Universitat Autònoma de Barcelona

THE INTEGRATION OF SIMULATION, OPTIMISATION, AND BAYESIAN INFERENCE FOR ENHANCING AIRPORT STAND ALLOCATION

Margarita Bagamanova

DOCTORAL DISSERTATION

THE INTEGRATION OF SIMULATION, OPTIMISATION, AND BAYESIAN INFERENCE FOR ENHANCING AIRPORT STAND ALLOCATION

by

Margarita Bagamanova

Supervised by:

Dr. Juan José Ramos González	Dr. Miguel Mújica Mota
Universtitat Autònoma de Barcelona	Amsterdam University of Applied Sciences

Submitted in partial fulfilment of the requirements for the degree of

Doctor in Philosophy

PhD programme in Electronic and Telecommunication Engineering

Department of Telecommunications and Systems Engineering

UNIVERSITAT AUTÒNOMA DE BARCELONA

2020

Executive summary

Airport stakeholders are routinely challenged with the allocation of scheduled flights to available stands and parking positions in the most cost-efficient way. At the same time, they need to comply with airline preferences and service contracts and ensure passenger comfort. In the last years, a goal to reduce pollutant emissions added even more complexity to the modern air transport network, which often suffers from congestion problems and related operational disruptions.

To address these problems and facilitate airports with related decision support, this research presents a novel approach to the most common airport problem – efficient stand assignment. The proposed concept of a disruption-aware stand allocation tool combines advantages of Bayesian inference, simulation, and evolutionary optimisation and provides qualitative assignment schedules with robustness to a certain level of flight schedule deviations. The presented stand allocation approach coupled with simulation is innovative as it allows 1) to tackle the burden of schedule disruptions on the airport capacity, 2) optimise stand capacity use from multiple management perspectives, 3) helps to release resources that are usually blocked by extensive buffer times between allocated flights, and 4) improves airport environmental footprint.

The research presented in this dissertation contributes to the body of literature with the following:

- A disruption-aware stand allocation methodology provides decision support to tackle the interests of passengers, airport stakeholders, and the environment in a balanced way.
- The presented approach facilitates mitigation of operational variability on airport stand capacity management and its environmental footprint.
- The methodology generates solutions that consider historical disruptions and specific emissions characteristics of each aircraft, which provides airport stakeholders with more realistic stand assignment planning.
- The developed stand assignment approach is coupled with simulation to provide a qualitative assessment of the generated solutions and consider stochasticity of the real-life system not captured by the assignment-generating framework.

The developed approach could be further extended to consider all steps of the aircraft turnaround process. Moreover, the disruption-aware stand assignment approach could be directly incorporated into the airport simulation model to increase robustness and realism of the generated solutions.

Resumen ejecutivo

Las partes interesadas del aeropuerto se enfrentan habitualmente al desafío de la asignación de vuelos programados a los puestos de estacionamiento y puestos disponibles de la manera más rentable. Al mismo tiempo, deben cumplir con las preferencias de la aerolínea y los contratos de servicio y garantizar la comodidad de los pasajeros. En años recientes, el objetivo de reducir las emisiones contaminantes agregó aún más complejidad a la red moderna de transporte aéreo, que a menudo sufre problemas de congestión e interrupciones operativas.

Para abordar estos problemas y facilitar los aeropuertos con herramientas de apoyo a la toma de decisiones, esta investigación presenta un enfoque novedoso para el problema aeroportuario más común: la asignación eficiente de puestos de estacionamiento. El concepto propuesto de herramienta de asignación de stand combina las ventajas de la inferencia bayesiana, la simulación y la optimización evolutiva y proporciona soluciones de asignación cualitativos con solidez para un cierto nivel de desviaciones del programa de vuelo. El enfoque de asignación de puestos presentado junto con la simulación es innovador ya que permite 1) abordar la carga de interrupciones del horario considerando la capacidad del aeropuerto, 2) optimizar el uso de la capacidad del puesto desde múltiples perspectivas de gestión, 3) ayuda a liberar recursos que normalmente están bloqueados por extensos tiempos de amortiguamiento entre vuelos asignados, y 4) mejora la huella ambiental del aeropuerto.

La investigación presentada en esta disertación contribuye al cuerpo del conocimiento con lo siguiente:

- Una metodología de asignación de stands que considere las interrupciones y brinde apoyo a la toma de decisiones para abordar los intereses de los pasajeros, las partes interesadas del aeropuerto y el medio ambiente de manera equilibrada.
- El enfoque presentado facilita la mitigación de la variabilidad operativa en la gestión de la capacidad del stand del aeropuerto y su huella ambiental.
- La metodología genera soluciones que consideran las interrupciones históricas y las características de emisiones específicas de cada aeronave, lo que proporciona a las partes interesadas del aeropuerto una planificación de asignación de stand más realista.
- El enfoque de asignación de stand desarrollado se combina con la simulación para proporcionar una evaluación cualitativa de las soluciones generadas y considerar la estocasticidad del sistema real no considerada en muchas investigaciones previas.

El enfoque desarrollado podría ampliarse para considerar todos los pasos del proceso de entrega de aeronaves. Además, el enfoque de asignación de puestos de observación de interrupciones podría incorporarse directamente al modelo de simulación del aeropuerto para mejorar aún más la robustez y el realismo de las soluciones generadas.

Resum executiu

Els interessats de l'aeroport tenen un repte rutinari amb l'assignació de vols programats a les grades disponibles i a les posicions d'aparcament de la manera més rendible. Al mateix temps, han de complir les preferències de les companyies aèries i els contractes de serveis i garantir el confort dels passatgers. En els darrers anys, l'objectiu de reduir les emissions contaminants va afegir encara més complexitat a la moderna xarxa de transport aeri, que sovint pateix problemes de congestió i alteracions operatives relacionades.

Per abordar aquests problemes i facilitar els aeroports moderns amb el suport relacionat amb la presa de decisions, aquesta investigació presenta un nou enfocament del problema aeroportuari més comú: l'assignació eficient d'estands. El concepte proposat d'eina d'assignació d'estands conscients de la interrupció combina els avantatges de la inferència bayesiana, la simulació i l'optimització evolutiva i proporciona programacions d'assignació qualitatives amb robustesa fins a un cert nivell de desviacions de l'horari de vol. L'enfocament d'assignació d'estands presentat juntament amb la simulació és innovador ja que permet 1) abordar la càrrega de les interrupcions horàries de la capacitat de l'aeroport, 2) optimitzar l'ús de la capacitat d'estands des de múltiples perspectives de gestió, 3) ajuda a alliberar recursos que generalment estan bloquejats per temps d'amortiment entre els vols assignats i 4) millora la petjada ambiental de l'aeroport.

La investigació presentada en aquesta tesi contribueix al conjunt de la literatura amb el següent:

- Una metodologia d'assignació d'estands que tingui en compte la interrupció proporciona suport a la decisió per abordar els interessos dels passatgers, les parts interessades de l'aeroport i el medi ambient de manera equilibrada.
- L'enfocament presentat facilita la mitigació de la variabilitat operativa en la gestió de la capacitat dels estands de l'aeroport i la seva petjada ambiental.
- La metodologia genera solucions que tenen en compte les interrupcions històriques i les característiques específiques de les emissions de cada avió, que proporciona als grups d'interès de l'aeroport una planificació més realista de l'assignació d'estands.
- L'enfocament d'assignació d'estands desenvolupat s'uneix a la simulació per proporcionar una avaluació qualitativa de les solucions generades i considerar l'estocàstica del sistema de la vida real no captada pel marc generador d'assignacions.

L'enfocament desenvolupat es podria ampliar encara més per tenir en compte tots els passos del procés de canvi d'avions. A més, l'enfocament d'assignació d'estands conscient de la interrupció es podria incorporar directament al model de simulació d'aeroports per millorar encara més la solidesa i el realisme de les solucions generades.

List of figures

List of tables

Acknowledgements

This dissertation is a result of a long journey during which many people impacted its course. First, I would like to thank my supervisor from the Autonomous University of Barcelona, Dr Juan Jose Ramos for providing me with this PhD opportunity and introducing me to the fascinating world of aviation and machine learning algorithms. Also, I would like to give a special thanks to Dr Miguel Mujica Mota, without whose coaching and guidance I would have never seen the results of this research nor been able to share them with the academic community. His incredible patience and teaching talent ensured the success of this endeavour and even its development into further research opportunities.

Additionally, I would like to express my gratitude to Dr Miquel Angel Piera for providing me with a real-life project experience and his knowledge of simulation, and to Monica Gutierrez – for being a stronghold of the technical and social part of my stay at the UAB. During my work at the UAB, I had a pleasure to meet many people who inspired and supported me with their ideas, humour, and hard-working attitude: Dr Jenaro Nosedal, Dr Marko Radanovic, Dr Thimjo Koca, Marsel Omeri, Dr Liana Napalkova, Dr Mercedes Narciso, and many others. Additionally, I would like to thank Jose Manuel Cordero and AIRPORTS project team for our productive collaboration that gave origin to this research. Furthermore, I would like to thank my colleagues from the research stay at the Amsterdam University of Applied Sciences: Dr Rodrigo Romero-Silva, Dr Abdel El Makhloufi, and Dr Paolo Scala for their valuable advice and inspirational guidance.

This dissertation presents research that would not be possible without the support of my family, my friends, and my beloved ones. I dedicate this thesis to you.

Contents

List of publications

This thesis is based on a collection of the following peer-reviewed and published research articles.

I. M. Bagamanova and M. Mujica Mota, "A multi-objective optimization with a delay-aware component for airport stand allocation," *Journal of Air Transport Management*, vol. 83, p. 101757, Mar. 2020, doi: 10.1016/j.jairtraman.2019.101757 [1].

II. M. Bagamanova and M. Mujica Mota, "Reducing airport environmental footprint using a disruption-aware stand assignment approach," *Transportation Research Part D: Transport and Environment*, vol. 89, p. 102634, Dec. 2020, doi: 10.1016/j.trd.2020.102634 [2].

In addition, the following complimentary publications have been elaborated during the research period where the seminal ideas were introduced (included as appendices):

A. M. Bagamanova, J. J. Ramos González, M. À. Piera Eroles, J. M. Cordero García, and Á. Rodríguez-Sanz, "Modelling Dependence of Arrival Sequencing and Metering Area," in *Proceedings of the European Modeling and Simulation Symposium*, 2017, pp. 287–295 [3].

B. M. Bagamanova, J. J. Ramos González, M. À. Piera Eroles, J. M. Cordero García, and Á. Rodríguez-Sanz, "Identifying and modelling correlation between airport weather conditions and additional time in airport arrival sequencing and metering area," *International Journal of Simulation and Process Modelling*, vol. 14, no. 3, p. 213, 2019, doi: 10.1504/IJSPM.2019.101003 [4].

C. M. Bagamanova and M. Mujica Mota, "No More Surprises: Stand assignment algorithm with likelihood of turnaround time deviation," in *Posters of the SESAR Innovation Days 2018*, Dec. 2018 [5]

D. M. Bagamanova and M. Mujica Mota. "Reduction of taxi-related airport emissions with disruptions-aware stand assignment: case of Mexico City International Airport", in *Proceedings of the 2020 Virtual Winter Simulation Conference*, 2020 (in press) [6].

1 Introduction

Air transportation plays a vital role in modern economic growth and development. Besides facilitating connectivity and integration on national and international scales, it also helps enabling the trade and employment opportunities. According to ICAO, in 2018, the air transport industry generated more than 65 million jobs, resulting in a global economic impact of 3.6% of the world's gross domestic product [7]. The air transport industry has been continuously growing over the past decades. Currently, it is formed by over 1303 airlines with a fleet of nearly 31 717 aircraft, 3 759 airports and 170 air navigation service providers [8].

In 2020, global COVID-19 pandemic dramatically decreased the number of flights and threatened the existence of many airlines, air transport routes, and sustainability of operations [9]. However, the air transportation is expected to recover fully, provided the vaccination, during 2021 and continue growing by in average 3.7% annually in terms of passenger traffic over the next 20 years [10].

The growing number of flights and passengers creates pressure on the airport capacity and already led to congestion in many airports and increasing delays around the world. Thus, in 2018 less than 76% of arrivals in Europe happened within 15 minutes of the scheduled time; average departure delay constituted almost 15 minutes. Approximately 46% of these delays originated from the previous flight legs, with airport turnaround problems contributing to 31% of the delays [11]. EUROCONTROL predicts that by 2040 more than 16 European hub airports will be operating on their capacity limit, bringing the average level of delays from 12 to 20 min per flight in Europe. Airport capacity shortage is expected already by 2040, which means that the global economy will lose more than 1.5 million flights. [12].

Besides the capacity problem, modern aviation also faces sustainability issues. In 2018, air transportation was accountable for approximately 2% of all human-induced CO_2 emissions and 12% of all transport-related CO_2 [13]. The constant growth of air traffic resulted in an increase in 32% of CO_2 emissions in the past five years [14]. Although most of the aviation emissions occur during the cruise phase when the aircraft is airborne, landings, take-offs, and taxiing contribute significantly to airport footprint, especially considering inhabitants of the airport surroundings that are severely impacted by the noise [15]. According to Fleuti and Maraini, more than one-third of all aircraft emissions outside the cruise phase can be generated during aircraft taxiing [16]. Therefore, it is necessary to mitigate emissions from operations not only in the air but also on the ground.

Many countries have recognised the impact of international aviation on the global climate and have resolved to minimise this impact while ensuring the sustainable growth of the industry. To guide air transportation in this initiative, ICAO and its member states decided to adopt a global market-based measure scheme in the form of the Carbon Offsetting and Reduction Scheme for International Aviation [17]. In this long-term initiative, ICAO identifies that to reduce the climate-changing impact of aviation it is necessary to act holistically: not only to switch to sustainable aviation fuels but also to improve the technological and operational efficiency of aviation. [18].

1.1 Airport capacity

Airports are complex systems with many actors involved (see Figure 1), and their functioning and development require holistic analysis and strategical perspective to coordinate and harmonise activities of all parties involved. The number of flights that an airport could accommodate during the operational day is mostly affected by two elements: runway capacity and gate capacity [19]. Runway capacity depends on the meteorological conditions, control procedures, aircraft size, and the mix of aircraft types using the runway. Gate capacity refers to the ability of the specified number of gates to accommodate aircraft ground-handling operations during the operational day. The term “gate” here refers to a designated single-aircraft parking space that can be adjacent to a terminal building or can be located remotely on the apron. In the literature, such facilities are also called stands. The maximum number of ground-handling operations that can be accommodated by airport gate/stand capacity depends on:

- Parking space arrangement
- Aircraft turnaround time, ground service and passenger loading characteristics
- Gates/stands category and size characteristics
- Gate/stand occupancy time per flight
- Scheduling practices

Gate/stand occupancy time has a significant impact on overall airport gate/stand capacity and depends on many factors, such as the following among others:

- Aircraft type
- Passenger volume per flight
- Amount of cargo per flight
- The efficiency of ground handling operators and apron personnel
- Exclusive use by a particular airline or by any airline

Many of existing airports have outdated layouts that were built to consider size and crosswind characteristics of aircraft of many decades ago. To close the capacity gap and meet future air traffic demand, airports need to rebuild and expand their infrastructure. However, for many of them, it is physically not possible to build new runways owing to proximity to urbanisation areas and limited available land, so they must look for the alternative solutions. Some measures could temporarily increase runway capacity. For instance, synchronising arrival-departure sequencing, reducing runway intersections, crossings, could increase runway capacity by at least 20% for some airports [19].

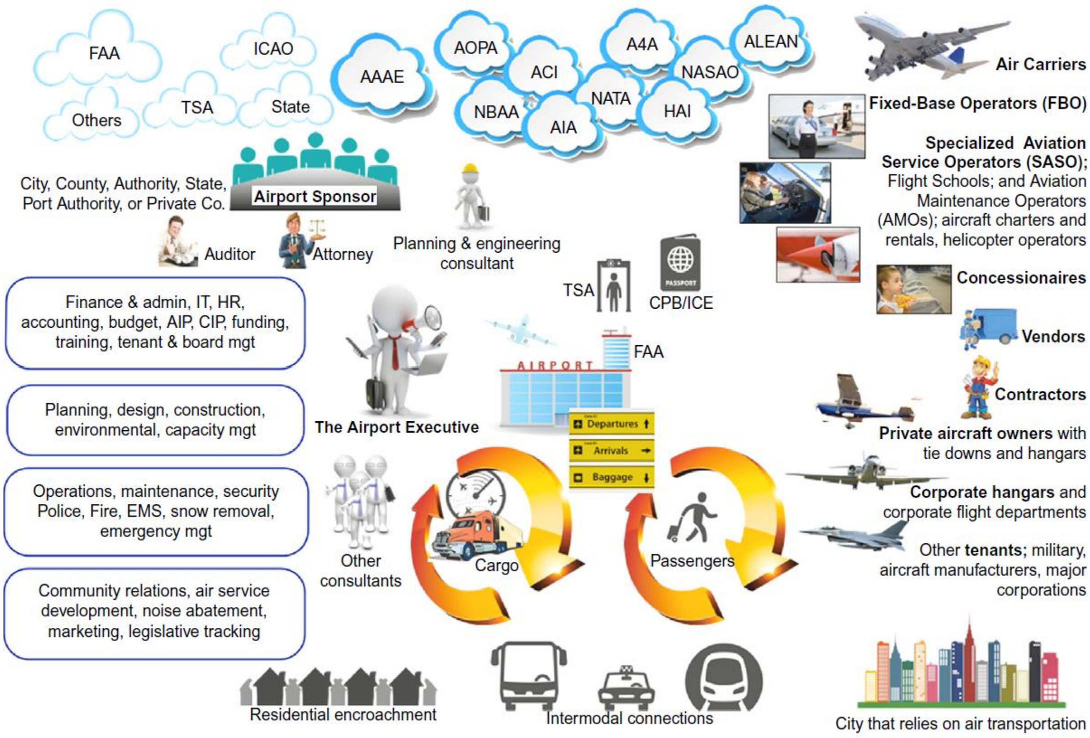


Figure 1. The complexity of the airport system [20].

EUROCONTROL identifies that successfully implemented optimisation of the use of existing airport capacity and initiatives like SESAR in Europe and NextGen in the United States could help reduce the capacity gap by 28% by 2040 [12]. Improvement projects (like EARTH [21], EAD [22] in Europe, AIP in USA [23]) and industrial transformations have already started, changing standards, legislation, technologies, and way of operations around the world. Their common goal is to improve safety, predictability, and resilience of airport operations, which means that many airports need to enhance their turnaround process. Some work has been already done towards these improvements in initiatives like Airport Collaborative Decision Making (A-CDM), which has been already implemented in 26 European airports [24]. Nevertheless, as the delay statistics show, there is still much improvement needed in the efficiency of airport operations.

One of the operations that suffer most from air network delays is the airport stand management. Airport managers must frequently iterate stand allocation plans to deal with the arising stochasticity and consequences of airside control procedures. Furthermore, traffic growth, sustainability regulations, and application of runway capacity enhancement measures, such as gate-holding and departure sequencing, add extra pressure to the airport stakeholders. Deteriorating airport performance can lead to additional fuel burn and emissions, a decrease of passenger service quality and disruptions that could propagate further in the air transport network. Thus, to ensure the qualitative performance of stand management, it is necessary to address the stand assignment, considering the efficient use of airport capacity, sustainability, and punctuality goals.

1.2 The stand allocation problem

A stand assignment typically affects three main stakeholders: airport operators, airlines, and passengers. For airport operators, it is vital to provide the required service level to the passengers and airlines and maximise revenue, while minimising costs of operations. Airlines usually seek for short ground-handling times, cost minimisation, and easy access to the terminal facilities. Passengers look for the convenience of transfer, punctual flights, and comfortable stay at the airport facilities. These interests are often contradictory and addressing them while ensuring the allocation of all aircraft to the suitable stands under stochastic conditions encapsulates the stand allocation problem (SAP).

In the literature and the scope of this dissertation, SAP is defined on assigning aircraft to airport stands. A stand is a parking space on the airport apron which allows storage and service of aircraft by the airport operators. Some sources call such spaces “gates” and refer to the SAP as the gate allocation problem (the GAP) [25]. Sometimes, GAP also refers to the flight scheduling problem, where flights must be allocated to the available boarding gates [26]. These boarding gates represent terminal facilities used by the passengers to embark and disembark from the aircraft. For clarity, this dissertation adopted the early definition of SAP proposed in [27], which focuses on aircraft parking positions - stands.

Often, the SAP must be tackled on several levels: seasonal, tactical, and operational [26]. First, the ability to accommodate flights from the proposed schedule must be examined during the seasonal flight schedule revisions. During this stage, planners decide to accept or decline the requests from the airlines to visit the airport. Accepted requests are later transformed into a flight schedule. On the tactical level, given a current flight schedule, daily plans must be developed before the upcoming operational day. Lastly, on the operational level, the created assignment often must be altered to accommodate schedule updates and operational disturbances during the day (reactive scheduling or reassignment).

The SAP is similar to the job-scheduling problem [28] and can be defined on a set of m stands and a set of k aircraft [25]. Usually, each aircraft $n = 1, \ldots, k$ serves two consecutive flights, so its arrival and departure times correspond to the arrival and departure times of those flights. Let the variables a_n and d_n denote arrival and departure times of aircraft n. If assignment variable x_{ni} has a value 1 when aircraft n is assigned to stand i and 0 otherwise; a feasible stand assignment must satisfy the following constraints [25]:

$$\sum_{i=1}^{m} x_{ni} = 1 \qquad (1)$$

$$x_{ni}x_{ji}(d_j - a_n)(d_n - a_j) \leq 0 \tag{2}$$

$$x_{ni} \in \{0; 1\} \tag{3}$$

In this formulation, constraint (1) guarantees that there are no unassigned aircraft. Constraint (2) ensures non-overlapping assignments. For instance, if aircraft n and j are both assigned to stand i, aircraft n must depart before the arrival of aircraft j or vice versa. Besides, there can be other constraints for stand assignment feasibility depending on the characteristics of a specific airport. For instance, some stands might be equipped only for certain aircraft types or be in a lease by a specific airline. Depending on the airport, the number of assignment constraints can reach hundreds and consideration of all existing requirements to match with arriving and departing aircraft can become a very challenging task, especially in the hub airports with hundreds and thousands of flights per day. Owing to this real-life quantity of constraints and decision variables, SAP is considered to be an NP-complete problem [29], which means that the task of stand allocation is too complex to be solved manually in an efficient way. Thus, for creating stand allocation complying with all requirements and avoiding errors, a body of literature on solving SAP has been developed over the last decades.

1.3 Scope of the dissertation and contributions

The focus of this dissertation lies within a problem of tactical stand allocation (further referred to as SAP). In the context of this work, the term *stand* means a physical airport asset that is used for parking an aircraft for ground handling-related services.

Although SAP has been studied for decades, several issues have not been efficiently addressed in the literature. This dissertation is focused on two of them:

- Air transport suffers from flight delays, which propagate through the network and deteriorate airport performance. Existing delay mitigation measures, such as buffer times insertion, reduce airport stand capacity and increase assignment complexity, which can be problematic in the congested airports [30].
- Air transportation aims for carbon-neutral operations to mitigate climate change. Nevertheless, existing stand assignment approaches neglect differences in aircraft fuel burn rate and emissions toxicity, failing to provide airport stakeholders with realistic stand allocation footprint.

Seeing these problems, this dissertation aims to answer the following research question:

How can an airport stand allocation be improved to address better the existing performance, capacity, and pollution problems, considering the interests of various airport stakeholders?

This dissertation tackles these problems with a methodology that can facilitate efficient stand capacity management and contributes to state of the art in the following way:

- The developed stand allocation framework uses historical data for inference of potential delays and creates an assignment schedule that considers these deviations in the stand occupancy times. Such a feature allows generating a robust stand assignment, where the range of considered disruptions can be based on stakeholders' risk acceptance level.
- The presented framework considers emissions toxicity, fuel burn rates, emission factors, and taxi time of each aircraft to estimate the level of emissions produced during the taxiing phase in a realistic way.
- Emissions mitigation goal is included in the stand allocation optimisation objective, which facilitates airport stakeholders with a way to balance environmental protection, efficient use of stand capacity, and passenger service.
- The simulation is used for the consideration of real-life stochastic events that are not captured by the stand assignment algorithm. The information obtained from simulation experiments facilitates better-informed stand capacity management.
- The formulation of the developed methodology allows including extra allocation constraints to the assignment generator of Module II, as well as selection between two multivariate optimisation goals, which provides airport stakeholders with the flexibility to tackle different assignment priorities.

Research presented in this dissertation proposes a framework to address SAP as a stand-alone problem neither considering arrival and departure sequencing, nor collaborative departure management.

1.4 Dissertation structure

This dissertation continues as follows. Chapter 2 presents the body of literature on the topic of stand and gate allocation and reviews methods and scopes of the existing solution approaches. Then, Chapter 3 introduces the methodology of the developed disruption-aware stand assignment framework. The case-study application results and their dissemination in scientific publications are described in Chapter 4. Chapter 5 and 6 present the principal publications that constitute the compendium of papers and describe the developed methodology and its applications on the case-study airport. Conclusions and further research are discussed in Chapter 7.

2 State of the art on the gate and stand allocation

The stand allocation problem and the gate allocation problem (the same theoretical approaches used, so here are discussed as a similar problem) have been a topic of academic research for many decades. From a mathematical view, GAP/SAP has been formulated in many ways, such as integer, binary (BIP), or mixed-integer (MIP), general linear or nonlinear models. Specific formulations as binary or mixed binary quadratic models have also been suggested. Other well-known related problems in combinatorial optimisation such as quadratic assignment problem (QAP), clique partitioning problem (CPP), and scheduling problem have been used to formulate GAP/SAP. Periodical reviews of GAP/SAP formulations can be found at [25], [31].

The SAP solution approaches can be generally divided into two groups: expert systems and optimisation approaches [32]. An expert system is a software system that enhances the human expert's performance. Optimisation approaches mathematically define an objective function and maximise or minimise its value subject to certain constraints. In their turn, optimisation approaches can be divided into categories of exact, heuristic, and combined methods. Figure 2 illustrates an overview of the solution methods with some commonly applied algorithms [32].

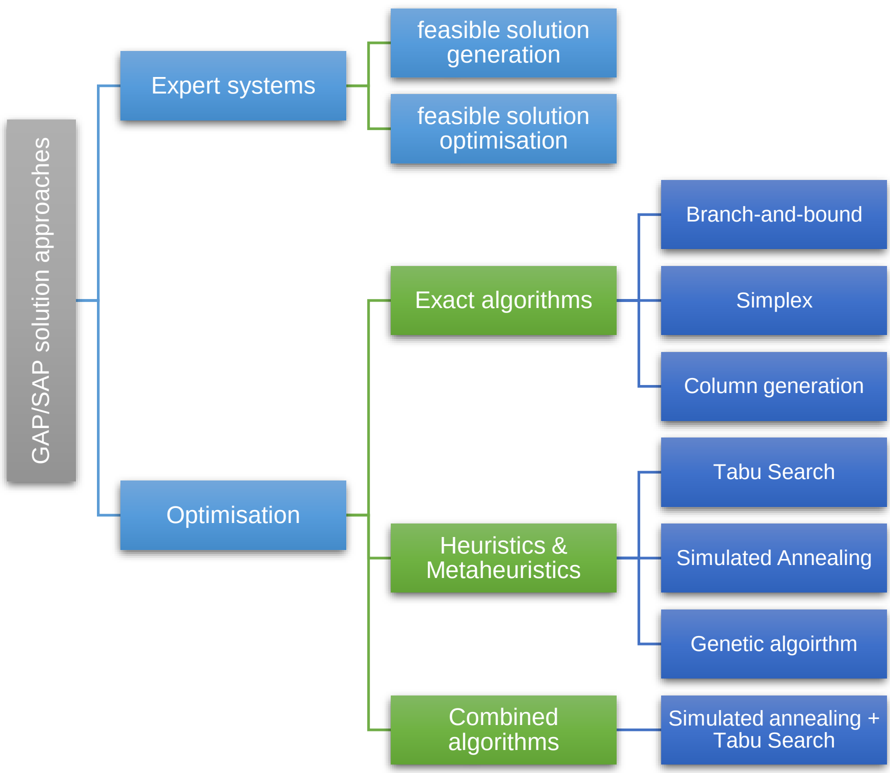


Figure 2. Methodological approaches and common algorithms for solving GAP and SAP

Exact algorithms aim to find an optimal solution from a mathematical point of view; however, these algorithms do not work well for large-scale problems due to longer computational time needed to the find the mathematical optimum [29]. Heuristic algorithms, theoretically, have a chance to find an optimal solution. This chance can be relatively low as heuristics often get stuck at local sub-optimum, so metaheuristics were developed to increase the chances of finding a globally optimal solution. In a real-life operational decision making, time is often a crucial matter. Hence, a quickly obtained near-optimal solution could be accepted, especially if it complies with various allocation constraints and balances assignment costs. Heuristic and metaheuristic algorithms provide good results in a reasonable time and thus, have become the most popular methodologies for solving GAP/SAP in the past years [25].

2.1 Expert systems

Expert systems commonly generate solutions using databases which contain rules developed by human experts in a specific domain. An expert system operator may adjust the existing rules and add new ones if necessary, to improve the system's ability to generate solutions with better quality.

In 1980-s, Brazile and Swigger [33] were one of the first to develop a constraint-satisfaction allocation expert system, which consisted of two levels. The first level generated an initial schedule, and the second level modified the schedule considering flight delays, weather changes, and facility failures. The authors of this expert system used knowledge of experienced airport managers to develop allocation rules and restrictions. As a result, the system was producing feasible solutions in a matter of seconds.

Later, Gosling [34] developed a system that considered available personnel and could also adjust gate assignments to schedule deviations and equipment failures. Furthermore, this system considered consequences of allocation decision on downstream operations. A similar approach was used by Srihari and Muthukrishnan [35] and Su and Srihari [36] in their version of the gate assignment advisor. The latter integrated their knowledge-based expert system with the airport operational database.

Cheng [37] was one of the first to add an optimisation functionality to a gate assignment expert system. In his implementation, aircraft were assigned in groups under multi-objective optimisation function, which considered the assignment cost, waiting times and use of contact stands. Jo et al. [38] developed a ramp scheduling system called RACES. This system divided the overall problem into sub-problems and solved them independently by trade-off scheduling method under the constraints stored in a knowledge database. The resulting near-optimal solution was obtained in a matter of seconds for a case study of 400 flights. Lam et al. [39] combined knowledge-based system in the form of an intelligent agent with an optimisation model and developed a tool that could respond to real-time changes in the gates and flights.

Nowadays, many commercial software companies offer decision-support systems for airport operations (e.g., AIS, AirTOP, Transoft, INFORM, and CAST). These frameworks typically contain a gate/stand management module for airport planners to organise the allocation of aircraft to the gates/stands. These software developers claim that generated solutions consider all allocation rules and restrictions, such as airlines preferences, ground operations, and equipment characteristics [40], [41]. They also assure that last-minute changes and disturbances can be handled as well. Nevertheless, the developers of these commercial solutions do not disclose details of their algorithms in the literature.

2.2 Static solutions

Earlier GAP/SAP optimisation approaches mainly focused on individual stand/gate allocation objectives and considered a static flight schedule with no deviations. Minimisation of passenger walking distance and time has been approached by Babić et al. [27] and solved by backtracking branch-and-bound algorithm. Mangoubi and Mathaisel [42] proved computational time advantages of greedy heuristics for minimisation of passenger walking distances compared to linear programming relaxation of an integer formulation. Vanderstraeten and Bergeron [43] applied heuristic maximisation of aircraft allocated at the contact stands. Xu et al. [44] minimised the overall time during which passengers walk to catch their connection flights with a Tabu Search meta-heuristic. Jaehn [45] maximised flight/gate preference scores by decomposing the problem into subproblems of smaller time intervals. Cheng et al. [32] proposed to use a combination of meta-heuristic algorithms Tabu Search and Simulated Annealing for solving the gate assignment problem while minimising passenger walking distance.

Other researchers combined different optimisation perspectives to reflect various interests of agents involved in the stand assignment. Thus, Jo et al. [38] minimised towing cost and the number of waiting aircraft by breaking down the continuous-time values into discrete as periods of arrival and departure per flight. Yan and Huo [46] developed a dual-objective gate assignment model, which considered the minimisation of both the passenger walking distance and their waiting time and solved it separately by column generation, simplex and branch-and-bound algorithm. Hu and Di Paolo [47] minimised transfer and baggage distances with genetic algorithm. Drexl and Nikulin [48] minimised passenger walking distances, number of ungated flights and maximised airport preference score of assigning individual aircraft to particular gates and solved with Pareto simulated annealing. Marinelli et al. [49] minimised walking distance and ungated flights for both passenger and cargo flights by fast converging Bee Colony optimisation. Guépet et al. [29] solved GAP with mixed-integer programming and time and spatial decomposition, greedy and ejection chain algorithm, maximising the number of passengers/aircraft at contact stands and minimising the number of towing movements. Ding et al. [50] minimised the number of ungated flights and the total walking distances by a greedy algorithm and use of Tabu Search meta-heuristic. Ding et al. [51] minimised the number of ungated flights and total walking distances or connection times for the over-constrained gate assignment problem with simulated annealing and a hybrid of simulated annealing and Tabu Search. Benlic et al. [52] considered nine objectives in the form of a weighted sum of costs associated with them and used Breakout Local Search to solve the model. Behrends and Usher [53] applied a genetic algorithm to minimise taxi and passenger movement times in their job shop scheduling model.

Although the mentioned works considered realistic optimisation objectives and problem sizes, they neglected one of the significant airport management problems – the stochastic flight schedule deviations, which limits the application of these approaches to the real-life operations.

2.3 Stochastic and robust approaches

Stochastic flight schedule deviations on the day of operations can significantly disrupt allocation plans and can inhibit providing a required level of service to passengers and airlines. Therefore, recent research has been more concentrated on increasing robustness of gate/stand allocation to the possible schedule deviations. In the context of gate/stand assignment, *robustness* means the ability of an allocation plan to remain feasible under minor disturbances in the flight arrival and departure times.

One of the most popular measures to increase stand allocation robustness is the insertion of buffer times between consecutive flights/aircraft assigned to the same gate/stand. Buffer time refers to a planned time interval during which the gate/stand is always kept empty in the plan between two consecutive flights. Figure 3 illustrates this concept applied in the assignment plan. Buffer time insertion has been demonstrated to improve schedule punctuality [54] and to successfully absorb minor stochastic flight delays [46], [55].

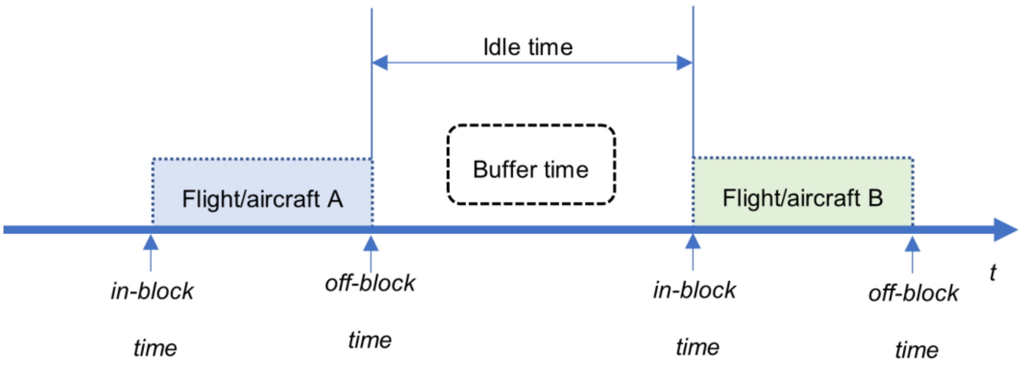


Figure 3. Illustration of idle and buffer times in the gate/stand assignment plan

Mangoubi and Mathaisel [42] were one of the first to use fixed buffer times for improving robustness while minimising passenger walking distances. Bolat [56] proposed to distribute idle times uniformly over the gates. Idle time here refers to a time interval between two successively assigned flights during which the gate is not used. Similarly, Diepen et al. [57] maximised idle times between each pair of consecutive flights with column generation algorithm.

Lim, Rodrigues, and Zhu [58] considered minimal flight deviations and proposed a gate scheduling model which was based on changing flight arrival and departure times. They used Tabu Search and genetic algorithm to solve the problem. Yan and Tang [59] tested different uniform buffer times for all planned flights and real-time reassignment rules in their penalty-based heuristic planning framework. They demonstrated that buffer times allow increasing robustness compared to manual assignments; however, their effect was limited under the conditions of high delays. Dorndorf et al. [60] also used uniform buffer times in their study of clique partitioning problem solved with an ejection chain algorithm to maximise the robustness of the resulting schedule to flight delays and the total assignment preference score and to minimise the number of unassigned flights and the number of tows. Şeker and Noyan [61] suggested defining optimal buffer times with mixed-integer programming (MIP) and solved the problem with Tabu search heuristic. Maharjan and Matis [62] also used buffer times and proposed a binary integer multi-commodity gate flow network model to minimise the fuel burn cost of aircraft taxi and a function of inter-gate distance and passenger connection time. Guépet et al. [29] used buffer times to mitigate schedule disruptions and developed a MIP gate assignment model with the objectives to maximise the number of passengers/aircraft at contact stands and minimise the number of towing movements. Yu et al. [63] applied buffer times and tackled minimisation of waiting time for arriving aircraft, transfer passenger walking distances and towing costs through MIP-based heuristics. They showed that long separation time between successive flights assigned to the same gate increases allocation robustness to schedule deviations; however, it requires more towing movements to ensure sufficient time gaps between the flights. Deng et al. [64] incorporated buffer times and minimised total time for passengers and balanced idle time for each gate by translating the multi-objective multi-commodity network flow model into the single-objective one.

Other researchers incorporated historical stochastic behaviour of the flights based on fitted probability distributions. Thus, Wei and Liu [65] proposed a fuzzy model with idle times as fuzzy variables and the objectives of minimising the total walking distance for passengers and maximising the robustness of assignment. Li [66] considered probability distribution functions on gate conflict between two aircraft and the objective to minimise the number of gate conflicts of any two adjacent aircraft assigned to the same gate, and to minimise the number of gates that airlines must lease or purchase for the smooth operation. Kim et al. [67] proposed a multi-objective model with goals to minimise the transit (walking) time of all passengers, taxi time (weighted to the number of passengers), and the duration of expected gate conflict, which was estimated based on historical delay probability distributions. Castaing et al. [68] also considered specific delay distributions in their model to minimise the expected impact of gate assignment conflicts. Prem Kumar and Bierlaire [30] proposed to improve gate allocation robustness by increasing buffer times on the percentile of the historical delay value and simultaneously addressed a multi-objective combination of passenger connection revenues and zone usage costs. Yu, Zhang, and Lau [63] proposed a robust gate assignment MIP-based heuristic model considering costs of gate conflicts, facility and personnel cost during tows, and passenger satisfaction level, where flight delay distributions were fit to the historical data and incorporated into buffer times. Van Schaijk and Visser [69] included regression models to generate probable flight delays based on historical data and minimised aggregated airline and airport cost of assigning flights to gates in their BIP model. Dijk et al. [70] also included probability distribution to enhance the effectiveness of buffer times between flights in a BIP model to minimise passengers' walking distance, tows, and to maximise the number of passengers allocated to contact stands and the potential airport commercial revenue.

Another way of dealing with stochastic flight delays was proposed in reassignment methodology, which is designed to cope with last-minute changes by adjusting the assignment during the operational day. Tang et al. [71] developed a gate reassignment framework and a tool that could replace manual reassignment process. Wang et al. [72] proposed a real-time reassignment model that deals with specific and uncertain flight delays and minimises their impact on the gate assignment based on the ant colony algorithm. Ali et al. [73] developed a passenger-centric model that analysed the impact of turnaround times, minimum connection times and stochastic delays on missed connections of self-connecting passengers, where real-time reassignment of gates was aiming to minimise spatial deviation from the optimised gate assignments.

With the growth of global awareness on ecological problems, more studies on airport environmental footprint appeared in the past years. The principal focus of such studies lied in the reduction of aircraft fuel consumption, which is the primary origin of aviation greenhouse gas emissions [13], [14]. Idle and taxi phases of aircraft movement were estimated to be the primary sources of fuel consumption and emissions at the airports [74]. Therefore, many researchers attempted to optimise operations in these two phases. Duinkerken et al. [75], Ithnan et al. [76], and Li and Zhang [77] suggested different taxiing modes, which included taxiing on one aircraft engine and using external engine power, for reduction of taxi-related emissions. Other researchers focused on scheduling aspects; however, the stand assignment was often omitted in their studies. Hence, Brinton et al. [78] developed a collaborative departure planning tool that also considers the emission level. Monroe et al. [79] estimated the environmental effects of eliminating short stop operations at active runway crossing. Sölveling et al. [80] proposed optimised scheduling of runway operations with consideration of aircraft fuel burn. Simaiakis et al. [81] proposed an optimised pushback rate control, which minimised congestion and idle waiting at the runway entry. Simaiakis and Balakrishnan [82] optimised queuing of the departure runway system based on pushback schedule and estimation of unimpeded taxi-out time distributions. Gate holding, de-rated take-offs [83], and departure metering [84] were also estimated to reduce aircraft emissions successfully. Zhang et al. [85] optimised aircraft taxi time and taxi emissions by considering taxiway conflicts and aircraft fuel consumption. Bertsimas and Frankovich [86] developed an airport operations model, in which gate assignment, taxiing, departure sequencing, and aircraft routes in the near-terminal airspace were optimised.

Overall, currently available GAP and SAP solutions, in their majority do not consider the environmental footprint of stand assignment operations. Those researchers who propose airport emissions mitigation, often focus on optimisation of runway capacity, throughput, and impact of waiting at the runway queue, often neglecting the interests of passengers, airlines, and other airport stakeholders. Furthermore, they often propose methods like gate-holding that increase stand occupancy times and reduce stand capacity.

3 Methodology

This chapter describes the developed disruption-aware methodology for tactical stand assignment that encompasses multi-criteria optimisation and robustness goals. This novel methodology is further referred to as DASA, which stands for a disruption-aware stand assignment. DASA architecture and functionality of each of its components are described in the next sections.

3.1 DASA framework architecture

The proposed DASA framework consists of two computational modules coupled with a simulation model that work together to consider stochasticity of operations and airport characteristics. Each component performs a designated functionality that addresses a specific type of airport stand management issues.

The design of the DASA is presented in Figure 4 and can be described as follows. Module I is a look-ahead component that analyses historical performance data and uses Bayesian modelling to construct probabilistic models describing the levels of schedule deviations. Module II generates an optimised stand assignment considering user-defined input data, restrictions, and optimisation priorities.

The framework workflow starts with the import of historical airport performance data to Module I in a table format. These data can include among others records for scheduled and actual in- and off-block times, flight arrival and departure times, weather conditions at the time of operations, characteristics of aircraft used, and local air traffic regulations. Next, the analytical component of Module I investigates the imported data and builds Bayesian distributional models for schedule deviations, based on the evidence present in the data. These models, together with the discovered interdependencies and corresponding parameters (regression coefficients), are then sent to Module II.

In Module II, if the user has specified the option of generating a robust allocation in the input, the target flight schedule is recalculated considering the information from the air traffic control and possible block occupancy time deviations. These block time deviations are calculated with an account of historical disruptions, based on Bayesian models from Module I and considering a user-defined delay probability level if specified. In the scope of this research, these estimated block times are called *probabilistic in-block time* (further referred to as PIBT) and *probabilistic off-block time* (referred to as POBT); total block time – *probabilistic block occupancy time* (referred to as PBOT).

As the next step, a recalculated flight schedule, where scheduled block times are replaced with PBOT, is passed to a metaheuristic optimisation algorithm. This algorithm performs a search for better stand assignment for the target flights, considering user-specified optimisation priorities expressed in a multivariate objective function. The algorithm was implemented with two variants of the primary optimisation function, described in section 3.3.1 and section 3.3.2.

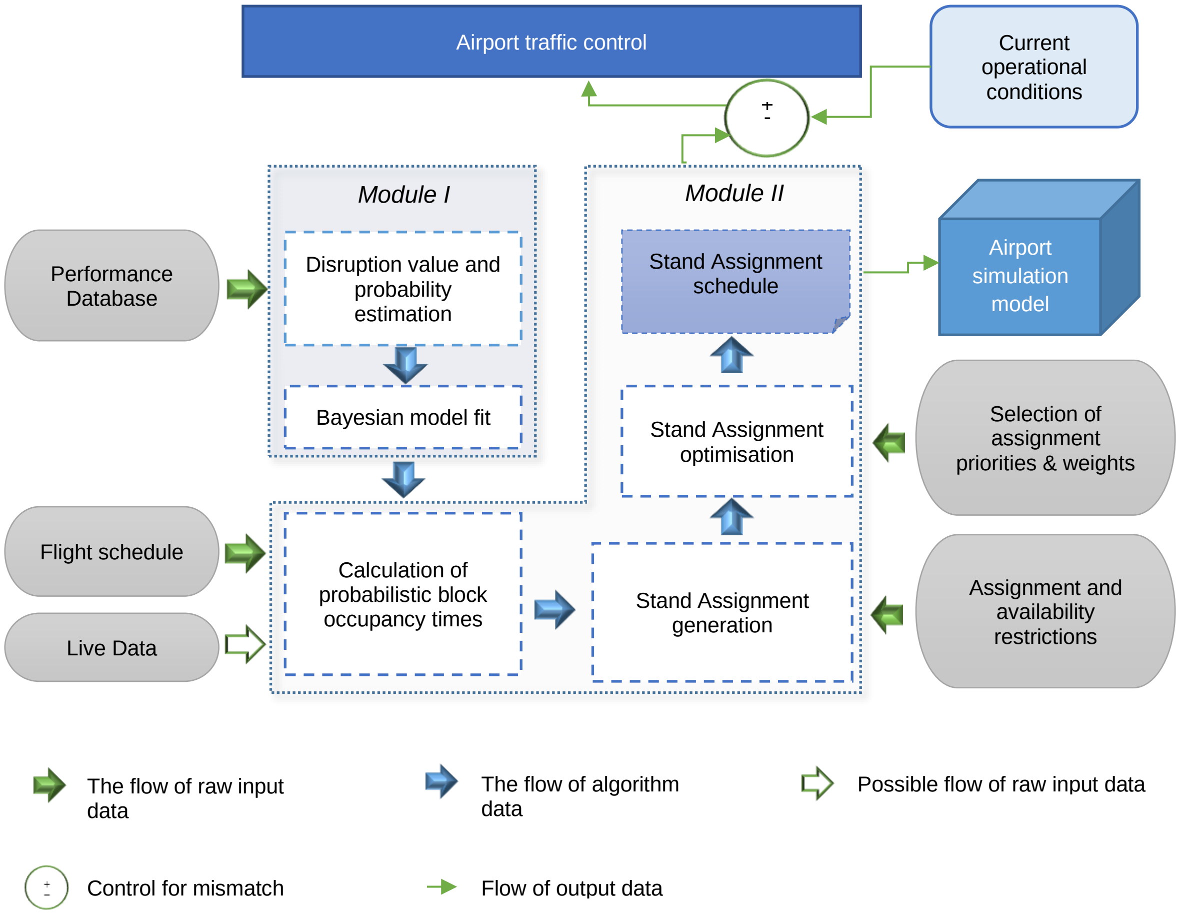


Figure 4. DASA framework architecture and primary data flows

The output of Module II can be further transferred to an airport simulation model to explore the quality of the generated solution in close-to-reality conditions. The simulation experiments can help to evaluate the created stand allocation for the various what-if scenarios, reflecting different airport performance situations. Elements of the developed algorithmic framework are described in the following sections.

3.2 Module I: modelling disruptions

The principal function of Module I is to explore historical performance data and build models that can predict possible schedule deviations based on the flight and airport environment characteristics. For this purpose, regression analysis can be used. However, exploring available historical data may require extended provisional analysis to select and test fitting of various regression models, and the outcome is strongly dependent on the expert's knowledge and experience in statistical modelling. To minimise the need of such time consuming and expertise-dependent analysis, Module I relies on the application of Bayesian modelling, which already proved its value for research in various scientific applications [4], [87]–[91].

One of the main reasons to use Bayesian modelling is that the data for the probability estimation can be used as it becomes available, so the models can be easily updated even after each operational day if such desired. Also, thanks to the Bayesian method, the likelihood estimations can be set to be independent of "outliers" or extreme data values influences through the incorporation of the prior probability values, which makes it a perfect approach for consideration of hidden, latent correlations between different performance variables without overfitting the model on human-induced extremes.

Bayesian distributional models

Bayesian distributional models allow the modelling from different perspectives of measurement considering their complex dependencies at the same time. In the heart of such a model lies the prediction of the response variable *y* at the data point *i* through the linear combination η of predicting factors, transformed by the inverse link function *f* adopting a specific distribution *D* for *y*: $y_i D(f(\eta_i), \theta)$.

The variable θ describes additional distribution-specific parameters that typically do not vary across data, such as the standard deviation σ in normal models or the shape α in Gamma or negative binomial models. The linear predicting factor can generally be written as: $\eta = X\beta + Zu$. In this equation, β and u are the regression coefficients at population-level and group-level respectively, and X and Z are the corresponding design matrices. The response y, as well as X and Z, form up the data, whereas β, u, and θ are the model parameters estimated with various sampling algorithms [92]. In such a way, by calculating level-corresponding coefficients, it is possible to obtain a multilevel distributional regression model for the target response variable.

Estimation of the probability of observing data point *i*, based on the evidence present in the data is done via calculation of a joint probability distribution for both the target variable and the set of its predictors using Bayes' rule. Following the Bayes' rule, the likelihood of observation *A*, occurring given the occurrence of observation *B*, can be written through the following equation: $P(A|B) = P(B|A)P(A)/P(B)$, where *P(B|A)* denotes the likelihood of *B* occurring when *A* occurs, *P(A)* and *P(B)* – are the probabilities of observing both observations independently of each other [93].

Bayesian modelling assumes that the model parameters can also be drawn from a distribution. In such a way, the posterior probability of the model parameters is linked to the evidence from the data as $P(d|y, X) = P(y|d, X) * P(d|X)/P(y, X)$, where $P(d|y, X)$ is the posterior probability distribution of the model parameters, given the input and output data, $P(y|d, X)$ is the likelihood of data, $P(d|X)$ is the prior probability of the model parameters, and $P(y, X)$ is a normalisation constant.

In case, if there is no prior knowledge on the possible model parameters, they are assumed to be drawn from a normal distribution. Nevertheless, the more data is available for fitting the model, the less influence such prior probabilities will have on the model quality.

Implementation

The Bayesian model learning in Module I is currently implemented via a connection with a free statistical tool R, particularly with its package *brms* [92]. This package allows to efficiently construct a Bayesian distributional model and export its parameters in any required form. However, as future work, the model building can also be implemented internally, as a part of the entire coding solution of the stand assignment framework for a real-time update of the model parameters with new performance data that becomes available during operations.

The output of Module I is a Bayesian distributional regression model, where for each predictor variable estimation of its effect on the response variable is made. With such a model, it is possible to generate future schedule disruption values and use them for calculation of PBOT for generating a disruption-aware stand assignment in Module II.

3.3 Module II: assignment generation and evolutionary optimisation

The primary function of Module II is to generate a stand assignment, considering the input constraints and disruption models, and then optimise it ensuring the better value of airport stakeholders' priorities expressed in a multi-objective function., the optimisation algorithm considers diverse perspectives, restrictions, and variables that can be important for airport stakeholders.

Specific restrictions for the stand assignment schedule can vary depending on the particularities of each airport. The following are restrictions implemented in the assignment generation:

- Domestic and international flights must be assigned to the specific gates. Usually, this depends on the internal specifications of the airport, e.g., international flights are usually allocated to gates that have access to the designated border control areas.
- Stand occupancy time for each aircraft is determined by its ground handling specifications and airline policy.
- No aircraft towing movements from one stand to another are considered in the current implementation of the algorithm. Each aircraft occupies its assigned stand for the time equal to its ground-handling time and then taxies to the runway for departure from the airport.
- A stand must correspond to the size of an aircraft (large aircraft require extra space due to larger wingspan). This is implemented through the identification of allowed stands for each flight on the stage of processing the input data in Module II.

- An assigned stand must correspond to airline preferences. This is implemented through the identification of preferred/contracted stands for each flight at the input data processing stage in Module II.
- When there are no stands available at the arrival, aircraft should wait until a position becomes available. This is implemented in the algorithm by assigning the flight to a "dummy" stand and incrementally delaying its PIBT until a suitable stand becomes available.

During the progress of the research, two variations of the main objective function were formulated. These formulations are discussed in the following sections.

3.3.1 Optimisation objective function: version I

The first formulation of the principal objective function can be described as in equation (4):

$$minimise \ \rightarrow \ F_I = w_1 * R_{open} + w_2 * R_{taxi} + w_3 * R_{hold} + w_4 * R_{service} \tag{4}$$

In this formulation, the following perspectives were considered:

- Airport management perspective R_{open}: to serve more passengers through the contact stands and minimise the use of open or remote parking positions:

$$R_{apron} = (Nfl|apron)/(TotalNfl) \tag{5}$$

where $Nfl \mid apron$ is the number of flights assigned to remote parking positions that are connected to the terminal building only via bus service; $TotalNfl$ is the total number of flights in the schedule to allocate.

- Airline and environmental perspective R_{taxi}: to minimise the taxi distance to the stand:

$$R_{taxi} = (AvSchTaxi)/(MaxAirportTaxi) \tag{6}$$

where $AvScheTaxi$ is the average taxi distance from stand to and from the runway in the allocated schedule; $MaxAirportTaxi$ is the maximum possible taxi distance at the airport for considered runway configuration.

- Air Traffic Control perspective R_{hold}: to minimise the number of aircraft waiting for stand availability:

$$R_{hold} = (Nfl|wait)/(TotalNfl) \tag{7}$$

where $Nfl \mid wait$ is the number of flights that must wait for the stand availability; $TotalNfl$ is the total number of flights in the schedule to allocate.

- Passenger comfort perspective $R_{service}$: to provide enough waiting space for passengers in the departure lounge:

$$R_{service} = (MaxAreaPax - ActAreaPax)/(MaxAreaPax) \quad (8)$$

where $MaxAreaPax$ is the maximum possible departure lounge area per passenger, calculated as the minimum number of passengers per flight divided by the area of the largest departure lounge in the airport; $ActAreaPax$ is the actual departure lounge area per passenger, available at the assigned gate for the assigned flight.

In equation (4), $w_{1,}, w_{2,}, w_{3,}$, and w_4 indicate priority weights for the corresponding perspectives. For practical implementations, the weights should be decided by negotiations of airport stakeholders. Prioritising one or more perspectives over the others may result in a certain cost for the neglected perspectives so that it can be used for estimation of various airport strategies and trade-offs between them.

3.3.2 Optimisation objective function: version II

The second variant of optimisation objective can be viewed as an emission-aware transfer passenger-centred modification. This change considers the environmental goal of modern air transportation and is defined as:

$$minimise \rightarrow F_{II} = w_{walk} * O_{walk} + w_{open} * O_{open} + w_{emis} * O_{emis} + w_{idle} * O_{idle} \quad (9)$$

In equation (5), the following perspectives were considered:

- Passenger service perspective O_{walk}: to minimise total walking distance for transfer passengers:

$$O_{walk} = \sum_{i=1}^{I} N_{paxi}\, d_{walk} \Big/ \sum_{i=1}^{I} N_{paxi}\, d_{maxwalk} \quad (10)$$

where N_{paxi} is the number of transferring passengers per i flight, d_{walk} is the walking distance to a connecting flight; $d_{maxwalk}$ is the walking distance between two gates located the furthest from each other, and I is the total number of flights with transfer passengers.

- Airport management perspective O_{open}: to serve more passengers through contact stands:

$$O_{open} = (N_{po} * N_{open}) \Big/ (N_p * N) \quad (11)$$

where N_{po} is the number of passengers in the aircraft assigned to remote stands, N_{open} is the number of aircraft assigned to remote stands; N_p is the total number of passengers on scheduled flights, and N is the total number of aircraft in the schedule.

- Environmental perspective O_{emis}: to minimise tax-related pollutant emissions:

$$O_{emis} = \sum_{n=1}^{N} \sum_{e=1}^{E} B_n H_e F_{ne} (T_n + DT_n) \Big/ \sum_{n=1}^{N} \sum_{e=1}^{E} B_n H_e F_{ne} (T_{hold} * N) C_t \tag{12}$$

where B_n is the fuel burn rate for aircraft n; H_e is the hazard weight assigned to the emission e; F_{ne} is the emission factor e for aircraft n per unit of fuel burnt; T_n is the taxi time for aircraft n; DT_n is the time penalty if aircraft n is assigned to a "dummy" stand; T_{hold} is the holding manoeuvre time; N is the total number of aircraft in the schedule; C_t is the holding emission factor increment, calculated as $C_t = f_{appr}/f_{taxi}$, where f_{appr} and f_{taxi} are the engine thrust levels for the approach and taxi phases, respectively. In practice, airport stakeholders can choose the values of H_e to emphasise the impact of certain pollutants according to their toxicity level.

- Air Traffic Control perspective O_{idle}: to minimise the number of aircraft waiting for stand availability:

$$O_{idle} = N_{idle} / N \tag{13}$$

where N_{idle} is the number of aircraft that have been assigned to a "dummy" stand and N is the total number of aircraft in the schedule.

- Variables $w_{walk}, w_{open}, w_{emis}, w_{idle}$ indicate priority weights for the corresponding perspectives, which can be decided upon by airport stakeholders.

The objective functions presented in equation (4) and equation (9) have conflicting goals owing to the nature of the stakeholders involved. Thus, for airlines, it can be cheaper to use remote stands; however, for airport managers, it can be more profitable to offer contact stands for the service. At the same time, passengers would prefer to have shorter walking distances and comfortable waiting for departure. However, such stands could be located relatively far from the runway exit and have a longer taxi distance to them. Such conflicting nature of different perspectives on the stand assignment is challenging to balance; therefore, it has been expressed as a multi-objective function to facilitate harmonisation of these contradicting goals and provide an estimation tool to compare various trade-offs in the assignment.

3.3.3 Evolutionary optimisation

After updating the target flight schedule with PIBT and POBT, an evolutionary optimisation algorithm performs generation and optimisation of the stand assignment according to the chosen multi-criteria objective (equation (4) or equation (9)). The algorithmic implementation is done in the form of a genetic algorithm [94], which has been successfully applied in solving SAP/GAP and many other air transport optimisation problems [65], [95]–[97].

One of the most important reasons for selection of genetic algorithm (GA) among other types of solution search algorithms is its ability to escape local optima by increasing the diversity of solutions, which for multi-objective optimisation is an important feature. Besides, GA implementation is relatively straight-forward and is easily adaptable for SAP formulation.

In the GA implemented in Module II, the stand assignment allocation is represented as N x M dimensional array, where N is the number of flights in the schedule; M is the number of various flight and aircraft characteristics to be considered in the assignment generation. Figure 5 illustrates an example of a chromosome's gene, representing a part of a stand assignment solution in GA.

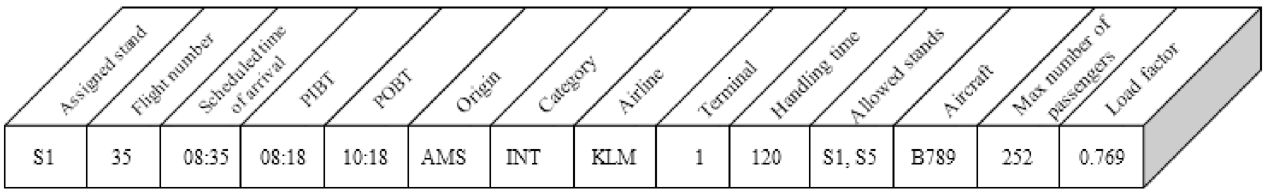


Figure 5. Example of a chromosome's gene content

The chromosome itself is an array of arrays, as shown on Figure 6, where each gene (cell) of a chromosome contains an assigned stand number and an array of information that can be used by the algorithm. For instance, each flight would contain an array of stands allowed for this flight and aircraft type. Similarly, other characteristics can be appended to the array of each flight.

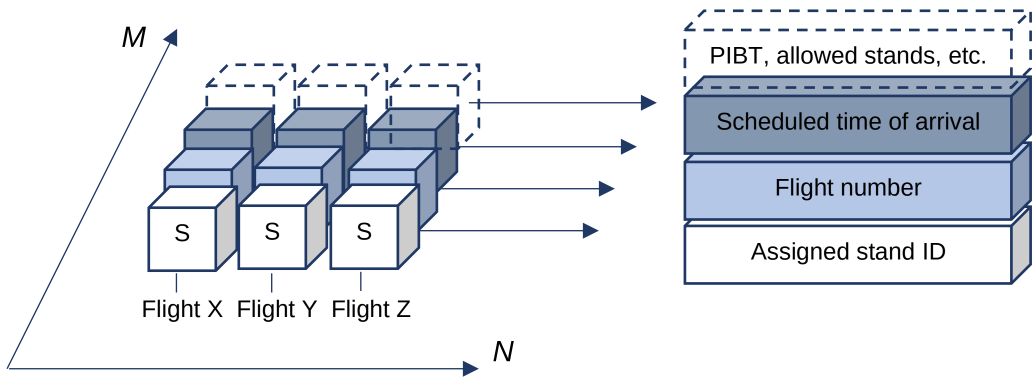


Figure 6. Stand assignment schedule coded for a genetic algorithm in Module II

The algorithmic process and operations are presented in Pseudocode 1. The workflow of the algorithm starts with importing target flights schedule, allocation constraints, Bayesian models from Module I, and user-defined priority weights for the objective function and other relevant variables (e.g., disruption probability, robust or static allocation, runway configuration for arriving and departing aircraft). These input data are used to generate an initial stand assignment solution (referred to as Adam chromosome in Pseudocode 1). Then, a set of new solutions is generated by making random changes in the copies of the initial allocation. After that, each solution is checked for feasibility (e.g., not more than one aircraft assigned to one stand at any given time) and updated if necessary. This step allows ensuring that only feasible solutions will be evolving through algorithm iterations and that the final solution will be feasible as well.

After the check for feasibility, the value of the objective function defined according to Equation 1 or Equation 2 is calculated for each generated solution (chromosome). A chromosome with the best (lowest) value is marked as the best solution. Next, the set of chromosomes overgoes through crossover, where some chromosomes randomly exchange their genes (assigned stands) with each other. Each chromosome has a 75% chance of exchanging its genes with another chromosome. Next, some chromosomes are subjected to a random change of some of their genes, which is referred to as *Mutation* in Pseudocode 1. The probability of mutation is determined for each chromosome and is equal to 10%.

```
GET Stop_Criteria
IMPORT
    FlightSchedule,
    Constraints,
    Priorities,
    Module_I.output
CREATE
    Adam chromosome, A
GENERATE
    Set(chromosomes), S = RandomChange(A)
WHILE CurrentSituation < > Stop_Criteria
    FOREACH  X  IN S DO
        EnsureFeasibility (X)
        Calculate objective function F(X)
        IF value F(X)> Best_Val THEN
            Best_val = value F(X)
            Best_Chromosome = X
        IF CurrentSituation = Stop_Criteria
            THEN STOP
        ELSE DO Crossover(Xi, Xj)
        IF MutationChance = TRUE
            Mutation (X)
EXPORT Best_Chromosome
```

Pseudocode 1. An evolutionary algorithm for stand allocation

As the next step, the entire set of chromosomes is re-evaluated for feasibility, and the corresponding objective function values are calculated. If the best new value is lower than the previous best, the marking of the best chromosome is updated accordingly. Then, the algorithm evaluates if the user-defined stopping criteria has been reached, and if true, the best stand assignment solution is exported in a format appropriate for the further use and saved on a local drive.

As a stopping criterion, a user can define a maximum number of algorithm iterations, a minimum value of the objective function, or total running time. Therefore, improvement of the solution quality is only restricted by the airport assignment constraints and user preferences.

The output of Module II can be further imported into an airport simulation model to explore the robustness of the generated solution and test the impact of various operational disruptions on the assignment quality and resilience. The current implementation of Module II is done in C#, which makes it possible to integrate the algorithm into various commercial general-purpose simulators (e.g., SIMIO).

3.4 Simulation

In recent decades, simulation has become a popular way of representation and studying of the complex dynamic systems. Generally speaking, a *simulation model* refers to a mathematical and/or digital representation of a real-life system [98]. Simulation models can provide knowledge about the behaviour of complex systems in time, which can be challenging to capture analytically. In the developed stand assignment framework, simulation provides the following added values:

- Simulation allows testing potential solutions in a close-to-reality environment.
- With simulation, emerging dynamics, such as runway and taxiway congestions, can be captured and their impact on the stand capacity management can be observed and estimated.
- Different disruptions can be incorporated into the simulation model to test the resilience of the system and stand assignment solutions.
- Various assignment policies can be investigated and compared in a simulation model without compromising the operations of a real-life airport.
- Although arrival and departure sequencing are not considered in the stand assignment generation, their influence on aircraft movement can be explored in the simulation model.

4 Overview of applications and experimental results

This section describes the results of the application of the developed DASA methodology, disseminated in the peer-reviewed publications and complimentary articles. DASA approach was published in several journal papers; in each of them, the simulation was used to validate the methodology in the close-to-reality conditions of a real airport.

The idea to predict schedule disruptions and use them for enhancing airport operations originated from the research performed for the project *"Airport Improvement Research on Processes & Operations of Runway, TMA & Surface"* (AIRPORTS). One of the goals of AIRPORTS was to develop a holistic airport performance monitoring framework which would facilitate analysis and prediction of the airport's KPIs based on historical data. In the scope of this project, a concept of Bayesian modelling has been investigated and evaluated for prediction of airport performance indicators. The first results of this evaluation have been published in a conference paper *"Modelling Dependence of Arrival Sequencing and Metering Area"* [3] and in a journal paper *"Identifying and modelling correlation between airport weather conditions and additional time in airport arrival sequencing and metering area"* [4]. In these papers, Bayesian modelling was used to predict the additional time (a delay) that an aircraft spends in arrival and sequencing metering area (ASMA) of an airport. The dataset containing weather conditions and the target variable – additional ASMA time, was used for building a predictive model, with all variables discretized in several levels to facilitate the modelling. The ASMA time model was further formulated and experimented with in a Coloured Petri net simulation environment [99]. The obtained Bayesian model was able to give relatively accurate predictions on additional time in ASMA, even with the limited amount of weather data. Combined with simulation, it was possible to explore different values of holdings and delays that aircraft suffer in the airport airspace. During the experiments, it has been noticed that the used simulation software was able to handle an only limited number of Bayesian model variables, hence limiting the further extension of the model with other airport's KPIs. Nevertheless, this paper served as a proof of concept that probabilistic and mainly, Bayesian modelling, can be successfully applied for prediction of disruptions in airport performance.

4.1 Results published in paper I

The idea to use Bayesian modelling for prediction of flight delays coupled with simulation tools was further developed into an optimisation approach for airport stand allocation. The created disruption-aware stand assignment (DASA) methodology was presented in the paper *"A multi-objective optimization with a delay-aware component for airport stand allocation"* [1]. This paper presented DASA with the optimisation objective (4) formulated as described in section 3.3.1 and discussed results of simulation experiments on a congested airport of Mexico, which operations often suffer from delays.

In this study, a weekly performance report of Mexico City International Airport was processed in DASA to build Bayesian models for arrival time deviations and to create a disruption-aware stand assignment. In total, a flight schedule of 564 arrivals with significant historical delays was allocated to 84 stands, representing the busiest day of operations from the selected report. The quality of allocation optimisation was confirmed with 30 minutes stopping criteria for optimisation, as can be seen in Table 1. DASA methodology was able to allocate all scheduled arrivals to the available airport stands and simultaneously reduce their taxi distance by nearly 6%.

Table 1. Optimisation results of Module II

Metric	1st generated solution	The solution generated after 30 min of running evolutional optimisation	Improvement from the 1st generated solution
Number of flights assigned to remote parking positions	259,0	217,0	16,2%
Total taxi distance for assigned schedule, km	900,9	848,6	5,8%
Number of flights assigned to wait for the availability of suitable stand	7	0	100,0%
The average area per passenger at the boarding gate, m^2	2,3	3,0	28,8%

The principal goal of experiments in the paper I was to prove the utility of a concept of disruption-awareness for increasing robustness of the stand assignment solution. In the scope of this article, the robustness was defined by two KPIs: the number of waiting aircraft (assignment conflicts) and average waiting time that these aircraft must wait for the stand availability. The DASA-generated allocation was tested in simulation experiments and compared to static stand allocations that did not consider schedule deviations. Table 2 shows the characteristics of compared operational scenarios. In total, three scenarios were simulated: on-time operations with static stand allocation (*Base Case*), stochastic arrival delays and static stand allocation (*Case 1*), and stochastic arrival delays and DASA-generated stand allocation (*Case 2*). The results of the performed simulation experiments can be seen in Figure 7. It is essential to mention that in these experiments if an aircraft was assigned to a stand that was still occupied, no reassignment actions were taken, and aircraft had to wait until the stand capacity was released.

Table 2. Characteristics of simulation experiments in the paper I

Scenario	Number of replications	Number of arrivals	Arrival time disruptions	Arrival time disruptions considered in the stand allocation plan
Base Case	30	564	No	No
Case 1	30	564	Yes	No
Case 2	30	564	Yes	Yes

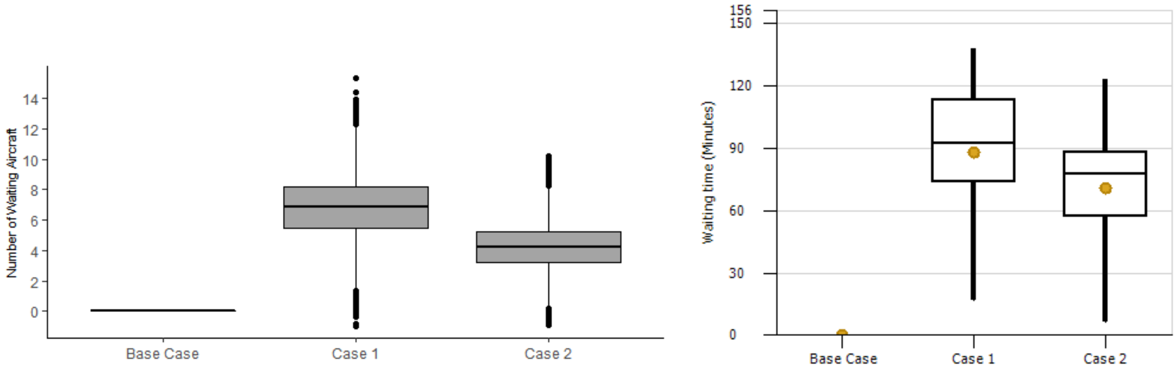


Figure 7. Experiments statistics for the number of waiting aircraft and average waiting time

As can be seen in Figure 7, the *Base Case* does not show any variability since all arrivals were on time. Once the arrivals became stochastic (*Case 1*), the number of assignment conflicts increased to 7, with an average waiting time of 88.4 min. DASA-generated disruption-aware assignment in *Case 2* resulted in 15.5% lower average waiting time compared to *Case 1*. Furthermore, *Case 2* showed 43% fewer assignment conflicts than *Case 1*. These results confirm improvement in the robustness when DASA is applied. Although disruption awareness did not eliminate assignment conflicts, it could significantly reduce their number and duration, which in realities of airport operations could lead to significant savings in cost and time.

4.2 Results published in paper II

Next journal paper, *"Reducing airport environmental footprint using a disruption-aware stand assignment approach"* [2] presented a redefined optimisation objective (9) that considered emission reduction goal, as well as interests of transfer passengers that were previously not taken into account. Module II emissions calculations were enhanced with aircraft engine specifications and consideration of toxicity levels of different pollutants. A search space reduction by changing weights of the multi-criteria objective function was also presented to increase the efficiency of the proposed emission reduction. The methodology was tested in a case study of Mexico City International Airport via simulation experiments and compared with simulated human-made allocation decisions.

In these experiments, DASA was used to allocate a one-week schedule with 3914 arrivals to 88 stands. DASA's Module II performed optimisation with the novel emission reduction objective (9), which considers the toxicity of the pollutants and passenger weighted use of contact stands and walking distance. The goal of this paper was to explore possible emission reduction that can be achieved by the application of the pollution-aware DASA. This modified DASA is further referred to as E-DASA.

For the study, E-DASA was tested in scenarios with and without arrival stochasticity, and also with and without application of reassignment of conflicted flights. The summary of these scenarios can be found in Table 3. In the performed experiments, random reallocation functionality was used to contrast E-DASA results. This feature was implemented to simulate an intervention of airport traffic controllers or gate managers, that can correct conflicting assignments manually during the day. The results of simulation experiments comparing E-DASA with different levels of manual reallocations are shown in Figure 8.

Table 3. Characteristics of simulation experiments in the paper II

Scenario name	Number of replications	Number of arrivals	Schedule disruptions	Schedule disruptions considered	Original assignment plan optimisation	Manual reallocation (no optimisation)
A	30	3914	-	-	Yes	-
B	30	3914	Yes	-	Yes	-
C	30	3914	Yes	-	Yes	Yes
D	30	3914	Yes	Yes	Yes	-
E	30	3914	Yes	Yes	Yes	Yes

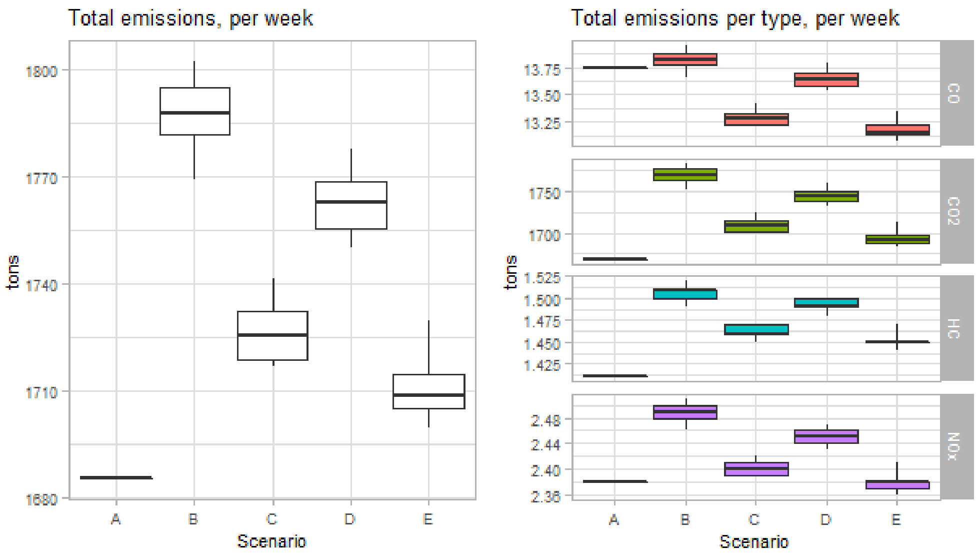


Figure 8. Experiments statistics for taxi-related emissions

As it can be seen in Figure 8, scenario *A* had the lowest emissions owing to allocation optimisation and absence of arrival stochasticity. When stochasticity was introduced into the simulation, the emission level increased by 6% (scenario *B*). This negative effect was first tackled by random reallocation of all conflicted flights during the day (scenario *C*), and then by considering probable delays in the allocation plan by E-DASA (scenario *D*). The latter appeared to have less effect than reassignment of all conflicts in scenario *C*. However, such last-minute relocation results in additional workload and complexity of organising turnaround operations for airport managers, and it is hardly applicable in any hub airport. Therefore, it is still necessary to create robust planning for stand usage. Scenario *E* illustrates the effect of combining E-DASA planning with last-minute reallocations. In this scenario, disruption-aware planning generated by E-DASA combined with human interventions for conflicted assignments produced the least amount of emissions among all performed stochastic experiments. Although the resulted values of emissions were specific to the airport of the case study, the methodology is generic and can be applied to any airport.

The paper also explored possible trade-offs in prioritising emission reduction against other goals. Maximum and minimum values of these goals were presented as a part of the approach for search space reduction. The resulting reduction in passenger walking distance was insignificant. However, prioritisation of emission reduction costs around 33% in terms of serving more passengers through contact stands for the case study airport.

4.3 Other publications

Additionally to the journal papers discussed above, the preliminary results of each of the DASA methodology formulations were presented in complimentary papers. The first concept of disruption-aware methodology with early results of improved stand allocation robustness was presented as a poster "*No more surprises: stand assignment algorithm with likelihood of turnaround time deviation"* [5]. Later, a study of the influence of terminal layout and stand distribution on emission reduction was presented in a conference paper "*Reduction of taxi-related airport emissions with disruptions-aware stand assignment: case of Mexico City International Airport"*[6]. In this article, a bi-objective formulation of E-DASA was used to find the stand allocation policy that would result in the lowest taxi emissions. The positive effect of having fewer allocation restrictions on emissions level has been demonstrated through simulation experiments on the layout of Mexico City International airport.

5 Journal paper I: A multi-objective optimization with a delay-aware component for airport stand allocation

This paper was published in the *Journal of Air transport Management* and presented the first development of disruption-aware stand allocation methodology with the objective function defined as in equation (4). In this article, the experiments prove the effectiveness of DASA approach for decreasing the number of stand assignment conflicts and related waiting time.

Journal of Air Transport Management 83 (2020) 101757

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

A multi-objective optimization with a delay-aware component for airport stand allocation

Margarita Bagamanova[a,*], Miguel Mujica Mota[b]

[a] Aeronautics and Logistics Departmental Unit, Engineering School, Autonomous University of Barcelona, Spain
[b] Aviation Academy, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Flight delays
Airport performance
Evolutionary algorithm
Terminal capacity management
Congestion
Capacity allocation
Decision support

ABSTRACT

Airport management is regularly challenged by the task of assigning flights to existing parking positions in the most efficient way while complying with existing policies, restrictions and capacity limitations. However, such process is frequently disrupted by various events, affecting punctuality of airline operations. This paper describes an innovative approach for obtaining an efficient stand assignment considering the stochastic nature of airport environment. Furthermore, the presented methodology combines benefits of Bayesian modelling and meta-heuristics for generating solutions that are more robust to airport flight schedule perturbations. In addition, this paper illustrates that the application of the presented methodology combined with simulation provides a valuable tool for assessing the robustness of the developed stand assignment to flight delays.

1. Introduction

Modern air transportation industry is encountered with a complex challenge. Air travel demand is rapidly growing every year and is estimated to double over the next two decades. The International Air Transport Association (IATA) reports that the amount of air travellers is expected to reach 8.2 billion travellers by 2037 (IATA, 2018). On the other side, air transport stakeholders are constantly pressured with changing standards and improvements in safety, passenger service and sustainability levels, as well as increasing need for modernisation of facilities.

Emerging technologies and overall demand to smoothen air passenger experience throughout the entire journey creates large investment pressure on airport stakeholders in the short time horizon. However, yet existing capacity constraints have to be also taken into consideration for the future investment areas (Symonds, 2018). If capacity development does not match the speed of traffic growth, congestion and economic problems will appear. Those problems would be a direct consequence of airlines not having access to necessary infrastructure for satisfying the increasing demand for air freight and passenger travel.

Everyday airport operations involve many aircraft and airport resources. Serving arriving aircraft and its passengers, preparing aircraft for departure and embarking its passengers require specific number of resources and corresponding equipment. At many airports these operations are performed by separate ground-handling dedicated companies, so the airport only must provide enough space for required time for making ground-handling possible. Nevertheless, while the number of aircraft passing through an airport grows from season to season, available space at an airport remains the same in most of the situations, thus, increasing the importance of facilities management by airport stakeholders. They must create an assignment schedule for the upcoming operational period, matching the existing parking positions with the requirements of airlines and passengers. However, in reality this assignment plan is often disrupted by deviations from scheduled times of arrival/departure of some flights, making the existing assignment schedule difficult to achieve and creating additional workload on decision-makers to re-make assignment schedule in time-constrained conditions.

In circumstances of limited capacity and high occupancy of terminal facilities, every deviation on arrival or departure time makes necessary to hold aircraft, affected by such deviations, waiting for availability of an appropriate parking position to be served at. This problem can sometimes be avoided by increasing buffer times between consecutive assignments to the same parking facility, however this reduces airport capacity. If the punctuality is disrupted, in the framework of a congested airport and limited airport apron space, it becomes vital to find tools that help to ease the burden of unpunctual arrivals and departures on the rest

* Corresponding author.
E-mail address: margaritabagamanova@gmail.com (M. Bagamanova).

https://doi.org/10.1016/j.jairtraman.2019.101757
Received 11 June 2019; Received in revised form 25 October 2019; Accepted 22 November 2019
Available online 28 November 2019
0969-6997/

of daily operations and mitigate the negative impact of operational stochasticity on the terminal capacity planning and scheduling to ensure its robustness.

The existing challenge of congested airports for mitigating the negative effect of arrival and departure delays on stand assignment schedule became a motivation to look for a leveraging methodology in the field of operational research. So, this paper presents a novel methodology for coping with the impact of unpunctual flights on the stand allocation schedule, which allows to benefit from historical data, optimization techniques and simulation for optimising the use of an airport stand capacity, considering the objectives of airport stakeholders and historical flight delays.

As mentioned, this paper focuses on the stand allocation problem (SAP); which refers to the problem of assigning flights to existing parking positions (stands) in such a way that all operational and technical constraints are satisfied. This problem is also called gate allocation problem (GAP). This term differs from SAP only in the definition of main subject: It considers only the gates used by passengers to go from the terminal building into the aircraft during the boarding procedure. In its nature, both SAP and GAP are approached by the same methodologies and are similar to a job-scheduling problem (Taillard, 1993), studied for decades. The complexity of the assignment is directly related to the number of flights to be assigned, which in airport routines can be to over 500 flights and more per day, depending on the size of an airport, which makes SAP/GAP an NP-hard problem due to real-life quantity of constraints and decision variables, such as aircraft size, airline business model, airport policy in stand assignment among others. As the number of flights for large airports can surmount thousand movements per day, the task of allocation becomes very complicated to be solved manually in an efficient way. Thus, for making stand allocation according to all required conditions and avoiding errors, SAP/GAP are often approached to be solved by the means of various algorithms, described in the next section. The article continues in the following way. Section 2 performs the literature review, in Section 3 presents the methodology and a case study is presented in Section 4. Section 5 exemplifies the application of the methodology in the case study and Section 6 concludes the paper and discusses the future work.

2. Literature review

According to the methodology used, the solving approaches can be divided into three categories: exact algorithms, heuristic algorithms and combined algorithms. While the first ones are aimed to find the best solution from a mathematical standpoint, the rest are designed to determine a qualitative near-optimal solution in a reasonable computational time (J Guépet et al., 2015). Due to the complex nature of the problem, exact solutions (e.g. a branch-and-bound algorithm) have difficulty in providing mathematically-optimal solutions within reasonable computational times for large-scale stand assignment problems. Therefore, recent studies mainly focus on developing heuristic algorithms, which do not guarantee the optimal solution but may provide near-optimal solutions in reasonable computational times. However, if the solution is not found by heuristic algorithm, it is not possible to determine whenever it is due to absence of any solution or due to inability of an algorithm to perform an abundant solution search (Pearl, 1986). Nevertheless, for real-life operational challenges finding the absolute optimum is not a vital requirement in everyday operations, as nearly optimal but quickly obtained solution would serve perfectly, especially when different costs of allocation have to be taken into account and the decision has to be made in short time.

Various optimization perspectives have been targeted as well as individually and as a group of objectives. Babić et al. (1984) were ones of the first to approach SAP/GAP using linear programming with objective to minimize walking distances for the passengers, assuming no flight delays are to happen. Later, Mangoubi and Mathaisel (1985) formulated a single objective function for passenger walking distances, considering randomness of walking distances while Yan and Tang (2007) included technical constraints for specific aircraft type and effects produced by flight delays on the stand allocation schedule into penalty-based heuristic planning framework. Solving SAP/GAP by decomposition into smaller time windows or flight sequences was successfully performed by Drexl and Nikulin (2008), Jaehn (2010), Şeker and Noyan (2012), Guépet et al. (2015), Marinelli et al. (2015), Yu et al. (2016), however, these authors did not consider flight delays.

When flight delays are considered, then SAP/GAP becomes more complex since it has to deal with their stochastic nature. For solving this type of problem, the insertion of buffer time between consecutive flights assigned to the same stand has been proved as a most effective solution for improving the schedule punctuality (Hassounah and Steuart, 1993). According to Yan and Chang (1998), Yan and Huo (2001), S. Yan, Shieh, and Chen (2002) these buffer times can be used to absorb not very significant stochastic flight delays (less than 30 min), that is why they proposed a simulation framework to analyse effects of flight delays on gate assignments and evaluate buffer times and gate assignment rules. Furthermore, extreme delays impacting on the gate assignment has been evaluated by Kontoyiannakis et al. (2009).

Despite of many years of research on SAP/GAP, the focus of solving algorithms has not changed much. The full amount of real-life problem constraints is not always considered, particularly, the stochastics of flight arrivals is often neglected, however, it should be considered since it carries a lot of uncertainty to be managed by airport stakeholders. Therefore, in contrast to the researches mentioned above and to fill the gap considering such an important factor as flight delay, this work presents an innovative approach where the probabilities of having certain delay levels are estimated for each flight and this information used for creating a qualitative stand assignment schedule. By approaching to the issue in this way, the stand assignment problem is formulated as close to the operational reality as possible in order to increase the applicability of the solutions generated by the developed algorithm.

3. Methodology

To overcome challenges induced by stochasticity of airport environment while considering the expectations of different actors involved in airport activities, this paper presents a methodology for stand assignment that deals both with operational uncertainties and multi-objective optimization goals. The methodological approach consists of different algorithms and processes that interact in such a way that it is possible to generate robust solutions that consider the historical delays, current capacity and required capacity. The implementation is done using two modules. Module I estimates probabilities of flight delays and their severity based on the historical data of operational periods and formulates the corresponding statistical models. Module II allocates flights to the stands using an evolutionary approach, considering the desired technical and operational restrictions for a target flight schedule and by calculating the stand occupancy time for each flight based on the delay models from Module I.

3.1. Architecture

The design of the presented approach can be described as follows: Module I is a look-ahead component which analyses the nature of historical delays and, using Bayesian inference techniques, calculates possible future delays; and Module II – generates an optimised stand assignment, considering various objective functions and management perspectives.

Main data flows between the modules and the principal functionality is presented in Fig. 1. The process starts with the analysis of data imported from an airport performance database, which can include among others the information about scheduled and actual flight arrival times, actual and scheduled block occupancy times, as well as the weather

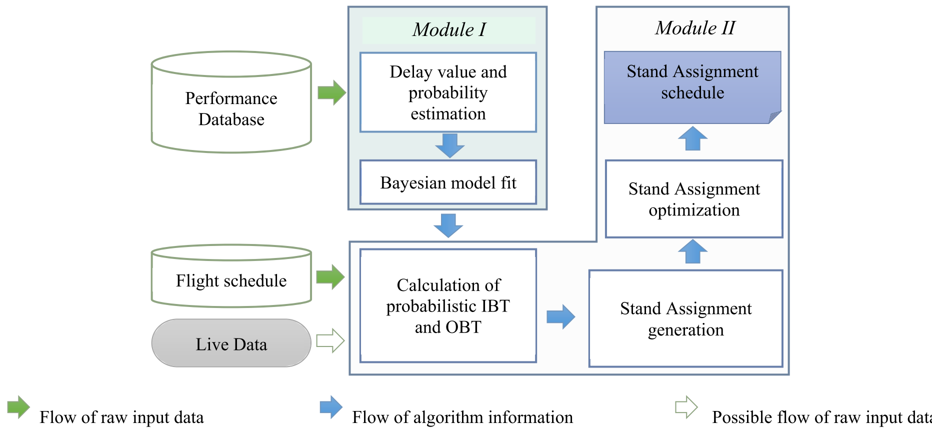


Fig. 1. Algorithm architecture and main data flows.

conditions at the time of operations and local air traffic regulations. In the scope of this paper only information about actual and scheduled time of arrival was available and therefore has been used for the initial implementations. Nevertheless, the analysis performed in Module I can also include (when available) information about airport operational environment such as weather conditions, runway configuration, flight delays in the airport of origin, which would significantly benefit Module I outcome.

Airport performance data is imported into Module I in a table format, where it is analysed and the probabilities of flight schedule deviations depending on several factors are estimated using inference based on Bayes' rule, which defines the probability of future event based on prior data of circumstances that might be related to this event. From these interdependencies and probabilities the corresponding Bayesian distributional regression models are built. These models together with the corresponding parameters (regression coefficients) are then transferred to Module II.

In Module II, the flight schedule is recalculated considering the standard information from the air traffic control tower log and the new information from Module I, estimating the possible arrival time deviation with the regression models. With this information, the estimated block occupancy times are calculated and used in the allocation algorithm (optimization phase). In the scope of this paper these estimated times are called *probabilistic in-block time* (further referred to as PIBT) and *probabilistic off-block time* (further referred to as POBT), together – *probabilistic block occupancy time* (further referred to as PBOT). As the next step, a re-calculated flight schedule, where scheduled time of arrival is replaced with PIBT and scheduled time of departure - with POBT, is processed by a metaheuristic search algorithm, which looks for a better stand assignment for the flights considering the probabilistic delays, while optimising the user-specified objective function. In this paper the objective function is a combination of four objectives, as it is described in section 3.1.2. The different elements of the presented two-module algorithmic approach are described in the following sections.

3.1.1. Module I: Bayesian models

As acquiring hidden knowledge from raw airport performance data may require long provisional analysis and fitting the corresponding models, the presented approach proposes a balanced solution to get necessary insights on the latent performance characteristics in a quite short period of time. The presented solution is based on the application of Bayes' rule in multilevel modelling, which already proved its value for research in various scientific applications (Brown and Prescott, 2014; Demidenko, 2013; Gelman and Hill, 2007; Pinheiro and Bates, 2000).

3.1.1.1. Multilevel models. The greatest benefit of multilevel models is that they allow the modelling from different perspectives of measurement at the same time, considering their complex dependencies. The heart of any multilevel model (further referred to as MLM) is the prediction of the response variable y at the data point i through the linear combination η of predicting factors, transformed by the inverse link function f adopting a certain distribution D for y: $y_i D(f(\eta_i), \theta)$.

The parameter θ describes additional distribution-specific parameters that typically do not vary across data, such as the standard deviation σ in normal models or the shape α in Gamma or negative binomial models. The linear predicting factor can generally be written as: $\eta = X\beta + Zu$. In this equation, β and u are the regression coefficients at population-level and group-level respectively and X, Z are the corresponding design matrices. The response y as well as X and Z form up the data, whereas β, u, and θ are the model parameters estimated with various sampling algorithms (Bürkner, 2017). In such a way by estimating level-corresponding coefficients it is possible to obtain a multilevel distributional regression model for the target response variable – *flight delay on arrival.*

3.1.1.2. Modelling on Bayes' rule. To estimate the regression coefficients for different performance parameters based only on historical data and use them for the inference of future data values it is necessary to estimate the joint probability distribution for both the target variable and the set of its predictors. This estimation could be done using Bayes' rule. Thus, following the Bayes' rule, the likelihood of observation A, occurring given the occurrence of observation B, can be written through the following equation: $P(A|B) = P(B|A)P(A)/P(B)$, where $P(B|A)$ denotes the likelihood of B occurring when A occurs, $P(A)$ and $P(B)$ – are the probabilities of observing both observations independently of each other (Stuart and Ord, 2010).

One of the main advantages of Bayesian modelling is that the data for the probability estimation can be used as it becomes available, so the models can be easily updated even after each operational day, if such desired. Also, thanks to the Bayesian method, the likelihood estimations are independent from "outliers" or extreme data values influences, which makes it a perfect approach for this study for consideration of hidden latent correlations between different performance variables.

In the presented case study, the estimation of delay probabilities is done on the arrival punctuality data for one week with the use of open statistical tool R, particularly with R package *brms* (Bürkner, 2017). This package performs efficient Bayesian model estimation for mixed data types and allows exporting the fitted distributional regression model parameters in any required form. It allows also to estimate an effect of each of the model parameters to the mean and variance of the response variable distribution (Bürkner, 2017). As an output of fitting the MLM with *brms* a fitted distributional regression model for the target variable – flight delay is received, where the effect of each of the chosen predictors has its own fitted linear regression function. When having such

MLM, it is possible to generate future (probable) delay values with various levels of probability, according to the chosen predictor variables, and use them for calculation of PBOT used in the schedule generated by Module II. By introducing this innovative approach it is expected to reduce the problems caused by delays thus making a more robust schedule; which is evaluated later in the paper with the case study.

3.1.2. Module II: evolutionary optimization

This module takes care of generating stand assignment, considering the input constraints, and then optimising it to ensure better value of quantitative expressions of airport stakeholders' objectives (multi-objective function). In order to make it practical for the potential stakeholders, an algorithm that considers the diverse variables that are important for the actors affected by the allocation in the airport: required level of service for passengers and airlines, cost and environmental impact, was developed. To ensure the evolvement of the optimization algorithm for the chosen airport data, it was decided to consider the following objective function, where the different objectives are described:

$$F = w_1 \times R_{apron} + w_2 \times R_{taxi} + w_3 \times R_{hold} + w_4 \times R_{service} \tag{1}$$

1 Airport management perspective: to serve more passengers through the contact stands and minimize the use of remote parking positions

$$R_{apron} = (Nflightsassigned|apron) \,/\, (TotalNflights) \tag{2}$$

Where:

- *Nflightsassigned|apron* – is the number of flights assigned to remote parking positions, that are connected to the terminal building only via bus service;
- *TotalNflights* – is the total number of flights in the schedule to allocate.

2. Airline and environmental perspective: to minimize the taxi distance to the stand

$$R_{taxi} = (AverageScheduledTaxi) \,/\, (MaxAirportTaxi) \tag{3}$$

Where:

- *AverageScheduledTaxi* – is the average taxi distance in the allocated schedule;
- *MaxAirportTaxi* – is the maximum possible taxi distance at the airport for considered runway configuration.

3. Air Traffic Control perspective: to minimize number of aircraft waiting for stand availability

$$R_{hold} = (Nflights|waiting) \,/\, (TotalNflights) \tag{4}$$

Where:

- *Nflights|waiting* – is the number of flights that must wait for the stand availability;
- *TotalNflights* - is the total number of flights in the schedule to allocate.

4 Passenger comfort perspective: to provide enough waiting space in the departure lounge

$$R_{service} = (MaxAreaPerPassenger - ActualAreaPerPassenger) \,/\, (MaxAreaPerPassenger) \tag{5}$$

Where:

- *MaxAreaPerPassenger*– maximum possible number of m2 per passenger, calculated for the flight with smallest number of passengers in the schedule assigned to the stand with the largest waiting lounge of the airport;
- *ActualAreaPerPassenger* – actual number of m2 per passenger, available at the assigned gate for the assigned flight

5 w_n– indicates priority weights for the corresponding perspectives, for practical implementations, the weights should be decided by negotiations of the different stakeholders of the airport. In this paper all the weights are equal to 1 in order to obtain a stand assignment equally balanced for all considered perspectives. Prioritizing one or more perspectives over the others may result in certain cost for under prioritized perspectives, so it can be used for estimation of various airport strategies and answer the questions such as how much will it cost in taxi distance to prioritise passenger comfort?

As presented, there are conflicting objectives due to the nature of the actors involved. For instance, airlines objectives would aim to minimize taxi distance preferring parking positions located as close as possible to the runway exits. This may cause a conflict with the uniform stand assignment policy of the airport operator. On the other hand, airport operators would prefer to use contact stand as often as possible to provide the best service for the airlines and spread the allocation so that there is even use of infrastructure. In addition, airlines would like to have sufficient space in the waiting lounges of the gates to provide the best level of service for passengers; again, this objective might conflict the environmental one as some gates with bigger lounge areas might not necessarily be located closest to the runway exit.

The particular restrictions to be considered in the stand assignment schedule can vary slightly depending on the particularities of each airport. The following are the restrictions implemented in the presented algorithm:

1 Spatial
   - Domestic and international flights must be assigned to the specific gates. Normally these are internal specifications of the airport e.g. international flights are assigned to gates that have access to the designated border control areas.
   - Enough space for passengers waiting to board must be provided. These values depend on the layout of each airport; for every gate there will be a specific area dedicated to the passengers waiting for boarding. This issue was approached by considering that each gate has a specific area and the number of passengers to board depends on the type of aircraft and its load factor. For instance, an A380 (450 passengers) is not preferred to be allocated to the stand next to a Boeing-777 (305 passengers) at the same time. Formula (5) was used for evaluating this condition.

2 Temporal
   - Flight delays must be considered in the assignment schedule (according to conditional delay probability distributions from Module I). In this paper, only arrival delays are considered due to unavailability of ground handling performance data and correspondence of arriving aircraft to departing aircraft.

3 Operational
   - Parking position must correspond to the size of an aircraft (large aircraft require extra space due to larger wing span). This is implemented through identification of allowed stands for each flight on the stage of processing the input data in Module II.
   - Aircraft with large number of passengers should be served at contact stands. This restriction is implemented to ensure smoother transfer experience to the passengers and it is ensured through the objective function – Formula (5).

- The use of contact stands is prioritized. This is implemented via the objective function component Formula (2).
- In case when there are no parking positions available at the moment of arrival, aircraft should wait on the apron until a position becomes available. This is implemented in the algorithm by assigning the flight to a "dummy" stand and incrementally delaying its PBOT until a suitable stand becomes available.

3.2. Evolutionary algorithm

For the optimization of stand assignment schedule with PIBT and POBT, a genetic algorithm (Goldberg, 1989) was developed. Although many solution search algorithms (among others Particle Swarm Optimization (Kennedy and Eberhart, 1995), Harmony Search Algorithm (Zong Woo Geem et al., 2001), Simulated Annealing (Kirkpatrick et al., 1983)) could be applied for the presented stand assignment problem, a genetic algorithm (GA) has been chosen for various reasons. One of the most important reasons is its ability to escape the local optima by increasing the diversity of solutions, which in the case of a multi-objective optimization is preferably to have this feature. Some local search algorithms as Tabu search and simulated annealing (S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, 2003) are very good, but as time passes by it becomes more difficult to them to escape the local optima. In addition, the authors have already worked with GA previously with good results. Nevertheless, it is considered to try another optimization algorithms to compare their performance in the future work.

There are many examples of successful implementations of genetic algorithms in air transport optimization problems. Some relevant research can be found at Ghazouani et al. (2015), Mujica Mota (2015), Abdelghany et al. (2017) among others. They differ around implementation, formulation of a problem and in computational techniques used. In this paper the stand assignment schedule is represented as a NxM dimensional array, where N refers to the number of flights to be assigned to the stands and M is the number of various characteristics to be considered in the assignment. Thus, each array cell (flight) has an array of characteristics which are considered by the constraints in section 3.1.2. Fig. 2 illustrates one chromosome with the correspondent information that can be used by the algorithm. An array implementation allows to add extra characteristics if necessary.

A complete genotype which represents a potential solution is illustrated in Fig. 3.

The different operations and selection of the different chromosomes of the algorithm are presented in Pseudocode 1.

The general flow of the algorithm starts with importing target flight schedule, terminal building characteristics, weights for objective functions components and MLM data from Module I. These data are used to create an initial stand assignment solution (referred to as "Adam chromosome" in Pseudocode 1), which serves as a base to create a set of solutions by making random changes of assigned stands into different ones. After that, the quality of the generated solutions (chromosomes) is evaluated by the objective function for each of the chromosomes and the one with the smallest value is saved and marked as best chromosome. Next, the set of chromosomes is subjected to random crossover (where several chromosomes exchange their assigned stands with each other), thus creating a set of new chromosomes with different stand assignment. This procedure is followed by randomly changing some of the chromosomes by the function called *Mutation*. The chance of mutation is generated randomly for each of the chromosomes. After that, the entire set is evaluated again by computing the objective function for changed and new chromosomes and the best chromosome marking is updated if needed. This is followed by evaluating if the algorithm has reached the stopping criteria defined by user, and if so, the algorithm stops and exports the best solution into the data file.

```
GET Stop_Criteria
IMPORT
   FlightSchedule,
   Constraints,
   ModuleI.output
CREATE
    Adam chromosome, A
  GENERATE
    Set(chromosomes), S = RandomChange(A)
WHILE CurrentSituation <> Stop_Criteria
  REPEAT
    FOR X = 1 TO count (S) DO
        Calculate objective function F(x)
        IF value F(x)> Best_Val THEN
           Best_val = value F(x)
           Best_Chrom = X
        IF CurrentSituation = Stop_Criteria
           THEN BREAK
        DO Crossover(Xi, Xj)
           IF MutationChance = TRUE
           Mutation (X)
EXPORT Best_Chrom
```

The number of algorithm iterations, total running time or certain objective function value can be established as the stopping conditions for the algorithm of Module II, according to user needs. Therefore, the solution quality improvement is only restricted by user preferences.

In the following section the implementation of the two-module approach in a real-case study is presented.

4. Case study: Mexico city international airport

Mexico City International Airport (IATA code: MEX) is the main airport in Mexico and one of the busiest airports in the world with its traffic increasing by 10% on average annually since 2012 (Moody's Investors Service, 2018). Such constant growth rapidly led to congestion and nowadays MEX is serving almost 50% more passengers than it was designed to. As a consequence, the local economy is already suffering from lack of aviation connectivity and due to that it is expected to lose up to $20 billion in future GDP by 2035 if airport capacity is not increased (International Airport Review, 2018).

According to Airport Council International (2018) Mexico City International Airport is on the 20th place in the ranking of world airports with biggest number of aircraft movements with approximately 450 thousands landings and take-offs annually. However around 20% of MEX departures in 2018 have suffered from substantial delay (more than

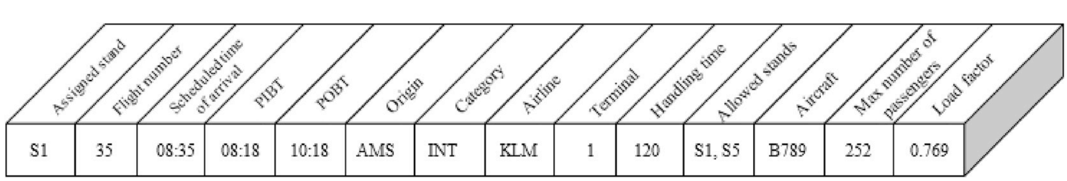


Fig. 2. Chromosome's content.

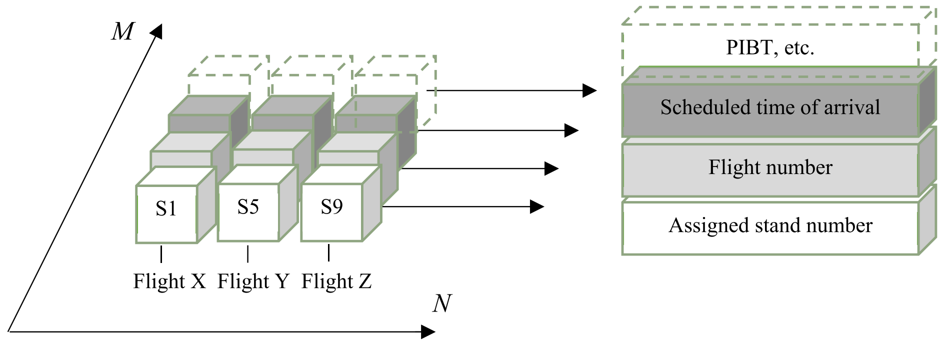


Fig. 3. Stand assignment schedule coded for genetic algorithm in Module II.

15 min from scheduled time of departure), which resulted in an average delay of approximately 50 min per flight (see Fig. 4). Since 2013 MEX operates at its maximum capacity (SCT, 2013), so the actual number of flights affected by such high delay is even bigger.

Furthermore, delayed or early arrivals complicate the situation further, creating additional burden on the existing terminal capacity. Such perturbations and dense flight schedule make MEX an attractive case to test the functionality of the proposed algorithm.

Mexico City International Airport has 2 terminals, both used for international as well as domestic flights, and there are 26 airlines operating at it. These terminals are separated by two parallel runways, not operating simultaneously due to not enough separation between each other. Some other relevant information on MEX can be found in Table 1. Since 2017 MEX has been declared with a capacity of 61 operations per hour with maximum of 40 landings (SCT, 2017).

Available at AICM (2018) an official on-time performance report has been considered for this case study. The analysed report consists of one-week operations from 28.05.2018 to 03.06.2018, both arrivals and departures, with actual time of arrival/departure and scheduled time of arrivals/departure, which allowed us to extract information about arrival delay. Nevertheless, there was no open access information about individual departure delays per each aircraft, therefore in the scope of this research only on the arrival delays are considered (i.e. difference between actual time of arrival and scheduled time of arrival), however, the algorithmic framework presented can be used as well with departure delays if the appropriate information is available.

Deviation from scheduled time of arrival (STA) per day and hour of analysed schedule for one week can be seen in Fig. 5. Positive values denote late arrival of a flight, and negative values denote early arrival. Further in this paper both types of deviation from scheduled time of arrival both positive and negative are referred to as delays. In the graph only the delays within 2 h interval are presented. However, the presence of quite a high number of early arrivals as well as severe delays (more than 30 min) can be clearly noticed. Out of 3917 flights arrived during the studied week 2091 flight arrived more than 15 min (red dashed line on the graph) later or earlier than scheduled, which constituted more than 53.4% of all studied passenger flights. Early arrivals (earlier than 15 min) constituted approx. 36.6% of week arrivals. Regarding the statistics of arrival delay per days of week, it is important to notice that only on Thursday 50% of the flights stayed in the limits of 15 min deviation from arrival schedule. Moreover, approx. 7.3% of the studied flights trespassed the limits of 60 min deviation from the arrival schedule, which for the airport operating beyond its maximum capacity is an extremely severe complication.

Table 1
Mexico City International Airport characteristics (AICM, 2019; IAS, 2019).

	Terminal 1	Terminal 2
Surface area	54,8 ha	24,2 ha
Contact aircraft parking positions	33	23
Remote aircraft parking positions	11	17
Airlines	20	6
Passenger throughput in 2017	2, 26 billion passengers	1, 75 billion passengers

Regarding hourly performance, as demonstrated on the lower part of Fig. 5, it is important to notice that there is a lot of variability in most of the hours of the day, which means that operational day is constantly under the pressure of disruptions. And for some hours, like from 4am to 9pm for instance, the deviation from scheduled time of arrival exceeds 60-min threshold.

When examining statistics for the 52 airlines, on the selected week schedule, it is important to notice (Fig. 6) that the top 10 airlines (which correspond to the ones with the biggest number of flights) have a rather stable performance. Only few of the top airlines exceeded average delay value of 15 min.

However, the overall punctuality is very poor. A comparison of Fig. 6 with Fig. 7 shows that 7 out of Top 10 airlines with the biggest number of flights in the studied schedule also appear in the rank of the ones with

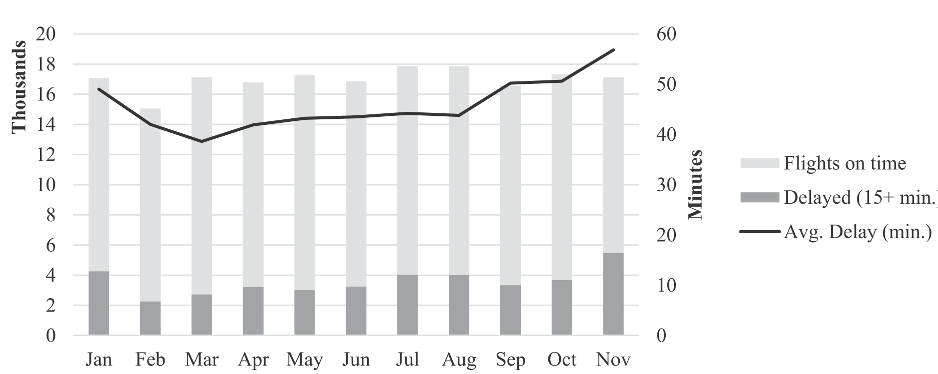


Fig. 4. MEX Departure performance statistics for 2018 (Flightstats, 2018).

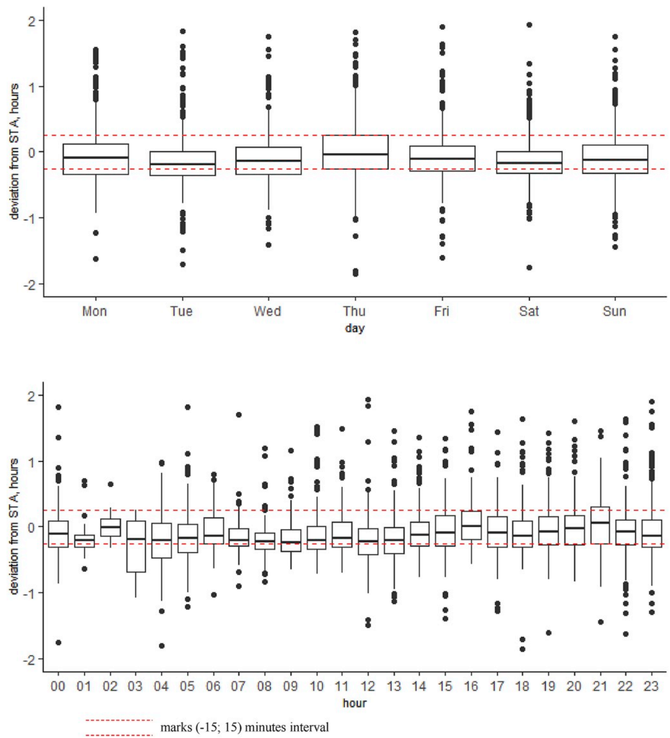


Fig. 5. Deviation from scheduled time of arrival (STA): statistics per day and hour.

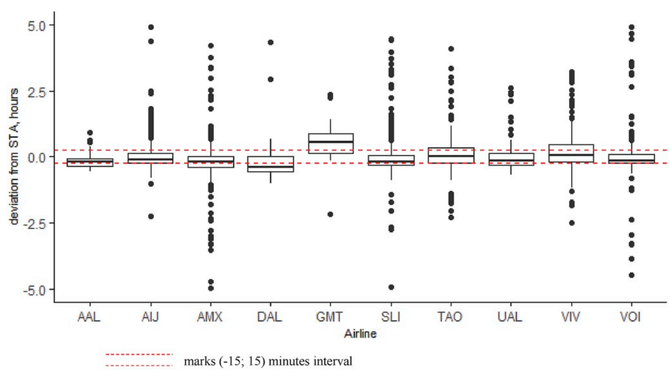


Fig. 6. Top 10 airlines with the biggest number of flights and deviation from scheduled time of arrival (STA).

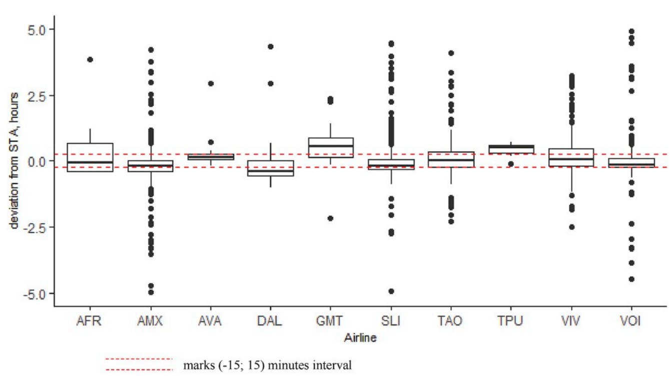


Fig. 7. Top 10 airlines with the biggest mean deviation from scheduled time of arrival (STA).

worst punctuality. Flights of AMX, SLI and VIV quite often exceed the value of 120 min deviation from their scheduled time of arrival, as it can be seen in Fig. 7. The largest average delay, which constitutes in a deviation of more than 1 h, occurred in the arrival of less than 3% flights out of a total flight schedule.

5. Two-module approach: the case of MEX

The delay model for this case study has been built on the correlation of delay level with the time of the day of arrival and the airline. After processing MEX one-week operational data in Module I considering two available variables - hour of arrival and airline name - as predictors of arrival delays, the corresponding MLM has been obtained and a sample of the resulting parameters are presented in Table 2. The complete table can be found in Appendix A. These parameters are linear regression coefficients for predictor variables *Airline and Hour* and allow to generate arrival delay value based on the corresponding airline name and hour of scheduled arrival. Detailed explanation of the layout, presented in Table 2, can be found at Bürkner (2017).

After generating delay values through the obtained MLM and randomly sampling from the obtained data, the resulting distribution with parameters from Table 3 was compared to the observed arrival delay. This comparison is presented in Fig. 8 (*yrep* represents simulated data, *y* – original historical data) and as it can be noticed from this figure, the distribution shape for simulated flight delays quite closely matches the distribution shape of historical flight delays which suggests that delay has a strong dependency on hour of the day and airline type.

Following the algorithmic implementation, the obtained MLM from Module I, along with a target 1-day flight schedule, is imported into Module II, where the 1-day flight schedule has been assigned to the available parking positions, as described in Section 3.

In order to evaluate the potential of the algorithm, the handling operations were assumed for domestic flights to have a block occupancy time of 60 min and international flights of 120 min. This has been considered by Module II and after running it with available input and constraints, the obtained results are illustrated in Table 3. This flight schedule has been generated considering the expected delays forecasted by Module I.

The obtained stand assignment has no instances of overlapping assignment of different flights to the same stand, which assume no need of direct ATC involvement in regular operational conditions (no rare weather phenomena, no unique air traffic regulations in the area).

Regarding the optimization part of Module II, the values for the

Table 2
Sample of Module I output.

Population-Level Effects:	Estimate	Est.Error	Q2.5	Q97.5
Intercept	−10,24	2,02	−14,15	−6,32
AirlineAFR	10,21	7,27	−2,63	27,35
AirlineAIJ	7,60	1,07	5,52	9,68
AirlineAMX	5,32	1,10	3,16	7,46
AirlineDAL	1,42	1,69	−1,93	4,69
AirlineGMT	16,60	2,66	11,51	21,89
AirlineSKU	210,35	223,77	−69,48	457,40
AirlineSLI	4,78	1,05	2,73	6,83
AirlineTAO	7,11	1,34	4,53	9,73
AirlineVIV	8,91	1,31	6,39	11,47
AirlineVOI	9,67	1,18	7,38	11,99
hour03	−5,74	4,11	−13,98	2,26
hour05	8,10	2,02	4,19	12,02
hour06	6,69	1,88	2,98	10,35
hour16	9,07	1,92	5,36	12,72
hour21	10,13	1,93	6,26	13,83
hour23	0,21	2,03	−3,84	4,19

Family: student.
Formula: Delay ~ Airline + Hour.
Samples: 3 chains, each with iterations = 3500; warmup = 1750; thin = 1; total post-warmup samples = 5250.

Table 3
Part of Module II output.

Origin	Flight	Category	Airline	Scheduled Arrival Time	Terminal	Stand	Handling Time	Equip	Aircraft Type
VSA	3249	INTERNATIONAL	VIVA AEROBUS	28-05-18 0:00	T-1	G39_A	120	EA32	Large
TGZ	3259	DOMESTIC	VIVA AEROBUS	28-05-18 0:00	T-1	G15	60	EA32	Large
MTY	185	DOMESTIC	VOLARIS	28-05-18 0:10	T-1	G12	60	EA32	Large
HMO	763	DOMESTIC	VOLARIS	28-05-18 0:10	T-1	G13	60	EA21	Large

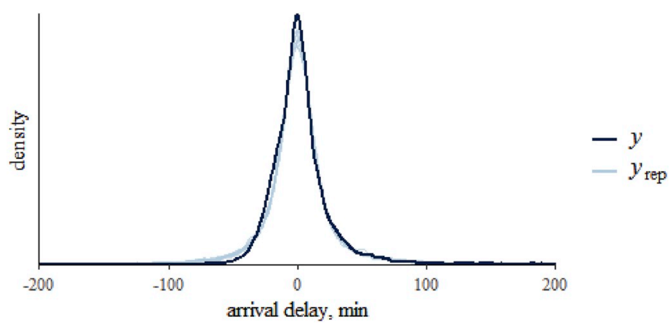


Fig. 8. Historical arrival delay distribution vs simulated MLM arrival delay distribution.

multi-objective function components can be found in Table 4:

As it can be seen in Table 4, after 30 min, the optimization algorithm of Module II gives quite a significant improvement for all the objectives considered. Additionally, it is important to mention that reduction in both total taxi distances and number of flights assigned to the remote parking positions, can lead to significant fuel consumption economy, which in its turn leads to economic and environmental savings for the airlines. And also, to improvement in passenger service, as less passenger time is spent on waiting to reach the terminal building and proceed further to passengers' destination.

6. Validation of the two-module approach

For further evaluating the quality of the presented two-module stand assignment approach, a validated simulation model of Mexico City International Airport has been used. More information about the model and its functionality can be found at Mujica Mota and Flores (2019). This simulation model considers the design and characteristics of MEX layout and runway dimensions, as well as corresponding taxi ways. This model can be considered as a digital twin of MEX and is suitable for initially testing such novel operational approaches.

With the simulation model of Mexico City, it has been tested how the integration of Bayesian modelling into the stand assignment impacts the number of aircraft that have to wait for their assigned stand to become available in the presence of stochastic arrival delays. The following three scenarios were chosen for comparison. These scenarios have one common characteristic - no intended buffer times between consecutive flights assigned to the same stand were included into the assignment schedule. Such feature allows to fully observe the impact of flight arrival deviation on the airport performance, while often in reality some arrival deviations are absorbed by the buffers.

- *Base Case.* It assumes an ideal performance situation: everything goes according to the flight schedule and all flights arrive and depart punctually. Stand assignment is generated and optimised with the use of Module II only (actual block occupancy times correspond to the scheduled block occupancy times).
- *Case 1.* Flights arrive considering its stochasticity generated from a delay distribution model, learnt in Module I. Thus, it is possible to have a positive or negative delay (early arrival). Stand assignment is generated and optimised with the use of Module II only.
- *Case 2.* Flights arrive with stochastic deviation, generated from delay distribution model, learnt in Module I. Stand assignment is generated with the use of both Module I and Module II (PBOT is used).

To get representative data from the experiments runs for the chosen three scenarios, 30 replications of the scenarios have been executed with duration of 30 h. This provides enough time to execute an entire flight schedule of one operational day of 564 arrivals with possible arrival time deviations. For all three scenarios the number of aircraft that must wait for the stand availability have been tracked, as well as the waiting time for such availability.

Fig. 9 displays the statistics of stand availability waiting time for the performed experiments runs. As it can be seen, the Base case does not present any variability since all the operations are on time. Once the variability of the real system is introduced (Case 1), it results in significant waiting time with a mean value of 88.4 min per aircraft. Regarding Case 2, when arrival deviations are considered in the stand assignment

Table 4
Optimization results of Module II.

Metric	1st generated solution	Solution generated after 30 min of running evolutional optimization	% improvement compared to 1st generated solution
Number of flights assigned to remote parking positions	259,0	217,0	16,2%
Total taxi distance for assigned schedule, km	900,9	848,6	5,8%
Number of flights assigned to wait for availability of suitable stand	7	0	100,0%
Average area per passenger at the boarding gate, m^2	2,3	3,0	28,8%

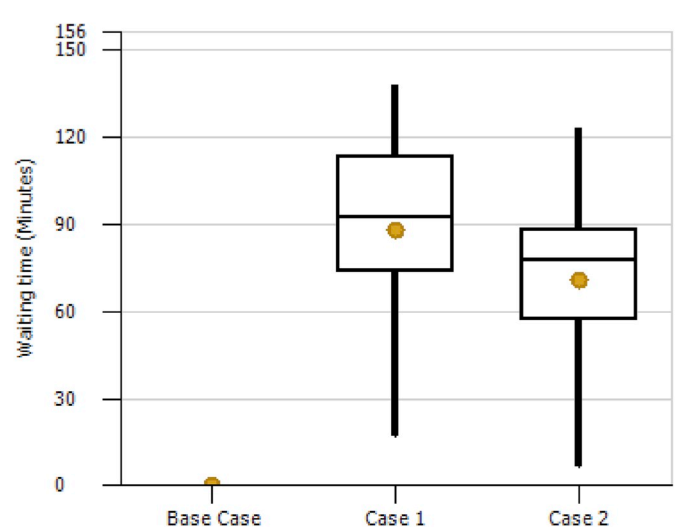


Fig. 9. Simulation experiments statistics for waiting time.

and the two-module approach is used to produce an optimised and delay-aware stand assignment, the average waiting time is decreased by 15.5% compared with Case 1 with mean of 71.2 min per aircraft. Furthermore, as it can be noted from Fig. 9, in Case 2, 75% of waited aircraft had a waiting time of less than 88 min, while in Case 1 more than 50% of waited aircraft had to wait longer than that. This comparison demonstrates improvement in the waiting situation for more than half of the total number of waiting aircraft compared to the stand assignment of Case 1.

Regarding the number of aircraft that had to wait for a stand (Fig. 10), it can be noticed that the average value of Case 2 is 39% less than of Case 1, and what is more important from managerial point of view is the variability; as it can be seen from Fig. 10, the dispersion of number of aircraft is considerably reduced, since it changed from up to 15 to 10. This means that in the real system it will be less likely that some AC would not have a stand. And if translated to the number of arrivals per day, Case 2 illustrates that in average only 4 aircraft out of 564 flights were forced to wait for a stand in average, while for Case 1 the value corresponds to 7 making a reduction of 43%. These numbers extrapolated to a yearly operations mean a lot of time, fuel, and money spent by the airlines, airport and passengers.

The results from the approach suggests that by considering some characteristics of the flights and environment it is possible to decrease in half the number of potential stand occupancy conflicts (for the example presented). As it was illustrated with the case study, the algorithmic framework has the potential to produce better schedules considering the historical delay, different perspectives and the technical restrictions present in the system. As mentioned before the presented approach is an innovative combination of Bayesian inference, optimization and simulation that has not been previously presented.

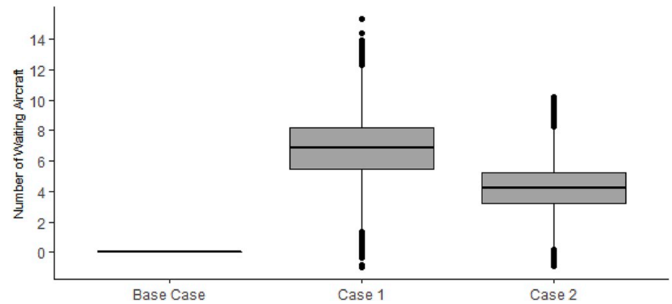


Fig. 10. Simulation experiments statistics for number of aircraft waiting for stands.

7. Conclusions and further research

This paper presents an innovative delay-aware approach that combines Bayesian methods and multi-objective heuristic optimization with variability for aiming at solving one of the most complex airport capacity management problems – stand allocation in airports. The implementation is done via a two-module approach in which each module performs key functionality that will provide value for the final solution. It generates robust solutions; in the first iteration of Module I, it provides airport stakeholders with a problem-aware stand assignment. Then, new probabilistic values of stand occupancy times are considered by an optimization algorithm of Module II. In order to validate the effects of the presented approach application on airport performance, simulation was also introduced into study to include the variability of the real system.

The combination of all the elements makes a very robust approach that can be implemented for any type of airport just by specifying the particular restrictions. Together with simulation, the methodology facilitates delay risk management and delay impact assessment on the slot adherence. In the case presented, the methodology showed a clear decrease of the number of aircraft waiting for the stand availability by 40% and decrease of total taxi distance by nearly 6%.

The implementation of the presented approach in the stand allocation planning process is an innovative one that for the first time together with simulation can help easing the burden of arrival and departure time deviations on the airport capacity, optimise airport capacity allocation from various management perspectives and can help to release capacity resources that are usually blocked by extensive buffer times between allocated flights.

Furthermore, the presented stand assignment methodology is formulated in such a way that any additional constraints can be added to the optimised assignment in Module II, which provides with the flexibility to tackle various assignment strategies and goals. In addition, Module I can be also enriched with departure punctuality historical data of a real airport or weather conditions ensuring a more holistic view during the delay model estimation. On the other hand, if the corresponding requirements are included in Module II, the stand allocation problem could also be tackled taking into account the slot adherence instead of comparison of scheduled and actual arrival and departure times.

As future work, other variables would be considered in Module I for providing more accuracy on the expected delay, and the use of information obtained from the Simulation model will be incorporated in the optimization loop in order to provide even more robust solutions as the authors have already applied in other ATM problems with good results.

Acknowledgements

The authors would like to thank Autonomous University of Barcelona, the Aviation Academy of the University of Applied Sciences for supporting this study and the Dutch Benelux Simulation Society (www.dutchbss.org) and EUROSIM for the dissemination of the findings of this study.

Appendix A

Output of Module I.

Population-Level Effects:	Estimate	Est.Error	Q2.5	Q97.5
Intercept	−10,24	2,02	−14,15	−6,32
AirlineABX	12,92	6,86	−0,38	26,21
AirlineACA	2,40	2,51	−2,78	7,10
AirlineAFR	10,21	7,27	−2,63	27,35
AirlineAIJ	7,60	1,07	5,52	9,68
AirlineAJT	−1,91	11,23	−25,46	18,22
AirlineAMX	5,32	1,10	3,16	7,46

(continued on next page)

(*continued*)

Population-Level Effects:	Estimate	Est.Error	Q2.5	Q97.5
AirlineANA	2,31	5,00	−7,08	12,45
AirlineARE	118,79	66,99	35,23	217,63
AirlineASA	1,02	2,88	−4,52	6,78
AirlineAVA	14,71	2,47	9,83	19,62
AirlineAZA	1,17	5,29	−9,41	11,58
AirlineBAW	0,85	4,31	−7,89	9,00
AirlineCHH	−3,57	9,39	−19,80	14,89
AirlineCKS	80,03	79,31	−45,68	269,10
AirlineCLU	36,43	53,99	−39,37	119,48
AirlineCLX	8,12	5,93	−2,71	20,43
AirlineCMP	−0,70	1,76	−4,18	2,69
AirlineCPA	9,14	13,15	−5,69	44,40
AirlineCSN	9,23	9,03	−7,15	29,95
AirlineDAL	1,42	1,69	−1,93	4,69
AirlineDLH	2,51	4,06	−5,23	10,37
AirlineESF	23,35	3,39	16,60	29,95
AirlineGEC	−7,74	8,99	−26,17	8,46
AirlineGMT	16,60	2,66	11,51	21,89
AirlineGTI	190,34	17,29	159,57	221,68
AirlineIBE	2,88	2,99	−2,99	8,58
AirlineICL	46,61	12,75	22,16	71,47
AirlineJBU	−8,53	2,17	−12,77	−4,30
AirlineJOS	10,05	4,63	1,18	19,35
AirlineKLM	10,23	4,39	1,02	18,53
AirlineLAN	23,80	4,55	14,10	32,35
AirlineLPE	1,79	4,08	−6,48	9,57
AirlineMAA	58,98	31,69	5,18	112,37
AirlineQCL	7,92	10,40	−11,98	29,05
AirlineQTR	9,24	7,54	−5,23	24,95
AirlineRPB	−0,11	4,94	−9,51	9,99
AirlineSKU	210,35	223,77	−69,48	457,40
AirlineSLI	4,78	1,05	2,73	6,83
AirlineSWA	2,88	1,86	−0,78	6,56
AirlineTAI	−0,70	3,00	−6,50	5,41
AirlineTAM	12,66	4,51	3,46	21,46
AirlineTAO	7,11	1,34	4,53	9,73
AirlineTNO	8,32	2,99	2,68	14,37
AirlineTPU	9,33	4,87	0,50	20,09
AirlineUAE	−1,27	7,42	−14,49	14,96
AirlineUAL	3,48	1,42	0,71	6,34
AirlineUPS	−1,87	4,64	−10,80	7,48
AirlineVIV	8,91	1,31	6,39	11,47
AirlineVOC	20,09	4,20	11,75	28,27
AirlineVOI	9,67	1,18	7,38	11,99
AirlineWJA	6,27	2,59	1,26	11,28
hour00	0,27	2,43	−4,29	4,71
hour01	0,23	2,30	−4,33	4,68
hour02	0,72	4,00	−7,53	8,18
hour03	−5,74	4,11	−13,98	2,26
hour04	−5,39	2,48	−10,30	−0,62
hour05	8,10	2,02	4,19	12,02
hour06	6,69	1,88	2,98	10,35
hour07	2,87	1,95	−0,99	6,64
hour08	0,11	1,89	−3,69	3,77
hour09	1,68	1,90	−2,09	5,42
hour10	4,12	1,90	0,37	7,82
hour11	1,92	1,92	−1,86	5,69
hour12	1,27	1,92	−2,49	5,06
hour13	2,04	1,91	−1,81	5,71
hour14	2,41	1,95	−1,46	6,18
hour15	4,95	1,90	1,10	8,58
hour16	9,07	1,92	5,36	12,72
hour17	8,61	1,93	4,76	12,25
hour18	5,46	1,96	1,60	9,31
hour19	5,33	1,94	1,41	9,11
hour20	6,42	1,94	2,61	10,15
hour21	10,13	1,93	6,26	13,83
hour22	0,23	2,06	−3,74	4,21
hour23	0,21	2,03	−3,84	4,19

Family: student.
Formula: Delay ~ Airline + Hour.
Samples: 3 chains, each with iterations = 3500; warmup = 1750; thin = 1; total post-warmup samples = 5250.

References

Abdelghany, A., Abdelghany, K., Azadian, F., 2017. Airline flight schedule planning under competition. Comput. Oper. Res. 87, 20–39. https://doi.org/10.1016/j.cor.2017.05.013.

AICM, 2018. OPERACIONES DE LLEGADA Y SALIDA SEMANAL CON HORARIO ASIGNADO (Weekly operations of arrival and departure with scheduled times) [WWW Document]. URL. https://www.aicm.com.mx/wp-content/uploads/2019/04/OperacionesdeLlegadaySalida25al31Marz2019.pdf (accessed 4.10.19).

AICM, 2019. Estadisticas del AICM (AICM Statistics [WWW Document]. URL. https://www.aicm.com.mx/categoria/estadisticas/ (accessed 4.13.19).

Airport Council International, 2018. ACI World releases preliminary 2017 world airport traffic rankings [WWW Document]. URL. https://aci.aero/news/2018/04/09/aci-world-releases-preliminary-2017-world-airport-traffic-rankings-passenger-traffic-indian-and-chinese-airports-major-contributors-to-growth-air-cargo-volumes-surge-at-major-hubs-as-trade-wars-thre/ (accessed 12.10.18).

Babić, O., Teodorović, D., Tošić, V., 1984. Aircraft stand assignment to minimize walking. J. Transport. Eng. 110, 55–66. https://doi.org/10.1061/(ASCE)0733-947X(1984)110:1(55).

Brown, H., Prescott, R., 2014. Applied Mixed Models in Medicine. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/9781118778210.

Bürkner, P.-C., 2017. Brms : an R package for bayesian multilevel models using stan. J. Stat. Softw. 80 https://doi.org/10.18637/jss.v080.i01.

Practical introduction to simulation optimization. In: Chick, S., Sánchez, P.J., Ferrin, D., Morrice, D.J. (Eds.), 2003. Proceedings of the 2003 Winter Simulation Conference.

Demidenko, E., 2013. Mixed Models : Theory and Applications with R, second ed. Wiley.

Drexl, A., Nikulin, Y., 2008. Multicriteria airport gate assignment and Pareto simulated annealing. IIE Trans. 40, 385–397. https://doi.org/10.1080/07408170701416673.

Flightstats, 2018. Airport on-time performance reports [WWW Document]. URL. https://www.flightstats.com/v2/monthly-performance-reports/airports (accessed 12.10.18).

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/hierarchical Models. Cambridge University Press.

Ghazouani, H., Hammami, M., Korbaa, O., 2015. Solving airport gate assignment problem using Genetic Algorithms approach. In: 2015 4th International Conference on Advanced Logistics and Transport (ICALT). IEEE, pp. 175–180. https://doi.org/10.1109/ICAdLT.2015.7136615.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Longman Publishing Co., Inc.

Guépet, J., Acuna-Agost, R., Briant, O., Gayon, J.P., 2015. Exact and heuristic approaches to the airport stand allocation problem. Eur. J. Oper. Res. 246, 597–608. https://doi.org/10.1016/j.ejor.2015.04.040.

Hassounah, M., Steuart, G., 1993. Demand for aircraft gates. Transp. Res. Rec. 1423, 26–33.

IAS, 2019. International airport “benito juarez” Mexico city (AICM) – technical specifications [WWW Document]. URL. http://aviationsystem.com.mx/mex/ (accessed 4.13.19).

IATA, 2018. IATA forecast predicts 8.2 billion air travelers in 2037 [WWW Document]. IATA Press Release No. 62. URL. https://www.iata.org/pressroom/pr/Pages/2018-10-24-02.aspx (accessed 1.17.19).

International Airport Review, 2018. Mexico City airport "bursting at the seams" at almost 50% over capacity [WWW Document]. URL. https://www.internationalairportreview.com/news/68356/mexico-city-bursting-seams/ (accessed 12.10.18).

Jaehn, F., 2010. Solving the flight gate assignment problem using dynamic programming. Z. für Betriebswirtschaft 80, 1027–1039. https://doi.org/10.1007/s11573-010-0396-9.

Kennedy, J., Eberhart, R., 1995. Particle Swarm optimization. Proc. IEEE Int. Conf. Neural Netw. IV, 1942–1948. https://doi.org/10.1109/ICNN.1995.488968 [WWW Document]. Neural Networks.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220, 671–680. https://doi.org/10.1126/science.220.4598.671.

Kontoyiannakis, K., Serrano, E., Tse, K., Lapp, M., Cohn, A., 2009. A simulation framework to evaluate airport gate allocation policies under extreme delay conditions. In: Proceedings - Winter Simulation Conference. IEEE, pp. 2332–2342. https://doi.org/10.1109/WSC.2009.5429205.

Mangoubi, R.S., Mathaisel, D.F.X., 1985. Optimizing gate assignments at airport terminals. Transp. Sci. 19, 173–188. https://doi.org/10.1287/trsc.19.2.173.

Marinelli, M., Dell ’orco, M., Sassanelli, D., 2015. A metaheuristic approach to solve the flight gate assignment problem. Transp. Res. Procedia 5, 211–220. https://doi.org/10.1016/j.trpro.2015.01.013.

Moody’s Investors Service, 2018. Research: announcement: Moody’s: air traffic growth boosts sector but Mexico City airport cancellation would be negative - moody’s [WWW Document]. URL. https://www.moodys.com/research/Moodys-Air-traffic-growth-boosts-sector-but-Mexico-City-airport–PR_382448 (accessed 12.10.18).

Mujica Mota, M., 2015. Check-in allocation improvements through the use of a simulation–optimization approach. Transp. Res. Part A Policy Pract. 77, 320–335. https://doi.org/10.1016/J.TRA.2015.04.016.

Mujica Mota, M., Flores, I., 2019. Model-based methodology for identifying capacity limitations in airports: Mexico city airport case. SNE Simul. Notes Eur. 29, 13–25. https://doi.org/10.11128/sne.29.on.10462.

Pearl, J., 1986. Heuristics: intelligent search strategies for computer problem solving. Telematics Inf. 3, 305. https://doi.org/10.1016/S0736-5853(86)80081-8.

Pinheiro, J.C., Bates, D.M., 2000. Mixed-Effects Models in S and S-PLUS, Statistics and Computing. Springer-Verlag, New York. https://doi.org/10.1007/b98882.

SCT, 2013. DECLARATORIA de saturación en el campo aéreo del Aeropuerto Internacional de la Ciudad de México Benito Juárez [WWW Document]. URL. http://www.dof.gob.mx/nota_detalle.php?codigo=5296188&fecha=17/04/2013 (accessed 5.3.19).

SCT, 2017. Capacidad de operaciones [WWW Document]. URL. https://www.aicm.com.mx/wp-content/uploads/2017/12/OF1747CapacidadDeOperaciones.pdf (accessed 12.11.18).

Şeker, M., Noyan, N., 2012. Stochastic optimization models for the airport gate assignment problem. Transp. Res. Part E Logist. Transp. Rev. 48, 438–459. https://doi.org/10.1016/j.tre.2011.10.008.

Stuart, A., Ord, J.K., 2010. Kendall’s Advanced Theory of Statistics, Volume 1, Distribution Theory, sixth ed. Wiley.

Symonds, D., 2018. The future of airports. Passeng. Termin. World Showc. 6–16.

Taillard, E., 1993. Benchmarks for basic scheduling problems. Eur. J. Oper. Res. 64, 278–285. https://doi.org/10.1016/0377-2217(93)90182-M.

Yan, S., Chang, C., 1998. A network model for gate assignment. J. Adv. Transp.

Yan, S., Huo, C.M., 2001. Optimization of multiple objective gate assignments. Transp. Res. Part A Policy Pract. 35, 413–432. https://doi.org/10.1016/S0965-8564(99)00065-8.

Yan, S., Shieh, C.-Y., Chen, M., 2002. A simulation framework for evaluating airport gate assignments. Transp. Res. Part A Policy Pract. 36, 885–898. https://doi.org/10.1016/S0965-8564(01)00045-3.

Yan, S., Tang, C.H., 2007. A heuristic approach for airport gate assignments for stochastic flight delays. European Journal of Operational Research 180 (2), 547–567. https://doi.org/10.1016/j.ejor.2006.05.002.

Yu, C., Zhang, D., Lau, H.Y.K., 2016. MIP-based heuristics for solving robust gate assignment problems. Comput. Ind. Eng. 93, 171–191. https://doi.org/10.1016/J.CIE.2015.12.013.

Zong Woo Geem, Z.W., Joong Hoon Kim, J.H., Loganathan, G.V., 2001. A new heuristic optimization algorithm: Harmony search. Simulation 76, 60–68. https://doi.org/10.1177/003754970107600201.

6 Journal paper II: Reducing airport environmental footprint using a disruption-aware stand assignment approach

This article was published in the journal *Transportation Research Part D: Transport and Environment* and presented the emission-aware modification of DASA, as described in section 3.3.2. Experimental results, discussed in this paper, showed that emission-aware optimisation of stand allocation planning could tangibly reduce pollutant emissions with low deterioration in passenger service level. The most significant reduction could be achieved by combing disruption-aware planning with fast last-minute reallocation decisions.

Transportation Research Part D 89 (2020) 102634

Contents lists available at ScienceDirect

Transportation Research Part D

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

Reducing airport environmental footprint using a disruption-aware stand assignment approach

Margarita Bagamanova[a,c,*], Miguel Mujica Mota[b]

[a] Aeronautics and Logistics Departmental Unit, School of Engineering, Autonomous University of Barcelona, Bellaterra 08193, Spain
[b] Aviation Academy, Amsterdam University of Applied Sciences, Weesperzijde 190, Amsterdam 1097 DZ, the Netherlands
[c] Department of IT & Logistics, Amsterdam School of International Business, Amsterdam University of Applied Sciences, Fraijlemaborg 133, Amsterdam 1102 CV, the Netherlands

ARTICLE INFO

Keywords:
Pollutant emissions
Capacity optimisation
Flight delays
Capacity management
Congestion
Capacity allocation
Decision support
Sociotechnical system

ABSTRACT

Modern airport management is challenged by the task of operating aircraft parking positions most efficiently while complying with environmental policies, restrictions, schedule disruptions, and capacity limitations. This study proposes a novel framework for the stand allocation problem that uses a divide-and-conquer approach in combination with Bayesian modelling, simulation, and optimisation to produce less-pollutant solutions under realistic conditions. The framework presents three innovative aspects. First, inputs from the stochastic analysis module are used in a multivariate optimisation for generating variability-robust solutions. Second, a combination of optimisation and simulation is used to finely explore the impact of realistic uncertainty uncaptured by the framework. Lastly, the framework considers the role of human beings as the final control of operational conditions. A case study is presented as a proof of concept and demonstrates results achievable and benefits of the framework proposed. The experimental results demonstrate that the framework generates less-pollutant solutions under realistic conditions.

1. Introduction

Air transportation provides global freedom of movement for people and cargo. In 2018, approximately four billion passengers and 64 million tons of cargo travelled over 22,000 routes, generating more than 65 million jobs, and a GDP of approximately $2.7 trillion (IATA, 2019a). The demand for air transport passenger services is growing; according to IATA (2018), this trend is expected to continue, and by 2037, the number of passengers travelling by air is expected to double, reaching approximately eight billion passengers per year. The demand for air cargo transportation is also growing. Boeing (2018) predicted annual growth of 4.3% for air cargo operations in terms of revenue tonne-kilometres. The constant growth of demand for air transport services creates additional challenges for airport capacity management and airport environmental protection goals.

With the growth of air transport, related pollutant emissions have been increasing. Graver et al. (2018) reported that CO_2 emissions from aviation increased by 32% in the previous five years. Currently, global aviation generates approximately 2% of all human-induced emissions and 12% of all transport-related emissions (ATAG, 2019). These percentages are expected to increase (Graver et al., 2018), creating additional sustainability challenges for air transport stakeholders.

Aircraft fuel burn is considered to be the main source of air transport pollutant emissions, which include carbon dioxide, nitrogen

* Corresponding author.
E-mail addresses: mm.bagamanova@hva.nl (M. Bagamanova), m.mujica.mota@hva.nl (M. Mujica Mota).

https://doi.org/10.1016/j.trd.2020.102634

oxides, and noise. The largest proportion of these emissions occurs during the cruise phase; however, ground movement of aircraft, including landing, taxiing, and take-off, also contributes significantly to total emissions and affects inhabitants in the proximity of airports (ICAO, 2019a). Aircraft taxiing between the runway exit and the designated stand can generate over one-third of all aircraft emissions outside of the cruise phase, and mostly depends on the distance between the stand and the runway exit/entry points (Fleuti and Maraini, 2017). Thus, it is important to allocate scheduled flights to minimise taxi distance and related fuel burn and emissions.

The stand allocation schedule is often disrupted by changes in flight schedule. Such perturbations may lead to increased turnaround time and decreased airport terminal performance, thereby affecting the level of emissions. Owing to airport congestion, aircraft may have to wait on the ground with their engines on or perform holding manoeuvres in the terminal manoeuvring area (TMA), leading to additional fuel consumption and related emissions. Inefficient management of terminal facilities can propagate schedule perturbations to successive flights and connected airports, increasing the risk of additional emissions. Therefore, efficient management of airport facilities, including stands, is necessary to increase airport capability for addressing perturbations and reducing emissions generated during aircraft ground operations.

This study proposes a novel emission-aware stand assignment approach, based on a disruption-aware stand assignment approach (DASA) introduced in a seminal study by Bagamanova et al. (2020). The methodology proposed in this study combines the benefits of data-mining, evolutionary optimisation, and simulation for generating a stand assignment that minimises pollutant emissions and increases robustness to possible flight schedule deviations while ensuring passenger service quality. The presented emission- and delay-aware stand assignment approach (E-DASA) makes use of airport historical performance data, from which the algorithm learns probabilities of schedule deviations based on characteristics of scheduled flights using Bayesian distributional modelling. The probabilities are considered in calculating the most likely or user-defined level of deviation for each flight in the target flight schedule. The deviations are considered in generating the stand assignment, which is optimised to minimise emissions generated during aircraft taxiing.

The rest of this article is organised as follows. Section 2 reviews related research publications. Section 3 presents the E-DASA methodology and its novel emission-aware component. A case study is presented in Section 4. Conclusions and future research are presented in Section 5.

2. Related research

The stand assignment problem (SAP) approached in this study has been similarly considered by many researchers as the gate assignment problem (GAP). These problems have been researched for many decades from various perspectives in a single-objective as well as in a multi-objective formulation. First works did not consider airport system stochasticity and were more concentrated on minimisation of passenger walking distances (Babić et al., 1984; Hu and Di Paolo, 2007; Mangoubi and Mathaisel, 1985). With the development of the air transport industry, the researchers started to consider technical requirements for aircraft ground-handling (Yan and Tang, 2007) and additional objectives. More attention has been paid to minimisation of towing and stands usage cost (Jo et al., 1997; Prem Kumar and Bierlaire, 2014; van Schaijk and Visser, 2017), improvement of passenger service level and transfer facilitation (Ali et al., 2019; Benlic et al., 2017; Deng et al., 2018; Dijk et al., 2019; Kim et al., 2013a), and maximisation of contact stands use (Dijk et al., 2019; Guépet et al., 2015). Some researchers considered the taxiing phase on the airport ground as a part of their gate/stand assignment study. They concentrated on the minimisation of aircraft idle/taxi time on the ground and therefore, minimisation of taxiways congestion and airline costs. Maharjan and Matis (2012) attempted to minimise taxi-related fuel burn as a part of their binary integer multi-commodity flow network model for the gate assignment. Kim et al. (2013b) proposed gate assignment approach to minimise ramp congestion, as well as passenger transit time in the terminal. Behrends and Usher (2017) proposed to generate gate assignment to minimise aircraft taxi time and applied random selection, genetic algorithm and simulated annealing for optimisation.

Some authors considered real-life stochasticity in the form of schedule perturbations in stand assignments without focusing on taxi movement. They often mitigated disruptions by inserting a uniform buffer time between consecutive flights assigned to the same gate/stand (Deng et al., 2018; Guépet et al., 2015; Maharjan and Matis, 2012). Some researchers instead of applying the uniform buffer times for all assignments proposed to increase individual buffer times on a historical flight disruption value, based on a 95% percentile; thus considering a wider range of possible deviations (Kim et al., 2013a; Prem Kumar and Bierlaire, 2014). In general, inserting buffer times has been proved as an effective solution for minor deviations (up to 30 min) (Hassounah and Steuart, 1993; Yan et al., 2002; Yan and Chang, 1998; Yan and Huo, 2001). Although such buffer times helped to reduce the number of gate conflicts, they also resulted in an increment of assignment problem complexity, increasing the required computational time and leading to lower quality of the outcome of the considered objectives (Prem Kumar and Bierlaire, 2014). Considering future growth of demand, new techniques that go beyond the buffer solution must be developed; buffering significantly reduces airport terminal capacity and may be unfeasible at congested airports.

In the last years, more attention has been paid to the problem of pollutant emissions and their correlation with the growth of economic activities and transportation (Egilmez and Park, 2014; Fisch-Romito and Guivarch, 2019; Wang et al., 2019, 2018). According to Grampella et al. (2017), 1% increment in air traffic movements leads to 1.05% increment in total airport environmental effects. As air transport demand grows, the development of measures for its emissions mitigation becomes highly important for researchers.

Nikoleris et al. (2011) estimated that idling and taxiing states of aircraft movement are the greatest sources of fuel consumption and emissions in an airport, and therefore represent a significant research interest. Many researchers have investigated methods to mitigate pollutant emissions during taxiing through technical improvements. Duinkerken et al. (2013), Ithnan et al. (2013), and Li and Zhang (2017) estimated different taxiing approaches, which included using only one aircraft engine and external engine power, and showed that significant emission reduction can be achieved. Zhang et al. (2019) optimised aircraft taxi time by considering taxiway conflicts and aircraft fuel consumption.

Some researchers concentrated on waiting time emissions reduction by applying different runway congestion-related strategies such as pushback rate control (Simaiakis et al., 2014), runway departure sequence optimisation (Simaiakis and Balakrishnan, 2016; Sölveling et al., 2011), gate holding, de-rated take-offs (Ashok et al., 2017), and departure metering (Murça, 2017). Although the measures proposed by these works helped to reduce the negative environmental impact by almost a third, they also led to increasing stand occupancy times, thereby significantly reducing airport capacity which can be problematic in congested airports.

Many published methods successfully reduced the level of pollutant emissions, however, up to our knowledge, none of them specifically addressed a combination of environmental footprints of schedule disruptions, stand occupancy conflicts, nor human intervention stochasticity on the operational level. Hao et al. (2016) estimated that the lack of predictability in flight times contributes a 1% increase in the amount of fuel consumed, which proportionally increases the emission footprint. Thus, combined measures are necessary that simultaneously address both the level of pollutant emissions from aircraft ground movements, stand capacity usage optimisation and the stand assignment resilience to schedule disruptions under realistic conditions.

To fill the gap in this area, this study proposes an innovative approach that considers disruptions for each flight to create an efficient stand assignment with reduced environmental impact. Furthermore, this study introduces a technique to address the SAP using a divide-and-conquer approach, first identifying the most promising region to explore for the best solution using the optimisation element of the algorithmic architecture, and then focusing on the local exploration for solutions by introducing the stochasticity of the system in a simulation model. The proposed approach is illustrated with a case study in airport infrastructure, in which the stand assignment optimisation algorithm addresses assignment priorities in the scope of emissions. Furthermore, this study demonstrates how the proposed combination of optimisation techniques with Bayesian inference and human intervention can contribute to the airport sociotechnical system while minimising emissions from ground operations and what would be the impact of human interventions on the passenger service level.

3. Methodology

To reduce the negative impact of schedule disruptions on airport operations and efficiency of airport environmental policy, this study uses E-DASA methodology that addresses operational stochasticity and environmental footprint reduction objectives. This section gives a brief description of E-DASA algorithm, which is the base of this study. The approach presented in this study is an emission-aware instance of the general algorithmic architecture presented by Bagamanova et al. (2020) in their seminal study.

E-DASA consists of two components, each with its own functionality and algorithmic logic. Data flow and architecture of E-DASA are illustrated in Fig. 1.

Module I uses an inference technique to learn probabilities of flight disruptions by analysing historical airport performance data. These probabilities are estimated by application of Bayesian distributional modelling, where the target variable (flight arrival time deviation in the scope of this study) is described through its predictors (other variables present in the historical data). The predictor variables could be weather conditions, information about the airline, type of aircraft, aircraft emissions factor, and other variables available in the historical performance data.

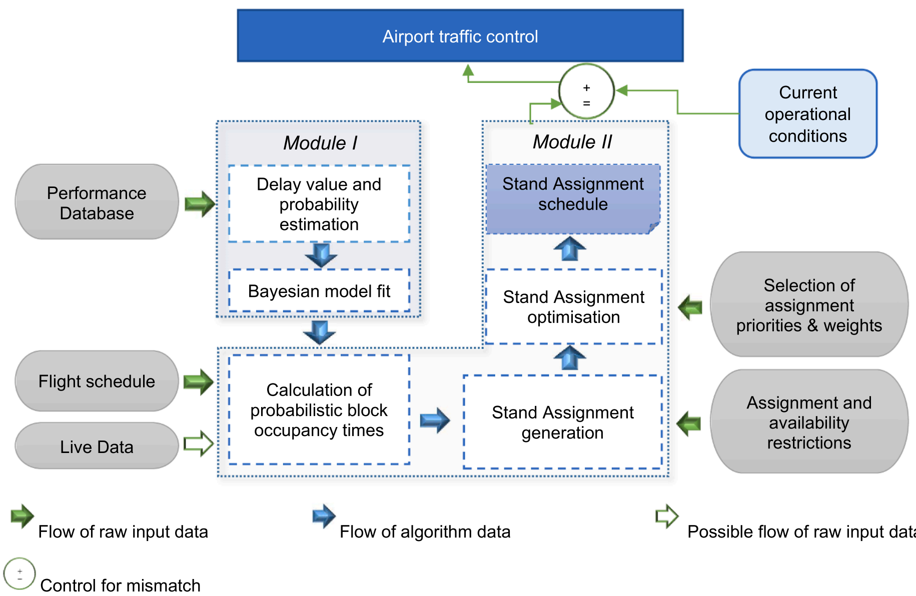


Fig. 1. Architecture of E-DASA.

Inference of schedule disruptions probabilities in Module I is implemented via Bayesian distributional regression modelling. In this technique, response distribution location and shape parameters (e.g. mean, scale, and/or shape) are estimated through predicting variables and response dependence expressed based on the Bayes rule (Stuart and Ord, 2010). In this way, a Bayesian distributional model with the response variable *y,* adopting a certain distribution *D,* and observation *i* can be expressed through $y_i \tilde{} D(\theta_{1i}, \theta_{2i}, \cdots)$, where θ_p are the parameters of the response distribution *D*. Each parameter θ_p is regressed on its own predictor factor η_p through the inverse link function f_p as $\theta_{pi} = f_p(\eta_{pi})^2$. The linear predicting factor η_p can generally be written as $\eta_p = X\beta_p + Zu_p$, where β_p and u_p are the regression coefficients at population-level and group-level, respectively, and *X* and *Z* are the corresponding design matrices (Bürkner, 2018).

When the probabilities of flight disruptions are learnt, their corresponding Bayesian distributional models are transferred as inputs to Module II. In this module, the target flight schedule is analysed, and the most probable flight deviations are calculated based on the Bayesian distributional models from Module I and the characteristics of each scheduled flight. The minimum probability level for generated flight deviations can be set up by the user in the algorithm's input settings, or exact flight deviation values can be drawn randomly from the distributional models for the user-defined probability interval. Such a feature enables the generation of different risk scenarios of stand assignments, which correspond to different levels of likelihood.

When flight deviations are computed, Module II generates a new flight schedule; block occupancy times for each flight are calculated as originally scheduled block occupancy time plus most probable flight schedule deviation value. Module II performs the assignment of an updated schedule to available airport parking positions, considering the probable schedule deviation and optimisation objectives. These objectives can be fine-tuned based on current user preferences. Such flexibility allows for different stand assignments, satisfying different user preferences and goals without the need for reprogramming the entire module.

Owing to stochasticity in the airport system, the solution generated by E-DASA may become unfeasible at some moment during operations. In this case, it is necessary to act to resolve assignment conflicts and maintain the required airport performance level. As it is illustrated in Fig. 1, E-DASA-generated stand assignment can be controlled for feasibility by airport traffic control (ATC) on the day of operations. If any of the planned assignments become infeasible (for instance, due to flight regulation en-route or temporal unavailability of a stand due to technical problems with its equipment or other sources of disturbances not captured by the framework), ATC can reassign the arriving flight to another suitable stand/apron area. How efficient such reassignment is in terms of passenger comfort and taxi-related emissions, depends on the available decision time and availability of fast-working decision support tools for ATC. Therefore, it is necessary to experiment with such interventions to see how they can impact stand assignment KPIs. E-DASA intends to produce a stand assignment with a certain resilience and in such a way that contributes to a better performance of the ATC sociotechnical system.

Modified optimisation component of Module II

The optimisation component presented in this study is an emission-aware modification of the general multi-objective approach first introduced by Bagamanova et al. (2020). We refer to this new algorithm as E-DASA. To consider environmental footprint reduction while providing competitive passenger service, the objective function of Module II optimisation in this study is defined as:

$$minimise(w_1 {*} O_{walk} + w_2 {*} O_{open} + w_3 {*} O_{emis} + w_4 {*} O_{idle}) \tag{1}$$

In this formula the following individual objectives are considered:

1. O_{walk} – the objective to minimise total walking distance for potential transfer passengers:

$$O_{walk} = \sum_{i=1}^{I} N_{paxi} d_{walk} / \sum_{i=1}^{I} N_{paxi} d_{maxwalk}$$

where N_{paxi} is the number of transferring passengers per *i* flight, d_{walk} is the walking distance to a potential connection flight; $d_{maxwalk}$ is the walking distance between two gates located the furthest from each other, and *I* is the total number of flights with transfer passengers.

2. O_{open} – the objective to minimise the number of aircraft assigned to remote stands and to serve more passengers through contact stands:

$$O_{open} = (N_{po} {*} N_{open}) / (N_p {*} N)$$

where N_{po} is the number of passengers in the aircraft assigned to remote stands, N_{open} is the number of aircraft assigned to remote stands; N_p is the total number of passengers on scheduled flights, and *N* is the total number of aircraft in the schedule.

3. O_{emis} – the objective to minimise taxi-related pollutant emissions:

$$O_{emis} = \sum_{n=1}^{N} \sum_{e=1}^{E} B_n H_e F_{ne} (T_n + DT_n) / \sum_{n=1}^{N} \sum_{e=1}^{E} B_n H_e F_{ne} (T_{hold} {*} N) C_t$$

where B_n is the fuel burn rate for aircraft n; H_e is the hazard weight assigned to the emission e; F_{ne} is the emission factor e for aircraft n per unit of fuel burnt; T_n is the taxi time for aircraft n; DT_n is the time penalty if aircraft n is assigned to a 'dummy' stand; T_{hold} is the holding manoeuvre time; N is the total number of aircraft in the schedule; C_t is the holding emission factor increment, calculated as $C_t = f_{appr}/f_{taxi}$, where f_{appr} and f_{taxi} are the engine thrust levels for the approach and taxi phases, respectively. In practice, airport stakeholders can choose the values of H_e to emphasise the impact of certain pollutants according to their toxicity level during the stand allocation.

4. O_{idle} – the objective to minimise the number of aircraft not assigned to any stand and related pollutant emissions:

$$O_{idle} = N_{idle}/N$$

where N_{idle} is the number of aircraft that have been assigned to a 'dummy' stand and N is the total number of aircraft in the schedule.

5. w_1, w_2, w_3, w_4 – indicate priority weights for the corresponding assignment priorities. For practical implementations, different airport stakeholders can decide the weights to reflect different priorities.

In the presented objective function (1), there are conflicting objectives due to the nature of the actors involved. For instance, airlines aim to minimise passenger walking distance (O_{walk}) by locating connecting flights as close as possible to each other. In contrast, airport operators prefer to use contact stands (O_{open}) as often as possible to provide the best service for the airlines and spread allocation for even use of infrastructure. An airport would like to minimise taxi-related emissions and taxi time to the stand (O_{idle}); this objective may conflict with airline preference, as some flights must be allocated to the certain terminal area due to border control procedures, requiring passengers to walk a greater distance to their transfer connection.

Every airport has a stand assignment policy, which implies certain restrictions for the use of stands. The following are the restrictions and assumptions considered in the presented algorithm:

- Domestic and international flights must be assigned to specific stands in the designated zones. These are internal specifications of the airport; e.g. international flights are assigned to stands that have access to designated border control areas.
- An assigned stand must correspond to the size of the aircraft (large aircraft require extra space owing to larger wingspan). This is implemented through the identification of allowed stands for each flight at the input data processing stage in Module II.
- An assigned stand must correspond to airline preferences. This is implemented through the identification of preferred/contracted stands for each flight at the input data processing stage in Module II.
- No aircraft towing movements from one stand to another are considered in the algorithm. Each aircraft occupies its assigned stand for the time equal to its ground-handling time and then taxies to the runway for departure from the airport.
- Flight delays must be considered in the assignment (according to conditional probability distributions from Module I). In this study, only arrival time disruptions are considered in the case study due to unavailability of ground handling data and correspondence of arriving aircraft to departing aircraft.
- When no parking positions are available at the moment of arrival, aircraft should wait on the apron until a position becomes available. This is implemented in the algorithm by assigning the flight to a 'dummy' stand and incrementally delaying its in-block time on DT_n until a suitable stand becomes available.
- For the calculation purposes, holding manoeuvre time T_{hold} should be larger than the maximum possible airport unimpeded taxi time.
- Engine thrust levels for the approach phase f_{appr} and the taxi phase f_{taxi} are equal to 30% and 7%, respectively, based on the ICAO LTO cycle settings (ICAO, 2019b).

The next step in the algorithmic implementation is to consider the stochasticity of the system by simulating the target flight schedule. This is performed using a discrete-event simulation (DES) model of the actual airport system discussed in Section 4. In the model, the obtained stand allocations are simulated under different schedule disruptions scenarios for seven days and the results are discussed.

Table 1
Mexico City International Airport characteristics (AICM and SCT, 2019; IAS, 2019).

	Terminal 1	Terminal 2
Surface area	54.8 ha	24.2 ha
Contact aircraft parking positions	33	23
Remote aircraft parking positions	11	17
Airlines	20	6
Passenger throughput in 2019	29.5 million passengers	20.8 million passengers

4. Case study: Mexico City international airport

Mexico City International Airport (IATA code: MEX) is the main airport in Mexico and 20th in the world ranking of airports by the largest number of aircraft movements, with approximately 450,000 landings and take-offs annually. Twenty-six airlines operate in two terminals at MEX, with international and domestic flights. Terminal buildings are separated by two parallel runways that are not operated simultaneously due to lack of separation distance between them. Such a design significantly restricts MEX capacity; since 2017 MEX has been assessed with a capacity of 61 movements per hour, with a maximum of 40 landings (SCT, 2017). Other relevant information about MEX considered in this study is presented in Table 1.

Fig. 2 illustrates the layout of MEX. Two runways run in parallel from southwest-northeast; runway configuration 05R is most often used for landings, with 05L used for departures. Both terminals have remote parking positions located near runway exits with the shortest taxi distance to them. Approximate location of these positions is shown in Fig. 2.

As observed in Fig. 2, Terminal 2 is located further away from the runways. The average taxi distance for stands in Terminal 1 is 4.2 km, and for stands in Terminal 2 is 5.6 km, which is 33% greater than the average taxi distance for Terminal 1.

4.1. MEX schedule perturbations and emissions

Since 2010, passenger traffic at MEX has grown by an average of 8.5% annually; the number of aircraft movements has grown an average of 4% annually (AICM, 2019). MEX suffers from noticeable schedule disruptions. In 2018, only 67% of all flights at MEX complied with the schedule (SCT, 2019). In 2018, more than 20% of departing flights were delayed, with an average delay of approximately 46 min (Flightstats, 2018).

According to Graver et al. (2018), in 2018 Mexico generated approximately 1.5% of global air passenger traffic-related emissions. The official MEX website does not disclose any information about the level of MEX emissions, or information concerning measures to mitigate the environmental impact of its operations. However, in 2017, Mexico officially joined the global air transport initiative for carbon–neutral operations on a state level, which means that all its airports, including MEX, must follow ICAO emission reduction policies and standards (ICAO, 2020).

Considering the elevated level of schedule perturbations and the recent entry of MEX into the global carbon emission reduction initiative, MEX is an ideal candidate for the current approach to estimate potential emissions reduction.

4.2. Implementation of E-DASA

To estimate the environmental effects of the application of E-DASA at MEX, an official on-time performance report for one week has been used in this study (AICM, 2018). This report consisted of actual and scheduled times of arrival for 3914 flights from 28 May 2018 to 03 June 2018; 53% of the flights were operated by airlines allocated to Terminal 2, and the rest of the flights were operated by airlines located in Terminal 1. In the studied week, the level of schedule disruptions was significant. More than 53% of scheduled flights arrived with a delay of more than 15 min, and more than 36% of flights arrived more than 15 min earlier than scheduled.

In additions to the one-week flight schedule retrieved from the MEX performance report, the following data have been used as input for correct schedule generation in Module II:

- Stand/aircraft size/type of flight correspondence matrix

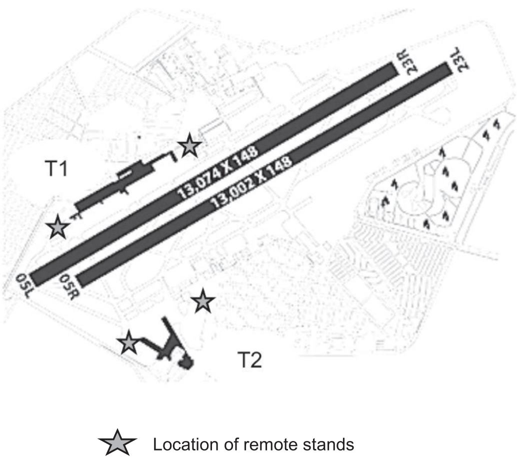


Fig. 2. MEX layout: runways and terminal buildings (Universal Weather and Aviation, 2019).

- Stand/airline correspondence matrix. As actual data about stands preferred/contracted by specific airlines were not available, it was assumed that any airline could use any stand, as long as it corresponded to the airline allocation terminal (retrieved from (AICM, 2017)), type of flight (domestic or international) and aircraft size
- Unimpeded taxi time per stand for the considered runway configuration 05R-landings/05L-departures
- Walking distances matrix for contact stands with walking distance penalisations for remote stands
- Ratio of connecting passengers per flight based on flight origin. Due to unavailability of actual data, these ratios were adapted from IATA (2019b).
- List of airlines offering connecting flights. As actual data were not available, it was assumed that airlines belonging to the same alliance provide such connecting flights.
- Emission factors CO, NOx, HC, and CO_2 and fuel burn rate per aircraft type.

Owing to the unavailability of actual data for the calculation of block occupancy times, the ground-handling times were assumed based on the slots scheduled for the corresponding airlines and aircraft types from AICM (2020). When there was no information available for an airline or aircraft type, ground-handling was assumed to be 120 min for international flights and 60 min for domestic flights. This assumption resulted in the values presented in Table 2. By assuming ground handling times equal to the officially published slots, it was intended to make the calculations as close to reality as possible despite the unavailability of actual data. In real-life operations, ground handling times depend on the aircraft type, airline, airport and available resources among others and can often become one of the sources of disruptions (Fricke and Schultz, 2009; Schultz and Fricke, 2016). In this article, the ground handling time disruptions were not considered, however it would be beneficial for E-DASA to include the probable turnaround time disruptions in the future work.

Furthermore, there were no data available on aircraft engines specifications for the studied flights. Therefore, the aircraft engines and corresponding emissions factors were adapted from ICAO Aircraft Engine Emissions Databank (ICAO, 2019b), as presented in Table 3. This databank contains rates of fuel burn and emissions with CO, NO and HC rates specified for different types of aircraft and various engines, and CO_2 rate calculated as a constant of 3.15 kg of CO_2 per one kg of fuel burnt.

As the considered aircraft emissions depend on the amount of fuel burnt and generated exhaust, CO, NOx, HC and CO2 emissions were calculated as *emissionfactor*fuelburnrate*numberofengines*. Assumptions presented in Table 3 were necessary for illustrative purposes; however, for a real-world application where actual data are available, values corresponding to the actual engines specifications should be used for more accurate results.

Due to congestion at MEX and its location in an urban area, it was decided to heavily penalise assignments to a 'dummy' stand. MEX aerodrome territory does not have sufficient space to safely allocate many waiting aircraft on the apron and holding manoeuvres greatly affect local noise and pollution levels. Thus, for the Module II optimisation algorithm T_{hold} was assumed to be 60 min (compared to the maximum MEX unimpeded taxi time of 12 min). To get an insight on overall MEX emissions, it was decided to assume hazard weight H_e to be equal to 1 for all considered emissions.

Following the workflow in Fig. 1, the target flight schedule was processed in Module I and the corresponding Bayesian distributional models were built for arrival time deviations, describing the likelihood of delays and early arrivals based on the assumed correlation of disruptions with airline and hour of scheduled arrival.

An extract of the obtained model parameters is presented in Table 4. The complete list of model parameters can be found in

Table 2
Assumed ground-handling times, minutes.

Aircraft type	Min of GH time	Average of GH time	Max of GH time
A388	120	120	120
AT42	40	41	70
AT76	55	55	70
B737	60	60	60
B73B	50	62	120
B73S	60	60	60
B73W	50	90	120
B748	120	120	120
B74F	120	120	120
B757	70	70	70
B767	50	92	120
B777	120	120	120
B788	40	109	120
B789	120	120	120
E170	60	64	120
E190	55	66	120
EA19	35	76	120
EA21	30	63	120
EA32	30	67	120
EA33	105	105	105
EA34	120	120	120
SU95	25	59	120

Table 3
Assumed emission factors per aircraft type.

Aircraft type	Number of engines	Engine type	Fuel burn, kg/s/engine	CO, kg/s/engine	NOx, kg/s/engine	HC, kg/s/engine	CO_2, kg/s/engine
A388	4	8RR046	0.3	0.004530	0.001530	0.00006000	0.945
AT42	2	PW124B	0.0988	0.0023771	0.000524628	0.000026	0.31122
AT76	2	PW124B	0.0988	0.0023771	0.000524628	0.000026	0.31122
B737	2	3CM032	0.109	0.002398	0.0004796	0.0002616	0.34335
B73B	2	3CM032	0.109	0.002398	0.0004796	0.0002616	0.34335
B73S	2	1CM004	0.1140	0.003922	0.0004446	0.0002599	0.3591
B73W	2	8CM051	0.1130	0.002124	0.0005311	0.0002147	0.35595
B748	4	11GE139	0.2160	0.004093	0.0009569	0.0001231	0.6804
B74F	4	2GE045	0.1990	0.003827	0.0009413	0.0003065	0.62685
B757	2	5RR038	0.1800	0.003659	0.0007920	0.00004860	0.567
B767	2	1GE012	0.1500	0.004230	0.0005100	0.0009420	0.4725
B777	2	8GE100	0.2960	0.003756	0.001803	0.0001214	0.9324
B788	2	11GE136	0.1990	0.004302	0.0008438	0.0001612	0.62685
B789	2	12RR055	0.2370	0.002003	0.001296	0.00001185	0.74655
E170	2	8GE108	0.06400	0.001162	0.0002950	0.000008320	0.2016
E190	2	11GE146	0.08800	0.003672	0.0003247	0.0003538	0.2772
EA19	2	3CM027	0.09400	0.002820	0.0003572	0.0005828	0.2961
EA21	2	3IA008	0.1363	0.001270	0.0007142	0.00001363	0.429345
EA32	2	3CM026	0.1040	0.002434	0.0004472	0.0004784	0.3276
EA33	2	14RR071	0.2700	0.006472	0.001258	0.0006642	0.8505
EA34	2	8RR045	0.2300	0.002291	0.001401	0.00002990	0.7245
SU95	2	11PJ002	0.10000	0.002755	0.0003820	0.00008200	0.315

Table 4
Sample of obtained regression models characteristics.

Population-Level Effects	Estimate	Estimation Error	Q2.5	Q97.5
Intercept	−10,24	2,02	−14,15	−6,32
Airline AFR	10,21	7,27	−2,63	27,35
Airline AIJ	7,60	1,07	5,52	9,68
Airline AMX	5,32	1,10	3,16	7,46
Airline VOI	9,67	1,18	7,38	11,99
Scheduled arrival hour 03	−5,74	4,11	−13,98	2,26
Scheduled arrival hour 05	8,10	2,02	4,19	12,02
Scheduled arrival hour 06	6,69	1,88	2,98	10,35
Scheduled arrival hour 16	9,07	1,92	5,36	12,72
Scheduled arrival hour 23	0,21	2,03	−3,84	4,19

Appendix A. In Table 4, the "Population-Level Effects" contains levels of the predictor variable; "Estimate" and "Estimation Error" columns contain mean and standard deviation of the effect of the corresponding predictor; columns "Q2.5" and "Q97.5" show limits of 95% confidence interval for the mean effect value. Negative estimate values correspond to flight arrivals earlier than scheduled, positive values correspond to flights delays.

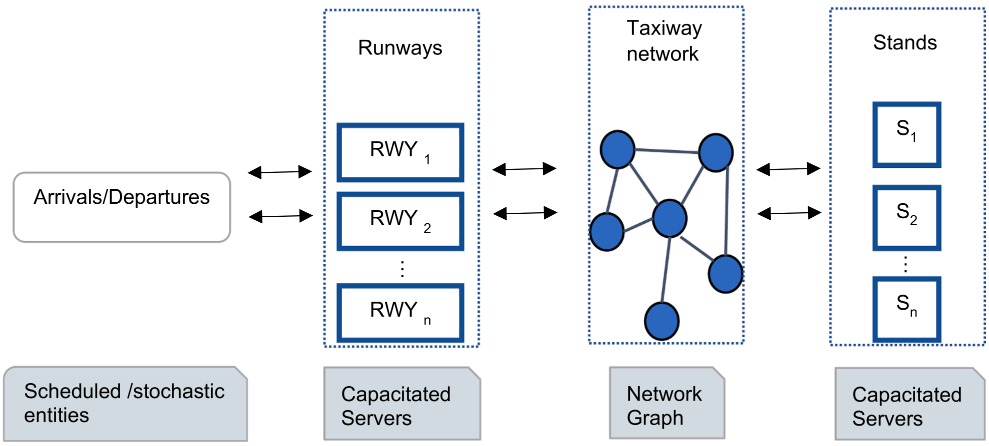


Fig. 3. MEX simulation model framework.

After learning Bayesian distributional models for schedule disruptions, Module II used distributional models to generate the new stand occupancy times, allocated aircraft based on the data matrixes, and optimised the obtained allocation, following the objective expressed in Formula (1). In this study, the new stand allocation schedule was generated considering arrival time deviations with a minimum probability of 60%.

To further estimate the stand assignment generated by E-DASA in close-to-reality conditions, a series of experiments in a simulation model of MEX was executed. A general description of the used MEX simulation model can be found in the next section.

4.3. Simulation model

The MEX simulation model was built under the paradigm of *Discrete Event System* (DES) (Ramadge and Murray Wonham, 1989). This approach implies a dynamic system, whose discrete state values change abruptly due to occurring events (Silva, 2018). The model was built following the concept shown in Fig. 3.

The model consists of the runway system, taxiways, apron areas and stands, which are interconnected by a network of edges and nodes and replicate MEX layout shown in Fig. 2. Entities (aircraft) use those edges as paths to land and depart and to move towards the stands and from them. The edges are scaled to the real distance they represent; all aircraft movements calculations are based on Newton's laws considering the distance, speed, and acceleration of the aircraft. The ground handling operations are modelled by a time consumed by the aircraft at the corresponding stand (server). The taxi operations are modelled by the movement of the aircraft along the edge according to the corresponding taxiing speed limits. The movements of the passengers, buses, pushback tractors and other service vehicles are not modelled.

The MEX model is composed of entities, servers, attributes, and activities. An entity is an object that can move through the system and perform or be a subject of different activities. A server is an object that simulates certain activities of the entities by incurring on a delay that can be deterministic or stochastic. In MEX model, entities represent aircraft; servers represent remote and contact stands, as well as runways. Every entity has attributes which describe its characteristics and can be specified by the user, like the speed of movement, size, flight number, etc. Furthermore, every server has its attributes like processing time, capacity. Activity means a period of time of the specified length. In the model, the following are the activities considered: ground handling, aircraft movement on the runway, aircraft waiting in the arrival/departure queue.

MEX simulation model was implemented using a general-purpose DES commercial simulation software. Nevertheless, the presented framework can be implemented in any DES or multi-agent simulation software. A more detailed description and validation of the MEX simulation model can be found in Mujica Mota and Flores (2019).

4.4. Simulation experiments

The main goal of using a simulation model in this research is to capture sources of stochasticity that occur in the system that were not considered in the allocation algorithm, to make the solutions more realistic. For instance, E-DASA does not consider potential aircraft waiting at the stand due to occupancy of a taxiway or stop-and-go situations that may occur on the airport apron due to numerous aircraft taxiing simultaneously. Such conditions may result in longer taxi times and therefore more emissions. We use the E-DASA output as the input for the simulation model, which enables us to evaluate the potential of the algorithm in more realistic conditions.

4.4.1. Reducing the search space

The SAP is an NP-hard problem in its nature (Guépet et al., 2015); considering the possible combinations of optimisation objectives weights in Formula (1), the set of possible solutions is too large to be entirely tested in the simulation model. Thus, we reduce the search space, identify the most promising area, and then evaluate solutions located in this area under the stochastic conditions of the simulation model.

To restrict the set of possible solutions, the objective function weights w_1, w_2, w_3, corresponding to the minimisation of walking distance, remote stands, and emissions, respectively, were limited to discrete numbers 0 and 1 and the resulting stand allocations were simulated in MEX model. Only w_4, corresponding to the minimisation of unassigned aircraft, remained set to 1 through all scenarios as the stand allocation feasibility requires a minimum number of unallocated aircraft. The results of these simulations compared to the

Table 5
Stand allocations characteristics for different values of objective function weights.

Scenario	Number of arrivals	Number of replications	w_1	w_2	w_3	w_4	$\sum_{i=1}^{I} N_{paxi} d_{walk}$, pax*km	$N_{po}*N_{open}$, pax*stand	$\sum_{n=1}^{N}\sum_{e=1}^{E} B_n H_e F_{ne}(T_n + DT_n)$, tons (average)	N_{idle}
I	3 914	30	1	1	1	1	77 768	297 206 741	1 804.6	0
II	3 914	30	0	1	1	1	78 280	294 015 502	1 803.5	0
III	3 914	30	1	0	1	1	77 605	331 294 112	1 804.1	0
IV	3 914	30	1	1	0	1	77 656	294 949 215	1 823.8	0
V	3 914	30	0	0	1	1	78 232	378 324 408	**1 760.9**	0
VI	3 914	30	0	1	0	1	78 738	**282 749 621**	1 811.2	0
VII	3 914	30	1	0	0	1	**77 520**	339 779 232	1 821.7	0

stand allocation with all weights set to 1 are presented in Table 5. The lowest values for each objective are shown in bold.

From Table 5, all generated stand allocations had zero instances of unassigned aircraft. The lowest product of walking distance and transfer passenger number corresponds to scenario VII for all priorities set to 0 except w_1. The lowest product of passenger number and number of aircraft assigned to open stands was obtained in scenario VI, when w_1 and w_3 were set to 0. In scenarios VI and VII, the level of emissions resulted in a high value.

The difference between the level of emissions in scenario I and scenario V shows that including passenger comfort priorities in the stand allocation optimisation increases the level of emissions by 2.5%. When minimisation of emissions is completely omitted from the goal, as in scenario IV, the resulting stand allocation produces 3.6% more emissions than in scenario V. The lowest emissions value corresponds to scenario V; thus, the solutions generated with this set of weights in the objective function (1) represent the greatest interest for simulation.

The results presented in Table 5 suggest that prioritising only on emission reduction results in a more environmentally friendly stand allocation than with a complex objective. However, as airport stand allocation planning involves many interested parties, such a simplification is not acceptable for airport stakeholders. Nevertheless, generating a stand assignment with a simplified objective function can be useful for analysis of allocation limitations in terms of environmental footprint or any other chosen priority.

4.4.2. *Stochastic search*

Walking distance and contact stand priority weights set to 0 results in solutions with less pollutant footprint; thus, we discarded other possible combinations of weights and focused on the solutions generated under w_1 and w_2 set to 0. The corresponding stand allocations generated by E-DASA were evaluated for MEX emissions reduction potential with the following simulation experiments.

Table 6 summarises five scenarios executed in the MEX simulation model. These scenarios represent different approaches for stand allocation, where planning is optimised to minimise the emissions level. For each scenario, the corresponding CO, HC, NOx, and CO_2 emissions were tracked in the simulation model.

The presented scenarios can be described as follows:

1. Scenario A - a base case, representing ideal on-time arrivals with no disruptions. This scenario shows the level of emissions that can be achieved by pure allocation optimisation without the influence of schedule perturbations.
2. Scenario B - shows emissions that occur under disrupted arrivals if the allocation plan does not consider schedule disruptions and aircraft use only originally planned stands. This scenario includes stochastic arrival time deviations generated with distributions from Module I. If the planned stand is not available at the arrival, aircraft must wait on the apron for the planned stand to become available.
3. Scenario C – the allocation plan does not consider disruptions. This scenario reproduces involvement of ATC (airport traffic control) that manually reassigns aircraft to a random suitable stand if the planned stand is not available at aircraft arrival due to disruptions. This scenario includes stochastic arrival time deviations generated with distributions from Module I.
4. Scenario D – the application of E-DASA that considers probable disruptions in the allocation plan; all aircraft must follow this plan. This scenario includes stochastic arrival time deviations generated with distributions from Module I. If the planned stand is not available at the arrival, aircraft must wait on the apron for stand availability.
5. Scenario E – the application of E-DASA with the involvement of ATC that manually reassigns aircraft to any other available suitable stand if the planned stand is not available at aircraft arrival due to disruptions. This scenario includes stochastic arrival time deviations generated with distributions from Module I.

To replicate close-to-reality airport operations, scenarios C and E simulate possible ATC intervention in daily operations to resolve assignment conflicts. Such interventions often occur in the stochastic airport environment, and often ATC has limited time to determine another stand from the available stands. Due to such time limitations, these decisions are often made without consideration of assignment optimisation, which can impact airport footprint. In such a way, scenarios C and E consider the impact of unoptimised manual reassignments performed by ATC.

4.5. *Experiments results and discussion*

Each scenario presented in Table 6, was run for 178 simulation hours, which is equivalent to seven days of simulated flight schedule plus extra hours for possible arrival time deviations. The stand assignment schedules generated with E-DASA did not require specific

Table 6
List of simulation experiment scenarios.

Scenario name	Number of replications	Number of arrivals	Schedule disruptions	Schedule disruptions considered	Original assignment plan optimisation	Manual reallocation (no optimisation)
A	30	3914	–	–	Yes	–
B	30	3914	Yes	–	Yes	–
C	30	3914	Yes	–	Yes	Yes
D	30	3914	Yes	Yes	Yes	–
E	30	3914	Yes	Yes	Yes	Yes

buffer times between consecutive flights assigned to the same stand, which are often used by airports to absorb arrival deviations and inefficiencies of the turnaround operations (Fricke and Schultz, 2009; Schultz and Fricke, 2016). Excluding mandatory buffer times from allocation allows full observation of the effects of schedule disruption on the emissions level.

It is important to note, that for the simulation purposes it was assumed that an aircraft starts to emit as soon as it leaves the stand and begins the taxi procedure. However, in real-life operations aircraft often start their engines only after being pushed back to the taxiway by a towing tractor, which can be electric or use diesel or LPG. Nevertheless, for the proof of concept objective, which was the goal of the simulation experiments, such detailed modelling of the taxi procedure was considered not essential and therefore was omitted.

The results of executed simulation scenarios for weekly total emission levels statistics are shown in Fig. 4. The results for the total number of aircraft assigned to remote stands weighted to the passengers' number and total walking distance weighted to the transfer passengers' number are shown in Fig. 5.

Scenario A, the base case, does not show any variability in emissions, as all operations were on time and no aircraft waited for stand availability. Emissions produced in this scenario are the lowest among all experiments. When the stochasticity of arrivals is introduced into the simulation in scenario B, total emissions increased by 6%, and there was considerably more variation in total produced emissions. In this scenario, the schedule perturbations were left unattended, and many aircraft waited for the planned stand to become available. It can be concluded that not considering schedule disruptions and not reallocating conflicted flights to another stand results in increased airport pollution.

When manual reassignment of conflicted aircraft by ATC was introduced into the simulation, it decreased unnecessary waiting time. As a result, overall emissions decreased 3.5% compared to scenario B. However, there was still much variation in emissions in scenario C and the average emissions level was 2.4% higher than in scenario A.

The E-DASA allocation, tested in scenario D, was able to decrease emissions by 1.5% compared to the disruption-unaware stand allocation plan in scenario B. However, it could not decrease emissions as well as ATC-assisted reallocation in scenario C. Emissions in scenario D were 2.1% higher than in scenario C. The lowest emissions level in conditions of disrupted arrivals was demonstrated in scenario E. In this scenario, disruption-aware planning generated by E-DASA, combined with ATC assistance for conflicted assignments, reduced emissions by 4.5% compared to scenario B.

Prioritising emissions mitigation penalised passenger walking distance and usage of contact stands, as it can be seen in Fig. 5. The best scenario in terms of emissions (scenario E) resulted in longer walking distances for transfer passengers and lower usage of stands equipped with air bridges. This illustrates the contradictory optimisation objectives considered in Formula (1) that make this situation a challenge for airport decision-makers. In the real-life stand allocation planning, each airport should decide priority weights for each optimisation perspective of the multi-objective function (1). As it is illustrated, in some cases passenger comfort might be sacrificed for improving the environmental situation, but it might positively impact the price of air ticket for passengers owing to the reduction of carbon-offset (Jou and Chen, 2015).

The experimental results demonstrate the advantage of disruption-aware planning for real-life emission reduction. Scenario E illustrated that when E-DASA is not able to address all the stochasticity, the intervention of ATC helps in performing the reallocation with a certain passenger service penalty. These measures allow reducing airport carbon emissions by almost four thousand tons annually, which is equal to the annual CO_2 emissions of 873 typical passenger vehicles (US EPA, 2018).

5. Conclusions and future work

This study presents an innovative approach that combines Bayesian modelling, a multi-objective heuristic optimisation, and simulation for solving airport stand allocation problems. We used a divide-and-conquer approach to reduce the search space, aiming to minimise allocation-related emissions for airports. The presented work utilised simulation to include the variability of real systems and possible stop-and-go conditions that might occur on the airport apron with numerous aircraft taxiing simultaneously. Furthermore, it was demonstrated that the complexity of the stand allocation problem could be reduced by making an initial deterministic optimisation for identifying promising regions that can be further finely explored making use of simulation techniques.

An illustrative case study confirmed the effectiveness of the methodology presented aiming at reducing allocation-related pollutant emissions. The lowest emissions levels could be achieved by relaxing the stand assignment priorities, and by combining the outcome of the framework with airport traffic control intervention if needed. In such a way, the experiments demonstrated that the integration of the presented approach into a sociotechnical airport management system can reduce nearly four thousand tons of emissions per year for the case study presented. The methodology is generic and can be applied to any airport irrespective of the layout, however, it would be more beneficial for large international hubs where the different elements play an important role in the decision process of the allocation of gates.

Besides the contribution of this study, it opens opportunities for further research. For instance, other variables may be considered in Module I to provide increased accuracy in expected schedule deviations like meteorological information and ground-handling disruptions. One of the limitations of the study that can be investigated further is that we did not disaggregate the pushback operation from the complete taxi-out process. The consideration of the pushback will allow the algorithm to prioritise stands that are more environmentally friendly or/and provide a source of aircraft fuel burn reduction e.g. use electric vehicles, ground electricity, preconditioned air. Moreover, it would be important to investigate how changing emissions hazard weights in the objective function would impact the quality of stands assignments.

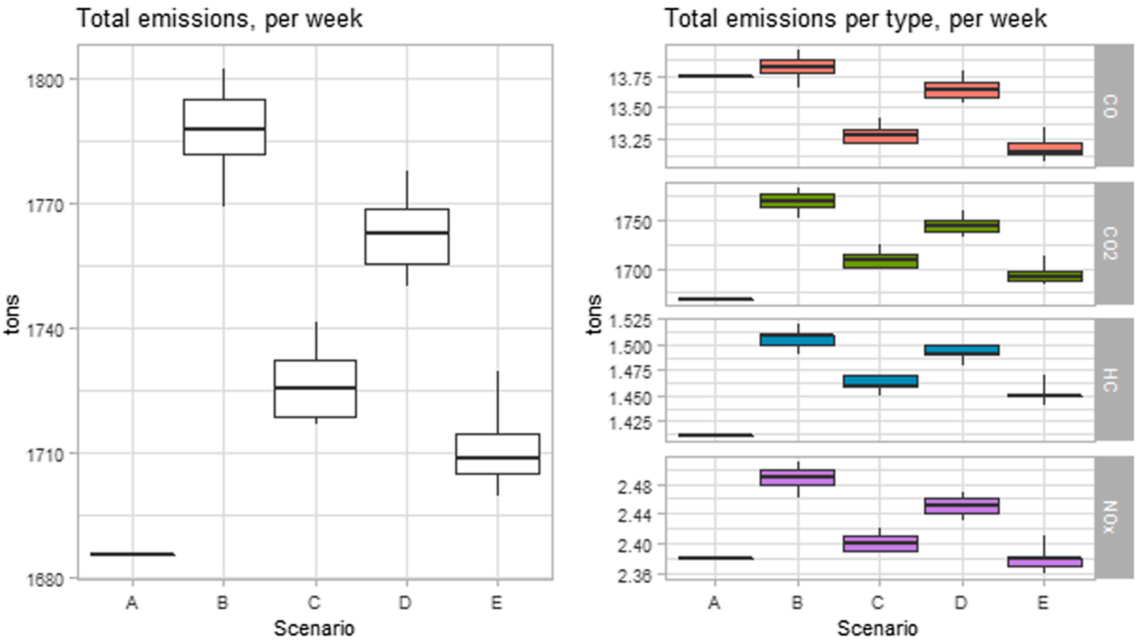


Fig. 4. Experimental results for taxi-related emissions.

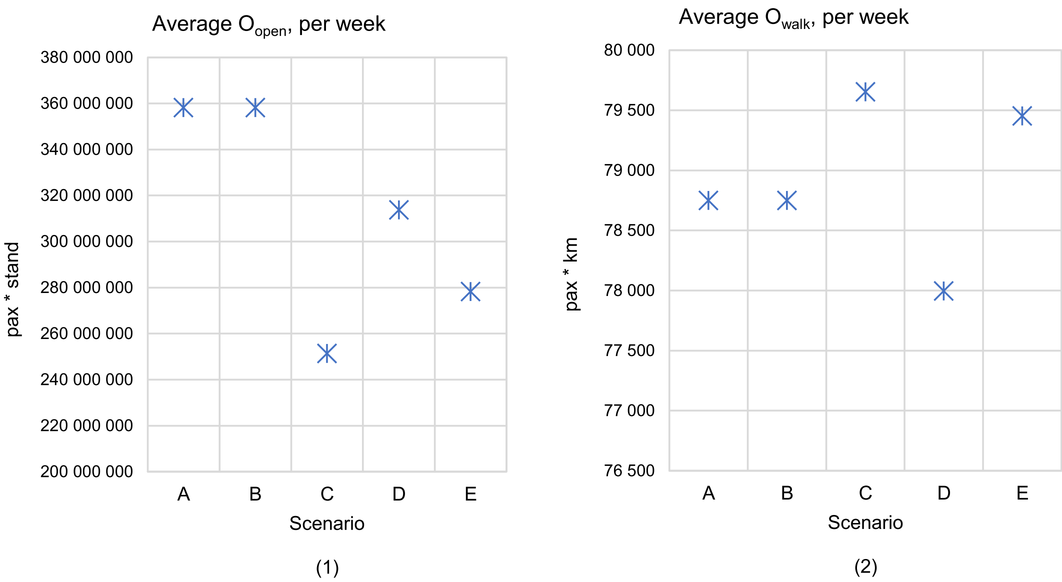


Fig. 5. Experimental results for the passenger-weighted number of aircraft assigned to remote stands O_{open} (1) and transfer passenger-weighted walking distance O_{walk} (2).

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.

Acknowledgements

The authors would like to thank the Autonomous University of Barcelona and the Aviation Academy of the Amsterdam University of Applied Sciences for supporting this study, and the Dutch Benelux Simulation Society (www.dutchbss.org) and EUROSIM for the dissemination of the findings of this study. Furthermore, we would like to express additional gratitude towards the reviewers and editors of this paper for their valuable comments and suggestions that helped to improve the article. The research presented in this paper did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Module I output

Population-Level Effects	Estimate	Estimation Error	Q2.5	Q97.5
Intercept	−10,24	2,02	−14,15	−6,32
Airline ABX	12,92	6,86	−0,38	26,21
Airline ACA	2,40	2,51	−2,78	7,10
Airline AFR	10,21	7,27	−2,63	27,35
Airline AIJ	7,60	1,07	5,52	9,68
Airline AJT	−1,91	11,23	−25,46	18,22
Airline AMX	5,32	1,10	3,16	7,46
Airline ANA	2,31	5,00	−7,08	12,45
Airline ARE	118,79	66,99	35,23	217,63
Airline ASA	1,02	2,88	−4,52	6,78
Airline AVA	14,71	2,47	9,83	19,62
Airline AZA	1,17	5,29	−9,41	11,58
Airline BAW	0,85	4,31	−7,89	9,00
Airline CHH	−3,57	9,39	−19,80	14,89
Airline CKS	80,03	79,31	−45,68	269,10
Airline CLU	36,43	53,99	−39,37	119,48
Airline CLX	8,12	5,93	−2,71	20,43
Airline CMP	−0,70	1,76	−4,18	2,69
Airline CPA	9,14	13,15	−5,69	44,40
Airline CSN	9,23	9,03	−7,15	29,95
Airline DAL	1,42	1,69	−1,93	4,69
Airline DLH	2,51	4,06	−5,23	10,37
Airline ESF	23,35	3,39	16,60	29,95
Airline GEC	−7,74	8,99	−26,17	8,46
Airline GMT	16,60	2,66	11,51	21,89
Airline GTI	190,34	17,29	159,57	221,68
Airline IBE	2,88	2,99	−2,99	8,58
Airline ICL	46,61	12,75	22,16	71,47
Airline JBU	−8,53	2,17	−12,77	−4,30
Airline JOS	10,05	4,63	1,18	19,35
Airline KLM	10,23	4,39	1,02	18,53
Airline LAN	23,80	4,55	14,10	32,35
Airline LPE	1,79	4,08	−6,48	9,57
Airline MAA	58,98	31,69	5,18	112,37
Airline QCL	7,92	10,40	−11,98	29,05
Airline QTR	9,24	7,54	−5,23	24,95
Airline RPB	−0,11	4,94	−9,51	9,99
Airline SKU	210,35	223,77	−69,48	457,40
Airline SLI	4,78	1,05	2,73	6,83
Airline SWA	2,88	1,86	−0,78	6,56
Airline TAI	−0,70	3,00	−6,50	5,41
Airline TAM	12,66	4,51	3,46	21,46
Airline TAO	7,11	1,34	4,53	9,73
Airline TNO	8,32	2,99	2,68	14,37
Airline TPU	9,33	4,87	0,50	20,09
Airline UAE	−1,27	7,42	−14,49	14,96
Airline UAL	3,48	1,42	0,71	6,34
Airline VIV	8,91	1,31	6,39	11,47
Airline VOC	20,09	4,20	11,75	28,27
Airline VOI	9,67	1,18	7,38	11,99
Airline WJA	6,27	2,59	1,26	11,28
Scheduled arrival hour 00	0,27	2,43	−4,29	4,71
Scheduled arrival hour 01	0,23	2,30	−4,33	4,68
Scheduled arrival hour 02	0,72	4,00	−7,53	8,18
Scheduled arrival hour 03	−5,74	4,11	−13,98	2,26
Scheduled arrival hour 04	−5,39	2,48	−10,30	−0,62
Scheduled arrival hour 05	8,10	2,02	4,19	12,02
Scheduled arrival hour 06	6,69	1,88	2,98	10,35
Scheduled arrival hour 07	2,87	1,95	−0,99	6,64
Scheduled arrival hour 08	0,11	1,89	−3,69	3,77
Scheduled arrival hour 09	1,68	1,90	−2,09	5,42
Scheduled arrival hour 10	4,12	1,90	0,37	7,82
Scheduled arrival hour 11	1,92	1,92	−1,86	5,69
Scheduled arrival hour 12	1,27	1,92	−2,49	5,06

(continued on next page)

(*continued*)

Population-Level Effects	Estimate	Estimation Error	Q2.5	Q97.5
Scheduled arrival hour 13	2,04	1,91	−1,81	5,71
Scheduled arrival hour 14	2,41	1,95	−1,46	6,18
Scheduled arrival hour 15	4,95	1,90	1,10	8,58
Scheduled arrival hour 16	9,07	1,92	5,36	12,72
Scheduled arrival hour 17	8,61	1,93	4,76	12,25
Scheduled arrival hour 18	5,46	1,96	1,60	9,31
Scheduled arrival hour 19	5,33	1,94	1,41	9,11
Scheduled arrival hour 20	6,42	1,94	2,61	10,15
Scheduled arrival hour 21	10,13	1,93	6,26	13,83
Scheduled arrival hour 22	0,23	2,06	−3,74	4,21
Scheduled arrival hour 23	0,21	2,03	−3,84	4,19
Family: student				
Formula: Delay ~ Airline + Hour				
Samples: 3 chains, each with iterations = 3500; warmup = 1750; thin = 1; total post-warmup samples = 5250				

References

AICM, 2020. Slots [WWW Document]. URL https://www.aicm.com.mx/en/business/slots (accessed 5.19.20).

AICM, 2019. Estadisticas del AICM (AICM Statistics) [WWW Document]. URL https://www.aicm.com.mx/categoria/estadisticas/ (accessed 4.13.19).

AICM, 2017. Which terminal will my flight depart from? [WWW Document]. URL https://www.aicm.com.mx/en/flights/what-terminal-will-my-flight-depart-from (accessed 8.26.20).

AICM, 2018. Operaciones de llegada y salida semanal [Weekly arrival and departure operations] [WWW Document]. URL https://www.aicm.com.mx/wp-content/uploads/2019/04/OperacionesdeLlegadaySalida25al31Marz2019.pdf (accessed 4.10.19).

AICM, SCT, 2019. AICM en cifras [AICM in numbers] [WWW Document]. URL https://www.aicm.com.mx/acercadelaicm/archivos/files/Estadisticas/Estadisticas2019Dic.pdf (accessed 5.19.20).

Ali, H., Guleria, Y., Alam, S., Schultz, M., 2019. A passenger-centric model for reducing missed connections at low cost airports with gates reassignment. IEEE Access 7, 179429–179444. https://doi.org/10.1109/ACCESS.2019.2953769.

Ashok, A., Balakrishnan, H., Barrett, S.R.H., 2017. Reducing the air quality and CO 2 climate impacts of taxi and takeoff operations at airports. Transp. Res. Part D Transp. Environ. 54, 287–303. https://doi.org/10.1016/j.trd.2017.05.013.

ATAG, 2019. Facts & figures [WWW Document]. accessed 1.26.20. https://www.atag.org/facts-figures.html.

Babić, O., Teodorović, D., Tošić, V., 1984. Aircraft stand assignment to minimize walking. J. Transp. Eng. 110, 55–66. https://doi.org/10.1061/(ASCE)0733-947X(1984)110:1(55).

Bagamanova, M., Mujica Mota, M., Mota, M.M., 2020. A multi-objective optimization with a delay-aware component for airport stand allocation. J. Air Transp. Manag. 83, 101757 https://doi.org/10.1016/j.jairtraman.2019.101757.

Behrends, J., Usher, J.M., 2017. An integrated solution for the aircraft taxi and gate assignment problems. Cogent Eng. 4 https://doi.org/10.1080/23311916.2017.1413722.

Benlic, U., Burke, E.K., Woodward, J.R., 2017. Breakout local search for the multi-objective gate allocation problem. Comput. Oper. Res. 78, 80–93. https://doi.org/10.1016/j.cor.2016.08.010.

Boeing, 2018. World Air Cargo Forecast 2018-2037 [WWW Document]. URL https://www.boeing.com/commercial/market/cargo-forecast/ (accessed 2.2.20).

Bürkner, P.C., 2018. Advanced Bayesian multilevel modeling with the R package brms. R J. 10, 395–411. https://doi.org/10.32614/rj-2018-017.

Deng, W., Zhao, H., Yang, X., Li, D., Li, Y., Liu, J., 2018. Research on a robust multi-objective optimization model of gate assignment for hub airport. Transp. Lett. 10, 229–241. https://doi.org/10.1080/19427867.2016.1252876.

Dijk, B., Santos, B.F., Pita, J.P., 2019. The recoverable robust stand allocation problem: a GRU airport case study. OR Spectr. 41, 615–639. https://doi.org/10.1007/s00291-018-0525-3.

Duinkerken, M.B., Selderbeek, T., Lodewijks, G., 2013. Reducing emissions of taxiing at airports. In: Modelling and Simulation 2013 - European Simulation and Modelling Conference, ESM 2013, pp. 331–335.

Egilmez, G., Park, Y.S., 2014. Transportation related carbon, energy and water footprint analysis of U.S. manufacturing: an eco-efficiency assessment. Transp. Res. Part D Transp. Environ. 32, 143–159. https://doi.org/10.1016/j.trd.2014.07.001.

Fisch-Romito, V., Guivarch, C., 2019. Transportation infrastructures in a low carbon world: an evaluation of investment needs and their determinants. Transp. Res. Part D Transp. Environ. 72, 203–219. https://doi.org/10.1016/j.trd.2019.04.014.

Fleuti, E., Maraini, S., 2017. Taxi-Emissions at Zurich Airport 1–22.

Flightstats, 2018. Airport On-time Performance Reports [WWW Document]. URL https://www.flightstats.com/v2/monthly-performance-reports/airports (accessed 12.10.18).

Fricke, H., Schultz, M., 2009. Delay impacts onto turnaround performance - optimal time buffering for minimizing delay propagation. In: Proceedings of the 8th USA/Europe Air Traffic Management Research and Development Seminar, ATM 2009. pp. 616–625.

Grampella, M., Martini, G., Scotti, D., Tassan, F., Zambon, G., 2017. Determinants of airports' environmental effects. Transp. Res. Part D Transp. Environ. 50, 327–344. https://doi.org/10.1016/j.trd.2016.11.007.

Graver, B., Zhang, K., Rutherford, D., 2018. CO2 emissions from commercial aviation.

Guépet, J., Acuna-Agost, R., Briant, O., Gayon, J.P.P., 2015. Exact and heuristic approaches to the airport stand allocation problem. Eur. J. Oper. Res. 246, 597–608. https://doi.org/10.1016/J.EJOR.2015.04.040.

Hao, L., Hansen, M., Ryerson, M.S., 2016. Fueling for contingencies: the hidden cost of unpredictability in the air transportation system. Transp. Res. Part D Transp. Environ. 44, 199–210. https://doi.org/10.1016/j.trd.2016.02.016.

Hassounah, M., Steuart, G., 1993. Demand for aircraft gates. Transp. Res. Rec. 1423, 26–33.

Hu, X.B., Di Paolo, E., 2007. An efficient Genetic Algorithm with uniform crossover for the multi-objective airport gate assignment problem. In: 2007 IEEE Congress on Evolutionary Computation, CEC 2007. IEEE, pp. 55–62. https://doi.org/10.1109/CEC.2007.4424454.

IAS, 2019. International Airport "Benito Juarez" Mexico City (AICM) – Technical specifications [WWW Document]. URL http://aviationsystem.com.mx/mex/ (accessed 4.13.19).

IATA, 2019a. Annual Review 2019.

IATA, 2019b. The value of air transport in Mexico [WWW Document]. URL https://www.iata.org/contentassets/af6706bf193d44efb2091037373cd74a/mexico-country-report-en.pdf (accessed 6.3.20).

IATA, 2018. IATA Forecast Predicts 8.2 billion Air Travelers in 2037 [WWW Document]. IATA Press Release No. 62. URL https://www.iata.org/pressroom/pr/Pages/2018-10-24-02.aspx (accessed 1.17.19).
ICAO, 2020. CORSIA Participation [WWW Document]. URL https://www.icao.int/environmental-protection/Lists/CORSIAParticipation/AllItems.aspx (accessed 2.12.20).
ICAO, 2019a. ICAO Environmental report. https://doi.org/10.5558/tfc2015-038.
ICAO, 2019b. The ICAO Aircraft Engine Emissions Databank [WWW Document]. URL https://www.easa.europa.eu/easa-and-you/environment/icao-aircraft-engine-emissions-databank (accessed 11.16.19).
Ithnan, M., Selderbeek, T., Beelaerts van Blokland, W., Lodewijks, G., 2013. Aircraft Taxiing Strategy Optimization 1–9.
Jo, G.S., Jung, J.J., Yang, C.Y., 1997. Expert system for scheduling in an airline gate allocation. Expert Syst. Appl. 13, 275–282. https://doi.org/10.1016/S0957-4174(98)00049-9.
Jou, R.C., Chen, T.Y., 2015. Willingness to pay of air passengers for carbon-offset. Sustain. 7, 3071–3085. https://doi.org/10.3390/su7033071.
Kim, S.H., Feron, E., Clarke, J.-P., Marzuoli, A., Delahaye, D., 2013a. Airport Gate Scheduling for Passengers, Aircraft, and Operation, Europe Air Traffic Management Research and Development Seminar.
Kim, S.H., Feron, E., Clarke, J.P., 2013b. Gate assignment to minimize passenger transit time and aircraft taxi time, in: Journal of Guidance, Control, and Dynamics. American Institute of Aeronautics and Astronautics, pp. 467–475. https://doi.org/10.2514/1.57022.
Li, N., Zhang, H., 2017. Calculating aircraft pollutant emissions during taxiing at the airport. Huanjing Kexue Xuebao/Acta Sci. Circumstantiae 37, 1872–1876. https://doi.org/10.13671/j.hjkxxb.2016.0381.
Maharjan, B., Matis, T.I., 2012. Multi-commodity flow network model of the flight gate assignment problem. Comput. Ind. Eng. 63, 1135–1144. https://doi.org/10.1016/j.cie.2012.06.020.
Mangoubi, R.S., Mathaisel, D.F.X., 1985. Optimizing gate assignments at airport terminals. Transp. Sci. 19, 173–188. https://doi.org/10.1287/trsc.19.2.173.
Mujica Mota, M., Flores, I., 2019. Model-based methodology for identifying capacity limitations in airports: Mexico City Airport Case. SNE Simul. Notes Eur. 29, 13–25. https://doi.org/10.11128/sne.29.on.10462.
Murça, M.C.R., 2017. A robust optimization approach for airport departure metering under uncertain taxi-out time predictions. Aerosp. Sci. Technol. 68, 269–277. https://doi.org/10.1016/j.ast.2017.05.020.
Nikoleris, T., Gupta, G., Kistler, M., 2011. Detailed estimation of fuel consumption and emissions during aircraft taxi operations at Dallas/Fort Worth International Airport. Transp. Res. Part D Transp. Environ. 16, 302–308. https://doi.org/10.1016/j.trd.2011.01.007.
Prem Kumar, V., Bierlaire, M., 2014. Multi-objective airport gate assignment problem in planning and operations. J. Adv. Transp. 48, 902–926. https://doi.org/10.1002/atr.1235.
Ramadge, P.J.C., Murray Wonham, W., 1989. The control of discrete event systems. Proc. IEEE 77, 81–98. https://doi.org/10.1109/5.21072.
Schultz, M., Fricke, H., 2016. Improving Aircraft Turnaround Reliability. ICRAT - Int. Conf. Res. Air Transp.
SCT, 2019. Traffic Statistics by Airline/Estadística Operacional de Aerolíneas - International Airport of Mexico City [WWW Document]. URL http://www.sct.gob.mx/transporte-y-medicina-preventiva/aeronautica-civil/7-inconformidades-de-aerolineas/71-demorasindice-de-puntualidad/aeropuerto-internacional-de-la-ciudad-de-mexico-aicm/ (accessed 11.15.19).
SCT, 2017. Capacidad de operaciones [WWW Document]. URL https://www.aicm.com.mx/wp-content/uploads/2017/12/OF1747CapacidadDeOperaciones.pdf (accessed 12.11.18).
Silva, M., 2018. On the history of Discrete Event Systems. Annu. Rev. Control 45, 213–222. https://doi.org/10.1016/j.arcontrol.2018.03.004.
Simaiakis, I., Balakrishnan, H., 2016. A queuing model of the airport departure process. Transp. Sci. 50, 94–109. https://doi.org/10.1287/trsc.2015.0603.
Simaiakis, I., Khadilkar, H., Balakrishnan, H., Reynolds, T.G., Hansman, R.J., 2014. Demonstration of reduced airport congestion through pushback rate control. Transp. Res. Part A Policy Pract. 66, 251–267. https://doi.org/10.1016/j.tra.2014.05.014.
Sölveling, G., Solak, S., Clarke, J.P.B., Johnson, E.L., 2011. Scheduling of runway operations for reduced environmental impact. Transp. Res. Part D Transp. Environ. 16, 110–120. https://doi.org/10.1016/j.trd.2010.09.004.
Stuart, A., Ord, J.K., 2010. Kendall's Advanced Theory of Statistics, Volume 1. Distribution Theory, 6th Edition, 6th ed. Wiley.
Universal Weather and Aviation, 2019. LIC. B. JUAREZ INTL (MMMX/MEX) - MEXICO CITY, DISTRITO FEDERAL, MEXICO. Airport diagram [WWW Document]. URL https://www.universalweather.com/airports-guide/MMMX-MEX-LIC-B-JUAREZ-INTERNATIONAL-AIRPORT-MEXICO-CITY-DISTRITO-FEDERAL-MEXICO/ (accessed 2.9.20).
US EPA, 2018. Emission Facts - Greenhouse Gas Emissions from a Typical Passenger Vehicle. US Environ. Prot. Agency 1–5.
van Schaijk, O.R.P., Visser, H.G., 2017. Robust flight-to-gate assignment using flight presence probabilities. Transp. Plan. Technol. 40, 928–945. https://doi.org/10.1080/03081060.2017.1355887.
Wang, Z., Sun, Y., Wang, B., 2019. How does the new-type urbanisation affect CO 2 emissions in China? An empirical analysis from the perspective of technological progress. Energy Econ. 80, 917–927. https://doi.org/10.1016/j.eneco.2019.02.017.
Wang, Z., Yang, Y., Wang, B., 2018. Carbon footprints and embodied CO2 transfers among provinces in China. Renew. Sustain. Energy Rev. 82, 1068–1078. https://doi.org/10.1016/j.rser.2017.09.057.
Yan, S., Chang, C.-M., 1998. A network model for gate assignment. J. Adv. Transp. 32, 176–189. https://doi.org/10.1002/atr.5670320204.
Yan, S., Huo, C.-M.M., 2001. Optimization of multiple objective gate assignments. Transp. Res. Part A Policy Pract. 35, 413–432. https://doi.org/10.1016/S0965-8564(99)00065-8.
Yan, S., Shieh, C.-Y.C.-Y., Chen, M., 2002. A simulation framework for evaluating airport gate assignments. Transp. Res. Part A Policy Pract. 36, 885–898. https://doi.org/10.1016/S0965-8564(01)00045-3.
Yan, S., Tang, C.-H.H., 2007. A heuristic approach for airport gate assignments for stochastic flight delays. Eur. J. Oper. Res. 180, 547–567. https://doi.org/10.1016/j.ejor.2006.05.002.
Zhang, M., Huang, Q., Liu, S., Li, H., 2019. Multi-objective optimization of aircraft taxiing on the airport surface with consideration to taxiing conflicts and the airport environment. Sustain. 11, 1–27. https://doi.org/10.3390/su11236728.

7 Conclusions and further research

This dissertation aimed at addressing several topics not covered in the existing literature on solving tactical airport stand allocation problem (SAP). Firstly, stand allocations should be able to remain feasible through the operational day despite flight delays and other operational perturbations. Existing SAP solutions use buffer times between flights successively assigned to the same stand. However, such a measure reduces stand capacity and can be problematic in congested airports. Secondly, air transport must reduce its environmental footprint, which means that aircraft movements must be optimised in such a way to produce as little emissions as possible. Existing SAP approaches do not consider emissions directly, but taxi time instead, neglecting differences of aircraft engines and their emission rates.

To address these issues, a multi-component stand allocation framework has been developed in which each component performs a critical function that provides value for the generated solutions. The framework tackles the problems of disruptions and emissions mitigation in the following manner:

- The first component, Module I, estimates probabilities of schedule disruptions and builds Bayesian models for flight delays based on the flight and airport environment characteristics. Predicted schedule deviations then considered during the stand assignment generation and optimisation in the second component, Module II.
- The second component, Module II, uses an evolutionary algorithm to optimise stand allocation for a better passenger service level, improved use of airport facilities, and reduced taxi-related emissions. For each aircraft, fuel burn rate, emission factors, and emissions hazard level are considered to provide a more realistic estimation of the stand allocation-related footprint.
- Generated stand assignments are tested in the airport simulation model, which allows considering stochastic events not captured by the algorithm and facilitates a realistic assessment of the stand assignment quality.

The efficiency of the disruption- and emissions-aware stand assignment has been validated in close-to-reality conditions. The presented framework has been applied to the previously validated simulation model of the Mexico City International airport, and the following contributions have been proven:

- The developed framework reduces assignment conflicts caused by flight delays owing to the consideration of historical disruptions in the assignment schedule.
- The framework creates a more realistic stand assignment schedule, which allows reducing the application of buffer times for absorbing schedule deviations and facilitates a more efficient stand capacity use.

- Consideration of aircraft taxi time, fuel burn rates, and emission factors allows creating a less-pollutant stand allocation, reducing aircraft taxi times and airport environmental footprint.
- Multivariate formulations of the optimisation goals facilitate balancing interests of various airport stakeholders in the stand allocation.
- The use of simulation supports a realistic assessment of the stand allocation quality and provides a better-informed stand capacity management.

Besides enriching literature on the stand and gate assignment problem, the presented work opens possibilities for further research. Some of the potential points for future research include:

- Module I could be enhanced with a direct connection to the airport performance database to facilitate updates of Bayesian models when new data become available.
- Consideration of the towing movements could be added to the algorithm, as well as cost aspects of assignments to specific stands.
- Soft constraints could be added to the algorithm, such as acceptance of overlapping assignment to a certain extent.
- Specification of each step of the ground-handling process could be added to the approach to consider the impact of different aircraft service operators.
- Simulation could be directly incorporated in the optimisation loop to facilitate a self-learning mechanism and provide even more close-to-reality robust solutions.
- The presented framework can serve as a base for the development of a decision-support tool for airport management.

The latter idea will be explored in the next years as a part of a follow-up research project “A-Boost” in collaboration with Airport Research Center GmbH (funded by the Federal Ministry of Transport and Digital Infrastructure of Germany, FKZ 19F2177B).

References

[1] M. Bagamanova and M. Mujica Mota, “A multi-objective optimization with a delay-aware component for airport stand allocation,” *J. Air Transp. Manag.*, vol. 83, p. 101757, Mar. 2020, doi: 10.1016/j.jairtraman.2019.101757.

[2] M. Bagamanova and M. Mujica Mota, “Reducing airport environmental footprint using a disruption-aware stand assignment approach,” *Transp. Res. Part D Transp. Environ.*, vol. 89, p. 102634, Dec. 2020, doi: 10.1016/j.trd.2020.102634.

[3] M. Bagamanova, J. J. Ramos González, M. À. Piera Eroles, J. M. Cordero Garcia, and A. Rodríguez-Sanz, “Modelling dependence of arrival sequencing and metering area transit time on airport meteorological conditions,” in *29th European Modeling and Simulation Symposium, EMSS 2017, Held at the International Multidisciplinary Modeling and Simulation Multiconference, I3M 2017*, 2017, pp. 287–295.

[4] M. Bagamanova, J. J. Ramos González, M. À. Piera Eroles, J. M. Cordero García, and Á. Rodríguez-Sanz, “Identifying and modelling correlation between airport weather conditions and additional time in airport arrival sequencing and metering area,” *Int. J. Simul. Process Model.*, vol. 14, no. 3, p. 213, 2019, doi: 10.1504/IJSPM.2019.101003.

[5] M. Bagamanova and M. Mujica Mota, “No More Surprises: Stand Assignment Algorithm with Likelihood of Turnaround Time Deviation,” *Posters. SESAR Innovation Days 2018*, Dec. 2018. https://www.sesarju.eu/sites/default/files/documents/sid/2018/posters/SESAR Innovation Days 2018 Posters.pdf (accessed Jun. 08, 2020).

[6] M. Bagamanova and M. Mujica Mota, “Reduction of taxi-related airport emissions with disruptions-aware stand assignment: case of Mexico City International Airport,” in *Winter Simulation Conference 2020*, 2020, pp. 516–527, [Online]. Available: https://informs-sim.org/wsc20papers/043.pdf.

[7] ICAO, “Aviation Benefits Report 2019,” 2019.

[8] ATAG, “Aviation benefits beyond borders,” 2018. Accessed: Feb. 12, 2020. [Online]. Available: https://www.aviationbenefits.org/environmental-efficiency/climate-action/operational-improvements/.

[9] International Civil Aviation Organization, “Economic Impacts of COVID-19 on Civil Aviation.” https://www.icao.int/sustainability/Pages/Economic-Impacts-of-COVID-19.aspx (accessed Dec. 08, 2020).

[10] IATA, “20 Year Passenger Forecast.” https://www.iata.org/contentassets/e938e150c0f547449c1093239597cc18/pax-forecast-infographic-2020-final.pdf (accessed Oct. 19, 2020).

[11] EUROCONTROL, “Performance Review Report An Assessment of Air Traffic Management in Europe during the Calendar Year 2018 Performance Review Commission,” 2019. Accessed: Oct. 20, 2020. [Online]. Available: http://www.eurocontrol.int/prc/publications.

[12] EUROCONTROL, “Challenges of growth - European aviation in 2040,” 2018. https://www.eurocontrol.int/publication/challenges-growth-2018 (accessed Oct. 20, 2020).

[13] ATAG, “Facts & figures,” 2019. https://www.atag.org/facts-figures.html (accessed Jan. 26, 2020).

[14] B. Graver, K. Zhang, and D. Rutherford, “CO2 emissions from commercial aviation, 2018,” 2019. https://theicct.org/publications/co2-emissions-commercial-aviation-2018 (accessed Feb.

12, 2020).

[15] ICAO, “ICAO Environmental report,” 2019. doi: 10.5558/tfc2015-038.

[16] E. Fleuti and S. Maraini, “Taxi-Emissions at Zurich Airport,” pp. 1–22, 2017, [Online]. Available: https://flughafen-zuerich.ch/~/media/flughafenzh/dokumente/das_unternehmen/laerm_politik_und_umwelt/taxi_study_zurichairport_20171207.pdf.

[17] ICAO, “Carbon Offsetting and Reduction Scheme for International Aviation (CORSIA).” https://file.veryzhun.com/buckets/carnoc/keys/3fa55da709101d0d937e78732a88cd9d.pdf (accessed Feb. 02, 2020).

[18] ICAO, “Climate Change Mitigation: CORSIA,” 2019.

[19] N. J. Ashford, S. H. Mumayiz, and P. Wright, *Airport Engineering: Planning, Design and Development of 21st Century Airports*. John Wiley & Sons, Inc., 2011.

[20] J. C. Price and J. S. Forrest, *Practical Airport Operations, Safety, and Emergency Management: Protocols for Today and the Future*. Elsevier Inc., 2016.

[21] SESAR JU, “Increased Runway and Airport Throughput.” https://www.sesarju.eu/index.php/projects/earth (accessed Oct. 21, 2020).

[22] SESAR Joint Undertaking, “Enhanced arrivals and departures (Wave 2).” https://www.sesarju.eu/node/3535 (accessed Oct. 21, 2020).

[23] Federal Aviation Administration, “Airport Improvement Program (AIP) – Airports.” https://www.faa.gov/airports/aip/ (accessed Oct. 21, 2020).

[24] EUROCONTROL, “Airport collaborative decision-making (A-CDM).” https://www.eurocontrol.int/concept/airport-collaborative-decision-making (accessed Oct. 21, 2020).

[25] G. S. Daş, F. Gzara, and T. Stützle, “A review on airport gate assignment problems: Single versus multi objective approaches,” *Omega*, vol. 92, pp. 102–146, Apr. 2020, doi: 10.1016/j.omega.2019.102146.

[26] E. Pesch, U. Dorndorf, and F. Jaehn, “Flight gate allocation: Modells, methods and robust solutions,” *7th Int. Conf. Pract. Theory Autom. Timetabling, PATAT 2008*, 2008.

[27] O. Babić, D. Teodorović, and V. Tošić, “Aircraft Stand Assignment to Minimize Walking,” *J. Transp. Eng.*, vol. 110, no. 1, pp. 55–66, Jan. 1984, doi: 10.1061/(ASCE)0733-947X(1984)110:1(55).

[28] E. Taillard, “Benchmarks for basic scheduling problems,” *Eur. J. Oper. Res.*, vol. 64, no. 2, pp. 278–285, 1993, doi: 10.1016/0377-2217(93)90182-M.

[29] J. Guépet, R. Acuna-Agost, O. Briant, and J. P. P. Gayon, “Exact and heuristic approaches to the airport stand allocation problem,” *Eur. J. Oper. Res.*, vol. 246, no. 2, pp. 597–608, Oct. 2015, doi: 10.1016/J.EJOR.2015.04.040.

[30] V. Prem Kumar and M. Bierlaire, “Multi-objective airport gate assignment problem in planning and operations,” *J. Adv. Transp.*, vol. 48, no. 7, pp. 902–926, Nov. 2014, doi: 10.1002/atr.1235.

[31] A. Bouras, M. A. Ghaleb, U. S. Suryahatmaja, and A. M. Salem, “The airport gate assignment problem: a survey.,” *ScientificWorldJournal.*, vol. 2014, p. 923859, Nov. 2014, doi: 10.1155/2014/923859.

[32] C.-H. Cheng, S. C. Ho, and C.-L. Kwan, “The use of meta-heuristics for airport gate assignment,” *Expert Syst. Appl.*, vol. 39, no. 16, pp. 12430–12437, Nov. 2012, doi: 10.1016/j.eswa.2012.04.071.

[33] R. P. Brazile and K. M. Swigger, “An Airline Gate Assignment and Tracking Expert System,” *IEEE Expert. Syst. their Appl.*, vol. 3, no. 2, pp. 33–39, 1988, doi: 10.1109/64.2103.

[34] G. D. Gosling, "Design of an expert system for aircraft gate assignment," *Transp. Res. Part A Gen.*, vol. 24, no. 1, pp. 59–69, Jan. 1990, doi: 10.1016/0191-2607(90)90071-D.

[35] K. Srihari and R. Muthukrishnan, "An expert system methodology for aircraft-gate assignment," *Comput. Ind. Eng.*, vol. 21, no. 1–4, pp. 101–105, Jan. 1991, doi: 10.1016/0360-8352(91)90071-D.

[36] Y. Y. Su and K. Srihari, "A knowledge based aircraft-gate assignment advisor," *Comput. Ind. Eng.*, vol. 25, no. 1–4, pp. 123–126, Sep. 1993, doi: 10.1016/0360-8352(93)90236-Q.

[37] Y. Cheng, "A knowledge-based airport gate assignment system integrated with mathematical programming," *Comput. Ind. Eng.*, vol. 32, no. 4, pp. 837–852, 1997, doi: 10.1016/s0360-8352(97)00001-6.

[38] G. S. Jo, J. J. Jung, and C. Y. Yang, "Expert system for scheduling in an airline gate allocation," *Expert Syst. Appl.*, vol. 13, no. 4, pp. 275–282, 1997, doi: 10.1016/S0957-4174(98)00049-9.

[39] S.-H. Lam, J.-M. N. Cao, and H. Fan, "Development of an intelligent agent for airport gate assignment," *J. Air Transp.*, vol. 7, no. 2, pp. 103–114, 2002, Accessed: Oct. 15, 2020. [Online]. Available: https://www.scopus.com/record/display.uri?eid=2-s2.0-25144489930&origin=inward.

[40] "CAST Stand & Gate Allocation." https://arc.de/cast-stand-gate-allocation/ (accessed Dec. 11, 2020).

[41] INFORM GmbH, "Gates and Stands Management." https://www.inform-software.com/aviation/gates-and-stands-management (accessed Dec. 11, 2020).

[42] R. S. Mangoubi and D. F. X. Mathaisel, "Optimizing Gate Assignments at Airport Terminals," *Transp. Sci.*, vol. 19, no. 2, pp. 173–188, May 1985, doi: 10.1287/trsc.19.2.173.

[43] G. Vanderstraeten and M. Bergeron, "Automatic assignment of aircraft to gates at a terminal," vol. 14, no. 1, pp. 15–25, Jan. 1988, doi: 10.1016/0360-8352(88)90034-4.

[44] J. Xu, G. Bailey, Jiefeng Xu, and G. Bailey, "The airport gate assignment problem: mathematical model and a tabu search algorithm," in *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, 2001, vol. 00, no. c, p. 10, doi: 10.1109/HICSS.2001.926327.

[45] F. Jaehn, "Solving the flight gate assignment problem using dynamic programming," *Zeitschrift für Betriebswirtschaft*, vol. 80, no. 10, pp. 1027–1039, Oct. 2010, doi: 10.1007/s11573-010-0396-9.

[46] S. Yan and C.-M. M. Huo, "Optimization of multiple objective gate assignments," *Transp. Res. Part A Policy Pract.*, vol. 35, no. 5, pp. 413–432, Jun. 2001, doi: 10.1016/S0965-8564(99)00065-8.

[47] X. B. Hu and E. Di Paolo, "An efficient Genetic Algorithm with uniform crossover for the multi-objective airport gate assignment problem," in *2007 IEEE Congress on Evolutionary Computation, CEC 2007*, Sep. 2007, pp. 55–62, doi: 10.1109/CEC.2007.4424454.

[48] A. Drexl and Y. Nikulin, "Multicriteria airport gate assignment and Pareto simulated annealing," *IIE Trans.*, vol. 40, no. 4, pp. 385–397, Feb. 2008, doi: 10.1080/07408170701416673.

[49] M. Marinelli, M. Dell 'orco, and D. Sassanelli, "A metaheuristic approach to solve the flight gate assignment problem," *Transp. Res. Procedia*, vol. 5, pp. 211–220, Jan. 2015, doi: 10.1016/j.trpro.2015.01.013.

[50] H. Ding, A. Lim, B. Rodrigues, and Y. Zhu, "New heuristics for over-constrained flight to gate assignments," *J. Oper. Res. Soc.*, vol. 55, no. 7, pp. 760–768, Jul. 2004, doi: 10.1057/palgrave.jors.2601736.

[51] H. Ding, A. Lim, B. Rodrigues, and Y. Zhu, "The over-constrained airport gate assignment problem," *Comput. Oper. Res.*, vol. 32, no. 7, pp. 1867–1880, Jul. 2005, doi: 10.1016/j.cor.2003.12.003.

[52] U. Benlic, E. K. Burke, and J. R. Woodward, “Breakout local search for the multi-objective gate allocation problem,” *Comput. Oper. Res.*, vol. 78, pp. 80–93, Feb. 2017, doi: 10.1016/j.cor.2016.08.010.

[53] J. Behrends and J. M. Usher, “An integrated solution for the aircraft taxi and gate assignment problems,” *Cogent Eng.*, vol. 4, no. 1, Dec. 2017, doi: 10.1080/23311916.2017.1413722.

[54] M. Hassounah and G. Steuart, “Demand for aircraft gates,” *Transp. Res. Rec.*, vol. 1423, pp. 26–33, 1993.

[55] S. Yan and C.-M. Chang, “A network model for gate assignment,” *J. Adv. Transp.*, vol. 32, no. 2, pp. 176–189, Mar. 1998, doi: 10.1002/atr.5670320204.

[56] A. Bolat, “Procedures for providing robust gate assignments for arriving aircrafts,” *Eur. J. Oper. Res.*, vol. 120, no. 1, pp. 63–80, Jan. 2000, doi: 10.1016/S0377-2217(98)00375-0.

[57] G. Diepen, J. M. van den Akker, J. A. Hoogeveen, and J. W. Smeltink, “Finding a robust assignment of flights to gates at Amsterdam Airport Schiphol,” *J. Sched.*, vol. 15, no. 6, pp. 703–715, Dec. 2012, doi: 10.1007/s10951-012-0292-y.

[58] A. Lim, B. Rodrigues, and Y. Zhu, “Airport Gate Scheduling with Time Windows,” *Artif. Intell. Rev.*, vol. 24, no. 1, pp. 5–31, Sep. 2005, doi: 10.1007/s10462-004-7190-4.

[59] S. Yan and C.-H. H. Tang, “A heuristic approach for airport gate assignments for stochastic flight delays,” *Eur. J. Oper. Res.*, vol. 180, no. 2, pp. 547–567, Jul. 2007, doi: 10.1016/j.ejor.2006.05.002.

[60] U. Dorndorf, F. Jaehn, and E. Pesch, “Modelling Robust Flight-Gate Scheduling as a Clique Partitioning Problem,” vol. 42, no. 3, pp. 292–301, 2008, doi: 10.1287/trsc.1070.0211.

[61] M. Şeker and N. Noyan, “Stochastic optimization models for the airport gate assignment problem,” *Transp. Res. Part E Logist. Transp. Rev.*, vol. 48, no. 2, pp. 438–459, Mar. 2012, doi: 10.1016/j.tre.2011.10.008.

[62] B. Maharjan and T. I. Matis, “Multi-commodity flow network model of the flight gate assignment problem,” *Comput. Ind. Eng.*, vol. 63, no. 4, pp. 1135–1144, Dec. 2012, doi: 10.1016/j.cie.2012.06.020.

[63] C. Yu, D. Zhang, and H. Y. K. Lau, “MIP-based heuristics for solving robust gate assignment problems,” *Comput. Ind. Eng.*, vol. 93, pp. 171–191, Mar. 2016, doi: 10.1016/J.CIE.2015.12.013.

[64] W. Deng, H. Zhao, X. Yang, D. Li, Y. Li, and J. Liu, “Research on a robust multi-objective optimization model of gate assignment for hub airport,” *Transp. Lett.*, vol. 10, no. 4, pp. 229–241, Jul. 2018, doi: 10.1080/19427867.2016.1252876.

[65] D. X. Wei and C. Y. Liu, “Fuzzy model and optimization for airport gate assignment problem,” in *Proceedings - 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2009*, 2009, vol. 2, pp. 828–832, doi: 10.1109/ICICISYS.2009.5358268.

[66] C. Li, “Airport Gate Assignment A Hybrid Model and Implementation,” *arXiv Prepr. arXiv0903.2528*, Mar. 2009, Accessed: Sep. 25, 2020. [Online]. Available: http://arxiv.org/abs/0903.2528.

[67] S. H. Kim, E. Feron, and J. P. Clarke, “Gate assignment to minimize passenger transit time and aircraft taxi time,” in *Journal of Guidance, Control, and Dynamics*, Mar. 2013, vol. 36, no. 2, pp. 467–475, doi: 10.2514/1.57022.

[68] J. Castaing, I. Mukherjee, A. Cohn, L. Hurwitz, A. Nguyen, and J. J. Müller, “Reducing airport gate blockage in passenger aviation: Models and analysis,” 2014, doi: 10.1016/j.cor.2014.02.011.

[69] O. R. P. van Schaijk and H. G. Visser, “Robust flight-to-gate assignment using flight presence probabilities,” *Transp. Plan. Technol.*, vol. 40, no. 8, pp. 928–945, Nov. 2017, doi:

10.1080/03081060.2017.1355887.

[70] B. Dijk, B. F. Santos, and J. P. Pita, “The recoverable robust stand allocation problem: a GRU airport case study,” *OR Spectr.*, vol. 41, no. 3, pp. 615–639, Sep. 2019, doi: 10.1007/s00291-018-0525-3.

[71] C.-H. Tang, S. Yan, and Y.-Z. Hou, “A gate reassignment framework for real time flight delays,” *4OR*, vol. 8, no. 3, pp. 299–318, Oct. 2010, doi: 10.1007/s10288-009-0112-1.

[72] H. Wang, Y. Luo, and Z. Shi, “Real-time gate reassignment based on flight delay feature in hub airport,” *Math. Probl. Eng.*, vol. 2013, 2013, doi: 10.1155/2013/646241.

[73] H. Ali, Y. Guleria, S. Alam, and M. Schultz, “A Passenger-Centric Model for Reducing Missed Connections at Low Cost Airports with Gates Reassignment,” *IEEE Access*, vol. 7, pp. 179429–179444, 2019, doi: 10.1109/ACCESS.2019.2953769.

[74] T. Nikoleris, G. Gupta, and M. Kistler, “Detailed estimation of fuel consumption and emissions during aircraft taxi operations at Dallas/Fort Worth International Airport,” *Transp. Res. Part D Transp. Environ.*, vol. 16, no. 4, pp. 302–308, 2011, doi: 10.1016/j.trd.2011.01.007.

[75] M. B. Duinkerken, T. Selderbeek, and G. Lodewijks, “Reducing emissions of taxiing at airports,” in *Modelling and Simulation 2013 - European Simulation and Modelling Conference, ESM 2013*, 2013, pp. 331–335.

[76] M. Ithnan, T. Selderbeek, W. Beelaerts van Blokland, and G. Lodewijks, “Aircraft Taxiing Strategy Optimization,” pp. 1–9, 2013, [Online]. Available: http://rstrail.nl/new/wp-content/uploads/2015/02/izzudin_ithnan.pdf.

[77] N. Li and H. Zhang, “Calculating aircraft pollutant emissions during taxiing at the airport,” *Huanjing Kexue Xuebao/Acta Sci. Circumstantiae*, vol. 37, no. 5, pp. 1872–1876, 2017, doi: 10.13671/j.hjkxxb.2016.0381.

[78] C. R. Brinton, L. Cook, and S. C. Atkins, “Collaborative airport surface metering for efficiency and environmental benefits,” *Proc. 2007 Integr. Commun. Navig. Surveill. Conf. 7th ICNS Conf. 2007*, no. April, 2007, doi: 10.1109/ICNSURV.2007.384164.

[79] G. A. Monroe, Y. C. Jung, and L. Tobias, “Analysis of environmental impact of eliminating arrival hold short operations for runway crossings at Dallas/ft. Worth Airport,” *AIAA Guid. Navig. Control Conf. Exhib.*, no. March, 2008, doi: 10.2514/6.2008-6826.

[80] G. Sölveling, S. Solak, J. P. B. Clarke, and E. L. Johnson, “Scheduling of runway operations for reduced environmental impact,” *Transp. Res. Part D Transp. Environ.*, vol. 16, no. 2, pp. 110–120, Mar. 2011, doi: 10.1016/j.trd.2010.09.004.

[81] I. Simaiakis, H. Khadilkar, H. Balakrishnan, T. G. Reynolds, and R. J. Hansman, “Demonstration of reduced airport congestion through pushback rate control,” *Transp. Res. Part A Policy Pract.*, vol. 66, no. 1, pp. 251–267, Aug. 2014, doi: 10.1016/j.tra.2014.05.014.

[82] I. Simaiakis and H. Balakrishnan, “A queuing model of the airport departure process,” *Transp. Sci.*, vol. 50, no. 1, pp. 94–109, Feb. 2016, doi: 10.1287/trsc.2015.0603.

[83] A. Ashok, H. Balakrishnan, and S. R. H. Barrett, “Reducing the air quality and CO 2 climate impacts of taxi and takeoff operations at airports,” *Transp. Res. Part D Transp. Environ.*, vol. 54, pp. 287–303, Jul. 2017, doi: 10.1016/j.trd.2017.05.013.

[84] M. C. R. Murça, “A robust optimization approach for airport departure metering under uncertain taxi-out time predictions,” *Aerosp. Sci. Technol.*, vol. 68, pp. 269–277, Sep. 2017, doi: 10.1016/j.ast.2017.05.020.

[85] N. Li, Y. Sun, J. Yu, J. C. Li, H. F. Zhang, and S. Tsai, “An Empirical Study on Low Emission Taxiing Path Optimization of Aircrafts on Airport Surfaces from the Perspective of Reducing Carbon Emissions,” *Energies*, vol. 12, no. 9, 2019, doi: 10.3390/en12091649.

[86] D. Bertsimas and M. Frankovich, “Unified optimization of traffic flows through airports,”

Transp. Sci., vol. 50, no. 1, pp. 77–93, Feb. 2016, doi: 10.1287/trsc.2015.0590.

[87] H. Brown and R. Prescott, *Applied Mixed Models in Medicine*. Chichester, UK: John Wiley & Sons, Ltd, 2014.

[88] E. Demidenko, *Mixed models : theory and applications with R*, 2nd ed. Wiley, 2013.

[89] A. Gelman and J. Hill, *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, 2007.

[90] J. C. Pinheiro and D. M. Bates, *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag, 2000.

[91] Á. Rodríguez-Sanz, F. Gómez Comendador, R. Arnaldo Valdés, J. Cordero García, and M. Bagamanova, “Uncertainty Management at the Airport Transit View,” *Aerospace*, vol. 5, no. 2, p. 59, 2018, doi: 10.3390/aerospace5020059.

[92] P.-C. Bürkner, “brms : An R Package for Bayesian Multilevel Models Using Stan,” *J. Stat. Softw.*, vol. 80, no. 1, pp. 1–28, 2017, doi: 10.18637/jss.v080.i01.

[93] A. Stuart and J. K. Ord, *Kendall’s Advanced Theory of Statistics, Volume 1, Distribution Theory, 6th Edition*, 6th ed. Wiley, 2010.

[94] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Longman Publishing Co., Inc., 1989.

[95] H. Ghazouani, M. Hammami, and O. Korbaa, “Solving airport gate assignment problem using Genetic Algorithms approach,” in *2015 4th IEEE International Conference on Advanced Logistics and Transport, IEEE ICALT 2015*, May 2015, pp. 175–180, doi: 10.1109/ICAdLT.2015.7136615.

[96] M. Mujica Mota, “Check-in allocation improvements through the use of a simulation–optimization approach,” *Transp. Res. Part A Policy Pract.*, vol. 77, pp. 320–335, Jul. 2015, doi: 10.1016/J.TRA.2015.04.016.

[97] A. Abdelghany, K. Abdelghany, and F. Azadian, “Airline flight schedule planning under competition,” *Comput. Oper. Res.*, vol. 87, pp. 20–39, 2017, doi: 10.1016/j.cor.2017.05.013.

[98] J. M. Durán, “What is a Simulation Model?,” *Minds Mach.*, vol. 30, no. 3, pp. 301–323, Sep. 2020, doi: 10.1007/s11023-020-09520-z.

[99] K. (Kurt) Jensen and L. M. (Lars M. Kristensen, *Coloured petri nets : modelling and validation of concurrent systems*. Springer, 2009.

Appendices

Appendix A: Modelling dependence of arrival sequencing and metering area transit time on airport meteorological conditions

M. Bagamanova, J. J. Ramos González, M. À. Piera Eroles, J. M. Cordero Garcia, and Á. Rodríguez-Sanz

Published in the *Proceedings of the 29th European Modeling and Simulation Symposium*, EMSS 2017, Held at the International Multidisciplinary Modeling and Simulation Multiconference, I3M 2017 [3].

MODELLING DEPENDENCE OF ARRIVAL SEQUENCING AND METERING AREA TRANSIT TIME ON AIRPORT METEOROLOGICAL CONDITIONS

Margarita Bagamanova[a], Juan José Ramos González[b], Miquel Àngel Piera Eroles[c], Jose Manuel Cordero Garcia[d], Alvaro Rodríguez Sanz[e]

[a],[b],[c] Department of Telecommunications and Systems Engineering
Autonomous University of Barcelona
Carrer de Emprius 2, 08202 Sabadell, Spain
[d] CRIDA A.I.E. (Reference Center for Research, Development and Innovation in ATM)
Edificio Allende, Avda. de Aragón, 402, 28022 Madrid, Spain
[e] Department of Aerospace Systems, Air transport and Airports
Technical University of Madrid
Plaza Cardenal Cisneros, 3, 28040 Madrid, Spain

[a]Margarita.Bagamanova@uab.cat, [b]JuanJose.Ramos@uab.cat, [c]MiquelAngel.Piera@uab.cat,
[d]jmcordero@e-crida.enaire.es, [e]alvaro.rodriguez.sanz@upm.es

ABSTRACT

Airports are considered complex system in which the coexistence of different actors competing and collaborating for the same resources under operational time uncertainties can cause a poor performance on the overall ATM (Air Traffic Management) system. In order to facilitate the process of decision making to mitigate the propagation of perturbations through the different airport processes a causal model relying on machine learning, using data mining algorithms has been implemented to predict feasible states. This paper introduces a new approach for modelling causal relationships, which can be used for further analysing of feasible scenarios by means of simulation techniques. The state space analysis of reachable airport states is a relevant approach to validate the causal model using a huge amount of historical data for predictive purposes.

Keywords: airport management, Coloured Petri nets, Bayesian networks, decision support tool.

1. INTRODUCTION

The airport is a complex transportation hub serving aircraft, passengers, cargo, and surface vehicles (Office of Technology Assessment 1984). It has three major components: airside, landside and the terminal building, which performs connection between them. Airside is an airport area, where aircraft operate: take off and land, move between the different runways and the terminal. Landside consists of roadways and parking facilities. Terminal complex mainly consists of buildings, serving passengers and air cargo. All these areas are strongly interconnected to each other through different procedures and operations, often fully or partially operated and controlled with the use of IT systems. These operational activities of airports with modernized IT systems are generating an immense amount of data, which can be used for better understanding of hidden dynamics both at the airside and at the landside. However, raw data is quite difficult to be analysed at a glance due to its large volume: for instance, Madrid-Barajas airport airside operations data for one hour of operation with maximum 46 aircraft landed and departed, would make a table of at least 25 different columns with aircraft identification information and data stamps of its movements (landing, taxi in, engine start, taxi out, take off, etc.) and services it went through. The data table of such size can be quite demanding to analyse manually. Therefore for the analysis commodity, these data can be expressed in the form of so called Key Performance Indicators. These Key Performance Indicators (KPIs) are quantitative expressions of effectiveness in achieving performance objectives (European Organisation for the Safety of Air Navigation 2014). As various areas of airport due to their nature can have various KPIs, they are usually merged into Key Performance Areas (KPAs), representing different areas of management interest. An instance of airport KPAs and KPIs is presented in Table 1. The list of KPAs and KPIs can be enlarged according to what targets management team desires to monitor and analyse.

Table 1: Example of KPAs and KPIs (Tabernier 2015)

KPAs	KPIs
Environment / Fuel Efficiency	Average fuel burn per flight.
Airspace Capacity	En-Route and Terminal Manoeuvring Area throughput (average movement per hour).
Airport Capacity	Runway throughput (average movement per hour).
Predictability	Variance of difference in actual & Flight Plan
Punctuality	% Departures < +/- 3 mins vs. schedule due to ATM causes.

Unfortunately, due to tight interdependencies between apparently isolated airport sub processes, airport performance is very sensible to any change in the programmed activities which increase drastically the complexity of airport performance analysis (European Organisation for the Safety of Air Navigation 2017a).
The understanding of the sources of occurred operational issues remains one of the main directions of air transport management scope. Note for instance that European Organisation for the Safety of Air Navigation (EUROCONTROL) aggregates the performance data obtained from European airports and in the form of publicly open documents reveals main European air transport performance problems.
According to one of such reports 2016 was a year with increased volume of flights delay, and furthermore the contribution of reactionary delay has increased up to 45% of total delay minutes (Walker 2017). A reactionary delay is a delay caused by late arrival of aircraft or crew from previous flights (European Organisation for the Safety of Air Navigation 2005). In such manner any delay occurred in the departure airport could lead to severe delays in the following successive flights and their airports of destination. Nevertheless this kind of delay is not the only reason of on-time performance decrease in 2016, as it could be seen on Figure 1.

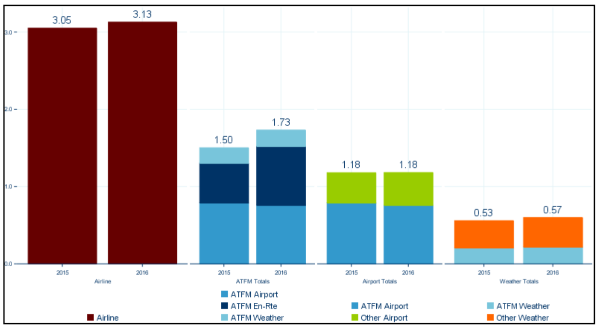


Figure 1: Primary Delay Causes in Europe 2015 vs. 2016, Minutes per Flight (Walker 2016)

Flight delays occurred due to weather conditions also constitute a considerable part of the common delay reasons structure. The fact that weather changes could not be controlled but could be predicted, motivates to obtain the way to efficiently prepare the airport system to any possible impact of weather conditions in order to reduce any negative consequence on its operational activities.
In this paper it is described an approach to model the possible dependence of one of the main airport performance indicators - Arrival Sequencing and Metering Area (further referred to as ASMA) transit time on the weather conditions of arrival airport. Section 2 describes mathematical tools that could be used for the modelling. Section 3 explains the use of Coloured Petri Nets formalism for modelling and simulation of ASMA transit time changes. Section 4 discusses some generated results, directions for further research and some concluding marks are given in Section 5.

1.1. Forecasting in Air Traffic Management

Various organisations perform forecasts for enplanements, airport operations, tracon operations and others. For instance, Federal Aviation Administration (USA) makes its forecast based on demand for aviation services. Econometric and time series modelling are typically used for this purpose. Beside of high potential powerfulness, econometric modelling includes many complex factors and parameters from internal and external infrastructure, which make its application quite difficult and skills demanding. On the other hand, time series modelling seems simpler as it consists of extrapolating knowledge from historical data into the future state. Nevertheless, such extrapolating requires solid statistical analysis and accurate historical data (Federal Aviation Administration 2016).
European Organisation for the Safety of Air Navigation (EUROCONTROL) provides customised analysis and modelling for any airport stakeholders with a use of calculations of performance indicators and different statistical metrics (European Organisation for the Safety of Air Navigation 2017b). International Civil Aviation Organization (ICAO) supports airports planning with medium and long-term forecasts of air traffic for global, regional and route-group levels (International Civil Aviation Organization 2017). These organisations provide open to public global and regional forecasts, however when it comes to the level of particular airport, these organisations could provide only an assistance in analysis and modelling, acting as an external consultant.

1.2. Causal Analysis

Many researchers offer different approaches for understanding and forecasting perturbations of various airport activities. For instance, Quadratic Response Surface (QRS) linear regression models and ensemble Bagging Decision Tree regression (BDT) models have been used to assess weather impact on maximum number of movements per time interval in few USA airports (Wang 2012). Queueing and integer programming models have been used to model the taxi-in process (Idris, Anagnostakis, Delcaire, Hansman, Clarke, Feron and Odoni 1999; Andersson, Carr, Feron and Hall 2000; Roling and Visser 2007). According to the conclusions of these works, the methods used have appeared to be quite helpful, but still not giving perfect approach for airport stakeholders. So the search needs to be continued.
Current modernisation initiatives Single European Sky ATM Research Programme (SESAR) in Europe and NextGen in USA impulse implementation of new operational concepts and technologies, aiming to transform current aviation network into highly efficient, robust and cost optimised system. In order to reach such efficiency it is necessary to understand and fully control any performance area of airport system. For measuring level of success in these tasks airport management can use performance indicators, which permit to compare actual and planned functionality of airport.
It is important to remember, that airports are not operating in isolated conditions, instead, airport

operational disruptions could generate severe reactionary delays through the full aviation network. Thus, it is important the research on new efficient tools for the causal analysis of operational deviations and its prediction, considering the operational conditions that affects each particular airport for the design of mitigation mechanism in the own airport but also at network level.

2. DATA RELATIONSHIP DISCOVERY

We have been provided with Key Performance Indicators data for year 2015, used by analysists of CRIDA (Reference Center for Research, Development and Innovation in Air Traffic Management) and the data from the METAR report, consisting of recorded meteorological conditions on the territory of Madrid-Barajas airport. Some of them are listed in Table 2.

Table 2: KPAs and KPIs

KPA	KPI
TMA	Percentage of flights with holding
	Separations - en NM
	Additional time in ASMA
AIRPORT	Real turnaround time compared to planned
	Additional taxi-out time
	Time between consecutive operations on a runway
	Regulated departures adjustment to CTOT
Capacity	Difference between capacity and demand
	Available capacity
Predictability	Punctual arrivals
	Punctual departures
	Arrivals’ standard deviation
	Departures’ standard deviation
Meteorology	Wind direction Variable wind direction
	Wind intensity Gusts of wind
	CAVOK
	Predominant visibility Minimal visibility
	Temperature Dew Point Atmospheric pressure
	Phenomenon Cloudiness

For various KPIs’ data has been provided in a different form. Some values have been measured for one hour interval, others for 20 minutes interval. Meteorological data consisted of observations for every 30 minutes. Furthermore, we have been commented by CRIDA analysts on the particular interest of discovering hidden causes of perturbations of time in ASMA of radius of 60 nautical miles (NM), expressed as additional ASMA transit time (current performance reports are performed for ASMA with radius of 40 NM).

It has been noted (Klein, Kavoussi and Lee 2009; European Organisation for the Safety of Air Navigation 2015) that weather impact on airport performance is quite significant, but yet not studied well enough. Therefore it has been chosen to study weather impact on one of the KPIs of Madrid-Barajas airport. For the scope of this paper study of weather conditions impact on airport functionality, the following available data has been considered of the first study interest:

- Additional ASMA transit time – a difference between actual time spent by aircraft in ASMA area with radius of 60 NM and average time, statistically measured for particular type of aircraft (for the modelling purpose shortly referred to as ASMA).
- Number of flights with holding patterns – number of flights, which have to take a special route around aerodrome in order to wait for an appropriate moment for landing. (H)
- Wind direction (Wind) and wind intensity (WI).
- Predominant visibility on the aerodrome territory (Vis).
- Dew point (DP).
- Atmospheric pressure (Pres).
- Weather phenomenon type – if fog or any other similar phenomenon occur (Fen).
- Cloudiness (Cloud).

Among the different analytical tools (Marsland 2015; Song 2007) to discover relationship structure between observed variables, the construction of Bayesian networks seems to provide a promising approach to better understanding of complex systems, such as airport, thanks to its capability to cope with high-dimensional problems of different data types (Marsland 2015; Song 2007; Xu, Laskey, Chen, Williams and Sherry 2007) and many powerful computer programs, that made any related computations easy and rather fast.

2.1. Bayesian Networks

Bayesian networks are commonly used for representation of a knowledge about an uncertain area (Song 2007). A Bayesian network is a graphical representation of relationships between different variables, where given variables are represented as nodes and their probabilistic dependencies of each other are represented as directed arcs connecting the nodes. In such manner the absence of direct arc between some two nodes means that these two nodes are conditionally independent of each other (Marsland 2015). When a node has an outgoing arc, it is called *parent*, the nodes with incoming arcs are called *children*. The joint probability distribution P_X of the chosen variables X is represented as a product of conditional probability distributions of each variable X_i (Nagarajan, Scutari, and Lèbre 2013):

$$P_X(X) = \prod_{i=1}^{p} P_{X_i}\left(X_i \middle| \Pi_{X_i}\right) \quad (1)$$

Through the conditional probability distribution, calculated for every variable of the studied data, it is also possible to conclude about posterior or future data values. This conclusion is expressed as likelihood function and could serve as the base for prediction model (Gelman, Carlin, Stern, Dunson, Vehtari and Rubin 2014).
The task of discovering a Bayesian network fitting the data consists of two phases: structure and parameter learning. Various algorithms have been developed for the first phase execution. However among all of them only two algorithms have been chosen for the purposes of this paper – Silander - Myllymäki (SM) (Silander and Myllymäki 2006) and Max-Min Hill-Climbing (MMHC) (Tsamardinos, Brown, and Aliferis 2006) algorithms. These algorithms combine constraint-based and score-based algorithms strong sides and are claimed to be highly effective in various situations (Nagarajan, Scutari, and Lèbre 2013). However the approaches, used by these algorithms, are quite different.

2.1.1. Silander – Myllymäki Algorithm
This algorithm was developed for discovering the globally optimal Bayesian network without any structural constraints (Silander and Myllymäki 2006). In order to find the optimal network structure for the specific data, the algorithm has to perform several steps:

1. Find the best parents for all $n2^{n-1}$ pairs of variables, taking the calculated scores for $n2^{n-1}$ as a choice criteria (the higher the score values, the better is the fitness of a candidate variable as a parent).
2. Find the best children node, which cannot be a parent to any other variable.
3. Based on the results of Step 2, find the best arrangement of the variables.
4. Find a best network, taking into account the results of Step 1 and 3 (Silander and Myllymäki 2006).

Despite of quite high quality of the possible SM algorithm results, it has some computational complications. Thus according to the experiments performed by the authors of SM algorithm, the memory requirement for discovering a network of 32 variables is about 16 GB, although distribution of the computation process among few computers could help to overcome this restriction (Silander and Myllymäki 2006). Still, as finding a globally optimal network is NP-hard (Chickering, Meek, and Heckerman 2004), the computational time for SM algorithm is rather long and could easily take 50 hours for a dataset of 30 variables (Silander and Myllymäki 2006). Therefore in order to speed up the discovering of Bayesian network, the use of faster performing algorithm has to be considered as well. One of the most popular algorithms (Nagarajan, Scutari, and Lèbre 2013) with this characteristic is Max-Min Hill Climbing (MMHC) algorithm.

2.1.2. Max-Min Hill-Climbing Algorithm
This algorithm combines principles from local learning and both constraint-based and search-and-score techniques. First, it reconstructs the skeleton of a Bayesian network, and then orients the arcs by performing a Bayesian-scoring greedy hill-climbing search (Tsamardinos, Brown, and Aliferis 2006).
This algorithm has many similarities with the Sparse Candidate (SC) algorithm, which was one of the first successfully performing approaches, applied to large datasets with several hundred variables (Friedman, Linial, and Nachman 2000). Both SC and MMHC perform stepwise reduction of a candidate parents set for each variable and then search for a network that maximise a chosen scoring function. However they do have one important difference. The SC algorithm performs the reduction and network search steps iteratively until there is no improvement in the scoring function value, MMHC performs the candidate parent estimation only once (Nagarajan, Scutari, and Lèbre 2013), therefore fastening the computational process by several times without significant loss in correctness (Tsamardinos, Brown, and Aliferis 2006).

2.1.3. Data Preparation and Learning the Network Structure
As the dataset, provided for analysis, consisted of data for different time intervals, first it has been necessary to transfer all KPIs to the same time interval for facilitation of analysis. It was considered to perform the analysis of data for the time interval of the size of one hour (most common interval of observation that have been seen in the KPIs´ dataset). All chosen for analysis KPIs´ with smaller time interval of observations have been aggregated till the level of one hour.
Additionally, it has been noticed, that provided KPIs values do not all have the same character of values. Some KPIs are observed as *continuous* variables, others – as *discrete*:

- Continuous variable - variable, that can take on any real value within certain interval (Joshi 1989); for instance, additional ASMA time is expressed in minutes.
- Discrete variable - can take on only certain values (Joshi 1989); for instance wind intensity.

Presence of such mixed data can potentially cause a problem in the step of defining a probabilistic model, fitting the data (Nagarajan, Scutari, and Lèbre 2013). Therefore it has been decided to perform a common used solution to avoid the mentioned problem – perform *discretization* or *binning* of the data. Discretization means assigning some particular integer value to the certain intervals of continuous data. There are different ways to define the intervals for data discretization: using expert knowledge on data, using heuristics, performing discretization and structure learning iteratively, etc. (Nagarajan, Scutari, and Lèbre 2013). Taking into account common practice of KPIs´ analysis by CRIDA experts, it has been decided to discretise continuous data as shown in Table 3.

Table 3: Intervals of Discretization

Category	Additional ASMA time, % of unimpeded ASMA time	Visibility, m	Wind direction, °	Actual temperature minus Dew point temperature	Cloudiness	Phenomenon, type	Pressure, QNH
1	(∞;-15)	<50	(22,5; 67,5]	0	FEW	BR	<1013
2	[-15; -10)	[50; 400)	(67,5; 112,5]	-	SCT	DZ	1013
3	[-10; -5)	[400; 8000)	(112,5; 157;5]	-	BKN	FG	>1013
4	[-5; 5)	≥8000	(157,5; 202,5]	-	OVC	RA	-
5	[5; 10)	-	(202,5; 247;5]	-	-	SN	-
6	[10; 30)	-	(247,5; 336,5]	-	-	-	-
7	≥30	-	(337,5; 22,5]	-	-	-	-
0	0	-	VRB	<>0	0	0	-

In Table 3 the following abbreviation have been used:

- VRB – variable wind direction.
- FEW – few clouds.
- SCT – scattered.
- BKN – broken clouds.
- OVC – overcast.
- BR – mist.
- DZ – drizzle.
- FG – fog.
- RA – rain.
- SN – snow.

After data preparation both SM and MMHC algorithms have been executed subsequently in the framework of R software.

As soon as both algorithms have performed their Bayesian network learning for the chosen airport performance data, the best network can be chosen based on the best value of the network scoring functions. Both algorithms have a possibility to evaluate the learnt network with three popular statistical scoring functions: BDeu (Bayesian-Dirichlet equivalent uniform), AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). These scoring functions are common tools for selection between different statistical models and represent goodness of fit of a model to observed data (Brockwell and Davis 1991).

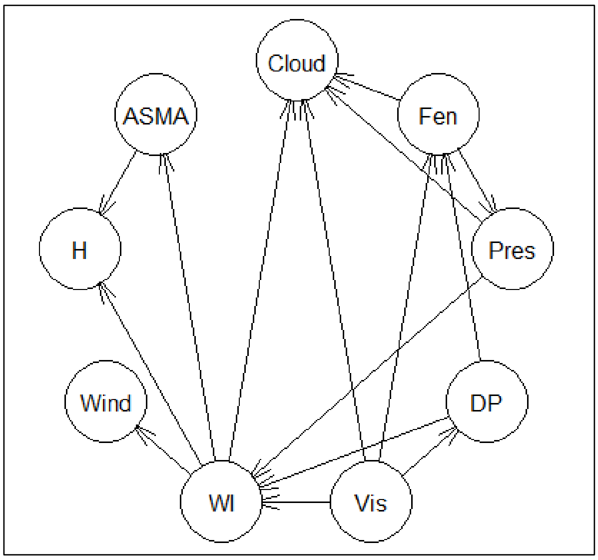


Figure 2: Bayesian network obtained with MMHC and SM Algorithms

In the case of chosen for this paper variables, both SM and MMHC algorithms have come to the same network structure, shown on Figure 2, therefore it was decided not to compare their score functions. Every arc of obtained network had probability of being true of not less than 95% and as MMHC algorithm has come to its results in a shorter computational time (less than one minute for Intel (R) i5-4300M CPU 2.60 GHz, 8 GB RAM), it has been considered to use its results for the further study.

2.2. Bayesian Inference

The knowledge obtained from Bayesian Networks about the data structure and its parameters is used for reasoning on further possible parameters of the chosen airport performance indicators. There are two main approaches for updating the posterior probabilities of data distribution: exact and approximate inference.

Variable elimination and Junction Tree are the two best-known approaches for exact inference task. First approach uses the network structure directly, taking into account the local distributions of the data variables. On contrary, the second algorithm transforms the network by clustering its nodes into a tree. However the feasibility of exact approach is restricted to small networks. Approximate inference algorithms create samples from the local distributions by the use of Monte Carlo simulations and then evaluate them. The sampling can be performed in different ways, implemented in several approximate algorithms (Nagarajan, Scutari, and Lèbre 2013).

The parameters learnt in this step take the form of regression coefficients, belonging to regression functions, describing the conditional dependence between studied variables. For this research it is considered to use the logic sampling approximate inference algorithm, already included in functionality of one of the R software packages for Bayesian Networks. The inferenced parameters of a network have been used for mathematical expression of relationships between observed variables in arc expressions of CPN model in order to perform simulation run and state space analysis.

3. MODELLING WITH CPN FORMALISM

A Coloured Petri Net (CPN) is a formalism, aimed to design, visualise and explore the behaviour of various systems. In order to model the system with CPN formalism it is necessary to define a set of parameters as (Jensen and Kristensen 2009):

- Set of colours – to represent the model entities (key performance indicators).
- Set of places nodes – to represent combinations of the model entities.
- Set of transition nodes – to represent systems' activities (weather changes, arriving aircraft, etc.).
- Set of Arcs – to relate transition and places nodes.
- Guard functions, which are associated to the transition nodes in order to insure their enabling only in case of satisfaction of conditions, described in the corresponding guard function.

For the net elements inscriptions CPN ML, a functional programming language, is included to the modelling framework. It provides the way to make different declarations and perform modelling of data manipulation (Piera and Musič 2010). This language is used in construction of arc functions and in declarations of intervals of possible values for model parameters.

For modelling the chosen KPIs of Madrid-Barajas airport, the colours, representing weather indicators, average additional ASMA time, and number of flights with holding pattern have been chosen. Furthermore it has been considered to introduce the colour, representing system time counter, for having a tool to track system dynamics in time without increasing model complexity.

Design of the developed CPN model is shown on Figure 3.

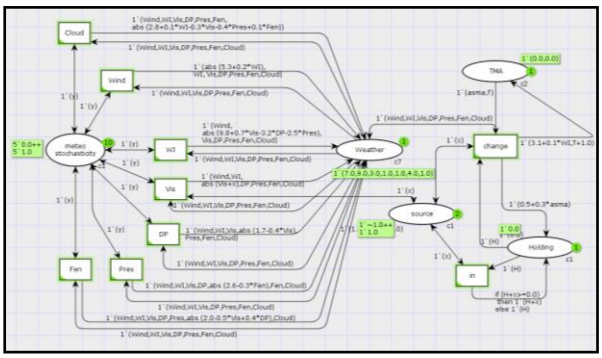


Figure 3: Weather Indicators CPN Model Design

The studied KPIs are distributed among three places as follows:

- Place Holding – number of flights with holding pattern (H).
- Place TMA – additional ASMA time and time counter.
- Place Weather – Wind, WI, Vis, DP, Pres, Fen and Cloud.

Furthermore two supporting places, ensuring the element of stochasticity, have been also added to the model. They are:

- Place Meteo stochasticity – provides tokens for stochastic weather changes.
- Place Source – provides tokens for stochastic changes in number of arriving flights with holding pattern.

In order to formulate the observed ASMA system behaviour in CPN Tools, it is required to define functions for the expressions of arc, connecting elements of the model. The arc functions have the following aspect, based on the maximum likelihood estimation parameters, obtained on the step of Bayesian inference:

$$C_i = \beta + k * C_j \tag{2}$$

Where

C_i = represents CPN colour i, a studied metric.

β = represents intercept value.

k = represents regression coefficient.

C_j = represent CPN colour j, on which CPN colour i is conditionally dependent. When there are more metrics, on which colour i is conditionally dependent, they are included with the corresponding regression coefficients.

After introducing all necessary system parameters, series of simulation runs can be executed in order to verify and validate the model.

4. SIMULATION AND RESULTS

Default tool of CPN Tools v. 4.0.1, which can be used for model verification, is the state space analysis. This analysis consists of generating all states and state changes of a model, that can be reached from the initial state (Jensen and Kristensen 2009). CPN Tools v. 4.0.1 allows to graphically represent all possible system states through *reachability tree* (RT) – a directed graph, where root node represents initial marking of the system, and the successive nodes represent the new states, that can be reached from the initial state, if the corresponding transitions have been fired (Jensen and Kristensen 2009). Few series of state space construction (reachability tree generation) have been performed with a use of CPN Tools v. 4.0.1 software in order to explore how parameters of the system – colours, change their values. The initial markings of the model, used for state space analysis are shown in Table 4. These values have been chosen from the available historical data for the same time period as for Bayesian inference, in order to compare the system dynamics, observed in the historical data and the changes, discovered through RT construction.

Table 4: Simulation Scenarios Initial Markings

Model parameters	Scenario 1	Scenario 2	Scenario 3
	Parameter value		
ASMA time	0	3	0
Flights with holding	0	2	0
Wind direction	7	0	7
Wind intensity	0	2	9
Visibility	4	2	3
Dew point	0	1	1
Pressure	3	1	1
Phenomenon	0	4	4
Cloudiness	0	2	1

In the RT generated for all three chosen scenarios, in every tree a branch with the same weather indicators changes has been found. This has allowed to compare how ASMA transit time has developed in these RT branches and in the historical data. Figure 4, 5 and 6 represent this comparison for each of three simulation scenarios respectively for the time period of 24 hours.

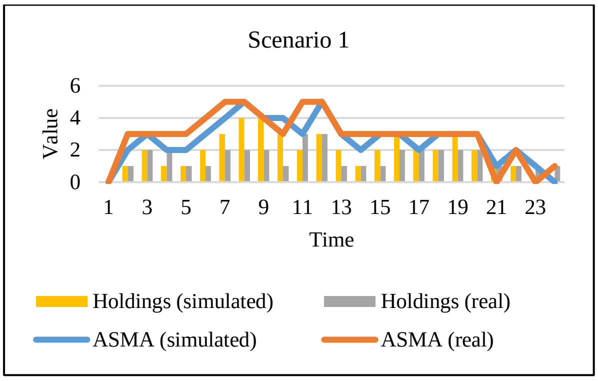


Figure 4: CPN Simulated ASMA Transit Time, Real ASMA Transit Time, CPN Simulated Holdings and Real Holdings Comparison for Scenario 1

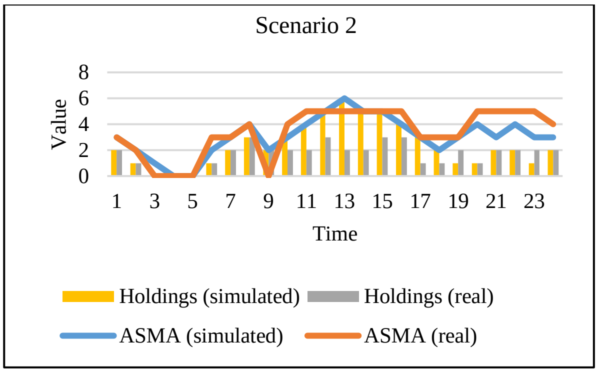


Figure 5: CPN Simulated ASMA Transit Time, Real ASMA Transit Time, CPN Simulated Holdings and Real Holdings Comparison for Scenario 2

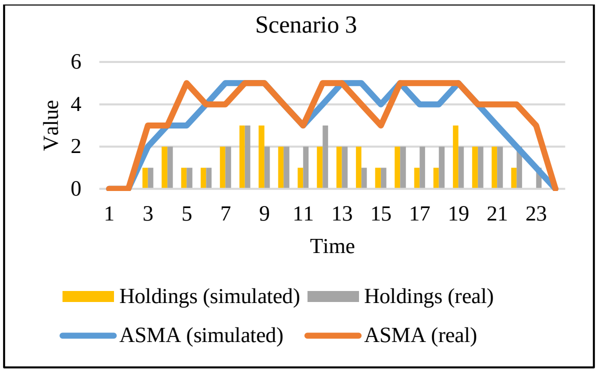


Figure 6: CPN Simulated ASMA Transit Time, Real ASMA Transit Time, CPN Simulated Holdings and Real Holdings Comparison for Scenario 3

All three simulation scenarios have demonstrated that additional ASMA time increases with the delay with the increase of number of flights with holding pattern, and also increases with the development of serious weather conditions (for instance, increasing wind intensity). This is illustrated on Figure 7. Although this correlation becomes not significant in the hours of low number of arriving aircraft (night time). The same behaviour was noted in Scenario 2 and 3 as well.

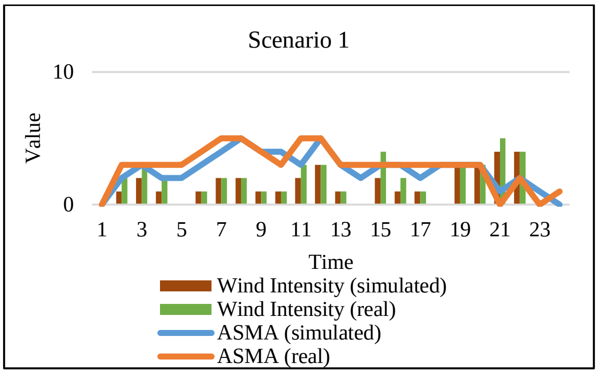


Figure 7: CPN Simulated ASMA Time and Real ASMA Time Comparison to CPN Simulated Wind Intensity and Real Wind Intensity for Scenario 1

Potentially, a set of variables, representing events, preceding the entering of the aircraft into the ASMA, can be added into the model in order to take into account influence of en-route regulations on number of flights with holding pattern.
Furthermore, it has been noticed that both number of flights with holdings and values of additional ASMA time do not increase infinitely. This phenomenon is considered to be probably related to the aerodrome capacity limit: an aerodrome can accept only finite number of aircraft per time interval (due to limited throughput of its runways). Nevertheless it is considered to perform more experiments in the future to better explore this phenomenon.
The explored through RT system dynamics raises the question of adding more metrics to the model, potentially representing en-route events for different flights and also other KPIs, not listed in Table 2, but available in the

ISBN 978-88-97999-85-0; Affenzeller, Bruzzone, Jiménez, Longo and Piera Eds.

databases of Madrid-Barajas airport. After adding the new metrics to the model, Bayesian inference and new series of simulation with CPN framework should be performed with various realistic initial markings.

5. CONCLUSIONS AND FURTHER RESEARCH

This paper describes an approach to explore relationships between ASMA transit time deviations, number of flights with holding pattern and weather indicators with the use of Bayesian Network. Mathematical expressions of the discovered relationship have been used in order to build a model, capable to show possible states of the system for different scenarios of ASMA transit time changes. These scenarios are considered to be used by airport decision makers in order to design other scenarios and be prepared for any deviation that could occur in the terminal maneuvering area and its surroundings in the future and be able to explore the possible causes of any deviations of ASMA transit times occurred in the past.
It is considered also that the model could be extended and more airport performance metrics could be added to it in order to perform more wide and complex analysis, considering bigger area of airport operational activities. The noise, representing stochasticity of weather conditions for aircraft on en-route phase, preceding arrival to the studied airport, could be also added and its influence could be observed during the further research. However the computational restrictions of the used software have to be taken into account, as if the model becomes more complex, it would take more time and computational resources in order to explore all possible state spaces and perform the analysis.

ACKNOWLEDGMENTS

This research is partially supported by the Ministry of Economy and Competitiveness. Project "Fire Guided Unmanned Aircrafts and Resources Distribution" (TIN2014-56919-C3-1-R) and the project AIRPORTS (CDTI). Opinions expressed in this paper reflect the authors' views only.

REFERENCES

Andersson K., Carr F., Feron E., and Hall W. D., 2000. Analysis and Modeling of Ground Operations at Hub Airports. 3rd USA/Europe Air Traffic Management R&D Seminar, June 13-16, Napoli.

Brockwell P. J. and Davis R. A., 1991. Time Series: Theory and Methods (2nd ed.). Springer.

Chickering D., Geiger D. and Heckerman D., 1995. Learning Bayesian networks: Search methods and experimental results. Fifth International Workshop on Artificial Intelligence and Statistics, pp. 112-128. January 4-7, Ft. Lauderdale (Florida, USA).

Chickering D., Meek C., and Heckerman D., 2004. Large-Sample Learning of Bayesian Networks. The Journal of Machine Learning Research, 5 (1999): 1287–1330.

European Organisation for the Safety of Air Navigation (EUROCONTROL), 2005. Performance Review Commission, Performance Review Report 2004 (PRR 8).

European Organisation for the Safety of Air Navigation (EUROCONTROL), 2014. SESAR Guidance on KPIs and Data Collection Version 1, D85, Ed. 00.01.01.

European Organisation for the Safety of Air Navigation (EUROCONTROL), 2015. How to proactively manage weather's impact on performance? Available from: https://www.eurocontrol.int/ [accessed 24 May 2017].

European Organisation for the Safety of Air Navigation (EUROCONTROL), 2017a. Managaing complexity in a fast-changing world. Available from: https://www.eurocontrol.int/ [accessed 24 May 2017].

European Organisation for the Safety of Air Navigation (EUROCONTROL), 2017b. Airport capacity and performance (ACAP) activities. Available from: https://www.eurocontrol.int/ [accessed 24 May 2017].

Federal Aviation Administration, 2016. Terminal Area Forecast Summary. Fiscal Year 2016 - 2045. Available from: https://www.faa.gov/ [accessed 10 June 2017].

Friedman N., Linial M. and Nachman I., 2000. Using Bayesian Networks to Analyze Expression Data. Journal of Computational Biology, 7: 601–620.

Gelman A., Carlin J.B., Stern H.S., Dunson D.B., Vehtari A. and Rubin D.B., 2014. Bayesian data analysis. FL, USA:Taylor & Francis Group, 3-8.

Idris H. R., Anagnostakis I., Delcaire B., Hansman R. J., Clarke J.-P., Feron E., and Odoni A. R., 1999. Observations of Departure Processes At Logan Airport to Support the Development of Departure Planning Tools. Air Traffic Control Quarterly, 7: 229–257.

International Civil Aviation Organization (ICAO), 2017. Forecasting. Available from: https://www.icao.int/ [accessed 17 June 2017).

Jensen K. and Kristensen. L.M., 2009. Coloured Petri Nets: Modelling and Validation of Concurrent Systems. Springer Berlin Heidelberg.

Joshi K.D., 1989. Foundations of discrete mathematics. New Age International Limited, 7.

Klein A., Kavousssi S. and Lee R.S., 2009. Weather Forecast Accuracy: Study of Impact on Airport Capacity and Estimation of Avoidable Costs. Eighth USA/Europe Air Traffic Management Research and Development Seminar. Available from: http://www.atmseminar.org/ [accessed 10 June 2017].

Marsland S., 2015. Machine Learning: An Algorithmic Perspective, Second edition. CRC Press.

Office of Technology Assessment, 1984. Airport System Development. OTA-STI-231, Washington, D. C.: U.S. Congress. Available from: https://www.princeton.edu [accessed 24 June 2017].

Piera M.A. and Musič G., 2010. Coloured Petri net scheduling models: Timed state space exploration

shortages. Mathematics and computers in simulation (82): 428-441.
Radhakrishnan N., Scutari M. and Lèbre S., 2013. Bayesian Networks in R with Applications in Systems Biology. Springer.
Roling P. C. and Visser, H. G., 2007. Optimal Airport Surface Traffic Planning Using Mixed-Integer Linear Programming. International Journal of Aerospace Engineering, 2008.
Silander T. and Myllymäki P., 2006. A Simple Approach for Finding the Globally Optimal Bayesian Network Structure. Networks, 445–452.
Song P. X.-K. 2007. Correlated Data Analysis: Modeling, Analytics, and Applications. Springer.
Tabernier L., 2015. SESAR Key Performance AREA, Performance targets and assessment. Available from: https://www.eurocontrol.int/ [accessed 15 June 2017].
Tsamardinos I., Brown L.E. and Aliferis C.F., 2006. The Max-Min Hill-Climbing Bayesian Network Structure Learning Algorithm. Machine Learning, 65 (1): 31–78.
Walker C., 2016. CODA Digest 2016. All-Causes and Cancellations to Air Transport in Europe – 2016. The European Organisation for the Safety of Air Navigation - EUROCONTROL. Available from: http://www.eurocontrol.int/ [accessed 24 May 2017].
Wang Y., 2012. Prediction of weather impacted airport capacity using RUC-2 forecast. Digital Avionics Systems Conference (DASC), 2012 IEEE/AIAA 31st. October 14-18. Williamsburg (Virginia, USA).
Xu N., Laskey K.B., Chen C., Williams S.C., Sherry L., 2007. Bayesian Network Analysis of Flight Delays. Available from: https://pdfs.semanticscholar.org/ [accessed 10 June 2017].

Appendix B: Identifying and modelling correlation between airport weather conditions and additional time in airport arrival sequencing and metering area

M. Bagamanova, J. J. Ramos González, M. À. Piera Eroles, J. M. Cordero García, and Á. Rodríguez-Sanz

Published in the *International Journal of Simulation and Process Modelling*, vol. 14, no. 3, p. 213, 2019 [4].

Identifying and modelling correlation between airport weather conditions and additional time in airport arrival sequencing and metering area

Margarita Bagamanova*, Juan José Ramos González and Miquel Àngel Piera Eroles

Universitat Autònoma de Barcelona,
Carrer de Emprius 2,
08202 Sabadell, Spain
Email: Margarita.Bagamanova@uab.cat
Email: JuanJose.Ramos@uab.cat
Email: MiquelAngel.Piera@uab.cat
*Corresponding author

Jose Manuel Cordero García

CRIDA A.I.E. (Reference Center for Research, Development and Innovation in ATM),
Edificio Allende, Avenida de Aragón 402,
28022 Madrid, Spain
Email: jmcordero@e-crida.enaire.es

Álvaro Rodríguez-Sanz

Universidad Politécnica de Madrid,
Plaza Cardenal Cisneros 3,
28040 Madrid, Spain
Email: alvaro.rodriguez.sanz@upm.es

Abstract: Different uncertainties during operational activities of modern airports can significantly delay some processes and cause chain-effect performance drop on the overall air traffic management (ATM) system. The decision-making process to mitigate the propagation of perturbations through the different airport processes can be improved with the support of a causal model, built with a use of data mining and machine learning techniques. This paper introduces a new approach for modelling causal relationships between various ATM performance indicators, which can be used to predict, by means of simulation techniques, the evolution of airport operations scenarios. The analysis of reachable airport states is a relevant approach to design mitigation mechanisms on those perturbations which drive the system to poor KPIs.

Keywords: ASMA time; holding; inbound traffic; weather impact; Bayesian networks; coloured Petri net; CPN; airport; decision support tool; TMA model.

Reference to this paper should be made as follows: Bagamanova, M., Ramos González, J.J., Piera Eroles, M.À., Cordero García, J.M. and Rodríguez-Sanz, Á. (2019) 'Identifying and modelling correlation between airport weather conditions and additional time in airport arrival sequencing and metering area', *Int. J. Simulation and Process Modelling*, Vol. 14, No. 3, pp.213–222.

Biographical notes: Margarita Bagamanova is a PhD candidate at the Universitat Autònoma de Barcelona (UAB). Her field of research is related to the optimisation of airport performance (causal models and forecasting) and practical application of machine learning techniques.

Juan José Ramos González is a researcher focusing on modelling, simulation and optimisation of dynamic systems, especially in the field of logistics. He received his PhD from the Universitat Autònoma de Barcelona (UAB) in 2003, where he is a Professor at the Department of Telecommunication and System Engineering. Currently, he is the Program Director of the European Master in Logistics and Supply Chain Management at UAB. He is a member of LogiSim, a recognised research group on Modelling and Simulation of Complex Systems and co-founder of two spin-off companies. He is co-author of four patents/IPR currently in industrial application. He is an expert in production technologies/logistics, intelligent transport systems, information and communication technology, industrial collaboration and technology transfer.

Miquel Àngel Piera Eroles holds a PhD in Computer Science. He is Professor at the Universitat Autònoma de Barcelona (UAB), Deputy Director of Engineering School, UAB Delegate for Technical Innovation Cluster on Aeronautical Management and Head of the Research Group on Modelling and Simulation of Logistic and Production Systems. He is an expert in modelling and simulation of production and logistic systems, object-oriented modelling and simulation, production planning and industrial informatics.

Jose Manuel Cordero García is an R&D Principal Engineer at CRIDA. His position mainly includes working in project management, both national and international, working with European partners and industry in project as SESAR (Single European Sky Advanced Research), and also includes being Head of IT Department.

Álvaro Rodriguez-Sanz holds a MEng in Aeronautical Engineering from Universidad Politécnica de Madrid (UPM) and an MSc in Airport Planning and Management from Cranfield University, UK. He has worked in airport development and air traffic operations projects for INECO and AENA, and in the strategic planning department of LATAM Airlines. Currently, he is Researcher and Assistant Lecturer for airport, air transport and air navigation subjects at UPM, where he is also a PhD candidate. His field of research is related to the optimisation of airport, air transport and air traffic operations (flow management, causal models and predictability analysis).

This paper is a revised and expanded version of a paper entitled 'Modelling dependence of arrival sequencing and metering area transit time on airport meteorological conditions' presented at the 29th European Modelling and Simulation Symposium, Barcelona, Spain, 18–20 September 2017.

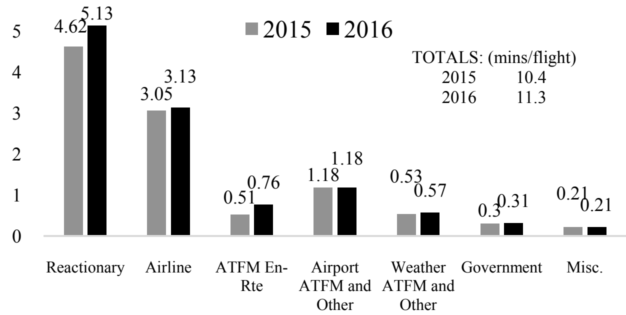
1 Introduction

Airports are very important elements of the modern air transportation network and any disrupted activity in an airport could easily create reactionary delays and overall performance drops in the whole network. A reactionary delay is a delay caused by late arrival of the aircraft or crew from previous flights (Performance Review Commission, 2004). In such manner any delay occurred at the departure airport could lead to significant delays in the following successive legs introducing perturbations on the scheduled processes at the destination airports.

Any delay means additional operational and environmental burden and cost for various agents, operating in an airport. This is why the understanding of the causes of occurred delays remains one of the main directions of analysis for air transportation stakeholders. For instance, European Organisation for the Safety of Air Navigation (EUROCONTROL) aggregates the data obtained from the European Civil Aviation Conference (ECAC) member airports and in the form of public open documents reveals main air transportation performance problems and official delay causes.

According to one of such reports, 2016 was a year with increased volume of flight delay. The average delay per delayed flight on arrival was approximately 29 minutes per flight in 2016 and increased by 1 minute compared to the previous year. The percentage of delayed arrivals in comparison to 2015 has also increased, by one percentage point to 38%. Furthermore, the contribution of reactionary delay has increased up to 45% of total delay minutes and delay caused by airport operations (Walker, 2016). And this kind of delay is not the only reason of on-time performance decrease in 2016, as it could be seen in Figure 1.

Figure 1 Breakdown of average delay in minutes per flight per delay in ECAC 2015 vs. 2016



Source: Walker (2016)

According to Walker (2016), flight delays occurring due to the weather conditions constitute approximately 5% of the average delay reasons structure. In contrast, weather conditions are recognised to be the largest cause of delays in the National Airspace System of USA, as stated by the Federal Aviation Administration (2017). In such a way, it is possible to consider that the weather impact has a strong regional character, varying for different climate characteristics. Nevertheless, the global air transportation network can still propagate the delays from one region to another, creating perturbations in local operations.

Delay breakdown provided in different documents usually only reflects official causes, reported by airlines, leaving aside the possible fuzzy influence of the weather conditions on flight punctuality. This motivates to investigate if indirect airport weather impact could produce noticeable alterations in airport performance.

This paper presents a new approach for modelling causal relationships between performance indicators of a virtual cylinder around the airport, termed as arrival sequencing and metering area (hereinafter referred to as

ASMA), and weather conditions, which can be used to predict the evolution of airport performance scenarios by simulation techniques. Section 2 describes similar research in ATM field. Scope and focus of the study are described in Sections 3 and 4. Methodology used for the causal analysis is presented in section 5. Section 6 describes the model, representing airport system behaviour and changing weather conditions. Verification of the developed causal model and simulation results are described in Section 7. At last, conclusions, derived from the simulations, and thoughts on the future research are given in Section 8.

2 State of art

Many researchers have tried different approaches to understanding and forecasting perturbations, affecting various airport activities. For instance, a complex regression model, combining temporal and spatial variables and random forest algorithms (Rebollo and Balakrishnan, 2014), has been quite successfully used to predict departure delays. Quadratic response surface (QRS) linear regression models and ensemble bagging decision tree regression (BDT) models have been used to assess weather impact on the maximum number of movements per time interval in few US airports (Wang, 2012). Influence of fog on landing operations was analysed statistically and classified into a few categories (Tuncay Özdemir et al., 2016). Sasse and Hauf (2003) have also used statistical analysis of historical data to study the thunderstorm effect on flight delays and concluded that the total delay depends on the intensity and duration of a thunderstorm and current airport capacity. Coffel and Horton (2015) have highlighted sensibility of the maximum allowable takeoff weight of an aircraft to extreme temperatures.

Lancia and Lulli (2017) have proposed a non-parametric inbound traffic modelling technique, based on modelling of arrivals distribution and various stochastic effects. Wang (2014) used machine learning methods in order to investigate the causal effect of weather conditions on ground stop operations in USA and achieved an accuracy of 85% of the developed model predictability.

Another weather impacted airport delay prediction model (Klein et al., 2010) showed a strong influence of severe weather conditions on flight delay development and cancellation policy. This model has used data on intended traffic demand together with weather components and their weights as main elements, not considering each of the weather phenomena, occurred at the moment of time, but only the most influential ones (according to the opinion of the authors). On the contrary, an approach presented in this paper aims to study weather conditions together, not giving higher importance to any particular phenomena and thus considering both individual and composite effects of different weather indicators.

According to the conclusions of the previously mentioned works, the developed models can help to predict delays in different weather conditions, however the models are quite sensitive to inaccurate weather forecast and their quite promising precision level is still leaving a space for improvement.

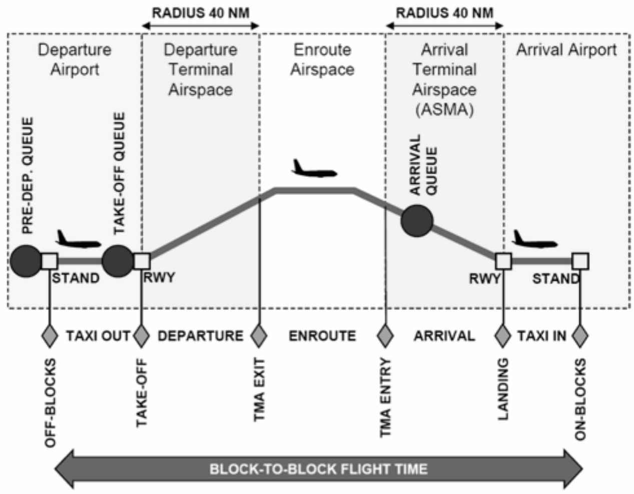
Thus, it is important to further research on new tools for an efficient causal analysis of operational deviations and its prediction, considering the operational conditions that affect each particular airport. This can help to design a decision support mechanism for mitigation of negative delay effects not only at each airport, but also at the air transportation network level.

This paper, in comparison to similar researches performed on weather conditions and flight delay correlation, presents a study of one of the composing elements of inbound traffic delays – additional ASMA time and contributes to the community with a methodology that combines existing and emerging technologies of data processing and allows to not only analyse but also predict delays in the approximations of an aerodrome.

3 Scope of the study

One of the principal airport structural components is a set of runways. The way they are used defines the airport capacity and throughput, the number of flights that can land and depart from an airport. When the current environmental or operational conditions do not allow to perform landing straight forward, an aircraft is often directed to enter the holding pattern and spend additional time in the ASMA around the airport. This airspace area is typically defined as a virtual cylinder. The radius value of the observed cylinder can be of any required for management purpose size, although is typically set to 40 nautical miles or 60 nautical miles (Capelleras, 2015).

Figure 2 ASMA place in the flight phases



Source: (Capelleras, 2015)

As the number of aircraft, remaining in the queue in ASMA, grows, fuel consumption and cost expenses push the priority to the arriving aircraft, creating disturbances for departure activities. In order to estimate the operational penalty at approach time, resulting from different control activities, time spent by an aircraft from entering the ASMA till landing is measured and referred to as ASMA transit time (Capelleras, 2015).

This transit time is then compared to *unimpeded time* – statistically calculated reference time, based on historical ASMA times in periods of low airport traffic level for a particular type of aircraft. The difference between the actual ASMA time of an aircraft and its unimpeded ASMA time is referred to as additional ASMA time (Capelleras, 2015).

This paper describes an approach to explore and model the possible correlation of one of the main airport performance indicators – additional ASMA time, on the weather conditions of arrival airport.

4 Dataset used for study

As it has been noted by different researchers, weather impact could cause significant deviation on airport operational performance, but it has not yet been studied well enough (Klein and Sadegh Kavoussi, 2009; European Organisation for the Safety of Air Navigation, 2015; Klein et al., 2010). Therefore, it has been chosen to explore the weather impact on one of the key performance indicators of the studied airport.

It has been chosen to analyse together additional ASMA time values, calculated for 60 minutes intervals in 2015 and the data from the METAR report, consisting of recorded on the territory of the studied airport meteorological conditions for the same time period.

In the scope of this paper, the following available data has been considered of the current study interest:

1 percentage of additional ASMA time of unimpeded ASMA time (adASMA), for ASMA of 60 NM radius

2 number of flights with holding patterns – number of flights, which have to take a special route around the aerodrome in order to wait for an appropriate moment for landing (H)

3 wind direction (Wind) and wind intensity (WI)

4 predominant visibility on the aerodrome territory (Vis)

5 difference between current temperature and dew point value (DP)

6 atmospheric pressure (Pres)

7 weather phenomena type – if fog or any other weather phenomena occur (Fen)

8 cloudiness (Cloud).

5 Methodology

Among the different analytical tools to discover relationships between observed variables, construction of Bayesian networks seems to provide a promising approach to better understanding of complex systems such as airports. In the past few decades, research on Bayesian networks has advanced significantly and many different algorithms have appeared. These new algorithms are meant to deal with new types of data, a larger number of variables and larger number of observations (Marsland, 2015; Song, 2007; Xu et al., 2007). Furthermore, with an increasing popularity of machine learning and data mining methods, use of Bayesian networks seems to be an essential part of modern data analysis, that can be easily incorporated into the organisational decision support system thanks to its capability to cope with high-dimensional problems of different data and many powerful computer programs, that made any related computations easier and rather fast.

5.1 Core concepts of Bayesian networks

A Bayesian network is a graphical representation of relationships between different variables in the form of directed acyclic graph. Analysed variables are represented as nodes and their probabilistic dependencies on each other are represented as directed arcs connecting the nodes. These conditional dependencies are often estimated by using known statistical and computational methods. Hence, Bayesian networks combine principles from graph theory, probability theory, computer science, and statistics.

The absence of direct arc between some two nodes means that these two nodes (variables) are conditionally independent of each other (Marsland, 2015). When a node has an outgoing arc, it is called a *parent*, the nodes with incoming arcs are called *children*. The joint probability distribution P_X of the chosen variables X is represented as a product of conditional probability distributions of each variable X_i (Nagarajan et al., 2013):

$$P_X(X) = \prod_{i=1}^{p} P_{X_i}\left(X_i \middle| \Pi_{X_i}\right) \quad (1)$$

The process of discovering Bayesian network, fitting the data, consists of two phases: structure and parameter learning. Many algorithms, depending on the data type, have been developed for learning the structure of a network, representing the data relationships. There are two main methods for learning Bayesian network. The first one is the search-and-score approach. It consists in searching in the space of all possible networks and identifying the one that maximises the scoring function, which indicates how well the model is fitting the analysed data. The second approach is constraint-based. It is estimated whether certain conditional independence between variables exist and returns only statistically equivalent networks consistent with statistical tests.

Nevertheless, for this paper, it has been chosen to use a hybrid algorithm – max-min hill-climbing (MMHC) algorithm. This algorithm combines strong sides of constraint-based and score-based algorithms, effectively learning both structure and parameters of a network with rather low computational expenses (Nagarajan et al., 2013).

5.2 *Max-min hill-climbing algorithm*

The MMHC algorithm was developed to fast and effective learning of Bayesian network for datasets with a large number of variables and observations. This algorithm first identifies the parents and children of each variable (PC set for each variable *X*) with a use of max-min parents and children (MMPC), then performs a greedy hill climbing search in the space of current network, changing the edges for obtaining the highest network score. MMPC is based on conditional independence tests and estimates the strength of association between variables, thus identifying if particular nodes are connected or not. Then the obtained skeleton is evaluated with greedy hill-climbing algorithm with a tabu list to identify a structure with the highest network score, reflecting the better fit of the network to the data. More detailed description can be found in Tsamardinos et al. (2006).

The MMHC pseudo-code looks as follows:

```
1  Input: data D
   Output: a network on the variables in D
   Restrict:
2  For every variable X ∈ V do
3      PC_X = MMPC(X, D)
4  End for
   Search:
5  Starting from an empty graph, perform greedy
   hill-climbing with operators add-edge, delete-edge,
   reverse-edge. Only try operator add-edge Y → X
   if Y ∈ PC_X.
6  Return the highest scoring network found
7  End
```

This algorithm has many similarities with the sparse candidate (SC) algorithm, which was one of the first successfully performing approaches, applied to large datasets with several hundred variables (Friedman et al., 2000). Both SC and MMHC perform a stepwise reduction of candidate parents set for each variable and then search for a network that maximise a chosen scoring function. However, they have one important difference. The SC algorithm performs the reduction and network search steps iteratively until there is no improvement in the scoring function value. MMHC first identifies the skeleton of a Bayesian network by performing conditional independence tests and then orients the arcs by performing a Bayesian-scoring greedy hill-climbing (Tsamardinos et al., 2006), estimating the candidate parents set only once (Nagarajan et al., 2013) and thus fastening the computational process by several times without significant decrease of correctness (Tsamardinos et al., 2006).

5.3 *Data preparation and structure learning*

The studied airport original data components do not have the same type. Some variables are observed as *continuous*, others – as *discrete*:

- *continuous* variable – variable, that can take on any real value within a certain interval (Joshi, 1989); for instance, additional ASMA time is expressed in minutes
- *discrete* variable – can take on only certain values (Joshi, 1989); for instance, cloudiness.

Presence of such mixed data can potentially cause a problem in the step of defining a network, fitting the data (Nagarajan et al., 2013). Therefore it has been decided to perform a commonly used solution to avoid the mentioned issue – perform *discretisation* or *binning* of the data. *Discretisation* means assigning some particular integer value to the certain intervals of continuous data. There are different ways to define the intervals for data discretisation: using expert knowledge on data, using heuristics, performing discretisation and structure learning iteratively, etc. (Nagarajan et al., 2013). The data has been discretised as shown in Table 1.

Table 1 Intervals of data discretisation

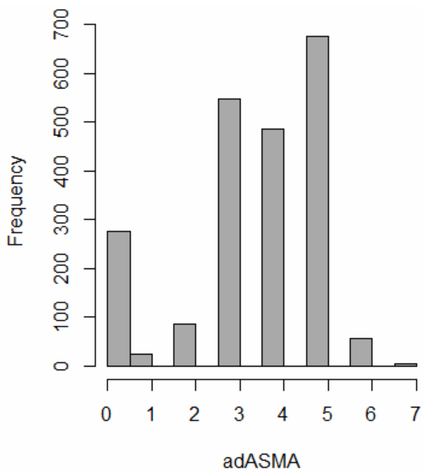
Category	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*0*
adASMA, %	(∞; –15)	[–15; –10)	[–10; –5)	[–5; 5)	[5; 10)	[10; 30)	≥30	0
H	1	2	3	4	5	6	7	0
Vis, m	<50	[50; 400)	[400; 8,000)	≥8,000	-	-	-	-
Wind, °	(22; 67]	(67; 112]	(112; 157]	(157; 202]	(202; 247]	(247; 336]	(337; 22]	VRB
WI, knots	[1; 5]	[6; 15]	[17; 25]	[25; 40]	[41; 47]	[48; 55]	[56; ∞)	0
DP, ºC	0	-	-	-	-	-	-	<>0
Cloud	FEW	SCT	BKN	OVC	-	-	-	no
Fen, type	BR	DZ	FG	RA	SN	-	-	no
Pres, QNH	<1,013	1,013	>1,013	-	-	-	-	-

In Table 1, the following abbreviations have been used:

- VRB – variable wind direction
- FEW – from 1/8 to 2/8 of the sky is covered by clouds
- SCT – scattered: from 3/8 to 4/8 of the sky is covered by clouds
- BKN – broken clouds: from 5/8 to 7/8 of the sky is covered by clouds
- OVC – overcast: the sky is fully covered by clouds
- BR – mist
- DZ – drizzle
- FG – fog
- RA – rain
- SN – snow.

Target variable of the study is adASMA and its distribution looks as shown on Figure 3. As it can be seen, the values of adASMA do not strictly follow the normal distribution.

Figure 3 Histogram of adASMA values



After the data preparation MMHC algorithm has been executed in the framework of R software and the structure, presented on Figure 3, has been obtained in less than one minute (for Intel (R) i5-4300M CPU 2.60 GHz, 8 GB RAM). Every arc of the obtained network has a probability of being true of not less than 95%.

This structure has been cross-validated with the dataset, used for its learning, which allowed us to consider that this network fits the data well enough. As it can be seen in Figure 4, there is a correlation between the value of additional ASMA time and wind intensity. In order to see the nature of this correlation, the plot, showing marginal effects of wind intensity on additional ASMA, was built. Figure 5 demonstrates that when the wind intensity increases on 1 knot, additional ASMA time increases on approximately 5 minutes.

Figure 4 Bayesian network, obtained with MMHC algorithm

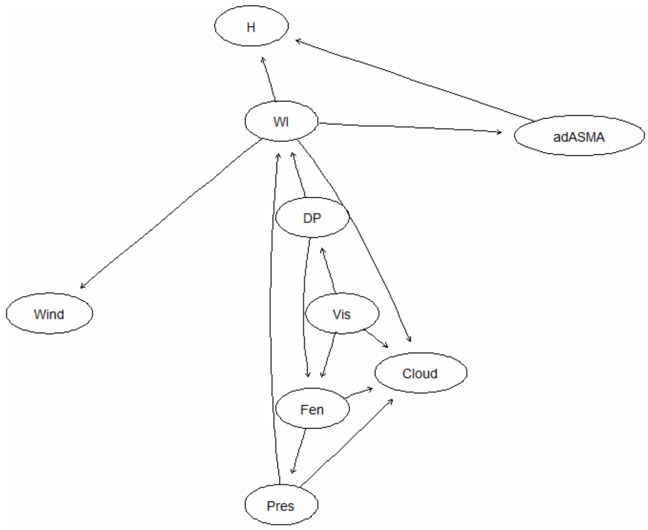
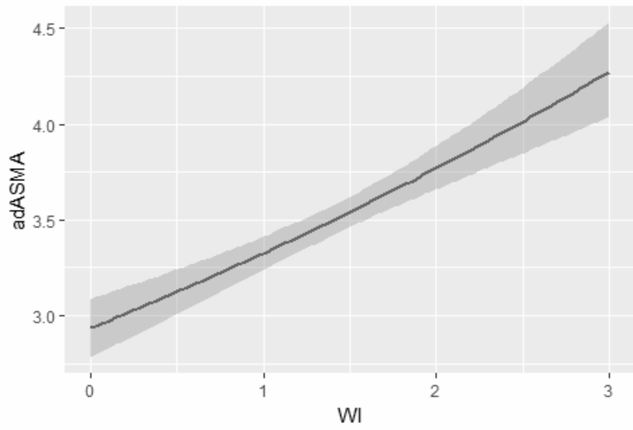


Figure 5 Marginal effects of WI on adASMA



5.4 *Bayesian inference*

The knowledge about the data structure and its parameters, obtained from Bayesian network, could be used for reasoning about the future probabilities of the studied indicators. Through the conditional probability distributions, calculated for every variable of the studied data, it is possible to conclude about posterior or future data values. This conclusion is formulated as likelihood function and can serve for development of prediction models (Gelman et al., 2014).

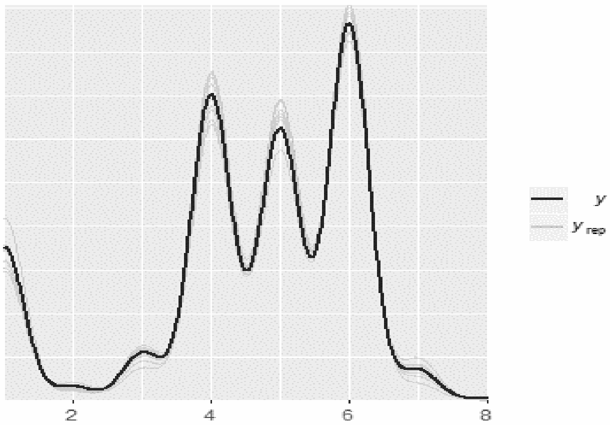
There are two main approaches for updating the posterior probabilities of data distribution: exact and approximate inference. Variable elimination and Junction Tree are the two best-known approaches for exact inference task. First approach uses the network structure directly, taking into account the local distributions of the data variables. On the contrary, the second algorithm transforms the network by clustering its nodes into a tree. However the feasibility of exact approach is restricted to small networks. Approximate inference algorithms create samples from the local distributions with the use of Monte Carlo simulations and then evaluate them. The sampling can be performed in different ways, implemented in several approximate algorithms (Nagarajan et al., 2013).

The parameters learnt in this step take the form of regression coefficients, belonging to regression functions, describing the conditional dependence between studied variables. In this paper it is considered to use the logic sampling approximate inference algorithm, already included in functionality of many R software packages for Bayesian Network learning. The inferenced network parameters have been used for the mathematical expression of relationships between observed variables in the arc expressions in the CPN model in order to perform simulation runs and state space analysis.

5.5 *Goodness of fit*

It is possible to estimate how well the learnt Bayesian network fits the data with graphical posterior predictive checking, which is a graphical display of comparison of the two distributions: of the observed data and of the predicted data. As presented in Figure 6, the fitness of Bayesian network and its parameters is quite high.

Figure 6 Comparison ofthe distributions of predicted values (y_{rep}) and observed values (y) of adASMA



Nevertheless, in order to check if it was possible to estimate and capture the changes of adASMA triggered by various events, in one way or another reflected in the data, it was decided to perform simulations in the framework of coloured Petri net (CPN).

6 Modelling with CPN formalism

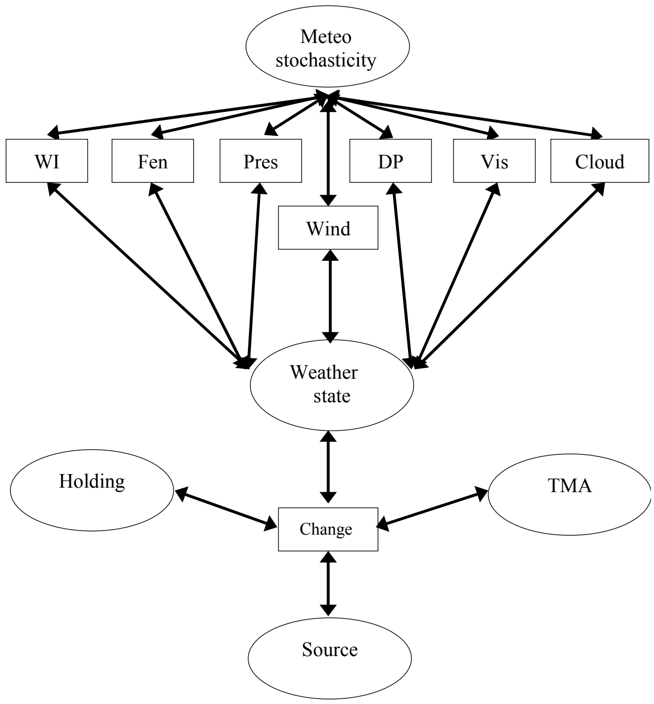
CPN is a modelling formalism, aimed to design, visualise, test and explore the behaviour of a system (Latorre-Biel et al., 2017). In order to model the system behaviour with CPN formalism, it is necessary to define a set of parameters:

- set of colours – to represent the model entity attributes (ASMA time and weather indicators)
- set of places nodes – to represent airport performance areas, such as ASMA and TMA, and airport weather state
- set of transition nodes – to represent systems' activities (weather changes, arriving aircraft, etc.)
- set of arcs – to relate transition and places nodes
- guard functions, which are associated with the transition nodes in order to insure their enabling only in case of satisfaction of conditions, described in the corresponding guard function (Jensen and Kristensen, 2009).

For the system elements inscriptions CPN ML, a functional programming language, is included in the modelling framework of CPN Tools. It provides a syntax to declare necessary variables and perform modelling of their changes (Piera and Mušič, 2011). This language is used in formulas of arc functions and in declarations of intervals of possible values for model parameters.

For modelling the chosen weather indicators and additional ASMA time of the airport, the colours, representing weather indicators, additional ASMA time, and number of flights with holding pattern, have been introduced into the model. Furthermore, it has been considered to introduce the colour, representing system time counter, for having a tool to track system dynamics in time without increasing model complexity. A schematic design of the developed CPN model is shown on Figure 7.

Figure 7 Weather indicators and ASMA CPN model design



The studied KPIs are distributed among the three places as follows:

- place holding – number of flights with holding pattern (H)
- place TMA – adASMA and time counter

- place weather state – Wind, WI, Vis, DP, Pres, Fen and Cloud

Furthermore, two supporting places, ensuring the element of random noise, have been also added to the model. They are:

- place meteo stochasticity – provides tokens for stochastic weather changes
- place source – provides tokens for stochastic changes in number of arriving flights with holding pattern.

We have introduced stochastic changes in order to see all possible evolutions of the system and see if the model is able to find any states similar to the ones, observed in the historical data.

In order to formulate in CPN tools the events, observed in ASMA area, it is required to define functions for the expressions of arc, connecting elements of the model. The arc functions are described in equation (2), based on the maximum likelihood estimation parameters, obtained in the step of Bayesian network learning:

$$C_i = \beta + k * C_j \tag{1}$$

where

C_i represents CPN colour i, a studied metric.

β represents intercept value.

k represents the regression coefficient.

C_j represent the CPN colour j, on which CPN colour i is conditionally dependent. When there are more metrics, on which colour i is conditionally dependent, they are included with the corresponding regression coefficients.

After introducing all necessary system parameters, series of simulation runs can be executed in order to verify and validate the model.

7 Model verification and simulation results

Reachability tree can be used for model verification relying on state space analysis. This analysis consists of generating all states and state changes of a model, which could be reached from the initial state (Jensen and Kristensen, 2009). CPN Tools v. 4.0.1 allows to graphically represent all possible system states through the *reachability tree* (RT) – a directed graph, where root node represents an initial marking of the system, and the successive nodes represent the new states, that can be reached from the initial state, if the corresponding transitions have been fired.

The initial model markings, used for various state space analysis scenarios, are shown in Table 2. These values have been chosen from the available historical data for the same time period as for Bayesian network learning, in order to compare the system dynamics, observed in the historical data and the changes, discovered through RT construction.

Table 2 Initial markings for simulation scenarios

Model parameters	*Scenario 1*	*Scenario 2*	*Scenario 3*
	Parameter value		
Additional ASMA time	0	3	0
Flights with holding	0	2	0
Wind direction	7	0	7
Wind intensity	0	2	9
Visibility	4	2	3
Dew point	0	1	1
Pressure	3	1	1
Phenomena	0	4	4
Cloudiness	0	2	1

In the reachability trees, generated for all three chosen scenarios, we have found branches with the same weather indicators evolutions as the ones, observed in the historical data. This has allowed to compare how additional ASMA time has developed in these RT branches and in the historical data. Figures 8, 9 and 10 represent this comparison for each of three simulation scenarios respectively for the time period of 24 hours.

Figure 8 CPN simulated adASMA, real adASMA, CPN simulated holdings and real holdings comparison for scenario 1

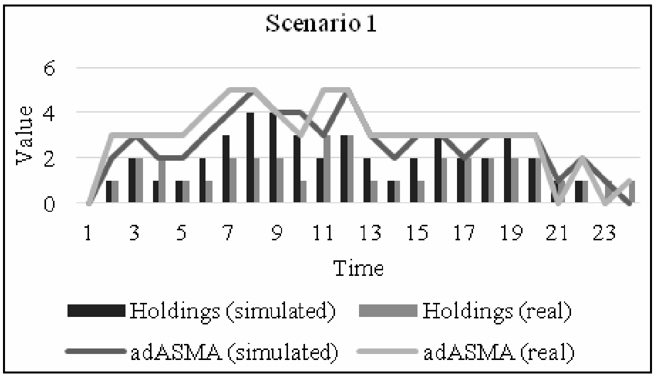


Figure 9 CPN simulated adASMA, real adASMA, CPN simulated holdings and real holdings comparison for scenario 2

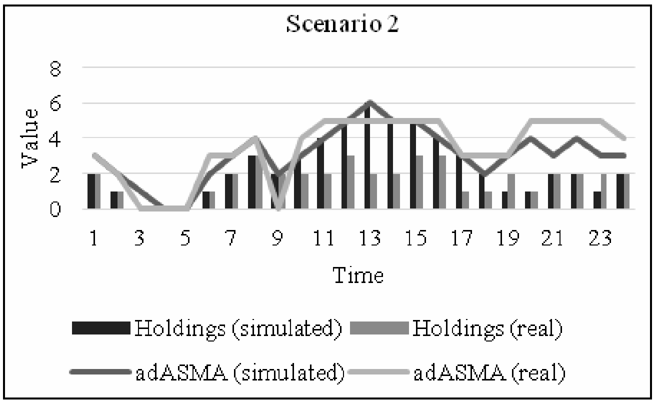
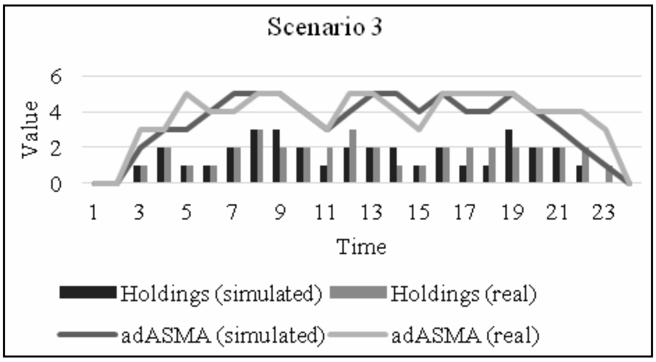
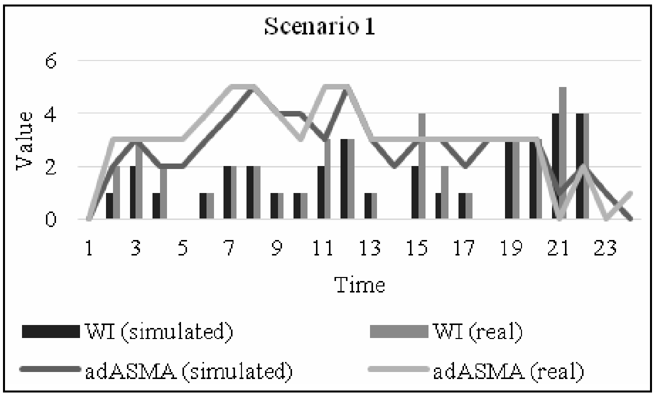


Figure 10 CPN simulated adASMA, real adASMA, CPN simulated holdings and real holdings comparison for scenario 3



All three simulation scenarios have demonstrated that number of flights with holding pattern increases at the same time as additional ASMA time increases. This correlation was expected to be seen, as it has been already discovered in the step of Bayesian network learning. Furthermore, it has been observed in the simulation runs and in the state space analysis that additional ASMA time also increases with the development of weather conditions - increasing wind intensity. This is illustrated in Figure 11.

Figure 11 CPN simulated adASMA and real adASMA comparison to CPN simulated WI and real WI for scenario 1



It has been also noted during simulations that impact of wind intensity on additional ASMA time has a certain periodicity and it is neglected in the time intervals of low inbound traffic level – in the early morning and late evening hours (time stamps 0 to 6 and 23 to 24). This correlation was not predicted intentionally and was discovered only by the means of Bayesian network itself, which allows to conclude that Bayesian network learning can still catch certain periodical effects, hidden in the data without significant interventions from the analyst side. The same behaviour of additional ASMA time and wind intensity correlation was noted in scenarios 2 and 3 as well.

It has been noticed that both number of flights with holdings and values of additional ASMA time do not increase infinitely. This phenomenon is considered to be probably related to the aerodrome capacity limit: an aerodrome can accept only finite number of aircraft per time interval (due to the limited throughput of its runways). Although it can also be related to the ATC (air traffic controllers) regulations. So it can also be useful to add to the model parameters, representing ATC intervention made on certain flights. It is considered to perform more experiments in the future with more parameters added to the model to better explore additional ASMA time behaviour. After adding the new metrics to the model, Bayesian inference and a new series of simulation with CPN framework should be performed with various realistic initial markings.

8 Concluding remarks and further research

This paper describes an approach to explore and model correlation between ASMA time, number of flights with holding pattern and weather conditions with the use of Bayesian network. Statistical expressions of the discovered relationships have been used in order to build a model, capable to show possible states of the system for different scenarios of ASMA time changes. These scenarios are considered to be possibly used by airport decision makers in order to test other scenarios and be able to reason on any deviation that could occur in the terminal maneuvering area and its surroundings in the future and be able to explore the possible causes of any deviations of ASMA transit times occurred in the past and its contributions to arrivals delays.

As the next step it is considered that it would be profitable for the model to add a set of variables, representing events, preceding the entering of the aircraft into the ASMA, in order to take into account influence of en-route regulations on inbound flights.

It is considered also that the model could be extended further and more airport performance metrics could be added to it in order to perform more wide and complex analysis, considering a wider area of airport operational activities. Noise, representing stochasticity of weather conditions for aircraft on en-route phase, preceding arrival to the studied airport, could be also added and its influence could be observed during the further research. However the computational restrictions of the used software have to be taken into account, as if the model becomes more complex, it would take more time and computational resources to explore all possible state spaces and perform the analysis.

Acknowledgements

This research is partially supported by the Ministry of Economy and Competitiveness, Project 'Fire Guided Unmanned Aircrafts and Resources Distribution' (TIN2014-56919-C3-1-R) and the project AIRPORTS (CDTI). Opinions expressed in this paper reflect the authors' views only.

References

Capelleras, L. (2015) 'Additional ASMA time performance indicator document', *Eurocontrol* [online] http://ansperformance.eu/references/library/pru-adasma-pi.pdf (accessed 25 December 2017).

Coffel, E. and Horton, R. (2015) 'Climate change and the impact of extreme temperatures on aviation', *Weather, Climate, and Society*, Vol. 7, No. 1, pp.94–102, doi: 10.1175/WCAS-D-14-00026.1.

European Organisation for the Safety of Air Navigation (2015) *How to Proactively Manage Weather's Impact on Performance?*, Eurocontrol [online] http://www.eurocontrol.int/news/proactively-manage-weather-impact (accessed 25 December 2017).

Federal Aviation Administration (2017) *FAQ: Weather Delay* [online] https://www.faa.gov/nextgen/programs/weather/faq/ (accessed 4 January 2018).

Friedman, N.I.R., Linial, M. and Nachman, I. (2000) 'Using Bayesian networks to analyze expression data', *Journal of Computational Biology*, Vol. 7, pp.601–620, doi: 10.1089/106652700750050961.

Gabry, J. (2018) *Graphical Posterior Predictive Checks Using the Bayesplot Package* [online] https://cran.r-project.org/web/packages/bayesplot/vignettes/graphical-ppcs.html#graphical-posterior-predictive-checks-ppcs (accessed 5 May 2018).

Gelman, A. et al. (2014) *Bayesian Data Analysis*, 3rd ed., CRC Press, FL, USA [online] https://www.math.muni.cz/~kolacek/docs/bayesian_data_analysis.pdf (accessed 25 December 2017).

Jensen, K. and Kristensen, L.M. (2009) *Coloured Petri Nets: Modelling and Validation of Concurrent Systems*, Springer, Heidelberg.

Joshi, K.D. (1989) *Foundations of Discrete Mathematics*, Wiley.

Klein, A. and Sadegh Kavoussi, R.S.L. (2009) 'Weather forecast accuracy: study of impact on airport capacity and estimation of avoidable costs', *Eighth USA/Europe Air Traffic Management Research and Development Seminar*.

Klein, A., Craun, C. and Lee, R.S. (2010) 'Airport delay prediction using weather-impacted traffic index (WITI) model', *AIAA/IEEE Digital Avionics Systems Conference – Proceedings*, pp.1–13, doi: 10.1109/DASC.2010.5655493.

Lancia, C. and Lulli, G. (2017) *Data-driven Modelling and Validation of Aircraft Inbound-Stream at Some Major European Airports* [online] http://mendeley.csuc.cat/fitxers/89e8d89afc80e9e2fef92cca34a605ff (accessed 11 December 2017).

Latorre-Biel, J.I., de la Parte, M.P. and Jiménez-Macías, E. (2017) 'Enriching the formalism of coloured Petri nets for modelling alternative structural configurations of a discrete event system: disjunctive CPN', *International Journal of Simulation and Process Modelling*, Vol. 12, No. 1, p.92, doi: 10.1504/IJSPM.2017.082805.

Marsland, S. (2015) *Machine Learning: An Algorithmic Perspective*, 2nd ed., Taylor Francis Inc, USA, doi: 10.1145/242224.242229.

Nagarajan, R., Scutari, M. and Lèbre, S. (2013) *Bayesian Networks in R*, Springer, New York, doi: 10.1007/978-1-4614-6446-4.

Performance Review Commission (2004) *An Assessment of Air Traffic Management in Europe during the Calendar Year 2004 Performance Review Commission* [online] https://www.eurocontrol.int/sites/default/files/publication/files/prr-2004.pdf (accessed 14 December 2017).

Piera, M.A. and Mušič, G. (2011) 'Coloured Petri net scheduling models: timed state space exploration shortages', *Mathematics and Computers in Simulation*, pp.428–441, doi: 10.1016/j.matcom.2010.10.014.

Rebollo, J.J. and Balakrishnan, H. (2014) 'Characterization and prediction of air traffic delays', *Transportation Research Part C: Emerging Technologies*, Vol. 44, pp.231–241, doi: 10.1016/j.trc.2014.04.007, Elsevier Ltd.

Sasse, M. and Hauf, T. (2003) 'A study of thunderstorm-induced delays at Frankfurt Airport, Germany', *Meteorological Applications*, Vol. 10, No. 1, pp.21–30, doi: 10.1017/S1350482703005036, John Wiley & Sons, Ltd.

Song, P.X-K. (2007) *Correlated Data Analysis: Modeling, Analytics, and Applications*, Springer Series in Statistics, Springer New York, New York, NY, doi: 10.1007/978-0-387-71393-9.

Tsamardinos, I., Brown, L.E. and Aliferis, C.F. (2006) 'The max-min hill-climbing Bayesian network structure learning algorithm', *Machine Learning*, Vol. 65, No. 1, pp.31–78, doi: 10.1007/s10994-006-6889-7.

Tuncay Özdemir, E. et al. (2016) 'Fog analysis at Istanbul Ataturk International Airport', *Weather*, Vol. 71, No. 11, pp.279–284, doi: 10.1002/wea.2747, John Wiley & Sons, Ltd.

Walker, C. (2016) *Coda Digest 2015*, pp.1–34 [online] http://www.eurocontrol.int/publications/coda-digest-2015 (accessed 25 December 2017).

Wang, Y. (2012) *Prediction of Weather Impacted Airport Capacity Using RUC-2 Forecast* [online] https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20150000119.pdf (accessed 20 December 2017).

Wang, Y. (2014) 'Analysis and prediction of weather impacted ground stop operations', *AIAA/IEEE Digital Avionics Systems Conference – Proceedings*, p.7A21–7A214, IEEE, doi: 10.1109/DASC.2014.6979510.

Xu, N. et al. (2007) 'Bayesian network analysis of flight delays', *TRB 2007 Annual Meeting*, Transportation Research Board, Washington, DC.

Appendix C: No more surprises: stand assignment algorithm with likelihood of turnaround time deviation

M. Bagamanova and M. Mujica Mota

Presented as a poster at the *SESAR Innovation Days* 2018.

No More Surprises: Stand Assignment Algorithm with Likelihood of Turnaround Time Deviation

Generating delay-aware stand assignment, optimized for the desired management perspectives

Margarita Bagamanova
Telecommunications and Systems Engineering
Universidad Autonoma de Barcelona
Barcelona, Spain
margarita.bagamanova@uab.cat

Miguel Mujica Mota
Aviation Academy
Amsterdam University of Applied Sciences
Amsterdam, the Netherlands
m.mujica.mota@hva.nl

Abstract—**Airport management is frequently faced with a problem of assigning flights to available stands and parking positions in the most economical way that would comply with airline policies and suffer minimum changes due to any operational disruptions. This work presents a novel approach to the most common airport problem – efficient stand assignment. The described algorithm combines benefits of data-mining and metaheuristic approaches and generates qualitative solutions, aware of delay trends and airport performance perturbations. The presented work provides promising solutions from the starting moments of computation, in addition, it delivers to the airport stakeholders delay-aware stand assignment, and facilitates the estimation of risk and consequences of any operational disruptions on the slot adherence**

Keywords- risk; airports; airspace;congestion;stand assignment; turn-round time; decision support

I. Introduction

In terms of rapid growth of air transport traffic and propagation of reactionary flight delays, it is essential to perform efficient management of airport facilities, maintaining costs as low as possible and keeping airport's KPIs on the required level. One of the most important problems that airport and airline managers have to be concerned about is efficiency of stand scheduling. Boost of air traffic and congestion of airport capacity have significantly increased the service complexity, which is further complicated by changes in the flight schedule on the day of operation. Poor terminal performance caused by inefficient stand scheduling can lead to decreasing of passenger service quality and increasing of turn-round time that can create a propagation of a delay to the successive flights and connected airports. Thus, it is necessary to make an optimal and effective use of terminal facilities, such as stands, to increase airport performance and to mitigate the propagation of negative effects through the air transportation network.

II. Problem Statement

A. Stand Assignment

The problem of stand allocation or stand assignment (further referred as SAP), as well as the similar problem of gate allocation, has been widely studied over the decades and numerous approaches have been applied to different sets of objectives, constraints and outcomes. SAP is a scheduling problem, which is NP-hard due to real-life quantity of constraints and decision variables. According to the methodology used, the solving approaches can be divided into three categories: exact algorithms, heuristic algorithms and combined algorithms. While the first ones aim to find the best solution from diverse perspectives, the rest are designed to determine a qualitative near-optimal solution in a reasonable computational time [1]. Due to the complex nature of the problem, exact algorithms (e.g. branch-and-bound algorithm) have difficulty in providing optimal solutions within reasonable computational times for large-scale stand assignment problems. Therefore, recent studies mainly focus on developing heuristic algorithms, which do not guarantee optimal solutions, but may provide near-optimal solutions in reasonable computational times. However, if a heuristic algorithm fails to find the solution, it is not possible to determine whenever it is due to the absence of any solution or due to the inability of an algorithm to move from local search region [2]. On the contrary, this work shifts the scope from the generation of better solutions to the assessment of the generated solutions not only from the objective function's value perspective, but also from the perspective of the risk of inconsistency of the generated schedule to the reality of operations.

Being a structural component of a very complex and tightly interconnected system, airports suffer from various types of uncertainties. This unpredictability is a natural part of the air transportation network, as many activities can suffer changes in the very last moment, affected by the weather conditions, governmental regulations, air traffic control and etc.

One of the main consequences of such uncertainty are flight delays and early arrivals. Some flights suffer from delay, originated in previous legs and propagating through the network as reactionary delays. Other flights can be coming to their destinations earlier than expected. Both of these deviations create additional load to the decision-making process. This work implies instead of predicting exact values of flight delays, estimate the probabilities of having a certain delay level for each flight and use this information for estimating the quality of the stand assignment schedule.

III. Methodology

We propose a concept of a stand assignment algorithm that deals both with environmental uncertainties and with optimization of facilities' usage. The algorithm consists of two modules. First module estimates probabilities of delays and their level based on the historical data of previous operational periods. The second module generates the assignment schedule, based on the desired technical and operational restrictions for a target flight schedule, and optimizes it with a genetic algorithm component. To calculate the stand occupancy time for each flight, we estimate in-block and off block times based on the target flight schedule and the delay probability, obtained in the first module.

To generate a stand assignment schedule for a specific operational day, the following data is used:

- Target flight schedule for assignment.
- Existing parking facilities and their technical and operational restrictions (compatible with specific aircraft types, individual use by certain airlines or for certain origins/destinations).
- Availability of stands.

Finally, assignment policy specific data, such as taxi time, walking distances for transfer passengers, etc., are added to the data set as well.

A. Algorithm Architecture

As it has been mentioned in the previous section, the algorithm consists of two modules: one - to estimate the probabilities of delay, and one - to generate a stand assignment schedule, optimized for specific management goal (minimizing transfer passengers walking distances, minimizing taxi time, etc.).

The first module is directly connected to a performance database, which allows re-estimating the delay probabilities in real time, considering also recently available information, e.g. about flight regulations and weather conditions. In this module, the historical delay values are analyzed for different combinations of factors (e.g. airline, aircraft type, operational hour, and weather conditions) and corresponding Bayesian distributional regression models are built. These models together with the corresponding parameters are then passed on to the second module.

In the second module, the target flight schedule is recalculated, according to the regression models obtained in the previous step, and the estimated delay values are added to the block occupancy times. After that, this recalculated flight schedule is passed to a metaheuristic solution search algorithm, which looks for a better stand assignment for the flights, optimizing the user-specified objective function or the weighted combination of them.

The number of iterations, total running time and objective function value can limit the calculation time, according to the user needs. Therefore, the solution quality improvement is only restricted by the user estimations.

B. Algorithm Output

On the exit of the second module component, metaheuristic search algorithm, the stand assignment schedule is obtained. Within the obtained schedule, for every flight, assigned to the stand, the deviation risk value is displayed. This risk value indicates that although the flight is assigned to the specific stand, there is a probability of N percent that this flight will suffer delays and affect the rest of the assignment schedule. By displaying such information, we intend to provide the airport managers with an insight to the most critical points in the schedule and facilitate the decision-making process with a quantitative estimation of possible operational scenarios. In such a way, it is possible to measure the impact of any air traffic regulations on the slot adherence and generate various stand assignment schedules for different performance scenarios with different levels of risk.

IV. Conclusions

In this work, we present a conceptual solution to the most common airport problem – efficient stand assignment. The presented algorithm combines benefits of data-mining tools and metaheuristic approaches and generates qualitative solutions, conscious to historical delay trends and performance drops. This two-module algorithm generates promising solutions from the first iterations, it provides airport stakeholders with an approach for delay-aware stand assignment and facilitates the estimation of impact of operational disruptions on the slot adherence.

Acknowledgment

Opinions expressed in this paper reflect the authors' views only.

References

[1] J. Guépet, R. Acuna-Agost, O. Briant, and J. P. Gayon, "Exact and heuristic approaches to the airport stand allocation problem," Eur. J. Oper. Res., vol. 246, pp. 597–608, 2015.

[2] J. Pearl, "Heuristics: Intelligent Search Strategies for Computer Problem Solving," Telemat. Informatics, vol. 3, no. 4, p. 305, Jan. 1986.

8th SESAR INNOVATION DAYS | 3-7 December 2018, Salzburg

SESAR JOINT UNDERTAKING

No More Surprises

Stand assignment algorithm with likelihood of turnaround time deviation

Objective:

To generate a robust stand assignment, optimized and balanced for various stakeholders

- **Airport** → Efficiently use the capacity, minimize waiting time
- **Airline** → Minimize non-profit time, depart on time
- **Government** → Reduce environmental impact, control the border
- **Passenger** → Receive high quality service

Methodology:

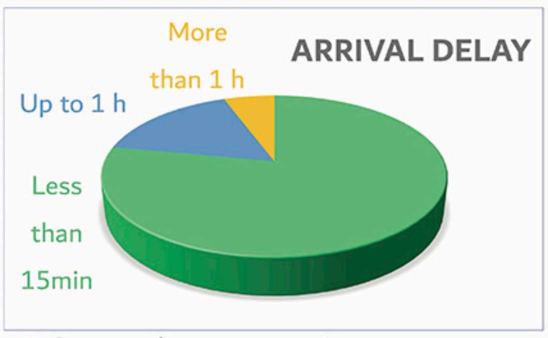
PROBABILISTIC MODELLING
Estimation of delay probability distribution, from historical data

DELAY – AWARE STAND ASSIGNMENT
Generation of stand assignment, considering probable delays

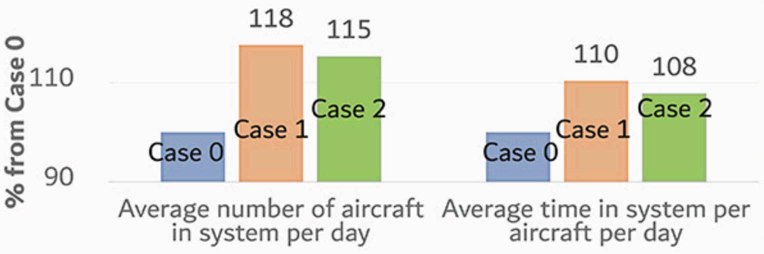
STAND ASSIGNMENT OPTIMIZATION
Optimization of generated assignment to maximize stakeholders benefits

- ✓ Optimization
- ✓ Robustness
- ✓ Real time adjustment
- ✓ Risk estimation
- ✓ Waiting time reduced on 2%/day

Case study: Mexico City International Airport



- Saturated runway capacity
- 2 terminals, 91 parking positions
- Approx. 45 million PAX per year (2017)



Universitat Autònoma de Barcelona
Amsterdam University of Applied Sciences
MARGARITA.BAGAMANOVA@UAB.CAT
M.MUJICA.MOTA@HVA.NL

Airports

UNIVERSITÄT SALZBURG
Engage

founding members
EUROPEAN UNION
EUROCONTROL

Appendix D: Reduction of taxi-related airport emissions with disruptions-aware stand assignment: case of Mexico City International Airport

M. Bagamanova and M. Mujica Mota

Published in the *Proceedings of the Winter Simulation Conference* 2020 (in press).

Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds

REDUCTION OF TAXI-RELATED AIRPORT EMISSIONS WITH DISRUPTION-AWARE STAND ASSIGNMENT: CASE OF MEXICO CITY INTERNATIONAL AIRPORT

Margarita Bagamanova

Aeronautical & Logistics Departmental Unit
Autonomous University of Barcelona
UAB University Campus
Bellaterra, 08193, SPAIN

Miguel Mujica Mota

Aviation Academy
Amsterdam University of Applied Sciences
190 Weesperzijde
Amsterdam, 1097 DZ, NETHERLANDS

ABSTRACT

Airport management is often challenged by the task of managing aircraft parking positions most efficiently while complying with environmental regulations and capacity restrictions. Frequently this task is additionally affected by various perturbations, affecting punctuality of airport operations. This paper presents an innovative approach for obtaining an efficient stand assignment considering the stochastic nature of the airport environment and emissions reduction target of the modern air transportation industry. Furthermore, the presented methodology demonstrates how the same procedure of creating a stand assignment can help to identify an emissions mitigation potential. This paper illustrates the application of the presented methodology combined with simulation and demonstrates the impact of the application of Bayesian modeling and metaheuristic optimization for reduction of taxi-related emissions.

1 INTRODUCTION

Modern airports are facing a global challenge of significant reduction of pollutant emissions and moving towards carbon-neutral operations while coping with rapid air traffic growth and maintaining the required level of service (ICAO 2019a). Emissions produced by airport activities influence local air quality at and around airports. One of the most important sources of emissions at the airport is aircraft operations, such as landing, taxiing, and take-off (ICAO 2019b). Thus, in addition to technological innovations and switching to sustainable aviation fuels, improvement in the efficiency of these operations is considered in the scope of global air industry measures (ATAG 2020).

The level of emissions produced during taxiing in the airport depends on the amount of fuel burnt and the time for which an aircraft has to move between its assigned parking position (stand) and runway entrance/exit points. In general, over a third of total aircraft emissions outside of the cruise phase can be generated during taxiing (Fleuti and Maraini 2017). Therefore, there is a need to allocate aircraft in such a way that taxiing distance and time are minimized, ensuring the reduction of fuel consumption and related emissions.

A stand assignment schedule can often be disrupted by last-minute changes in the flight schedule during the day. Such changes may lead to longer turnaround times and deteriorating airport performance. As a result, some aircraft might have to wait on the ground and some might have to wait in the air in the airport TMA, which culminates in higher fuel consumption and additional emissions.

Ineffective management of terminal facilities can create a propagation of schedule disruptions to the successive flights and connected airports, also affecting the level of emissions. Therefore, it is necessary to efficiently manage terminal facilities, such as stands, to mitigate the impact of scheduled perturbations and reduce the level of pollutant emissions, created during taxi, at the same time.

The stand allocation problem (also known as the stand assignment problem), tackled in this paper, was previously approached by many researchers. However, only some of them considered the stochasticity of airport operations in their methodology. Quite often to decrease the number of stand allocation conflicts and related aircraft waiting times, a stand has a certain idle time between two consecutive flights assigned to it. This idle time is called buffer time and has been proven as the best working measure for flight deviations up to 30 minutes (Hassounah and Steuart 1993; Yan and Chang 1998; Yan and Huo 2001; Yan et al. 2002). Nonetheless, such action can significantly reduce airport terminal capacity and therefore, should be avoided in modern congested airports.

Idle waiting and taxiing are estimated to contribute the most to the aircraft fuel consumption and airport emissions (Nikoleris et al. 2011). Therefore, the goals of idle waiting reduction and taxiing footprint optimization were approached by many researchers. Duinkerken et al. (2013), Li and Zhang (2017) estimated that using a single-engine approach, external electric engine, and towing sources for taxiing can significantly reduce emissions. Tsao et al. (2009) demonstrated that aircraft idle waiting time on the ground can be reduced by optimization of taxi-out and take-off sequences. Applications of pushback control, gate holding, and departure sequence optimization (Simaiakis and Balakrishnan 2016) applied by Khadilkar and Balakrishnan (2012), Simaiakis et al. (2014), and Ashok et al. (2017) showed a significant reduction of emissions related to taxiway and runway congestions.

Although the aforementioned methods proved to reduce environmental footprint, some of them also led to increasing stand occupancy times, thereby significantly reducing airport capacity which can become problematic in congested airports. Furthermore, these works have considered neither flight arrival time perturbations, nor taxiing from the runway to the stands (the taxi-in phase) that can substantially impact the taxiing time and related emissions (Hao et al. 2016). To fill the gap in this area and provide air transport management with a methodology to improve both efficiency and environmental impact of stand assignment operations, this paper presents how these two objectives can be combined in the stand assignment and demonstrates their achievement using simulation techniques.

This paper presents a bi-objective application of a stand assignment approach, that was previously introduced by Bagamanova and Mujica Mota (2020), for evaluating various stand assignment policies in terms of their sensitivity to schedule perturbations and environmental footprint. The presented methodology combines the benefits of data-mining and evolutionary optimization for generating a stand assignment that minimizes emissions and, by using simulation, the efficiency against possible schedule deviations and related emissions reduction is proved. The presented approach learns probabilities of schedule deviations depending on characteristics of the scheduled flights using Bayesian multilevel modeling (Bürkner 2017) from historical airport performance data. These probabilities are then used to calculate the most probable level of deviation for each flight in the target flight schedule. The calculated deviations are then considered in the generation of stand assignment, which is optimized to meet the goal of minimization of emissions generated during the taxi of an aircraft.

This paper continues as follows. Section 2 outlines the stand assignment methodology. Section 3 presents a case study and simulation experiments results. Conclusions and further research are presented in Section 4.

2 METHODOLOGY

The stand assignment method presented in this paper is composed of the two-module approach and experiments in a simulation model. The two-module approach generates optimized stand allocations based on the target flight schedule, historical data about schedule disruptions for the previous period, and user-defined assignment policies and optimization goals. After that, the obtained allocations are estimated in the simulation model that allows evaluating the environmental footprint quality of stand assignments generated in the two-module approach under the stochasticity of a real-life airport system.

2.1 Algorithm Description

This section gives a short description of the two-module approach that generates optimized stand assignments. A more general description can be found at Bagamanova and Mujica Mota (2020).

The two-module approach is composed of two elements. Module I takes care of estimating probabilities of schedule deviations from the airport historical data. These probabilities are expressed in the form of Bayesian distributional models and describe a likelihood of certain levels of schedule deviations for various flight characteristics available in the historical data (e.g. such as airline name, scheduled time of arrival, and day of the week). By considering probable disruptions in the assignment planning, it is intended to reduce the idle time that aircraft might have to spend waiting for the planned stand availability and related emissions.

Module II assigns the target flight schedule to the available stands, respecting user-defined assignment policy and restrictions, considering most probable or user-defined probability level schedule disruptions in the stand occupancy times. Then the generated assignment is optimized with a genetic algorithm according to user-specified optimization goals. The result of such optimization is not necessarily an optimal solution, however, randomness used in the genetic algorithm in the form of crossover and mutation operators allows us to obtain a good quality solution in a reasonable time (Bagamanova and Mujica Mota 2020). The resulting stand assignment considers the stochasticity in the form of stand occupancy times deviations generated from the schedule deviations distributional models.

2.2 Optimization Objective

To increase stand assignment efficiency and mitigate pollutant footprint, produced by aircraft movement on the ground and aircraft idle waiting for stand availability, the following bi-objective optimization goal function has been implemented in the optimization component of Module II of the two-module approach:

$$min(w_1 * O_{taxi} + w_2 * O_{hold}) \tag{1}$$

The objective function (1) consists of the following individual objectives:

1. Minimize taxi distance to and from the parking positions and therefore the related emissions:

$$O_{taxi} = \overline{d_{sched.taxi}} / Max\ d_{airport}$$

2. Minimize the number of aircraft waiting for stand availability and, therefore, the idle use of engines:

$$O_{hold} = \sum fl.hold \ / \sum fl.$$

Where:

- $\overline{d_{sched.taxi}}$ – the average taxi distance to and from the stand in the allocated schedule;
- $Max\ d_{airport}$ – the maximum possible taxi distance at the airport for considered runway configuration;
- $\sum fl.hold$ - the number of aircraft that must wait for the stand availability;
- $\sum fl.$ - the total number of aircraft in the schedule to allocate;
- w_n – priority weight for the corresponding objective. In the scope of this paper, all the weights are equal to 1 to obtain a stand assignment equally balanced for both considered objectives. For practical use, different stakeholders of the airport can decide the weights based on their preferences.

In the original implementation of the two-module approach by Bagamanova and Mujica Mota (2020), the optimization objective function in Module II also included maximization of the use of contact stands. This is a general preference for many airports as it allows to fully benefit from terminal building in terms of providing passenger experience and reduces the number of ground service vehicles moving on the apron. Yet, for the scope of this paper, such an objective was excluded as the primary goal is to generate a stand assignment with minimized emissions. Nevertheless, it might be interesting to investigate the environmental cost of prioritizing contact stand use in the optimization component in future work.

3 CASE STUDY: MEXICO CITY INTERNATIONAL AIRPORT

This section discusses the application of the two-module approach for encountering more environmentally efficient stand assignment policies for a case study airport.

3.1 General Information

Mexico City International Airport (IATA code: MEX) is the main airport in Mexico with approximately 450 thousand landings and take-offs annually. There are two terminal buildings, separated by two parallel runways. These runways are never operated simultaneously due to proximity to each other. Such layout restricts MEX capacity and since 2017 it has been officially limited to 61 movements per hour with a maximum of 40 landings (SCT 2017).

In the scope of this paper, it is considered that 26 airlines are operating in two terminals in MEX, performing both international and domestic flights. From the total 91 stands available at MEX, only 84 were considered in this paper, as the rest is not used for passenger flights. Hence, Terminal 1 is represented by 11 open stands and 33 contact stands, among which 16 stands are dedicated to domestic flights and 17 to international. Terminal 2 is represented by 17 open stands and 23 contact stands, where 13 are used for domestic flights, 10 – for international.

3.2 Schedule Disruptions and Emissions

On a global level, in 2018 Mexico generated approximately 1.5% of global air passenger transport-related emissions (Graver et al. 2019). MEX is located in the direct proximity of the urban zones of Mexico City, which makes the airport significantly affect air quality and noise levels of the city. According to SEDEMA (2018), MEX produces around 15% of the total pollutant emissions of Mexico city.

In 2017 Mexico has officially joined a global initiative for carbon-neutral air transport operations (ICAO 2020), which implies that all country airports have to follow ICAO emission reduction policies and standards. Despite these facts, up to the date of writing this paper official MEX website did not publish any official estimations of airport emissions level nor disclosed any measures to reduce the environmental footprint of its operations.

MEX frequently suffers from punctuality problems. In 2018 only 67% of all flights were performed on time (SCT 2019) with more than 20% of departing flights being delayed by 46 min on average (Flightstats 2018). Considering such a high level of perturbations and recent engagement in global pollutant footprint reduction initiative, MEX becomes a good target for application of the two-module approach to discover the hidden potential for emissions reduction related to stand assignment planning.

3.3 Implementation of the Two-Module Approach

As input data for this study, we used an official performance report for a period from 28.05.2018 to 03.06.2018, retrieved from International Airport of Mexico City (2018). This report consisted of more than 8,000 flights with actual and scheduled arrival times, flight numbers, airline names, and type of aircraft used. In the chosen week approximately 7% of arriving flights deviated for more than one hour from their schedule. More than 53% of scheduled arrivals suffered from a substantial delay of more than 15 min, which is a significant perturbation for a congested airport.

Due to the unavailability of actual data on turnaround times and arrival-departure aircraft correspondence, it was assumed to use only arriving passenger flights from the obtained report and define 60 minutes turnaround time for all flights in the performed experiments. Such limitations reduced the number of flights to 3,914 arrivals, where 31.7% were international flights and 68.3% - domestic.

The selected data of 3,914 flights have been processed in Module I and the Bayesian models for arriving time deviations were built, assuming the correlation of deviations with airline name and hour of scheduled arrival. The detailed description of the resulting parameters of regression models, composing the summative Bayesian model, and output of Module I can be found at Bagamanova and Mujica Mota (2020).

Lastly, Module II created an assignment, considering most probable scheduled deviations, assignment policy restrictions, and optimized it according to the objective function (1). As the two-module approach is considered to be a more effective replacement to traditionally used buffer times, for the generation of stand assignment in Module II no buffer times were intentionally added between consecutive flights assigned to the same stand. The resulting assignment statistics are shown in Figure 1.

Every airport has its own stand assignment policy restrictions, which implies certain use of the stands. The following are the restrictions considered in the presented algorithm:

- Domestic and international flights must be assigned to the specific stands in the designated zones. These are internal specifications of the airport e.g. international flights are assigned to stands that have access to the designated border control areas;
- Flight delays must be considered in the assignment (according to conditional probability distributions from Module I). In this paper, only arrival delays are considered due to unavailability of ground handling data and correspondence of arriving aircraft to departing aircraft;
- An assigned stand must correspond to the size of an aircraft (large aircraft require extra space due to larger wingspan). This is implemented through the identification of allowed stands for each flight on the stage of processing the input data in Module II.

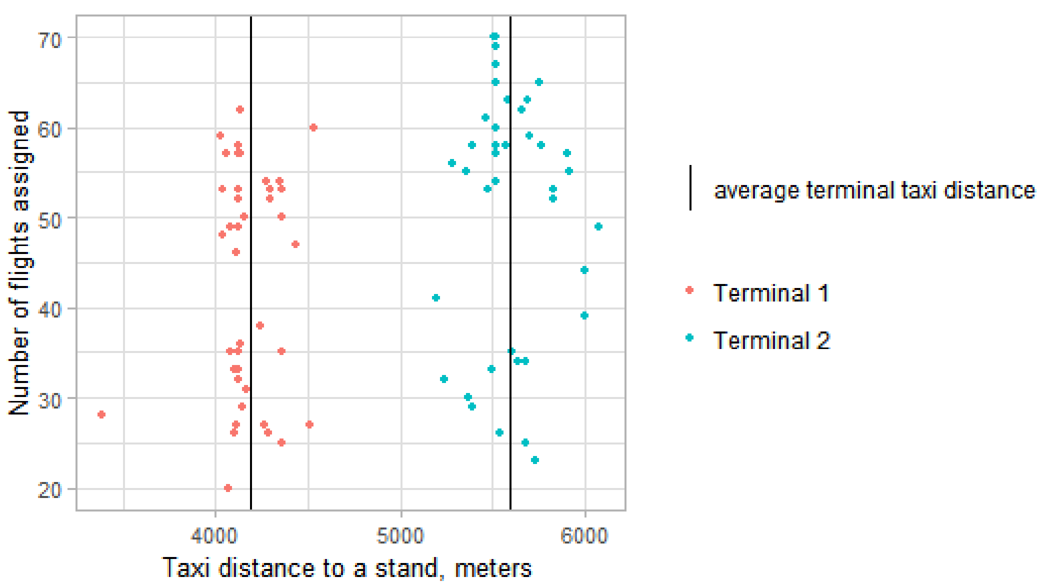


Figure 1: Assignment statistics for Module II generated stand allocation.

As can be observed from Figure 1, most of the flights were assigned to stands located not too far from the runways. In Terminal 1 approximately 61.1% of scheduled flights were assigned to a stand located closer than Terminal 1 average taxi distance of 4.2 km from the runway; for Terminal 2 - 61.3% of flights were assigned to the stands with less than average Terminal 2 taxi distance of 5.6 km. Naturally, some of the flights had to be assigned to further located stands due to assignment policy constraints, designated border control zones, and unavailability of closer located stands. Nevertheless, Figure 1 demonstrates the algorithm's success with the minimization of taxi distance.

One of the limitations of the data used for this study is the unavailability of actual historical MEX stand assignments. Therefore, for the moment, it is impossible to compare the quality of the two-module approach results with actual MEX stand assignments. Thus, to evaluate the quality of the obtained assignment and owing to the absence of actual historical stand assignments at MEX, the two-module approach assignment was tested in the environment of the MEX simulation model, as described in the next section. The detailed description and validation of this simulation model can be found at Mujica Mota and Flores (2019).

3.4 Simulation Experiments

The principal objective of using a simulation model in this study is to evaluate the effects of consideration of schedule deviations in the stand assignment on the taxi-related emissions in close-to-reality conditions and encounter ways to improve airport performance and emissions level. The simulation model used in this study allows us to incorporate stochastic elements (such as stop-go situations, waiting for push-back at the gate) that were not considered in the assignment generation, but do influence aircraft movements on the ground in the real life.

For each simulation replication the following performance indicators were tracked:

- total taxi distance for all aircraft of the allocated schedule: $d_{total\ taxi} = \sum_{i=1}^{N}(d_{in\ i} + d_{out\ i})$;
- total taxi time for all aircraft of the allocated schedule: $t_{total\ taxi} = \sum_{i=1}^{N}(t_{in\ i} + t_{out\ i} + t_{wait\ i})$;
- total amount of taxi-related pollutant emissions $e_{total\ taxi} = t_{total\ taxi} * F_{NO} + t_{total\ taxi} * F_{CO}$;

where:

- $d_{in\ i}$ – distance traveled by aircraft *i* from runway exit to a stand;
- $d_{out\ i}$ – distance traveled by aircraft *i* from a stand to runway entry point;
- $t_{in\ i}$ – time traveled by aircraft *i* from runway exit to a stand;
- $t_{out\ i}$ – time traveled by aircraft *i* from a stand to runway entry point;
- $t_{wait\ i}$ - time spent by aircraft *i* waiting for stand availability;
- F_{NO} and F_{CO} – emission factors for NOx and CO_2 respectively;
- $i \ldots N$ – number of aircraft.

Emission factors depend on the engine characteristics, type of fuel used, and aircraft weight among others (ICAO 2019b). Due to the unavailability of any actual data about engine specifications and aircraft weight for the studied flight schedule, the amount of total emissions $e_{total\ taxi}$ was calculated assuming constant taxi speed and the taxi emissions reference for Airbus A320 (engine CFM56) (European Environment Agency 2016). This aircraft type was chosen as it was used in 55% of the studied flights. Less than 1% of the studied flights were performed with a large type of aircraft and the rest of the flights were represented mostly by regional class. The adapted emission factors per minute of taxiing are shown in Table 1.

Table 1: Emissions factors per minute of taxiing.

Type	Factor, kg/min
Fuel consumption	14.52
NOx emission per min, F_{NO}	0.065196
CO_2 emission per min, F_{CO}	1.7604

Assuming certain emission factors in this paper is made to get a general estimation of the two-module approach application impact on airport emissions. Nevertheless, it is considered to perform a more detailed calculation in the future, accounting for different emission factors for all present types of aircraft, when more actual data on aircraft specifications become available.

At the time of performing this study, there was no information available about exact or historical stand assignments in MEX. Therefore, the two-module approach generated assignments were compared to a random last-minute assignment, generated directly during every simulation run. A random last-minute assignment allocates a flight during simulation to any suitable stand available at the moment of aircraft starting landing approach. That means that any suitable stand not occupied at the decision moment can be chosen regardless of its taxi distance to the runway. As the choice is made randomly, every simulation run results in different usage of stands. As there is no preliminary planned assignment in such last-minute allocation, it is considered that the effects of schedule disruptions on stand usage are minimized and there is less possibility for assignment conflicts. Although, it is not estimated at what environmental cost these effects are minimized. In this section, the effects of such last-minute random allocation on the taxi-related emissions are estimated and compared to a proactive allocation planning, performed by the two-module approach. Additionally, to trace the effects of schedule deviations on taxi-related emissions, simulation scenarios containing both on-time and disrupted arrivals were included in this study.

An overview of the defined stand assignment scenarios is presented in Table 2. These scenarios can be described as follows:

1. *Scenario A*. Base case. It represents an ideal situation with all flights arriving on time, stand assignment generated only with the use of Module II (i.e. optimized allocation without considering deviations).
2. *Scenario B*. Stand assignment generated only with the use of Module II (i.e. optimized allocation without considering deviations). The flights arrived with arrival time deviations, generated based on arrival time deviation distributions learned in Module I.
3. *Scenario C*. Stand assignment generated considering the expected delay with the use of both Module I and Module II. Flights arrived with arrival time deviations, generated based on arrival time deviation distributions learned in Module I.
4. *Scenario D*. Arriving flights are assigned to stands using last-minute random allocation. Flights arrived on time, according to the schedule.
   Scenario E. Arriving flights are assigned to stands using last-minute random allocation. Flights arrived with arrival time deviations, generated based on arrival time deviation distributions learned in Module I.

Table 2: Stand assignment scenarios.

Scenario name	Schedule disruptions	Schedule disruptions considered	Assignment optimization	Assignment generation
A	-	-	Yes	Module II
B	Yes	-	Yes	Module II
C	Yes	Yes	Yes	Two-module
D	-	-	-	Random last-minute
E	Yes	-	-	Random last-minute

The objective of this paper is to discover the hidden potential for the reduction of taxi-related emissions through stand assignment optimization. And as has been observed in the analysis of the generated assignment in section 3.3, the current distribution of domestic and international areas in the terminals has a considerable influence on the assignment results and therefore on the level of taxi-related emissions. Therefore, the relaxation of some restrictions of MEX was considered to verify if such action can bring any benefit to the environmental footprint of real-life stand assignment operations. Therefore, it has been decided to manipulate some of the available assignment restrictions and therefore come up with new assignment policies, that would not require major airport facilities reconstruction. The only requirement remaining strict for all simulated assignment policies is the requirement of assignment of large aircraft only to the specially equipped stands. The new assignment policies were compared to the original policy, which contains strict assignment constraints, through the series of experiments, simulating scenarios A-E under each of the defined policies. In such a way for every assignment policy, the performance of the two-module approach under on-time and disrupted arrivals were evaluated and compared to the random last-minute allocation. The defined assignment policies include the following:

1. *Group I* – base case experiments. Stand assignment generated according to the original set of assignment restrictions with strict adherence to the designated terminal and international/domestic zone.
2. *Group II* – aircraft are allocated to any available stand in the originally planned terminal. This means that both international and domestic flights can be allocated to the same stand.
3. *Group III* – aircraft may choose stands in any terminal but must obey the designated zone policy. This means that a domestic flight must be assigned to the domestic zone but can be assigned to the domestic zone of any terminal.
4. *Group IV* – aircraft can be assigned to any zone of any terminal. This is a layout restrictions-free assignment policy that allows getting closer to the minimum possible taxi distance and taxi-related emissions for the studied flight schedule.
5. *Group V* – Terminal 1 is fully designated for domestic flights. This means that even if a flight was originally planned to Terminal 2, in case if it is domestic it will be assigned to Terminal 1.
6. *Group VI* - Terminal 1 is fully designated for international flights. This means that even if a flight was originally planned to Terminal 2 if it is international it will be assigned to Terminal 1.

Using the same data to learn Bayesian distributional models for schedule disruptions and to generate simulation experiments stochasticity can be considered as a limitation of this paper. Nevertheless, the main goal of the proposed approach is to mitigate the negative impact of schedule disruptions on the airport environment, not to predict the exact delay or early arrival time for the scheduled flights. By considering a certain probability interval in the assignment planning, we intend to provide a tool for influencing stand allocation robustness. With a bigger probability interval, more perturbations can be considered; however, it might reduce stand resources capacity and thus, can be seen as a limitation for some congested airports. Smaller probability intervals would result in smaller stand blocking times but might increase the number of aircraft that might wait for the stand availability. This trade-off is not discussed in this paper, although will be explored in future research.

For each assignment policy, experiments A-E were executed with 30 replications each. Each replication had a duration of 7 days plus extra hours for arrival schedule deviations. The next section presents and discusses the results of the performed experiments.

3.5 Experiments Results

The performed experiments results were compared across scenarios to identify an assignment policy that allows to significantly reduce the emissions. The comparative statistics for the tracked indicators for experiments in Groups I – VI are presented in Figure 2 - Figure 4.

As can be seen in Figure 2, scenarios A and B have similar taxi distance values, as they used the same stand assignment; scenario B differs from scenario A only in presence of stochastic arrival time deviations. Scenarios D and E generally resulted in higher taxi distance value, as they did not optimize the assignment to minimize the taxi time. The lowest taxi distance was achieved in Group V, which corresponds to the assignment policy with Terminal 1 being fully dedicated to the domestic flights and Terminal 2 – to international. Under this policy, both the two-module approach and random last-minute allocation generated close values with a 0.2% difference. The lowest taxi distance corresponds to scenario E in Group V, which is 4.2% lower than in assignment generated by the two-module approach under original assignment policy in scenario C Group I. In these experiments, random last-minute allocation outperforming an optimized stand assignment can be explained by the fact that last-minute allocation in scenarios D and E was allowing overlapping assignments to the same stand if all other suitable stands already had been occupied. Overall, results demonstrated in Figure 2, reveal that by reorganizing the use of MEX terminal buildings and dedicating Terminal 1 entirely to domestic flights it is possible to reduce taxiing distance by 4.2% weekly.

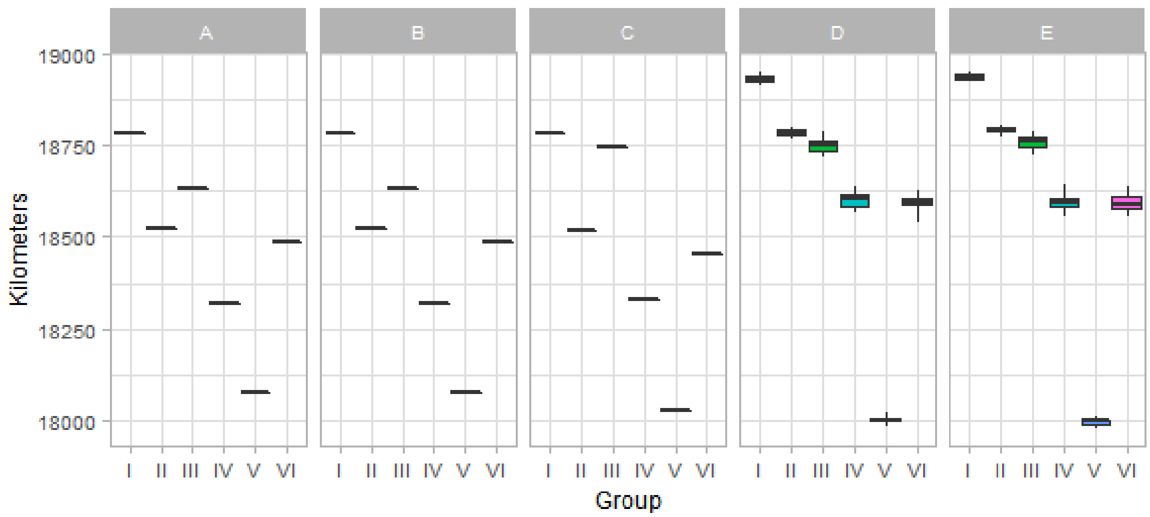


Figure 2: Comparison of total taxi distance for scenarios A - E through the groups of experiments.

When talking about total taxi time, shown in Figure 3, it can be noticed that scenario B shows more variability and higher mean values than scenario A due to the presence of stochastic deviations and aircraft waiting times. The lowest taxi time value corresponds to scenario C in Group V, which is the allocation generated by the two-module approach. The total taxi time obtained in this scenario is 9% lower than in scenario C of Group I.

Remarkably, the taxi time in scenario C through all the groups is always lower than in scenarios with random last-minute allocation and schedule disruptions not considered in the allocation (B, D, and E). Such observation allows concluding, that consideration of expected schedule deviations in the stand assignment is beneficial for airport operations as it results in shorter taxi times owing to decreased stand availability waiting times.

When the amount of total pollutant emissions is compared, the lowest value again corresponds to Group V for Scenarios C (see Figure 4). The amount of emissions in scenario C Group V is approximately 9% less than the amount produced under original assignment policy of Group I. Random last-minute allocation in scenario E in Group V, interestingly, resulted only in 3.8% higher emissions than in scenario C. However, such a random allocation demonstrated quite a high variability under all assignment policies.

It can be noticed that Figure 3 and Figure 4 have similar values, which could be explained by the assumption of uniform emissions factors for the entire study. Nevertheless, it could be interesting to repeat the experiments in the future with more specific emission factors, e.g. adapted from BADA (EUROCONTROL 2020), and analyze the correlation between total emissions and total taxi time for a purpose of combining them into a single optimization objective.

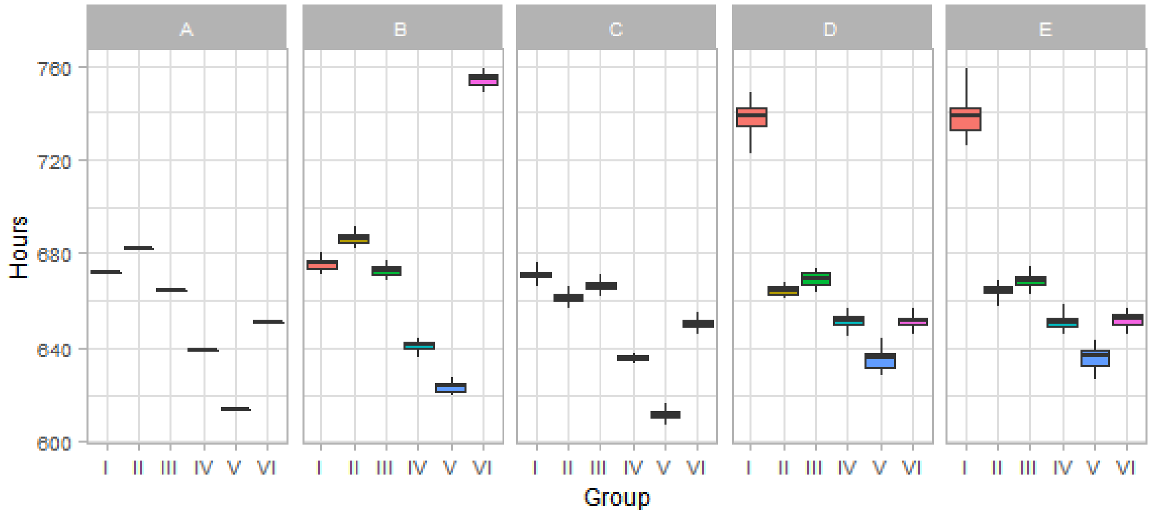


Figure 3: Comparison of total taxi time for scenarios A - E through the groups of experiments.

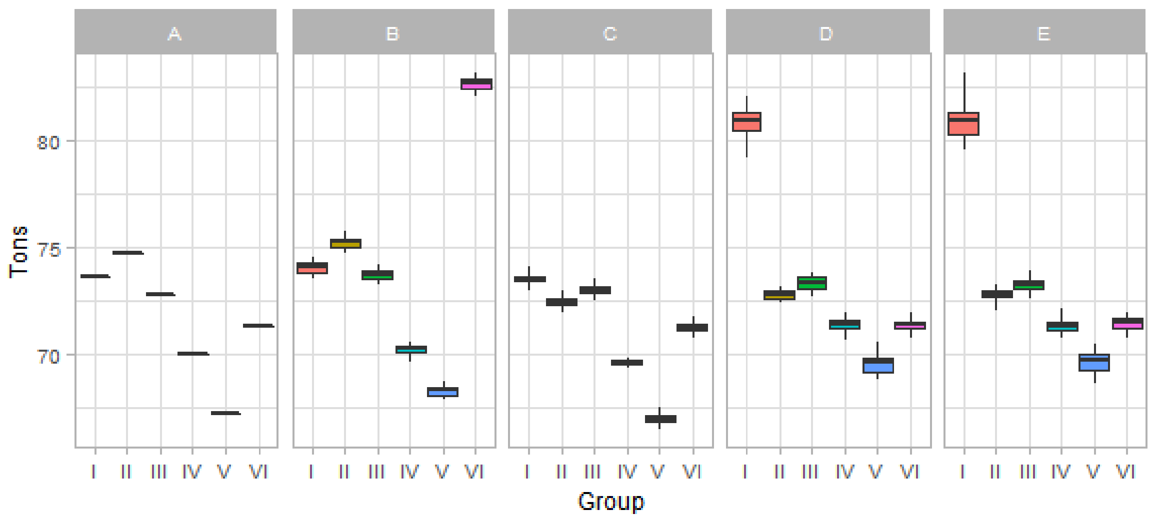


Figure 4: Comparison of total CO_2 + NOx emissions for scenarios A - E through the groups of experiments.

Summarizing the results, it can be concluded that the most beneficial stand assignment policy in terms of related emissions is the one in Group V. This means that rearranging the use of terminal buildings and dedicating Terminal 1 to domestic flights, can save MEX around 9% of total pollutant emissions weekly compared to existing terminal buildings designation under the operational conditions considered in the experiments.

4 CONCLUSIONS AND FUTURE RESEARCH

This paper presents an application of an innovative approach that combines Bayesian methods and a bi-objective heuristic optimization for solving the stand allocation problem in airports from the perspective of minimization of related emissions. To validate the impact of the presented approach on airport environmental footprint, the simulation was included in the methodology to introduce the effects of the stochastic nature of the real-life system. In the case presented, the methodology showed a clear benefit of consideration of possible schedule disruptions in the stand assignment planning for emissions mitigation. Furthermore, the application of the two-module approach with the relaxation of assignment restrictions revealed a hidden potential of mitigation of taxi-related pollutant emissions. For the case of Mexico City International Airport, the best-obtained results correspond to the dedication of entire Terminal 1 to domestic flights and Terminal 2 – to international flights Such rearrangement of terminal buildings could decrease taxi-related pollutant emissions by approximately 9% weekly, compared to the present use of terminals.

As future work, other variables, such as actual turnaround times and departure time deviations, and more historical performance data would be considered in Module I for providing more accuracy on the expected perturbations. When more historical data become available, it would also be beneficial to use different but comparable sets of data to learn the deviation models and perform the simulation experiments to better estimate the accuracy of the obtained deviation models. Enhancement of the two-module approach optimization component and simulation study with aircraft emissions specifications is also considered for a more precise estimation of emissions and their impact on the stand allocation.

Furthermore, it would be interesting to compare the quality of the two-module generated stand assignments with the historical (actual) stand assignments for the same airport and test the presented approach on other airport configurations and stand assignments policies. Additionally, it can be investigated if taxi distance reduction and emission reduction objectives can be combined into a single optimization objective and what would be the impact on emissions level if aircraft waiting times are considered instead of the number of waiting aircraft. Moreover, the use of information obtained from the simulation model will be incorporated into the optimization loop to provide even more qualitative solutions.

ACKNOWLEDGMENTS

The authors would like to thank the Autonomous University of Barcelona, the Aviation Academy of the Amsterdam University of Applied Sciences for supporting this study and the Dutch Benelux Simulation Society (www.dutchbss.org) and EUROSIM for the dissemination of the findings of this study.

REFERENCES

Ashok, A., H. Balakrishnan, and S. R. H. Barrett. 2017. “Reducing the Air Quality and CO 2 Climate Impacts of Taxi and Takeoff Operations at Airports”. *Transportation Research Part D: Transport and Environment* 54:287–303.

ATAG 2020. Aviation Industry Reducing Its Environmental Footprint: Aviation: Benefits Beyond Borders. https://www.aviationbenefits.org/environmental-efficiency/climate-action/, accessed 12th February.

Bagamanova, M., and M. Mujica Mota. 2020. “A Multi-Objective Optimization with a Delay-Aware Component for Airport Stand Allocation”. *Journal of Air Transport Management* 83:101757.

Bürkner, P.-C. 2017. “Brms: An R Package for Bayesian Multilevel Models Using Stan”. *Journal of Statistical Software* 80(1):1–28.

Duinkerken, M., T. Selderbeek, and G. Lodewijks. 2013. “Reducing Emissions of Taxiing at Airports”. In *the 2013 European Simulation and Modelling Conference*, edited by S. Onggo and A. Kavicka, 331-335. Ostend: The European Multidisciplinary Society for Modelling and Simulation Technology.

EUROCONTROL 2020. Base of Aircraft Data. https://www.eurocontrol.int/eec/public/standard_page/proj_BADA.html, accessed 13th April.

European Environment Agency. 2016. *European Monitoring and Evaluation Programme/European Environment Agency Air Pollutant Emission Inventory Guidebook*. https://www.eea.europa.eu/publications/emep-eea-guidebook-2016, accessed 8th April 2020.

Fleuti, E., and S. Maraini. 2017. *Taxi-Emissions at Zurich Airport*. https://flughafen-zuerich.ch/~/media/flughafenzh/dokumente/das_unternehmen/laerm_politik_und_umwelt/taxi_study_zurichairport_20171207.pdf, accessed 10th February 2020.

Flightstats. 2018. *Airport On-Time Performance Reports*. https://www.flightstats.com/v2/monthly-performance-reports/airports, accessed 10th January 2019.

Graver, B., K. Zhang, and D. Rutherford. 2019. *CO_2 Emissions from Commercial Aviation, 2018*. https://theicct.org/publications/co2-emissions-commercial-aviation-2018, accessed 12th February 2020.

Hao, L., M. Hansen, and M. S. Ryerson. 2016. “Fueling for Contingencies: The Hidden Cost of Unpredictability in the Air Transportation System”. *Transportation Research Part D: Transport and Environment* 44:199–210.

Hassounah, M., and G. Steuart. 1993. “Demand for Aircraft Gates”. *Transportation Research Record* 1423:26–33.

ICAO 2019a. ICAO Environmental Report. https://www.icao.int/environmental-protection/Documents/ICAO-ENV-Report2019-F1-WEB%20%281%29.pdf, accessed 25th March.

ICAO 2019b. The ICAO Aircraft Engine Emissions Databank. https://www.easa.europa.eu/easa-and-you/environment/icao-aircraft-engine-emissions-databank, accessed 16th November.

ICAO 2020. CORSIA Participation. https://www.icao.int/environmental-protection/Lists/CORSIAParticipation/AllItems.aspx, accessed 12th February.

International Airport of Mexico City. 2018. Operaciones de Llegada y Salida Semanal (Weekly Arrival and Departure Operations). https://www.aicm.com.mx/wp-content/uploads/2019/04/OperacionesdeLlegadaySalida25al31Marz2019.pdf, accessed 10th April 2019.

Khadilkar, H., and H. Balakrishnan. 2012. "Estimation of Aircraft Taxi Fuel Burn Using Flight Data Recorder Archives". *Transportation Research Part D: Transport and Environment* 17(7):532–537.

Li, N., and H. Zhang. 2017. "Calculating Aircraft Pollutant Emissions during Taxiing at the Airport". *Huanjing Kexue Xuebao/Acta Scientiae Circumstantiae* 37(5):1872–1876.

Mujica Mota, M., and I. Flores. 2019. "Model-Based Methodology for Identifying Capacity Limitations in Airports: Mexico City Airport Case". *SNE Simulation Notes Europe* 29(1):13–25.

Nikoleris, T., G. Gupta, and M. Kistler. 2011. "Detailed Estimation of Fuel Consumption and Emissions during Aircraft Taxi Operations at Dallas/Fort Worth International Airport". *Transportation Research Part D: Transport and Environment* 16(4):302–308.

SCT 2017. Capacidad de Operaciones (Capacity of Operations). https://www.aicm.com.mx/wp-content/uploads/2017/12/OF1747CapacidadDeOperaciones.pdf, accessed 11th December 2018.

SCT 2019. Estadística Operacional de Aerolíneas (Traffic Statistics by Airline) - International Airport of Mexico City. http://www.sct.gob.mx/transporte-y-medicina-preventiva/aeronautica-civil/7-inconformidades-de-aerolineas/71-demorasindice-de-puntualidad/aeropuerto-internacional-de-la-ciudad-de-mexico-aicm/, accessed 15th November 2019.

SEDEMA 2018. Inventario de Emisiones de La CDMX 2016 (Mexico City Emissions Inventory 2016). http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisiones-2016/mobile/, accessed 20th April 2020.

Simaiakis, I., and H. Balakrishnan. 2016. "A Queuing Model of the Airport Departure Process". *Transportation Science* 50(1):94–109.

Simaiakis, I., H. Khadilkar, H. Balakrishnan, T. G. Reynolds, and R. J. Hansman. 2014. "Demonstration of Reduced Airport Congestion through Pushback Rate Control". *Transportation Research Part A: Policy and Practice* 66(1):251–267.

Tsao, H. S. J., W. Wei, A. Pratama, and J. R. Tsao. 2009. "Integrated Taxiing and Take-Off Scheduling for Optimization of Airport Surface Operations". In *Proceedings of Indian Subcontinent Decision Sciences Institute Conference 2009*, January 3rd-5th, Bombay, India, 1–12.

Yan, S., and C. Chang. 1998. "A Network Model for Gate Assignment". *Journal of Advanced Transportation* 32:176-189.

Yan, S., and C. M. Huo. 2001. "Optimization of Multiple Objective Gate Assignments". *Transportation Research Part A: Policy and Practice* 35(5):413–432.

Yan, S., C.-Y. Shieh, and M. Chen. 2002. "A Simulation Framework for Evaluating Airport Gate Assignments". *Transportation Research Part A: Policy and Practice* 36(10):885–898.

AUTHOR BIOGRAPHIES

MARGARITA BAGAMANOVA is a lecturer and researcher at the Amsterdam University of Applied Sciences and a Ph.D. student at the Aeronautical & Logistics departmental unit, the Engineering School at the Autonomous University of Barcelona. Her research is focused on the application of simulation, multivariate optimization, and Bayesian inference techniques for the improvement of airport operations. She has publications in different peer-reviewed international journals and international conferences proceedings. Furthermore, she is an active member of the Dutch Benelux Simulation Society (EUROSIM) Steering Committee since 2018. Her email addresses are mm.bagamanova@hva.nl and margaritabagamanova@gmail.com.

MIGUEL MUJICA MOTA is an associate Professor in Aviation Management and senior researcher at the Aviation Academy of the Amsterdam University of Applied Sciences in the Netherlands. He holds a Ph.D. and an MSc. in informatics from the Autonomous University of Barcelona and a Ph.D. and MSc. in operations research from the National University of Mexico. He is the current president of EUROSIM the Federation of Simulation Societies in Europe and the Chair of the Dutch Benelux Simulation Society. He has participated in several international projects for industry and research projects funded by the European Commission, World Bank, or private industry. He is the co-author of four books and numerous scientific papers on simulation, operations research, aviation, manufacturing, and logistics. His research interests lie in the use of simulation, modeling formalisms, and heuristics for the optimization and performance analysis of aeronautical operations, manufacture, and logistics. His e-mail address is m.mujica.mota@hva.nl. His website is www.mmujicamota.com.