that sectors 6 to 12 (the ones that originally presented more traffic density, peak loads, and conflicts) are the ones where higher additional capacity was generated by the strategic deconfliction process, due to a great reduction in the number of conflicts (consequently reducing ATC workload too). Another revealing result is that none of the sectors was impacted negatively on the capacity metric.

It is worth mentioning that since the impact on capacity is highly sensitive to the sector structure, the combination of two TBO concepts, such as strategic deconfliction and dynamic sector configurations (i.e., sectors properly adapted to the actual traffic patterns), might contribute significantly to the improvement of ATM performance in terms of capacity, for example, achieving an even balance of ATC workload across sectors.

5.2. Delay, User Fuel Efficiency, and Emissions Analysis. According to the simulations, the number of trajectories amended through collaborative negotiations would be of 193 and 186 for the 7 NM and 10 NM versions, respectively, which in both cases means that more than 95% of the flights could

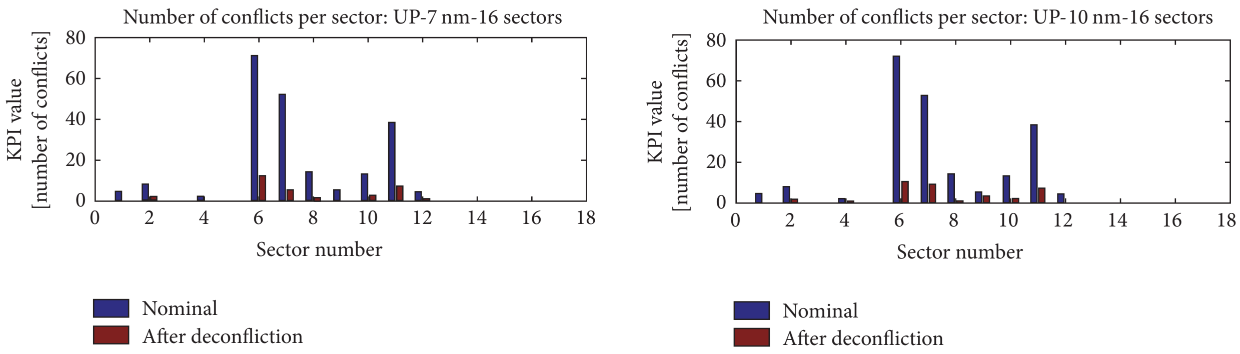


Figure 8: Number of conflicts per sector in the analysis cases of 7 NM/16 sectors and 10 NM/16 sectors.

fly their UP4DT in the en-route phase (simplified as a direct route with constant flight level in this research). For the rest of the traffic (less than 5% of the traffic), average delays close to 4 minutes per amended trajectory with maximum values of 17 minutes have been found.

From an ATM perspective (i.e., considering all the traffic), the metrics of delay (measured taking the entire duration of each flight as the reference), user fuel efficiency, and emissions were evaluated for the following analysis cases:

(a) 7 NM/1-AU/CI = 0 and 10 NM/1-AU/CI = 0, for each of the two versions of the algorithms and considering all aircraft in the scenario belong to the same AU

(b) 7 NM/100-AUs/CI = 0 and 10 NM/100-AUs/CI = 0, for each of the two versions of the algorithm and assuming that all aircraft in the scenario belong to 100 different AUs

The results are independent from the CI assigned to the flights and therefore are valid for the corresponding analysis cases with different CI allocations. The results for total delay, average delay per flight, and emissions remain the same in both cases and are reported in tabular form, while the metrics per AU change and they are reported in graphical form to allow comparison of impacts among the 100 AUs involved.

According to the results shown in Table 3, the impact of the TBO model used (STREAM algorithms) on the traffic demand in terms of delay would be close to 0.3% and in terms of extra fuel consumption and extra emissions would be close to a 0.6%, with respect to the baseline scenario in which all the flights could fly a direct route. A t-test showed that the differences between the cases of 7 NM and 10 NM and the baseline scenario are largely statistically significant in both cases (p values close to 0.0001), whereas the difference between the two is not (p value = 0.1286). According to F-tests, standard deviations can be considered equal for all the cases (p values close to 0.3290). The 10 NM version of the algorithm is observed to result in a slightly smaller average delay and fuel consumption than the 7 NM version. This may seem initially counterintuitive as the 10 NM version tends to produce longer deconflicted trajectories (larger deviations from the original route) in order to achieve larger buffered separations with respect to a predicted conflict. However, looking at the distribution of user fuel efficiency among the different users, it can be observed that there are some negative peaks for the 10 NM case, meaning that some AUs are actually using less fuel with the deconflicted trajectories than with the UP4DTs. Note that the deconflicted trajectories will always be costlier as per the flight's CI, but they may be more fuel-efficient than the UP4DTs. Looking at the distribution of delay and user fuel efficiency across the different users (Figures 9 and 10), it can be observed that both versions of the algorithm resulted in a limited number of AUs (9 to 11) bearing most of the delay and fuel consumption associated with the required deconfliction amendments. However, the specific users affected are different for each version of the algorithm, as the distribution depends on the specific trajectories modified in each case, suggesting that equity and fairness should be taken into account carefully during the application of strategic deconfliction, flight planning amendments, and trajectory negotiations, in order to ensure the balanced distribution of impact over different AUs.

5.3. User Cost Efficiency Analysis. Table 4 reports the average user cost efficiency per flight for each of the two versions of the algorithms and for the three different CI allocations considered (explained in Section 4.2), assuming that all aircraft were operated by the same AU.

Note that for CI = 0 the results are different from the results of user fuel efficiency metric. The reason is that the CI assigned to some flights may be sometimes slightly higher than zero to ensure that the deconflictions do not result in trajectories of lower costs, although they may result in trajectories with lower fuel burnt, as shown in the user fuel efficiency analysis above.

According to Table 4, the impact of STREAM algorithms to the en-route AUs' costs would result in an average increase between 0.5 and 1% in most cases and up to 5% in the most extreme cases, in which the cost of time could have a much higher value than fuel cost. The 10 NM version of the algorithm resulted in approximately 10% reduction in the average user cost efficiency with respect to the 7 NM version, for all CI cases considered. This is likely to be due to the combination of two factors.

On one hand, the 10 NM version results in a lower number of amended trajectories compared to the 7 NM version (due to positive domino effects, as explained in Section 4.6), and therefore there are more aircraft that can fly their minimum cost trajectories.

Table 3: Delay, user fuel efficiency, and emissions for the analysis cases of 7 NM/1-AU/CI = 0 and 10 NM/1-AU/CI = 0.

Analysis Cases	Total delay (seconds)	Average delay (seconds)	Change relative to total trip time in baseline (%)	User fuel efficiency (Kg)	Emissions (Kg pollutant)	Change relative to total user fuel efficiency and emissions in baseline (%)
7 NM, 1 AU, CI = 0	48,036	11.9791	+0.33%	4.8411	CO_2: 15.2448 NO_x: 5.9546 SO_2: 0.0041	+0.63%
10 NM, 1 AU, CI = 0	43,245	10.7843	+0.30%	3.9945	CO_2: 12.5787 NO_x: 4.9132 SO_2: 0.0034	+0.52%

Table 4: Variation in the average user cost efficiency per flight with the version of the algorithm used for different CI allocations.

Analysis cases	User cost efficiency (Kg/flight)	Change relative to baseline
7 NM, 1 AU, CI = 0	5.0137	+0.65%
10 NM, 1 AU, CI = 0	4.4628	+0.58%
7 NM, 1 AU, CI in [0, 400]	6.3387	+0.82%
10 NM, 1 AU, CI in [0, 400]	5.6531	+0.74%
7 NM, 1 AU, CI in [0, 10000]	38.2601	+4.98%
10 NM, 1 AU, CI in [0, 10000]	34.3810	+4.47%

On the other hand, the 10 NM version changed the original flight level of some tightly coupled trajectories in more cases than the 7 NM version. Since the new level may be more fuel-efficient or may reduce the delay with respect to the nominal trajectory, the net result is a lower impact on the overall additional costs associated with conflict resolution with respect to the 7 NM version.

The additional costs of the amended trajectories proposed by the 7 NM version are higher on average due to the difficulty in finding adequate horizontal resolution maneuvers for some airspace regions with many tightly coupled trajectories. For the same cruise altitude and speed, the modified trajectories follow a longer route which in some cases is costlier than applying a flight level change. In addition, a change of flight level in one flight may contribute to reduce the congestion of highly congested airspace altitudes, thus contributing to reduce the cost of amendments for those flights operating at these levels. These complex interactions should be further studied in future research in order to find more efficient strategic deconfliction and traffic planning strategies. In particular, on the light of the remarkable differences found in the performance results of the two versions of the CD&R algorithms, it seems evident that further studies are necessary to understand how the different types of resolution amendments, for example, heading angle changes, flight level changes speed variations [29, 30], and possibly controlled departure delays [31], may impact all the KPAs.

The distribution of user cost efficiency for different users shows that both versions of the algorithm tend to distribute most of the costs among the same number of users (9 to 11 users) in all cases (Figures 11–13), although the specific users affected are different for each version of the algorithm and CI selected. This additionally reinforces the consideration that the different strategic deconfliction measures may have a big impact on the equity and the fairness of the allocation of penalties and further research should be done to explicitly introduce these KPAs in the deconfliction algorithms. For this particular case, it can be noted that the 10 NM version of the algorithm appears to distribute the costs slightly more evenly among the set of users in the scenario (i.e., more peaks with lower cost values instead of fewer peaks with higher cost values).

5.4. Fairness and Equity Analysis among Flights and among Airspace Users. Tables 5 and 6 show, respectively, the fairness and equity among flights and the fairness and equity among airspace users results obtained with each of the two versions of the algorithms, for the cases in which 1 AU or 100 AUs are considered, and with CI allocations to different ranges (i.e., CI = 0, CI in [0, 400] and CI in [0, 10,000]). It can be observed that under all scenarios the values of the fairness KPI are almost one, reflecting the fact that the CD&R does not discriminate AUs much in terms of relative penalty allocation. Equity metrics on the other hand show that costs have not been distributed evenly among the flights and/or AUs (only 5% of the flights were amended, thus leading to an inequitable distribution of amendments). T-tests confirmed that fairness is equal in all the scenarios (p values close to 0.01) while in terms of equity the 10 NM scenarios are systematically more equitable than the 7 NM scenarios (p values close to 0.17 in all the cases). With the purpose to illustrate a benchmarking analysis between the two cases, the values of the metrics have been detailed up to six decimals in order to highlight the—small—differences in fairness and equity resulting from the application of the two algorithm versions in the same context.

The 10 NM version of the algorithm is observed to consistently result in a fairer and more equitable distribution than the 7 NM version, regarding how the resolution costs are allocated among the flights whose trajectory is amended

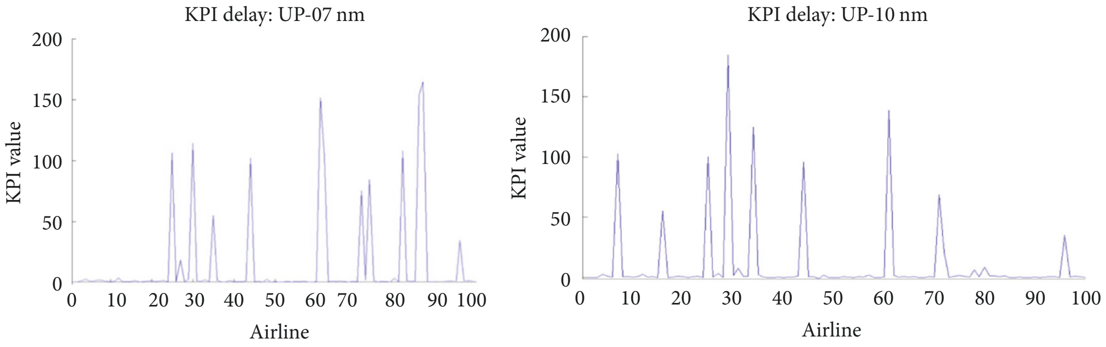


Figure 9: Average delay (in seconds) per airspace user for the analysis cases of 7 NM/100-AUs and 10 NM/100-AUs.

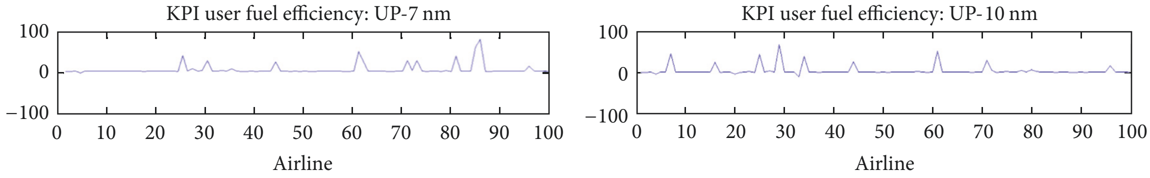


Figure 10: User fuel efficiency (Kg/flight) per airspace user for the analysis cases of 7 NM/100-AUs and 10 NM/100-AUs.

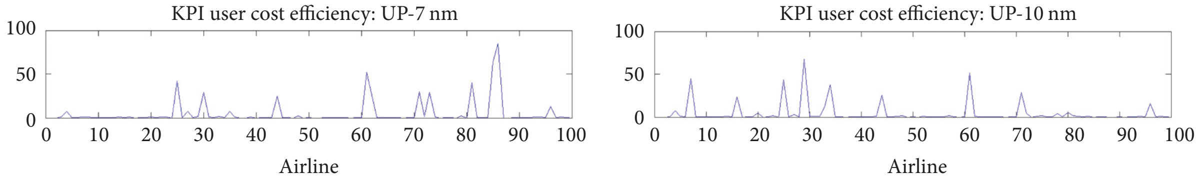


Figure 11: User cost efficiency (Kg/flight) in analysis cases 7 NM/100-AUs/CI = 0 versus 10 NM/100-AUs/CI = 0.

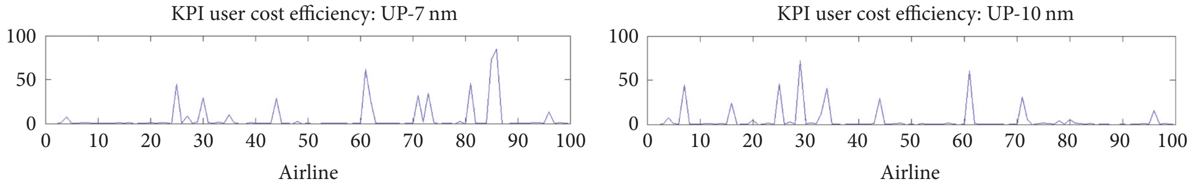


Figure 12: User cost efficiency (Kg/flight) in analysis cases 7 NM/100-AUs/CI in [0, 400] versus 10 NM/100-AUs/CI in [0, 400].

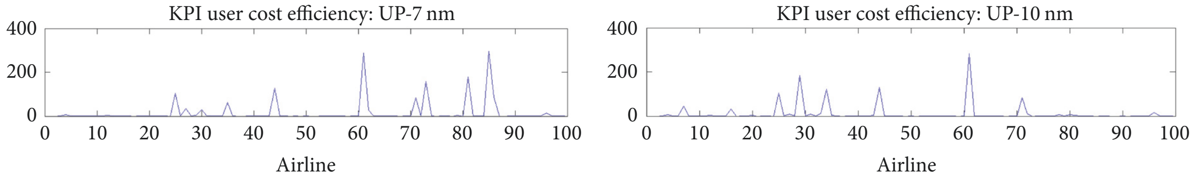


Figure 13: User cost efficiency (Kg/flight) in analysis cases 7 NM/100-AUs/CI in [0, 10000] versus 10 NM/100-AUs/CI in [0, 10000].

Table 5: Variation in fairness and equity among flights with different versions of the algorithm and different CI allocations.

Analysis cases	Fairness among flights (nondimensional)	Equity among Flights (nondimensional)
7 NM, 1 AU, CI = 0	0.9958	0.077
10 NM, 1 AU, CI = 0	0.9969	0.117
7 NM, 1 AU, CI in [0, 400]	0.9949	0.065
10 NM, 1 AU, CI in [0, 400]	0.9962	0.099
7 NM, 1 AU, CI in [0, 10000]	0.9868	0.036
10 NM, 1 AU, CI in [0, 10000]	0.9900	0.060

Table 6: Variation in fairness and equity among airspace users with different versions of the algorithm and different CI allocations.

Analysis cases	Fairness among airspace users (nondimensional)	Equity among airspace users (nondimensional)
7 NM, 100 AUs, CI = 0	0.999997	0.100
10 NM, 100 AUs, CI = 0	0.999998	0.136
7 NM, 100 AUs, CI in [0, 400]	0.999996	0.093
10 NM, 100 AUs, CI in [0, 400]	0.999997	0.132
7 NM, 100 AUs, CI in [0, 10000]	0.999994	0.054
10 NM, 100 AUs, CI in [0, 10000]	0.999995	0.086

and also among the AUs for whom at least one flight has been modified by the algorithms. This was expected from the user cost efficiency results above.

The equity increases as far as the number of AUs decreases, which is a direct reflection of the more even distribution of the costs among the AUs, also observed in the user cost efficiency results. The fairness results show the same tendency, although the differences between the two versions of the algorithms are smaller. This is due to the effect of the cost penalty threshold on the metric, implying that the same additional cost is impacting differently each flight and/or AU. The same distribution of costs may be more or less fair depending on the cost penalty threshold of the different flights and AUs. For both the 7 NM and 10 NM cases, it can be stated that the solutions found are not very equitable (only a few flights are amended), but they are highly fair (small relative negative impact to the operational efficiency for all the AUs).

6. Conclusions and Future Work

This paper has shown the capabilities of a new assessment methodology and simulation infrastructure with the potential to facilitate the early evaluation of new trajectory-based concepts and their potential impact on ATM performance and airline operational efficiency. The performance assessment framework relies on a high fidelity trajectory predictor, a strategic deconfliction of 4D trajectories with the capability of collaborative traffic optimization, and a set of metrics to be applied to the dynamic simulations. The ability to model trajectories with a high degree of fidelity to evaluate metrics such as fuel consumption, environmental impact, or impact on AUs operating costs can provide valuable insights about the operational impact of new concepts before conducting other more expensive and complex validation exercises such as real-time simulations. The evaluation of the metrics for different realistic scenarios, together with a deeper understanding of the complex interactions among them, may also help to define the requirements that need to be fulfilled to ensure a certain level of ATM performance in the future.

The simulation and assessment of a case study using a strategic and collaborative 4D trajectory deconfliction method developed in previous research (i.e., the STREAM algorithms) indicate that, under certain assumptions in line with future ATM concept definitions, the anticipation of traffic separation task through strategic deconfliction strategies could have an important and positive impact on all the KPAs. In particular on the metrics measuring ATM system performance (capacity, safety, and environment), the airspace users' operational efficiency, and equity/fairness. Under certain ideal conditions, the strategic deconfliction strategies have reduced notably the number of conflicts in the en-route airspace, thus having a positive impact on safety (lower probability of conflicts implying less separation infringements and consequently less potential mid-air collisions) and on the ATM en-route capacity (assuming a reduction of the air traffic controller's workload, due to less conflict management tasks). No negative impact on the capacity of any en-route sector has been identified in the case study.

The assessment made in this research has also highlighted the importance of understanding the complex interactions among aircraft trajectories, in order to identify at any moment the most convenient ATM constraints to traffic that optimize the whole ATM system performance. In particular, it has been observed that a more advanced design and management of airspace (e.g., dynamic sectorisation) and trajectory organization (e.g., optimizing the use of level changes to resolve aircraft conflicts) may have a large positive impact on performance. Therefore, further studies shall be conducted to explore and understand the trade-offs between the different combinations of trajectory planning resolution amendments (considering heading angle changes, flight level changes, speed variations, and possibly controlled departure delays), as well as the airspace design and configuration that should be given at each period. In addition, equity and fairness metrics indicated how sensitive the distribution of costs might be to any change applied to the strategic trajectory deconfliction or to different airspace users' preferences. Any change of these parameters may have a strong impact on the equity and fairness metrics. Future research shall therefore look at equity and fairness metrics during the traffic planning processes (e.g., computing the metrics during the process) as well as during postoperations analyses.

The above results could be questionable in the presence of uncertainties such as wind that may deviate flights from their 4D expected trajectories or weather events such as thunderstorms that may need to have more advanced collaborative planning tool. Future work will include random perturbations to the *"true"* trajectories during execution aimed at modeling errors (e.g., wind prediction errors) within the predicted trajectories used by the algorithms for planning purposes. A robustness analysis will be conducted with the performance framework proposed in this paper, to identify different strategic deconfliction/traffic planning strategies that can provide more stable and robust traffic plans.

Disclosure

The opinions expressed herein are those of the paper's authors and do not represent the official position either of the SESAR Joint Undertaking or of Eurocontrol.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The STREAM project has been funded by Eurocontrol acting on behalf of the SESAR Joint Undertaking with funds from the European Union.

References

[1] SESAR Consortium, *European ATM Master Plan*, 2nd edition, 2012.

[2] SESAR Consortium, "SESAR Concept of Operations Step 1, Edition 01.00.00," 2012.

[3] CAO ATMRPP, "Appendix A to TBO Concept Document (WP652)," 2015.

[4] ICAO., "Global Air Traffic Management Operational Concept (Doc 9854)," 2005.

[5] S. Ruiz, *Strategic Trajectory De-confliction to Enable Seamless Aircraft Conflict Management [PhD Thesis]*, Universidad Autónoma de Barcelona, 2013.

[6] S. Ruiz, M. A. Piera, J. Nosedal, and A. Ranieri, "Strategic deconfliction in the presence of a large number of 4D trajectories using a causal modeling approach," *Transportation Research Part C: Emerging Technologies*, vol. 39, pp. 129–147, 2014.

[7] EUROCONTROL, "European Route Network Improvement Plan (ERNIP) - Monitoring Report - AIRAC 1403," 2014.

[8] European Commission, "laying down a performance scheme for air navigation services and network functions," Commission Implementing Regulation (EU) 390/2013, 2013.

[9] L. Asante and F. Nieto, "Complexity in the optimization of ATM performance metrics," in *Proceedings of ATACCS 2012*, London, UK, 2012.

[10] G. B. Chatterji, "Fuel burn estimation using real track data," in *Proceedings of the 11th AIAA ATIO Conference, AIAA Centennial of Naval Aviation Forum.*, 2012.

[11] J. López-Leonés, M. Polaina, P. Sánchez et al., "User-centric cost-based flight efficiency and equity indicator," in *Proceedings of the ESAR Innovation Days*, 2017.

[12] C. Barnhart, D. Bertsimas, C. Caramanis, and D. Fearing, "Equitable and efficient coordination in traffic flow management," *Transportation Science*, vol. 46, no. 2, pp. 262–280, 2012.

[13] I. D. Poza, M. Vilaplana, and C. Goodchild, "Assessing fairness and equity in trajectory-based operations," in *Proceedings of the 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO), Aviation Technology, Integration, and Operations (ATIO) Conferences*, American Institute of Aeronautics and Astronautics, September 2009.

[14] J. D. Welch, J. W. Andrews, B. D. Martin, and E. M. Shank, *Macroscopic Workload Model for Estimating En Route Sector Capacity*, ATM Seminar, Barcelona, Spain, 2007.

[15] R. A. Paielli, "A linear altitude rule for safer and more efficient enroute air traffic," *Air Traffic Control Quarterly*, vol. 8, no. 3, pp. 195–221, 2000.

[16] D. K. Schmidt, "On modeling ATC work load and sector capacity," *Journal of Aircraft*, vol. 13, no. 7, pp. 531–537, 1976.

[17] R. Christien, A. Benkouar, T. Chaboud, and P. Loubieres, "Air traffic complexity indicators & ATC sectors classification," in *Proceedings of the 21st Digital Avionics Systems Conference.*, pp. 2D3-1–2D3-7, Irvine, CA, USA, 2002.

[18] J. Djokic, B. Lorenz, and H. Fricke, "Air traffic control complexity as workload driver," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 6, pp. 930–936, 2010.

[19] A. Majumdar and W. Y. Ochieng, "Factors affecting air traffic controller workload: Multivariate analysis based on simulation modeling of controller workload," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1788, pp. 58–69, 2002.

[20] ICAO, "International standards and recommended practices, Environmental protection," in *Aircraft Engine Emissions*, vol. 2 of *plus amendments 3, 4, 5, 6 and 7*, 3rd edition, 2008.

[21] EUROCONTROL, "Standard inputs for EUROCONTROL cost benefit analysis," 2009.

[22] J. López-Leonés, *Definition of an aircraft intent description language for air traffic management applications [M.Sc. thesis]*, Department of Aerospace Engineering, University of Glasgow, Glasgow, UK, 2008.

[23] J. A. Besada, G. Frontera, J. Crespo, E. Casado, and J. Lopez-Leones, "Automated aircraft trajectory prediction based on formal intent-related language processing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1067–1082, 2013.

[24] S. Ruiz and M. Soler, "Conflict pattern analysis under the consideration of optimal trajectories in the European ATM," in *Proceedings of the 11th USA/Europe Air Traffic Management Research and Development Seminar, ATM 2015*, ATM Seminar, Lisbon, Portugal, June 2015.

[25] National Weather Service, "The Global Forecast System (GFS) - Global Spectral Model (GSM), version 11.0.6," 2003.

[26] EUROCONTROL, "Base of Aircraft Data (BADA)," https://www.eurocontrol.int/services/bada, 2017.

[27] EUROCONTROL, "ATFCM users manual," 2015.

[28] K. Bilimoria and H. Q. Lee, "Properties of air traffic conflicts for free and structured routing," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit 2001*, August 2001.

[29] K. D. Bilimoria, "A geometric optimization approach to aircraft conflict resolution," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2000.

[30] H. Erzberger, "Automated conflict resolution for air traffic control," in *Proceedings of the 25th International Congress of the Aeronautical Science*, Presented at the 25th International Congress of the Aeronautical Science, Hamburg, Germany, September 2006.

[31] J. Nosedal, M. A. Piera, S. Ruiz, and A. Nosedal, "An efficient algorithm for smoothing airspace congestion by fine-tuning take-off times," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 171–184, 2014.


Journal of Engineering
The Scientific World Journal
International Journal of Rotating Machinery
Journal of Sensors
Advances in Multimedia
Advances in Civil Engineering
Journal of Control Science and Engineering
Journal of Robotics
Journal of Electrical and Computer Engineering
Advances in OptoElectronics
VLSI Design
International Journal of Navigation and Observation
Modelling & Simulation in Engineering
International Journal of Aerospace Engineering
International Journal of Chemical Engineering
International Journal of Antennas and Propagation
Active and Passive Electronic Components
Shock and Vibration
Advances in Acoustics and Vibration