Journal of Air Transport Management 113 (2023) 102471

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Impact of cost approximation on the efficiency of collaborative regulation resolution mechanisms

Gérald Gurtner, Tatjana Bolić *

University of Westminster, 35 Marylebone Road, London, NW1 5LS, United Kingdom

ARTICLE INFO

Keywords:
Cost of delay
User-Driven Prioritisation Process
Air traffic flow management
ATFM delay
ATFM slots
Agent-based model

ABSTRACT

Air Traffic Flow Management (ATFM) measures are often used to alleviate capacity demand imbalances. Usually, these measures impose a departure delay on flights crossing the airspace in question, where delay is assigned on a "First Planned First Served" (FPFS) base, which minimises total delay. Since delays typically have a different impact on individual flights in terms of cost, a procedure based on cost minimisation could reduce delay-related costs. User-Driven Prioritisation Process (UDPP) is developing a set of solutions aimed at allowing airlines to rearrange their flights within their own slots assigned by the FPFS rule. Inter-airline cost reducing approaches are still missing, but several works have been launched in this direction, using either central optimisations or market-based mechanisms.

We analyse the impact of cost approximations procedures, integral part of certain inter-airline mechanisms, overall likely to be used by airlines or the Network Manager (that manages ATFM measures). Using cost models and simulations to collect the true cost of delay profiles, we show that the impact of cost approximations have been severely underestimated, leaving little room for new mechanisms to improve over UDPP. Moreover, we show that the errors made by the airlines on their own costs, expected at least for some airlines, further deteriorate the situation, including UDPP. However, we find that approximation procedures create a strong resilience to these errors, showing how both UDPP and inter-airline procedures may benefit from not having the airlines communicate their detailed costs. Thus, we find that any design of new mechanisms could include a cost approximation procedure in order to increase its resilience.

1. Introduction

The air traffic is picking up after the reduction due to COVID-19 pandemic, with the summer 2022 traffic being just about 13% less than in 2019 (Performance Review Unit, 2022). Based on the latest EUROCONTROL's forecast (STATFOR, 2022) it is likely that the traffic levels will come back to the 2019 levels by 2025. Even though the air traffic is still lower than before the pandemic onset, the delays have returned to the 2019 levels (Performance Review Unit, 2022), signalling the continuing demand-capacity disbalance in European network. When the disbalance is foreseen, the flow manager of that ANSP proposes and agrees the activation of the ATFM regulation with the Network Manager (NM). NM then promulgates the information about the regulation to all interested parties (e.g., airlines, other ANSPs the affected flights need to cross), and assigns the ATFM slots to the flights in the regulation. In practice, NM assigns delays to flights, based on the First Planned First Served principle, in order to smooth the demand. The ATFM slot is assigned to flights about three hours before departure, and has duration of 15 min (allowing for the actual take-off to be in the interval of five minutes before new take-off time and 10 min after).

Airspace Users (AUs)[1] often need further flexibility for their operations in such an environment. On the one hand, due to unforeseen events (e.g. ATFM regulation delays), they sometimes need to reconfigure their operations, cancelling a flight, swapping or reordering several flights, etc., in order to reduce the impact of these events on their operational costs. On the other hand, the Air Navigation Service Providers (ANSPs), airports, and thus the network, need predictability and reliability to deliver the required capacity in an efficient manner. Thus, there is a need for mechanisms creating flexibility for AUs while ensuring an adequate level of predictability for the network. The User-Driven Prioritisation Process (UDPP) is a project that collects several algorithms and tools of varying maturity, designed to enable the AUs to prioritise their flights as a function of their business needs, and in order to reduce the effects of ATFM regulations on key flights, all in coordination with the NM. UDPP enhances the possibility to swap

* Corresponding author.
E-mail addresses: G.Gurtner@westminster.ac.uk (G. Gurtner), t.bolic@westminster.ac.uk (T. Bolić).

[1] Note that we use "AUs" and "airlines" interchangeably in the text.

https://doi.org/10.1016/j.jairtraman.2023.102471
Received 5 April 2023; Received in revised form 26 July 2023; Accepted 30 August 2023
Available online 14 September 2023

slots[2] of airline's own regulated flights, increases departure flexibility at airports, and elaborates prioritisation mechanisms for airlines to reorder their own flights in a regulation. The UDPP so far created one SESAR Solution[3] (SESAR Joint Undertaking, 2021), that is implemented at the following airports: Paris Charles de Gaulle, Dresden, Frankfurt, Hamburg, Munich, Nurnberg, Stuttgart, and Berlin-Brandenburg airports (SESAR Joint Undertaking, 2021). Other UDPP Solutions, comprising more complex algorithms are still under development (SESAR Joint Undertaking, 2021).

The UDPP development provides effective cost efficiency improvements to airlines, as shown in Pilon et al. (2019), based mainly on rearranging own flights. However, UDPP, an intra-airline process, has limitations. For example, low-volume users – airlines with low number of flights in the regulation – can hardly benefit from these processes as the low number of flights usually precludes the swaps. Further, it is currently not possible to swap the slots between different airlines,[4] which could enable the low-volume users to participate in the process, and offer additional possibilities to other AUs.

In this work we explore certain characteristics of new slot allocation mechanisms that are alternative to, or an extension of the current UDPP. The mechanisms in question are described briefly in Section 3, their main characteristics being the possibility of slot swapping *between airlines*, which is facilitated by AU's sharing their input "preferences" with a central optimiser, the most basic form of which is expressing preferences for each slot and each flight. Here, we consider the "preferences" intrinsically linked to the cost of delay incurred when a regulation is solved (i.e., after the initial ATFM slot distribution, with corresponding delay).

The slot swapping mechanisms under investigation here provide a process by which the AUs may communicate their costs for each flight, so that the central optimiser (which may be thought of as being a part of the Network Manager) finds the allocation between slots and flights that lowers the costs for all the users, as costs do not depend only on the amount of delay but also on what the delay implies in terms of operations (e.g. stranded passengers, curfew violation, or just delayed operations). From the optimisation point of view, UDPP can be thought of as a local optimisation, where airlines lower their costs by swapping their own flights within their own initially assigned slots. The initial ATFM slot assignment is based on First Planned First Served (FPFS) allocation, which is usually applied in practice today. Further, in this paper we assume that airlines know their own true costs to some extent, the pros and cons of which are discussed in conclusions (Section 7).

Here we explore the following setup – a central optimiser that receives the AUs' preferences, which poses two legitimate questions:

1. How to convey an equivalent of the cost function to the central optimiser using preferences?
2. How to ensure that airlines communicate their true costs?

The second question tackles the well-known problem of gaming (von Neumann and Morgenstern, 1992). If a mechanism is put in place where airlines declare their own costs — or an equivalent, there is no *a priori* mechanism ensuring the true costs will be communicated. With a trivial example of an absolute weighting mechanism, where airlines would assign a weight to each flight according to its importance, and where the weights would be compared across airlines without any renormalisation process, airlines could put very high weights to have a better allocation for their flights. More generally, in the context of cost declaration, they could be tempted to inflate their cost for the same reason. This is a known issue, that we do not analyse per se in this paper, but we point out the situations where it can play a role.

In this work we explore instead the first question, i.e., *how to convey the cost functions using preferences*. Even if a given mechanism has enough incentives for the airlines to share their true costs, it may be unrealistic to expect the airlines would communicate their full cost functions, mainly for practical reasons. Hence, inter-airlines slot swapping mechanisms using a central optimiser will likely use approximation processes at their heart.

The three mechanisms used in this paper are set up to have AUs communicate a small number of parameters to the central optimiser, representing (hopefully) a decent approximation of their costs. This in turn raises the question of the *quality of the approximation* and its *impact on the efficiency of the mechanism*. The objective of this paper is to explore these questions, and more specifically:

- explore the impact of approximation and errors on the cost of delay of flights involved in ATFM regulations;
- compare benchmarks with and without cost approximation processes, and draw general conclusions about the possible further improvements over intra-airline optimisation processes;
- show how errors from airline and approximation processes interplay and impact the efficiency.

The article is structured in the following way. First, the literature review on the issues around UDPP and cost of delay estimation is presented. Then the mechanisms explored (Section 3) and the indicators used to assess the impact of the mechanisms on the airlines are described. Section 4 presents the methodology applied in the analysis through the presentation of the agent-based model (ABM), the data used for setup and estimations, and the approximation archetypes used for the cost functions. In Section 5 the results in terms of the impact of the approximation process on the mechanisms and the role of errors in cost functions are presented. In Section 6 we discuss the implication of the findings for the future design of inter-airline mechanisms. The last section concludes by listing the potential lines of research left in this area.

2. Literature review

This paper explores a specific topic of reducing costs of delay within a wider area of the ATFM regulations. The ATFM regulations are used to enable demand-capacity balancing (DCB) in European network. Different DCB aspects have been addressed in literature. For example, reconciliation of AU flexibility in flight planning and the need for demand predictability by ANSPs, through the application of peak-load route charges (Bolić et al., 2017) that re-distribute planned traffic so as to respect available network capacities, while minimising the costs. Further, Castelli et al. (2011) designed market mechanism for ATFM slot allocation, while Ivanov et al. (2019) attempt to re-design European Air Traffic Management to allow Network Manager to coordinate capacity and demand management decisions, using economic instruments for both supply and demand sides. These works attempt to mitigate the DCB problems in the entire network, while we are addressing the improvements to the mechanism of the slot assignment within an ATFM regulation.

Already mentioned UDPP is under development as a response to this particular need for collaboration in ATFM regulation resolution mechanisms. The default solution applied in Europe is known as First Plan First Served (FPFS), implemented through the CASA algorithm (Tibichte and Dalichampt, 1997), and aims at minimising the total delay incurred by the flights in regulations. Even though the management of ATFM regulations is based on the simple FPFS, it offers flexibility to AUs in terms of re-routing, and sometimes flight swapping (EUROCONTROL et al., 2022). Even though the flight swapping

[2] Note that slots we are referring to are ATFM regulation slots, that are assigned to flights on the day of operations, not strategic airport slots.

[3] SESAR Solutions are new operational concepts, procedures and/or relevant technologies.

[4] Although airlines can in principle do it via manual process, first calling each other to agree on which slots to swap, and then inform the NM to check for network constraints.

capabilities exist, the UDPP solutions would allow the AUs to swap their own flights[5] through a series of innovations: flight departure reordering, selective flight protection (SFP), and flight margin Pilon et al. (2019, 2016). Flight departure reordering enables AUs to reorder their own flights within the regulation, as long as the new time allocated to a flight is not earlier than its expected arrival time (ETA). With SFP, an airline can request a new time allocation for a (specific) flight, where the requested time is outside their assigned ATFM slots, but not before the flights's ETA. Through the flight margin, an airline can indicate the time after which a particular flight would experience significant cost increase. The intent is to assign the slot for that flight to a time period before indicated margin. These innovations allow AUs to reduce their operational costs caused by the imposed ATFM delays, without explicitly sharing (and thus revealing) their costs, as the costs are considered confidential. As such, the UDPP is intended for intra-airline slot swaps, and can provide limited to no benefits to the low-volume users. Low-volume users could benefit from slot swaps between different airlines that we term inter-airline swaps.

Some attempts have been made at defining inter-airline swap mechanisms. A credit mechanism ("Flexible credits") has been proposed by Ruiz et al. (2019), based on a simple "delay vs. credit" scheme. Another study investigated a set of algorithms that included credits for low-volume users and auctions (Mocholí et al., 2020), adding behavioural economics principles to the analysis. Behavioural economics advances the quality and rigour of simulation models by including essential understanding of human behaviour and decision-making drawing from several disciplines (psychology, neuroscience, economics and decision science), with the goal to improve the classical economic modelling approaches (Angner, 2016). We are not using behavioural economics in this work, but we will mention it where we believe it could have an impact in the overall analysis.

When discussing improvements of the ATFM regulation resolutions, be those intra or inter-airline, it is important to define which characteristic or characteristics should be taken into account. So far the amount of delay and the cost of delay have been mentioned in this paper. The current slot assignment in the ATFM regulations is based on delay minimisation, as this is both easily tracked, and accepted as a fair and equitable approach by AUs. However, the same amount of delay can imply widely different operational costs. For example, a half an hour delay of the last flight of the day most probably causes just passenger inconvenience. Half an hour delay causing a group of connecting passengers to miss their next flight will likely have much larger operational costs, as those passengers need to be provided for. This can include re-booking on another flight, care (meals and refreshments), accommodation, and compensation (as mandated by Regulation 261 European Commission, 2004 in Europe. Regulation 261 describes all the provisions). For these reasons, the processes like UDPP are being considered, to offer an additional degree of freedom to airlines to minimise not only delay, but also their operational costs stemming from delay.

The costs of delay are an important part of operations as the airlines need to account for them twice. First, when planning schedules (in a strategic phase) buffers are usually included, to absorb as much delay as possible. Next, accounting for actual delays (i.e. tactical) on the day of operations (Cook, 2016). Even if the costs of delay represent an important part of the operations, their calculation is not straightforward, nor immediate. Cost of delay models have been developed in 2011 by Cook and Tanner (2011), updated in 2014 (Cook and Tanner, 2015), and again in 2020 (Cook et al., 2021). They originated from an observation on the lack of consensus over the true cost of delays in Europe. This line of work has been widely used in Europe by different stakeholders and regulatory bodies, including the Performance Review Body, and helped the transition from a delay-driven paradigm to the cost-driven one in Europe. Evler et al. (2020) worked on the integration of uncertainty in the cost functions, the tactical and strategic effects of which has been studied in Gurtner and Cook (2021), leading to the definition of stochastic cost functions (which are not used in this study).

In order to use the costs of delay in the ATFM regulation resolutions, AUs need to know those costs. Depending on the AU business model, their dispatch unit can be better or worse equipped to calculate these costs. Often where the calculation is not possible, some type of approximation is used, like striving to maintain on-time performance (Gurtner and Bolić, 2021). The difficulty of assessing the costs tactically, and the need for confidentiality are usually the reasons behind the use of cost approximation functions for ATFM regulation resolutions. This in turn poses the questions we are addressing: *the quality of approximation* and its *impact on the efficiency of the mechanism.* The assessment regarding these questions even for a single AU might be difficult. The UDPP processes, as developed over several years, mainly aim at *helping* the AUs to assess their costs, essentially using the equivalent of step-wise cost functions by setting priorities, 'time-not-after', margins, etc.

In order to include different AUs, under different conditions, we turn to the agent-based model (ABM) simulations. The ABM allows to simulate different behaviours for the airlines, since one can implement arbitrary complex decision-making processes in these models (Bonabeau, 2002) for each entity modelled (i.e., agent). The ABMs are particularly well suited when the number of agents is high and are typically used to simulate non-cooperative and cooperative economic games (Mueller and Pyka, 2016). This work uses the Mercury ABM model (Gurtner et al., 2021), as it can model a high number of agents and cost of delay models are explicitly implemented in it.

To be able to perform a proper assessment of a certain proposed solution, a benchmark to which to compare the results is needed. As such, the simplified version of current process (EUROCONTROL et al., 2022) is represented by the FPFS mechanism. Apart from being simple, the current process of ATFM slot allocation is very dynamic and flexible. The last two characteristics would introduce unnecessary complexity in this particular analysis, which induces us to use the simple, first step of the current process, FPFS. The second benchmark to investigate the efficiency of inter-airline mechanism is the intra-airline one – UDPP (detailed explanation of the mechanisms used in the article can be found in Section 3).

The assessment of ATFM regulations resolution mechanisms can be more or less comprehensive, using one or more key performance indicators (KPIs). The following KPIs are usually mentioned when SESAR Solutions for AUs are assessed: predictability and punctuality, flexibility, access and equity, cost efficiency, and robustness (SESAR Joint Undertaking, 2020). Owing to the formulation of the research question, only the cost efficiency is taken into account here. The efficiency indicator used is described in Section 4.3.

3. Inter-airline allocation mechanisms

The basic design principles needed for inter-airline ATFM slot swapping are described here, as well as the mechanisms used in this article. We use a reference, using the First Planned First Served rule, simulating the current situation in Europe. We also use a baseline, UDPP, representing the next improvement that will be deployed in Europe and two benchmarks, MINCOST and NNBOUND, used to estimate the maximum possible efficiency of slot-swapping mechanisms (with a constraint to equity in NNBOUND).

3.1. General design principles

While UDPP allows only intra-airline slot swapping (see Section 3.2.2), the mechanisms that allow inter-airlines slot swapping are investigated here. We consider a situation where when a regulation gets activated, N flights are affected. M slots are created to spread out the incoming traffic. We consider here the case where $M \geq N$, i.e., a

[5] That we term intra-airline swaps.

sufficient number of slots are created to accommodate all flights, but some slots may be empty due to flights' arrival times.

A slot allocation mechanism (or mechanism in short) is defined as the process where *every* flight i of the N flights in the flight set is allocated to *one and only one slot* k among the M slots available, The allocation itself is the mapping μ between the flight set and the slot, i.e. $\mu(i)$ is the slot allocated to flight i. There is a maximum of $M!/(M-N)!$ allocations possible, not taking into account constraints on arrival times.

Each flight belongs to an airline a from the set of airlines A, and has a target arrival time (TAT) t_i^0. We assume that an omniscient observer could know the (extra) cost $c_i(\delta t_i)$ incurred for any delay $\delta t_i = t_i - t_i^0$ on this flight, t_i being the time allocated to the flight by the mechanism. This cost is called the "true" cost of delay for flight i and is directly linked to the slot given by the allocation, so that we can note it $c_{i,k}$: the cost of flight i if it were allocated to slot k.[6] In principle, it would be desirable to minimise the cost for the airlines, such that the best allocation μ^* verifies:

$$\mu^* = \underset{\mu}{\operatorname{argmin}}\ C(\mu), \tag{1}$$

where:

$$C(\mu) = \sum_{i}^{N} c_{i,\mu(i)} \tag{2}$$

is the total cost of delay incurred by the flights with the allocation μ. In the following, we assume that no flight can be allocated to a slot with a time earlier than the flight's TAT, i.e., δt_i is always positive. This assumption could be relaxed in future studies but does not fundamentally change the procedure.

In order to find this optimal allocation, airlines need to communicate the information on their flights to a central optimiser that then finds the best allocation. When using a central optimiser, the reliability of the information sent by the airlines is important, as distortions between the 'true' cost of a given flight and the cost used by the central optimiser might appear, and may be caused by the following issues:

1. errors made by the airlines on their true costs,
2. an approximation process used to communicate the costs to the optimiser in a simplified way,
3. gaming effects, i.e. airlines communicating 'dishonest' costs, to gain an advantage over other airlines,
4. behavioural effects, e.g. airlines clinging to slots they already own.

The first two issues are the main focus of this article, while the last two are out of scope of this particular study.

Issue #1 is naturally related to the limited capacity of the airlines to compute their true costs (Gurtner and Bolić, 2021). In this regard, airlines are very heterogeneous. While some are highly data-driven and may know their costs in detail, others may only have limited awareness of their costs, relying on proxies like the number of connecting passengers, and rules of thumb, like applying a multiplier for reactionary delay. In the following, we will denote as c' the costs that airlines compute for themselves, that we call *noisy* cost, distinct from c, their *true* costs. We will also refer to the 'internal' costs, those that airlines think are their true costs, be it c or c'.

Issue #2 is related to the way the costs are communicated to the central optimiser. In an ideal world, airlines should be able to communicate their costs *in extenso* to the central optimiser, meaning that for each slot in the regulation, the cost of delay incurred by a given flight if it were allocated to that slot would be communicated. However, due to the limited and heterogeneous computing capacity discussed above, some degree of simplification is unavoidable in any reasonable, mid-term implementable mechanism. Thus, in this article we chose to use simplified archetypal step-functions (see Section 4.2.2 for a detailed description) to convey the information from the cost function. Indeed, analytical cost functions have a small number of values that parameterise them, which makes it:

- easier to share the function with interested stakeholders,
- easier to design anti-gaming schemes by constraining the parameters of the analytical functions. One example would be forcing airline to spend virtual credits to set those parameters.[7]

In the following we will use $\tilde{c}$ to denote the approximated cost, usually based on c or c'.

Issue #3 is linked to the fact that in any competitive environment, airlines would naturally try to get the "best deal" within the rules of the mechanisms. Centralised cost minimisation needs cost information, and its efficiency relies on the quality of the information sent by the airlines. Some airlines, however, may be tempted to distort their own costs in order to increase the likelihood of getting good slots in the allocation. In general, this behaviour should be expected, leading to a decrease of the efficiency of the mechanism as a whole, in an instance of the so-called tragedy of the commons.

Issue #4 is linked to the fact that some decisions, for some mechanisms, may be taken by humans. Humans are known to be prone to bias when it comes to decisions, like the fact that they may want to keep initial slots instead of new ones (i.e. endowment effect Thaler, 1980). As a result, the information communicated to the central optimiser may, once again, decrease in quality and thus decrease the efficiency. These effects are subtle, very context dependent, and quite hard to capture, and are outside of the scope of this article, as is the issue #3.

Even though issues #3 and #4, as well as the design of new mechanisms, are out of the scope of this article, they are important to take into account for any future development, and should be further investigated.

3.2. Mechanisms

While we expect the issue of the cost function approximation to arise for nearly any centrally optimised inter-airline mechanism, here we are interested in isolating its effect by using simple, "naive" mechanisms that should serve as references or bounds for any future inter-airline mechanisms. Thus, we start with the description of our baseline, a simple FPFS mechanism that mimics the current situation, and a reference mechanism, UDPP, an intra-airline optimisation scheme than any inter-airline mechanism should try to 'beat'. They are followed by the presentation of two mechanisms representing 'upper bounds': MINCOST and NNBOUND.

3.2.1. First planned, first served

The current default mechanism for ATFM slot allocation follows the so-called First Planned First Served rule, by which each f.light is allocated to the first available slot after its TAT, starting with the earliest flight. It is easy to show that this procedure minimises the total delay incurred by the flight, so that:

$$\mu_F = \underset{\mu}{\operatorname{argmin}}\ \sum_{i}^{N} \delta t_{i,\mu(i)} \tag{3}$$

[6] Likewise, we denote $t_{i,k}$ and $\delta t_{i,k}$ respectively the allocated time and the allocated delay.

[7] For example: all airlines could have a certain amount of credits at the beginning of the week/month/year; when a flight is caught in regulation, a default cost function could be considered for it, unless the airline spends credits so that the cost approximated cost function fits the real one more closely. Hence, if the number of credits is small enough, it prevents airlines to inflate their costs.

This mechanism is regarded as the baseline in this study. Note that minimising delay is different from minimising cost in general, due to the non-linearity of the cost function, which is why there is interest for designing new allocation mechanisms in the first place (UDPP included).

3.2.2. UDPP

The UDPP mechanism, as defined in this paper, can be viewed simply as a cost minimisation process internal to the airlines, plus the selective flight protection (SFP) process, which in some cases may slightly change the slot sequence of other airlines (but they should never lose from it). The intra-airline optimisation is performed by each airline, reordering the flights within their slots allocated by the FPFS process by assigning priorities (i.e., ordering of flights), and adding eventual protections. The final allocation is thus the aggregation of the allocations computed by each airline:

$$\mu_U = \bigcup_a \underset{\mu_a}{\operatorname{argmin}} \sum_{i \in F_a} c_{i,\mu_a(i)}, \tag{4}$$

where F_a are the flights of an airline a, and μ_a is a mapping across the slots initially allocated to a (i.e., from FPFS). This formula can also be applied with noisy costs c'_i or approximated costs $\tilde{c}_i$.

On top of that, we apply the SFP process. This process allows an airline to set a "time-not-after" for a flight, making the priority assignment obsolete (i.e. a flight can be assigned either a priority, or "time-not-after"). The algorithm finds the closest slot to this time, even if it does not 'belong' to this airline. The airline then has to free one of its slots, which has to be before the "time-not-after", which will then be allocated to another airline. This procedure effectively performs slot swaps among airlines, breaking the intra-airline process. However, it can be shown that impacts on other airlines are (1) very small, and (2) almost always beneficial to the other airlines (DFLEX Consortium, 2014).[8] This is due to the airlines relinquishing a slot when requesting SFP for a flight. In the merge to obtain the final allocation, the relinquished slot is assigned to another airline, and in more than 95% of the cases, these are earlier than the FPFS assigned slots for those airlines, thus beneficial. Hence, as a kind of first approximation, we consider the initial slot reordering with SFP as a quasi intra-airline process. In the implementation we are using, the SFP is performed at the same time as the intra-airline optimisation with a linear integer optimisation programme. The algorithm finds the best allocation *and* protections in one go, minimising the expected costs for the airline.

The UDPP mechanism, under this form, yields its maximum efficiency. In practice, UDPP would use a range of processes (most of them still under development) aimed at *helping* the airlines to find the allocation they desire. This is an important point that will be discussed in conclusions. In any case, for inter-airline mechanisms, this is the mechanism 'to beat', i.e., any new mechanism should be able to show that it is more efficient than this simple intra-airline optimisation.

Note that this mechanism is by design immune to gaming (except for the SFP process, which should be very minor due to the limited number of slots any airline has in any one regulation). Behavioural effects may still play a role, but very minor. Finally, the approximation process should not play any role *if airlines are willing and/or have instruments to compute their costs with enough details*. Otherwise, they may still use approximations, which we believe is essentially what they do when using more advanced UDPP processes like flight margins (see Section 2 for details).

[8] This report, prepared by the DFLEX project, on UDPP, shows the results collected through the live trials, see for example table 7.

3.2.3. MINCOST

This mechanism is designed to provide the maximum efficiency that one can hope for in terms of cost reduction. Indeed, it simply uses Eqs. (1) and (2), which can be implemented with internal (c or c') or approximated costs $\tilde{c}_i$: .

$$\mu_M = \underset{\mu}{\operatorname{argmin}} \sum_i^N c_{i,\mu(i)}. \tag{5}$$

The output of this allocation can be regarded as the best outcome one can hope for with the information given by the airlines.[9]

Note that by design, this mechanism is extremely sensitive to gaming,[10] as well as behavioural effects. The role of the approximation is also important in this mechanism, because airlines, even honest ones, are unlikely to communicate their full costs.

3.2.4. NNBOUND

Finally, we use a modified version of "MINCOST", where we minimise the total cost while adding constraints:

$$\mu_N = \underset{\mu}{\operatorname{argmin}} \sum_i^N c_{i,\mu(i)}, \tag{6}$$

subject to:

$$\forall a, \sum_{i \in F_a} (c_{i,\mu(i)} - c_{i,\mu_F(i)}) \leq 0, \tag{7}$$

i.e., no airline loses with respect to FPFS from the cost point of view. These extra constraints reduce the number of possible allocations, by for instance, forbidding single-flight airlines exchanging their slot for a worse one. On the positive side, the constraints inject a 'natural' equity in the solution, as the solution in MINCOST can be extremely unfair, where the gains could be assigned to only some airlines, and losses to all the others.

4. Methodology

This section describes the methodology that consists of the simulation model, implementation of the cost functions and definition of the efficiency indicator. As already mentioned, to be able to perform an assessment across different AUs, and under different conditions, we use agent-based model (ABM) simulations. The cost models are implemented in the ABM model, taking into account internal (true or noisy) and approximated costs. As the final the piece of the methodology, the cost efficiency indicator is defined. The section concludes with the experimental setup description.

4.1. Agent-based modelling

In order to assess the impact of cost approximation, we employ simulations using a simple agent-based model. The model consists of two agent types: airlines and Network Manager. In the simulations, we consider a series of ATFM regulations[11] impacting a set of airlines, at airports. Particular airlines may or may not be present in each regulation, as the presence depends on the regulation activation period and airlines' schedules. Each airline has a number of flights involved in the regulation. We assume that each airline, implemented as an autonomous agent, has the following pieces of information:

- the initial estimated time of arrival (TAT) of their flights,

[9] At least from the pure economic point of view.

[10] Indeed, the mechanism relies on airlines setting their costs in an absolute manner, i.e., that the costs are directly compared across airlines. Hence, airlines could decide to inflate their costs, hoping that this would bring them better priorities for their flights, effectively trying to game the system.

[11] For simplicity, we term these just regulations from here on.

- the full cost of delay for the flights, i.e., the extra cost incurred by the airline if their flights are delayed by receiving ATFM slots. This is represented by their internal costs, either true costs c or noisy costs c'.

When the simulation starts, another autonomous agent, that we call the Network Manager (NM), draws a regulation at random from our regulation dataset (see 4.2.1 for details on this dataset). The NM agent then requests the flight 'preferences' from all airlines. The 'preferences' may take different forms and are computed in different ways depending on the mechanism used. The description of the implementation of UDPP preferences in ABM can be found in the next section, Section 4.1.1, and the MINCOST and NNBOUND in Section 4.1.2.

The NM agent uses the preferences information to run one of the mechanisms and obtain the final ATFM slot allocation. The model then computes the *true* cost of delay for the final allocation, as well as the cost for the initial allocation, based on the true cost functions of the airlines. This information is saved, and the simulation moves to the next iteration, drawing a new regulation. The gains in saved costs are aggregated at the end of the simulations to compute the indicator defined in 4.3.

Note that for all these mechanisms, each iteration is independent of the previous one. Thus the simulations can be thought of as an iterated game without memory (see 4.4 for detailed description of the simulation setup).

4.1.1. UDPP preferences

As described in Section 3.2.2, with UDPP, airlines have to submit the following parameters for each of their flights:

- a priority (indicating ordering within own slots), OR,
- the indication that the flight is protected, and the "time-not-after" (TNA) requested for this flight.

We assume in this article that the airline takes the best decision possible for these parameters, based on their internal cost functions. This best decision is computable exactly because the impact of the decision is all internal, i.e. that they do not impact flights from other companies (up to the SFP, see below).

Putting aside SFP for a moment, computing the priorities for the flights is an easy task: the airline just has to compute the best allocation among the assigned slots (out of FPFS), then set the priorities. The priorities reflect the ordering of the flights in the slot allocation. The computation of the best allocation can be done via any method, even brute force for small numbers of slots/flights.

A slightly more complex process is needed to set the protections and TNAs for the flights. Indeed, the airline has to make a trade-off essentially between the importance of a flight A and the other flights, that will likely be worse off, since the airline has to free a (usually good) slot in order to set a protection for one of its flights.

The implementation we use performs the prioritisation and protection at the same time, which can be reduced to a linear integer problem. The objective of the optimiser it to minimise total cost across the flights of the airline. For each flight, the decision variables are priorities and TNAs. Several constraints are included, in particular the fact that flights cannot depart before their initial departure time, that slots can hold one flight only, that a flight can hold only one slot, etc.

The result of the optimisation is a set of priorities and TNAs (if chosen), computed in parallel to obtain the best flight/slot allocation that the airline can ask for, without knowledge of other airlines' preferences. When airlines submit only priorities, the final regulation solution allocation is just a merge of their submitted priorities. When the protection of one or more flights is also added, then the final allocation can change slightly the slots the airlines submitted. As mentioned in Section 3.2.2, in a large majority of cases (over 95%), these changes are beneficial to the airlines.

4.1.2. MINCOST and NNBOUND preferences

MINCOST and NNBOUND require from the airline agents that the entire cost functions be sent directly. Hence, the airline either sends its full internal costs c (or c') to the NM or sends approximated costs $\tilde{c}$. The approximation process implemented by the airline to produce $\tilde{c}$ is described in 4.2.2.

4.2. Cost functions

In order to provide preferences to the NM agent, each airline has to use cost functions for all their flights. These cost functions are computed based on real historical data and can be approximated and/or modified by adding noise.

4.2.1. Cost models

In order to obtain realistic estimations for this analysis, the ABM needs to use various data. To prepare this data, we used the Mercury simulator (Gurtner et al., 2021), which has been developed over several years as a stochastic mobility simulator, tracking passengers and aircraft during one day of operations. It features several processes, like passenger connections, turnaround processes, flight cancellations etc., and includes a detailed cost model for airlines, including cost of fuel, curfew infringement, etc. Inclusion of passenger itineraries (e.g., taking direct or connecting flights) is an important part in cost of delay calculations, due to the Regulation 261 aimed at passenger rights protections. Hence, Mercury is calibrated on historical data, the most relevant of which for the present study are:

- schedules for scheduled flights in Europe for the 12th of September 2014 (source: Innovata and Demand Data Repository 2),
- airport ATFM regulations for the entire 2014 year,
- passenger itineraries for the 12th of September 2014 (source: Paxis).

For this study, a reduced geographical scope compared to the full European area was used, considering only flights involved in regulations at 21 airports. These 21 airports are selected on three criteria:

- their size,
- diversity of traffic in terms of number of airlines present,
- the possibility to easily infer the relationship between a capacity-related regulations and the airport.[12]

The list of airports is available in Annex.

Mercury uses the cost of delay developed in Cook and Tanner (2015) in order to estimate the airlines' costs. Thus in Mercury, the airline cost function takes into account maintenance, crew (in an aggregated way), missed passenger connection (explicitly), turnaround delay (explicitly), and curfew costs (explicitly). As a result, very detailed cost functions for flights are obtained, such as those shown in Fig. 1. For example, the top left graph shows two jumps, one at 75 min of delays and another one at about 160 min. Such jumps usually indicate the presence of a group of passengers that would lose a connection to the onward flight, and trigger the compensations foreseen by Regulation 261. Another reason for a jump might be due to the need to change the crew when they reach the end of the shift, and another crew needs to be transferred to take the flight.

These functions, although not perfect, are considered as the *true* costs c of the airlines in our methodology.

Mercury is used to produce an intermediate dataset of regulations, as described in Section 4.4.

[12] For instance, when there is a capacity issue at an airport around London, AFTM regulations tend to be issued for upstream sectors, and are not linked exactly to the airport. It is thus difficult to infer from the data if regulations are linked to the capacity problem at one or more airports, or in the terminal maneuvering area.

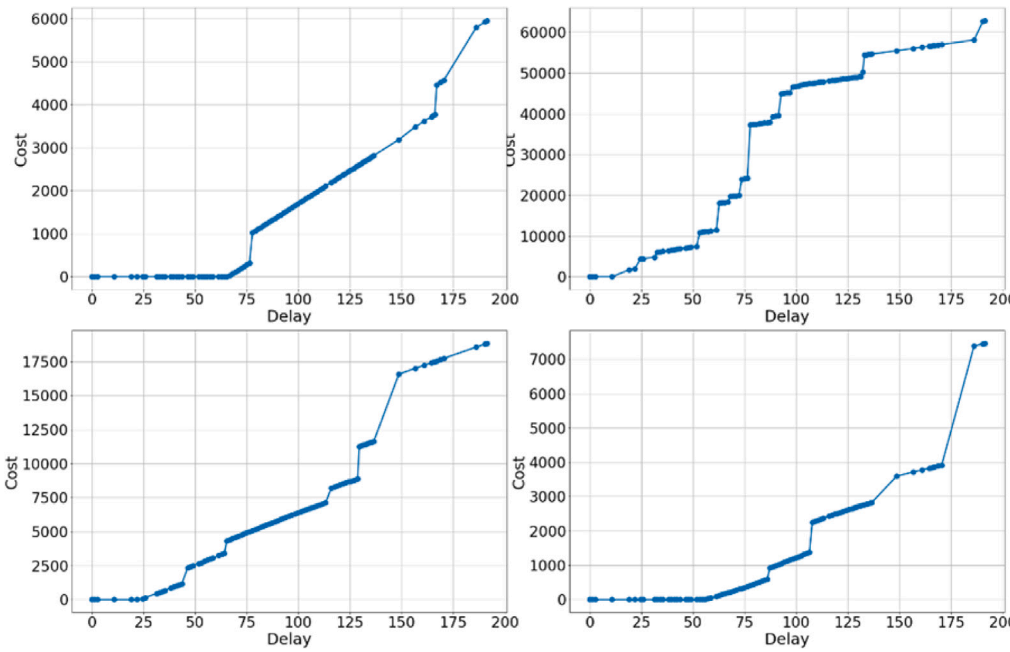


Fig. 1. Examples of cost functions used by Mercury and the agent-based model.

Table 1
Parameters in archetype functions.

Archetype	Parameters
Jump	Margin, magnitude of jump, slope
Jump2	Margin, magnitude of jump
Jump3	Margin, magnitude of jump, slope, offset
Double_jump	Two margins, two magnitudes of jump, slope
Double_jump2	Two margins, two magnitudes of jump
Double_jump3	Two margins, two magnitudes of jump, slope, offset

4.2.2. Cost approximation

For reason stated previously, we want to open the possibility to 'degrade' the information on the cost functions and use approximated costs instead. In order to do this, we start from the *internal* cost functions, i.e., either the *true* costs c or the *noisy* ones c', and perform a regression (with BFGS algorithm Zhao (2021)). The function to be fitted is chosen beforehand by the modeller among a series of archetypes.

The choice of the archetypes to be used was driven by trying to extend the UDPP and thus using similar parameters; using a 'priority' and a 'margin' (after which the flight should not be allocated), as in advanced flavours of UDPP, naturally leads to a step function. Step functions are quite intuitive, as they represent the point in time when the situation changes from good (before the jump) to critical (after the jump). For example, passengers missing their connection, or curfew infringement could be considered critical situations.

Here, we focus on approximation functions of similar shapes, that are represented by the archetypes shown in Fig. 2. The simplicity of their shapes means that only a few parameters are needed to describe the function, and to be communicated to the NM. The parameters for different archetype functions are summarised in Table 1.

Obviously, more complex functions will fit the true cost functions better, at the potential expense of the number of parameters. Other functions with less parameters than for instance the 'double_jump3' may also be good candidates. An example might be the logit function, having the following characteristics: the saturation of the cost function and the high trend in the beginning. These different functions may fit better the true cost functions using regressions, but may be much less intuitive to use in operations by airlines. Hence, we believe that the functions in Fig. 2 represent very plausible candidates for simplified cost functions. Added complexity might not be worthwhile, since it may then be easier to send the full cost function to the optimiser.

The result of the regression, yielding a function based on an archetype with the optimal parameters of the regression, represents the *approximated* cost of the airlines, $\tilde{c}$. This cost can then be used as input to the UDPP optimiser to find the flight priorities or sent to the NM agent for NNBOUND and MINCOST.

4.2.3. Noise on cost functions

As stated previously, the goal of this work is to explore the impact the errors and various approximations in costs reporting made by the airlines has on the efficiency of the mechanisms. In practice, there are very little to no information available regarding the kind of errors made by the airlines, for obvious reasons. From the conversations with airline representatives, we know that costs calculations and/or approximations across airlines are very heterogeneous, sometimes differing even over flight types of a single airline. In order to simulate the approximations and/or errors, a simple procedure of adding (controlled) noise on the true cost functions is used.

More specifically, let us consider the (true) cost c_k of flight i in slot k, where we omit index i. The noise is added by applying a normal distribution on this vector. However, if the noise is applied on each slot independently, the resulting costs might be unrealistic, especially when high levels of noise are used.

To avoid this undesirable effect, an 'autocorrelation' parameter τ is introduced. Thus, the noise is only applied on every τ component. To have smoother cost functions, we apply the noise cumulatively. In other words:

$$c'_k = c_k + n_k \tag{8}$$

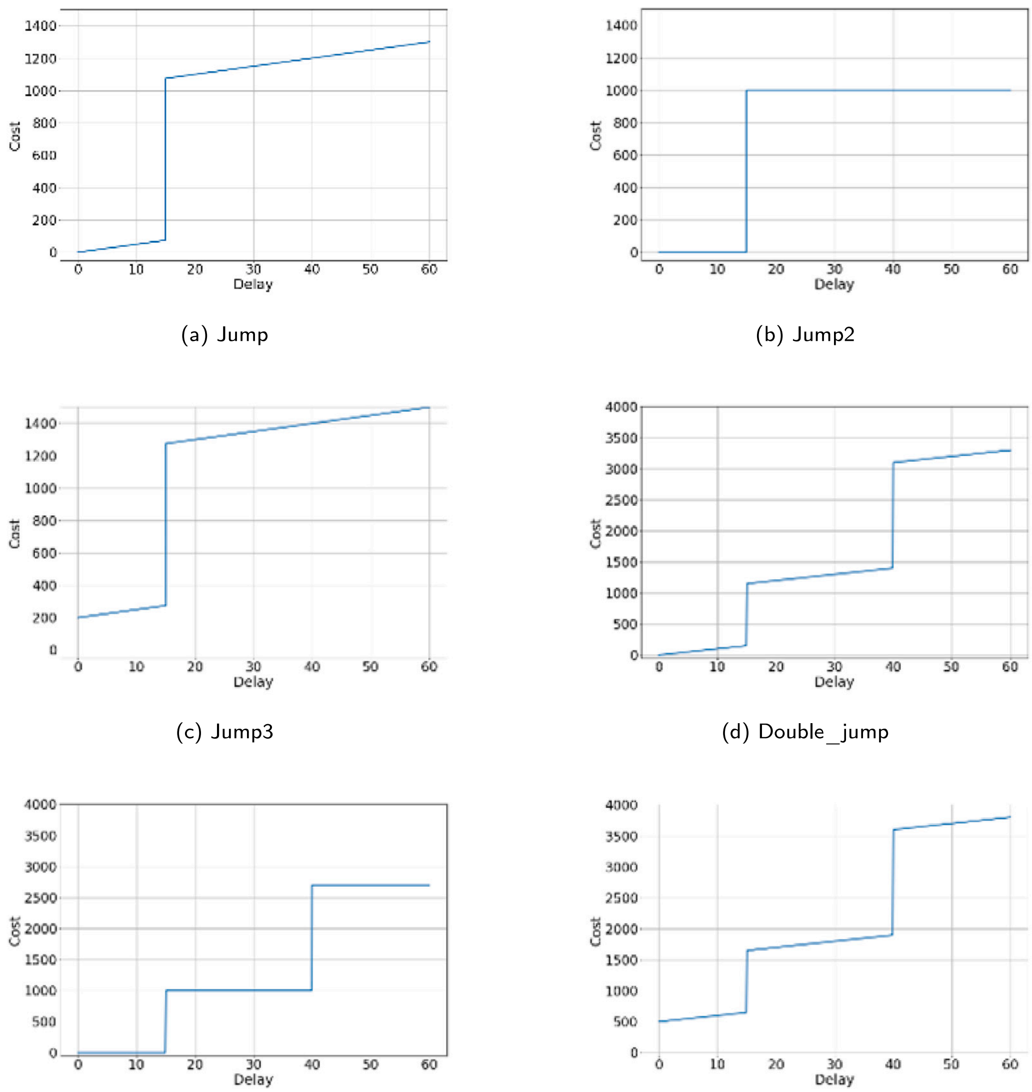
with:

$$n_k = n_{k-1}(1+\epsilon) \quad \text{if} \quad k \bmod \tau = 0 \tag{9}$$

$$n_k = n_{k-1} \quad \text{otherwise} \tag{10}$$

with $\epsilon \sim N(0,\sigma)$. We set the parameter $\tau = 10$, which results in cost functions similar to those presented in Fig. 3.

This noisy cost c' can then be used either as input of the UDPP optimiser to determine the priorities, or sent directly to the NM for NNBOUND and MINCOST. It can also be used as input to the approximation process described previously.



(a) Jump (b) Jump2

(c) Jump3 (d) Double_jump

(e) Double_jump2 (f) Double_jump3

Fig. 2. Archetype functions for the cost function approximation.

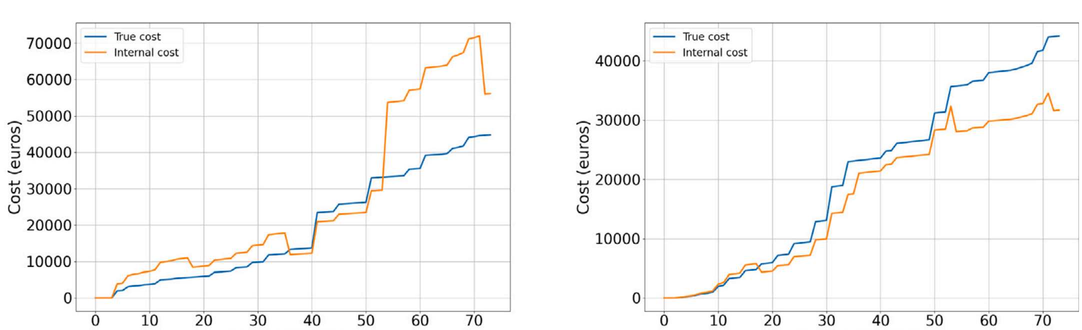


Fig. 3. Examples of true cost functions (blue) and noisy ones (orange). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Efficiency indicator

In order to measure the impact of the mechanisms, a metric we term *efficiency* is used. The efficiency of mechanism X is defined as:

$$E_X = -\frac{\sum_i \left(c_{i,\mu_X(i)} - c_{i,\mu_F(i)} \right)}{\sum_i c_{i,\mu_F(i)}} \tag{11}$$

In other words, it measures the relative gain in total *true* cost with respect to FPFS. Other indicators have also been tested, like the savings per flight, and the total cost saved. These metrics are not reported here because they bring us to the same conclusions.

4.4. Experimental setup

In order to study the impact of cost approximation for inter-airline ATFM slot exchanges, Mercury is used first to create an intermediate dataset of regulations for the ABM. It is run repeatedly (in this study, 1000 times per airport in the dataset), simulating different realisations of the chosen day of operation. From these iterations, every time a regulation is activated, the following information is collected:

- slots available in the regulation (times, durations),
- flights involved in the regulation, their TAT, and corresponding airline,
- true cost functions for each of these flights.

This information is saved as a part of the regulation dataset (Gurtner, 2023) and is used by the ABM model described above (i.e., to sample regulations) in two sets of experiments. The first set tests the impact of the cost approximation, while the second evaluates the noise effects on the internal cost calculation.

However, before performing the first set of experiments with the ABM, we look at how well different archetypes fit the true cost functions. This is performed independently of the mechanisms, by sampling the functions found in the regulation dataset, fitting different cost archetypes, and computing the coefficient of determination (R^2) as an estimation of the performance of the fitting procedure. Two aggregated values are then computed for each archetype: the ratio of positive coefficients and the average of the coefficients restricted to the positive ones.

After that, the following procedure is used to estimate the efficiency of each mechanism using different approximations:

- for each mechanism (FPFS, UDPP, MINCOST, and NNBOUND) and for each archetype, 200 iterations of the model are run,
- for each iteration, the efficiency of the mechanism is computed,
- the efficiency indicator is averaged over all simulations for each mechanism.

Finally, the effect of different levels of noise on the cost functions are tested, particularly its interplay with the approximation process. For this we select a list of values for the level of noise, i.e., for the value of the standard deviation σ of the normal distribution described in Section 4.2.3. Then, for each value of σ:

- we perform 1000 iterations for each mechanism (UDPP, MINCOST, and NNBOUND) *with* approximation and we average the efficiencies.
- we perform 1000 iterations for each mechanism (UDPP, MINCOST, and NNBOUND) *without* approximation and we average the efficiencies.

5. Results

This section describes the results of the impact of the approximation process on mechanisms, followed by the results on the impact of internal, noisy cost functions on the mechanisms.

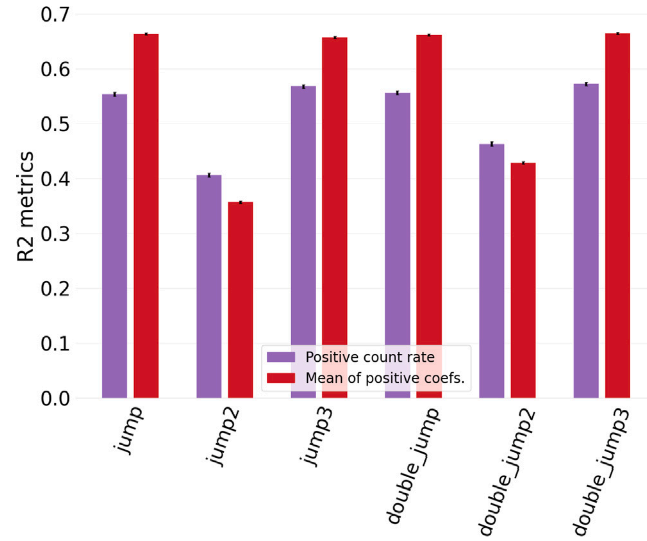


Fig. 4. Rate of positive coefficients of determination (red) and average coefficient of determination restrictive to the positive ones (violet) when regressing the cost functions with the different archetypes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.1. Impact of approximation process

5.1.1. Regressing cost functions

We start by examining the results of the fitting procedures themselves. In Fig. 4 we show the aggregated metrics based on the coefficients of determination R^2, as explained in the previous section.

A first conclusion from this figure is that none of the proposed functional forms are excellent at approximating the true cost functions, as can be seen from the relatively low value of R^2 indicators. It seems that the realistic cost functions are hard to capture using a small number of parameters, at least with the chosen archetype functions. The maximum values of R^2 can be seen for the mean of the positive coefficients, which are below 0.7, while jump2 and double_jump are below 0.5.

A second conclusion is that including the slope in the functional form (see jump, jump3, double_jump and double_jump3) seems to significantly improve the regressions. Including an offset barely improves the situation. Finally, there is no notable benefit in using two steps instead of one, except maybe without offset and no slope (see the improvement of double_jump2 over jump2).

5.1.2. Mechanism efficiency with approximated costs

Previous section shows that the chosen archetype functions are not able to represent well the true cost functions, which poses the question of the impact of such approximations on the mechanism efficiency. Fig. 5 depicts the efficiency of the mechanisms using true cost and approximated cost functions. The true costs for each mechanism are representing by the bar labelled "True".

It is clear from this figure that the approximation has a big impact on the efficiency in general. Using 'Jump2' drops the efficiency of MINCOST and NNBOUND by a factor of 3–4, highlighting the destructive power of erroneous information. Other archetypes seem to be faring better (except for double_jump2). Similarly to Fig. 4, the 'Jump' archetype seems to represent the best compromise (performing as well as the others, but with less parameters).

More importantly, the mechanism with the highest efficiency is MINCOST, and in the case when the approximated cost functions are used, the mechanism efficiency drops to that of the 'True' UDPP. This is a crucial point, because while we could imagine that airlines are sufficiently good at computing their costs internally at one point in

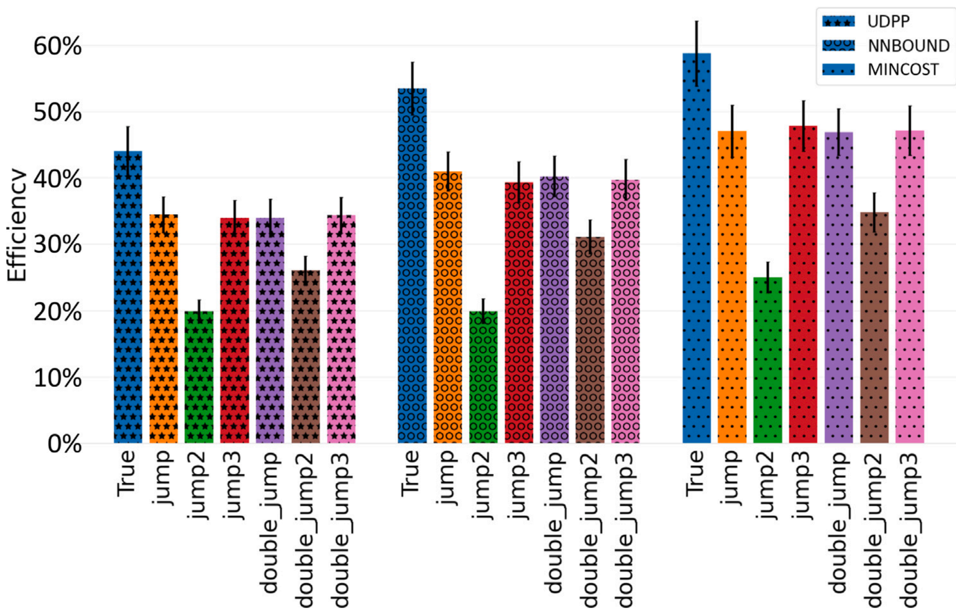


Fig. 5. Average efficiency of the mechanisms when different approximations are used, with respect to using true cost functions ('True'). Error bars are standard errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the future, without the crude approximations presented here, it will be difficult to design an inter-airline mechanism without approximation, for reasons we already stated. In other words, the gap between the orange bar ('Jump') in Fig. 5 for MINCOST, which represents the highest efficiency if airlines use approximations, and the blue bar ('True') for UDPP, which represents the likely efficiency of airlines if they were able to compute their costs perfectly, is where any inter-airline mechanisms using approximations should lie, which is a pretty small interval. NNBOUND, for instance, seems to be slightly below UDPP, while being pretty good in theory: around 41% of efficiency against 46% for MINCOST.

But is it really realistic to expect the airlines to use the UDPP so efficiently? Probably not, because as we already stated, many airlines currently do not have the means to compute the exact costs of their flights when a regulation hits. Hence, they may also rely on approximations and/or heuristics. However, even with the approximated costs, UDPP fares pretty well. There is only around 10 percentage points between the orange bars in UDPP and MINCOST. This leaves a small margin of improvement for a future inter-airline design, unless they can avoid this kind of approximation (we will discuss this point in conclusions).

The reason for this situation seems to be the apparent high efficiency of UDPP, reaching around 45% without approximation, or 35% with it. Even in the best case, one can only gain about 20 more percent. Is it worth trying to find new *inter-airline* mechanisms, if the *intra-airline* one is so efficient? In order to answer this, we also need to look at the possible initial errors made by the airlines, before approximation takes place.

5.2. Impact of noise on efficiency of mechanisms

In this section, we show the impact of the additional noise added on the flights' cost functions, with the procedure explained in Section 4.2.3. Fig. 6 shows the variations of the efficiency of different mechanisms when the added noise increases. All approximations are made with the 'jump' archetype, as it seems to represent the best trade-off from the efficiency and complexity point of view.

As expected, the efficiency decreases with the level of noise. However, this trend differs across mechanisms. It is striking to realise that, while mechanisms with approximated functions fare worse than with

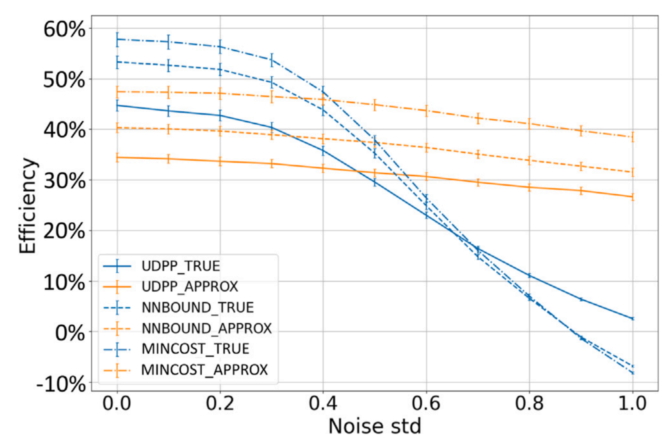


Fig. 6. Average efficiency of the mechanisms as a function of the level of noise, represented as standard deviation of the normal distribution. Results with true cost functions are in blue, approximated ones in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

true costs when the noise is small, the situation is reversed for high levels of noise. More specifically, the MINCOST mechanism for instance decreases from 58% when there is no noise, to −8% when the noise is high, if true costs are used. On the other hand, the efficiency of the same mechanism goes from 48% down to 38% when approximated functions are used on top of noise. This trend is the same for all mechanisms. Note that UDPP is more robust with respect to noise than other mechanisms, especially with true costs, but not to the extent of UDPP with approximated costs.

This trend is very surprising and puzzling. Indeed, one could have naively expected that performing an approximation – not a very good either – could be considered as some kind of noise, added on top of the noise on the true costs. Hence, we could have expected that the two types of noises are added, and thus the situation with approximated costs would be always worse than the one with true functions. However, it seems that the fitting procedure creates some kind of resilience to errors, at least with this type of noise, and manages to catch the main features of the underlying cost function, even when the latter is hidden by high levels of noise.

The reason for this behaviour is yet unknown and would require more work to be fully understood. It is possible that the symmetry in the noise plays a role at least. Indeed, asymmetric noises may very well bias the regression process and lead to systematic under- or over-evaluation of some parameters. Hence, other types of noises should be explored. However, it is an interesting fact that could have a crucial role in the design of future mechanism, because approximated cost functions could serve as a barrier against the randomness of the cost functions coming from some airlines. We discuss this point further in the conclusions.

6. Discussion

The results presented on the approximation process can be reasonably expected, qualitatively at least. A cost optimisation process, at its heart, needs a reliable approximation on the cost in order to work properly. However, thanks to a realistic setup and a careful comparison of different mechanisms, we argue that this issue has been severely underestimated in the past and jeopardises the entire concept of inter-airlines flight swapping, and even, to some extent, the very concept of intra-airlines flight swapping, i.e. UDPP.

Let us first consider the case where airlines are able to compute perfectly the cost functions of their flights. Then UDPP should be able to work to its maximum efficiency, with around 45% of reduction in cost with respect to FPFS according to applied methodology. Even without gaming effects, any approximation with step functions seems to yield a lower efficiency, with maybe an exception for a full global optimisation, which allows to gain only a few percentage points over UDPP at most, and which represents an upper bound on the efficiency of any mechanism. Since it can be expected for airlines to use their true costs (once again, if they know them) in UDPP, it is thus doubtful that anyone could build a mechanism with a sufficient efficiency based on a central optimisation process with cost approximation that beats intra-airline swapping.

There are two obvious solutions to this problem:

- improving the cost approximation archetypes, or
- designing a mechanism where airlines communicate their true cost functions.

The first solution might be feasible but a success is not guaranteed. Indeed, due to the diversity of cost functions, finding a single archetype that would work for everyone could prove difficult. The second solution assumes the airlines would be willing to communicate their trues costs, which in itself could create gaming problems. Some players might be tempted to inflate their costs, effectively gaming the mechanism in use and decreasing its efficiency. To avoid the gaming, the mechanism using the true cost might need to incorporate a compensation scheme whereby a certain amount needs to be paid (in a virtual currency or real money) to submit higher levels of costs.

Even with a good theoretical mechanism, it is not obvious that the airlines would find the communication of their true costs acceptable. According to the feedback from airlines that was gathered by the authors (for example see Gurtner and Bolić (2021)), they may be comfortable with sending a few parameters for their flights (like in UDPP) but would not be comfortable with sending too many details about their costs. Our results show that this desire from airlines is deeply incompatible with efficient inter-airlines mechanisms, at least for the mechanisms involving a central optimiser, as the optimisation works best when the complete information is used. Note that while the neutrality of the NM may help convincing airlines to disclose their true costs, there are other ways for an optimiser to use the costs without privacy issues, based on blockchain technologies. The SESAR ER4 SlotMachine (Publications Office of the European Union, 2022) project for instance explored this possibility.

While the hunt for an efficient mechanism that would avoid gaming issues and use an adequate approximation scheme continues, we would like to point out an even bigger issue: the potential lack of cost computation capabilities of some airlines and the subsequent errors in communicating their costs.

Over many years and several projects,[13] the authors received feedback on the way airlines compute their own costs. The conclusion is that airlines are very heterogeneous in this regard. Some airlines have very advanced cost models that they use in a tight loop to learn from past situations, using machine learning and AI, while others rely operationally on sets of indicators that are not directly related to cost (e.g. number of missed connections, on-time performance). Thus, it is safe to assume that airlines, in general, have an imperfect knowledge of these costs, at least nowadays.

An interesting line of work in this regard is developed within the UDPP solutions. Interestingly, the main effort in UDPP was geared more towards the *usability* of intra-airline swapping capabilities, in terms of the tools that would ease the exchange of information between the airlines and NM that manages the hotspot activation and delay apportionment through the slot allocation. For instance, UDPP allows airlines to give priorities to flights to change their allocation, which can always be modified by the airlines (up to a certain deadline, which is within the duration of the hotspot). Another example is the “flight margin” mechanism developed by UDPP lately, which allows airlines to set “margin-not-before” (i.e., time-not-before) and “margin-not-after” for their flight to produce allocations that make more sense for airlines, which can be continuously modified until a certain cut-off time. Hence, the emphasis on the development of such “helping” tools for UDPP show to which extent the airlines *do not* compute their costs explicitly.

This raises of course the problem of the efficiency of UDPP itself, but also of any other mechanism to be designed. In this article we used a simple setup to add noise on top of a theoretical, perfect true cost function, assuming the airline only knows the noisy function. The results are striking, first in terms of drop of efficiency, even though it is impossible for now to pinpoint the actual degree of noise that would mimic the real errors made by airlines. Regardless, our main finding in this area is that an approximation process may shield the mechanism efficiency against errors in cost computation. Indeed, we found that the drop in efficiency due to noise is much smaller when using the approximation as an intermediate step than with the true costs. Could this save the future design of inter-airline central optimisation mechanisms?

It may, but only if airlines in general are particularly bad at computing their own true costs, and provided that the results obtained here are generalisable to the kind of errors they make. Indeed, in this case, airlines *should* approximate their costs to *gain* in efficiency, which means that, suddenly, the gap between UDPP and the theoretical maximum (MINCOST) widens (to roughly 10 percentage points). Thus, there is room again for inter-airline mechanisms, even though the extra effort of designing and implementing a whole new mechanism should be questioned, when the efficiency of UDPP is already around 30% (representing an average of 600€ of savings per flight in regulation). In particular, taking into account gaming and behavioural effects from the airlines (which should be done routinely for these mechanisms, as shown during the BEACON project Gurtner and Bolic, 2023), it is likely that new mechanisms will only improve the efficiency by a few percentage points. Note that in any case, if errors from airlines are expected to be high, approximations should be used for both intra and inter-airline optimisations.

The results of the present article could serve as a guide for the future designs of mechanism. Indeed, new centralised mechanism usually start by how the costs/priorities should be prescribed and communicated to

[13] BEACON, Domino, European airline delay cost reference values (Cook and Tanner, 2015).

a central body. This initial choice relies on some degree of approximation, and is thus prone to the effects we have explored in this article. For instance, Balakrishnan et al. (2022) bin the importance of flights into high/medium/low priorities, which can be considered as a crude approximation on the y-axis of the cost function (the cost itself), with no information on the x-axis (the time, margins etc.). Based on our results, we thus know that the maximum cost reduction would be well below 45% (see Fig. 5), and that the corresponding UDPP process with true costs could be around 40%. Hence, at the design stage, by exploring the degree of approximation set by the mechanism, one already knows that this mechanism would likely not beat UDPP with true costs (the authors report 8%–20% of improving over costs). However, if it is known that some airlines have low capabilities in cost computation, we know that this type of approximation would likely shield them against their own errors. Hence, such a mechanism could be very beneficial to such airlines but far less efficient (than UDPP) for airlines with high computational capabilities.

7. Conclusions

In summary:

- approximation processes are probably unavoidable for any regulation resolution mechanism with a central optimiser,
- approximation processes close the gap in efficiency between intra- and inter-airline mechanisms,
- noise on the cost function also closes this gap,
- approximations reduce the effect of noise to some extent,
- new inter-airline mechanisms will have a hard time beating UDPP solutions, and if they do, it would probably be by a few percentage points, and using approximation to cancel out the effect of noise on the cost functions.

Finally, we close this article by considering where future lines of research lie in light of these results.

First, we should recall that we are using data of a limited temporal (mostly, one day of operation, except for the regulations) and spatial scope (only 21 airports). We are planning to extend the scope of the data used, in particular in the temporal dimension. However, if using more data will provide more accurate answers, it is likely that the trends shown in the article will stand, as they are already computed on a very high variety of cost functions and regulations. Different types of noises should also be tested to see if the resilience of the approximation process is confirmed. In particular, we believe that airlines may be very good at estimating some parts of the function (e.g. the position of a big jump) and bad for other parts, which in the future can be reflected in the way the noisy functions are produced. We are also planning to try to estimate the errors made by the airlines, by comparing what dispatchers would do with common information (number of passengers connecting, etc.) to what an agent knowing the true costs (on a very specific and controlled experiment) would do.

Second, we limited ourselves to central optimisation schemes here, where by design we need the airlines to cooperate and send their cost functions in one way or another. Market-driven mechanisms may be able to solve some of the issues we face, because airline costs are not revealed explicitly but used to buy and sell slots or priorities for instance. Such a mechanism – a primary auction for slots – has been tested in BEACON project (Mocholí, 2022) with a simple agent-based model. The results are encouraging from the efficiency point of view, but raise important issues in terms of usability, because we enter the realm of trading strategies, which may be far too complex for airlines' operational processes. Moreover, the degree to which these mechanisms are sensitive to errors on the cost functions is unknown.

Third, we have not taken into account any gaming or behavioural effects in this study. MINCOST in particular is easy to 'cheat', by inflating costs, and should only be considered as a benchmark. These behavioural and gaming effects in general can be expected to decrease the efficiency of the mechanisms, which in turn decreases the window of possible improvement over UDPP. Indeed, gaming effects will be absent of the intra-airline optimisation (UDPP), by definition, while behavioural effects can be expected to be small as well in this case.[14] This means that the UDPP mechanism does not lose in efficiency, while any inter-airline one would, thus further decreasing the window of opportunity for a mechanism to improve on UDPP. However, these effects can be mitigated, gaming ones with a good compensation scheme, behavioural ones with automation for instance. Future mechanisms should thus be designed with this fact in mind.

Finally, we assumed in this article that the primary goal of inter-airline mechanisms should be purely economic, i.e., that costs should be decreased as a whole. It is however important to note that airlines, at least at the management level, are highly sensitive to differential gains, i.e., how much they save with respect to their rivals. This can be captured by other Key Performance Areas, like Equity or Fairness, where indicators can be built to estimate how unfair the result of an allocation is. In this regard, the MINCOST mechanism fares obviously very badly, as it is not even guaranteed that airlines do not lose from the final allocation with respect to FPFS. UDPP is also known to be bad for low-volume users, as pointed out in the introduction. NNBOUND, on the other hand, may provide a good basis for a more equitable mechanism. In any case, maybe the future development of inter-airline allocation should focus on the search for an alternative objective, like keeping the efficiency at the level of UDPP but improving only the equity.

CRediT authorship contribution statement

Gérald Gurtner: Conceptualization, Methodology, Software development, Simulations. **Tatjana Bolić:** Data curation, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Used data is published and available, see Gurtner (2023).

Funding acknowledgement

The work presented here is the result of the BEACON project, funded by the European Union's Horizon 2020 – SESAR research and innovation programme under Grant Agreement No. 893100. More information can be found here: https://www.beacon-sesar.eu/.

Annex. List of airports used in simulations

The list of airports can be found in Table 2.

[14] More specifically, we expect effects captured by prospect theory – a conceptual framework that extends utility maximisation and allows to capture various, 'irrational' behaviours and biases – for instance to be small. Indeed, for intra-airline optimisation the most important thing is the priority among flights. Prospect theory, with its monotonously increasing prospect function, would not change the priorities among flights. Other biases like endowment may have an effect during the first period where such a mechanism is in place, until the airlines learn how to the use it.

Table 2
List of airports used for simulations.

Airport	ICAO code
Frankfurt	EDDF
Paris Charles de Gaulle	LFPG
Istanbul Atatürk	LTBA
Munich	EDDM
Oslo	ENGM
Vienna	LOWW
Zurich	LSZH
Helinki	EFHK
Fiumicino	LIRF
Copenhagen	EKCH
Paris Orly	LFPO
Stockholm Arlanda	ESSA
Berlin Tegel	EDDT
Dublin	EIDW
Madrid	LEMD
Barcelona–El Prat	LEBL
Antalya	LTAI
Brussels	EBBR
Düsseldorf	EDDL
Palma de Mallorca	LEPA
Hamburg	EDDH

References

Angner, E., 2016. A Course in Behavioral Economics. Springer, p. 336.

Balakrishnan, H., Chin, C., Gopalakrishnan, K., Li, M.Z., 2022. Airline disruption management with delay ledgers. In: International Conference on Research in Air Transportation (ICRAT).

Bolić, T., Castelli, L., Rigonat, D., 2017. Peak-load pricing for the European air traffic management system using modulation of en-route charges. Eur. J. Transp. Infrastruct. Res. 17 (1).

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. 99 (suppl_3), 7280–7287.

Castelli, L., Pesenti, R., Ranieri, A., 2011. The design of a market mechanism to allocate Air Traffic Flow Management slots. Transp. Res. C 19 (5), 931–943.

Cook, A., 2016. European Air Traffic Management Principles, Practice and Research. Taylor & Francis Group, p. 279.

Cook, A.J., Tanner, G., 2011. European Airline Delay Cost Reference Values. Technical Report.

Cook, A.J., Tanner, G., 2015. European Airline Delay Cost Reference Values - Updated and Extended Values (Version 4.1). Technical Report, University of Westminster, URL https://www.eurocontrol.int/sites/default/files/publication/files/european-airline-delay-cost-reference-values-final-report-4-1.pdf.

Cook, A.J., Tanner, G., Bolic, T., 2021. D3.2 Industry Briefing on Updates to the European Cost of Delay. Technical Report, BEACON Consortium, URL https://www.beacon-sesar.eu/wp-content/uploads/2022/10/893100-BEACON-D3.2-Industry-briefing-on-updates-to-the-European-cost-of-delay-V.01.01.00-1.pdf.

DFLEX Consortium, 2014. Demonstration Report (D2). Technical Report.

EUROCONTROL, Niarchakou, S., Sfyroeras, M., 2022. ATFCM Operations Manual.

European Commission, 2004. Regulation (EC) No 261/2004 of the European Parliament and of the Council of 11 February 2004 Establishing Common Rules on Compensation and Assistance to Passengers in the Event of Denied Boarding and of Cancellation or Long Delay of Flights, and Repealing Regulation (EEC) No 295/91 (Text with EEA relevance) - Commission Statement. Technical Report, URL http://data.europa.eu/eli/reg/2004/261/oj.

Evler, J., Schultz, M., Fricke, H., Cook, A., 2020. Development of stochastic delay cost functions. In: 10th SESAR Innovation Days.

Gurtner, G., 2023. Data – Air traffic management hotspots in Europe with airline cost functions. URL https://doi.org/10.5281/zenodo.8334836.

Gurtner, G., Bolić, T., 2021. D6.1 Intermediate Concept Assessment Report. Technical Report, URL https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5f47c29e9&appId=PPGMS.

Gurtner, G., Bolic, T., 2023. D5.2 Final Tactical Model and Results. Technical Report, URL https://www.beacon-sesar.eu/wp-content/uploads/2023/02/893100-BEACON-D5.2-Final-tactical-model-and-results-V.01.00.06.pdf.

Gurtner, G., Cook, A., 2021. The hidden cost of uncertainty for airspace users. J. Air Transp. Manage. 91, 102002.

Gurtner, G., Delgado, L., Valput, D., 2021. An agent-based model for air transportation to capture network effects in assessing delay management mechanisms. Transp. Res. C 133, 103358.

Ivanov, N., Jovanović, R., Fichert, F., Strauss, A., Starita, S., Babić, O., Pavlović, G., 2019. Coordinated capacity and demand management in a redesigned air traffic management value-chain. J. Air Transp. Manage. 75, 139–152.

Mocholí, D., 2022. D4.2 Final Model Results. Technical Report, URL https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5f1c817e1&appId=PPGMS.

Mocholí, D., Alcolea, R., Herranz, R., 2020. SESAR Engage KTN – Catalyst Fund Project Final Technical Report: Exploring future UDPP concepts through computational behaviouraleconomics. Technical Report, Nommon Solutions and Technologies SL, URL https://engagektn.com/wp-content/uploads/2021/04/C9-Exploring-future-UDPP-concepts-final-technical-report.pdf.

Mueller, M., Pyka, A., 2016. Economic behaviour and agent-based modelling.

Performance Review Unit, 2022. Single European sky data portal. URL https://www.eurocontrol.int/prudata/dashboard/.

Pilon, N., Guichard, L., Cliff, K., 2019. Reducing impact of delays using airspace user-driven flight prioritisation. In: Nineth SESAR Innovation Days.

Pilon, N., Ruiz, S., Bujor, A., Cook, A., Castelli, L., 2016. Improved flexibility and equity forairspace users during demand-capacity imbalance – an introduction to the user-driven prioritisation process. In: Schaefer, D. (Ed.), Proceedings of the 6th SESAR Innovation Days.

Publications Office of the European Union, 2022. CORDIS Results Pack on AI in Air Traffic Management. Publications Office.

Ruiz, S., Guichard, L., Pilon, N., Delcourte, K., 2019. A new air traffic flow management user-driven prioritisation process for low volume operator in constraint: Simulations and results. J. Adv. Transp. 2019, 1–21.

SESAR Joint Undertaking, 2020. European ATM Master Plan : Digitalising Europe's Aviation Infrastructure : Executive View : 2020 Edition. Publications Office.

SESAR Joint Undertaking, 2021. SESAR Solutions Catalogue 2019. Publications Office, URL: https://www.sesarju.eu/sites/default/files/documents/reports/SESARSolutionsCatalogue2021.pdf.

STATFOR, 2022. EUROCONTROL Forecast Update 2022–2028. techreport, EUROCONTROL.

Thaler, R., 1980. Toward a positive theory of consumer choice. J. Econ. Behav. Organ..

Tibichte, A., Dalichampt, M., 1997. ATFM Modelling Capability, EEC Note No. 28/97. Technical Report, EUROCONTROL Experimental Centre.

von Neumann, J., Morgenstern, O., 1992. Theory of Games and Economic Behavior. Princeton University Press.

Zhao, W., 2021. A Broyden–Fletcher–Goldfarb–Shanno algorithm for reliability-based design optimization. Appl. Math. Model. 92, 447–465.