Issue 25(1), 2025, pp. 1-23
ISSN: 1567-7141
TU Delft OPEN

EJTIR

SI: Sesar Innovation Days 2023
https://ejtir.tudelft.nl
DOI: 10.59490/ejtir.2025.25.1.7459

Modelling and Analysis of Autonomous Airport Surface Movement Operations based on Multi-Agent Planning

Malte von der Burg[1], Alexei Sharpanskykh[2]

[1] corresponding author M.F.vonderBurg@tudelft.nl, Faculty of Aerospace Engineering, Delft University of Technology, The Netherlands; https://orcid.org/0000-0002-0190-3274
[2] Faculty of Aerospace Engineering, Delft University of Technology, The Netherlands; https://orcid.org/0000-0001-6879-1134

Keywords
Multi-agent system
Multi-agent motion planning
Autonomous airport operations
Airport surface movement operations
Automation
Air traffic control

Publishing history
Submitted: 29 March 2024
Revised date(s): 05 August 2024, 14 October 2024
Accepted: 14 October 2024
Published: 09 January 2025

Cite as
von der Burg, M., & Sharpanskykh, A. (2025). Modelling and Analysis of Autonomous Airport Surface Movement Operations based on Multi-Agent Planning. *European Journal of Transport and Infrastructure Research, 25*(1), 1-23.

Abstract

Both EASA and SESAR JU define a vision and roadmap towards an autonomous air traffic management system. Furthermore, past and ongoing SESAR JU projects investigate how to increase the efficiency and predictability of current operations by means of automation. In this paper, we explore the operational implications that result from fully-automated airport surface movement operations modelled with high realism. A hierarchical multi-agent system model was developed to coordinate and control all movements on the airport surface. It comprises the Airport Operations Agent to handle the flight schedule and runway configuration, the Routing Agent to compute conflict-free trajectories, and the Guidance Agents to instruct and monitor the Aircraft Agents while these execute the planned routes. The model incorporates the decisive processes and elements of airport surface movement operations such as pushback, engine-start, inbound and outbound holding, compliance to CTOT-slots, and wake turbulence separation for takeoffs. To compute conflict-free trajectories for all taxiing agents, we tailored and extended state-of-the-art multi-agent motion planning algorithms: the two-level routing algorithm combines Priority-Based Search (PBS) with Safe Interval Path Planning (SIPP). We defined different sizes of aircraft, accounted for a minimal safety distance between them, and calibrated their speed limits in curves with historic ADS-B data. Using the real-world flight schedules of two of the busiest days at Amsterdam Airport Schiphol, including different runway configurations, we examine the performance of the autonomous taxiing system with respect to the historic operations. For the considered simulation conditions, we show that the MAS yields 30% lower taxi times that vary less and are more predictable and increases runway capacity.

1 Introduction

The air traffic demand is predicted to exceed 10 billion yearly passengers by 2050 (IATA, 2021), more than twice the amount of 2019. However, it is expected that infrastructural expansions of airports are insufficient to facilitate this growth (Eurocontrol, 2018). Therefore, large airports are facing challenges to improve the efficiency of their operations and, on top of that, to reduce their environmental footprint to achieve the industry-wide goal of net-zero emissions by 2050 (IATA, 2021).

When the congestion at airports increases, the taxi time of an aircraft, i.e. the time that it travels over the airport surface from runway to gate or vice versa, becomes harder to predict. This may affect the respective flight, but may also lead to network-wide knock-on effects (Eurocontrol, 2021). Moreover, when Air Traffic Control Officers (ATCOs) have to handle more and potentially less predictable traffic, their workload is amplified (Chua et al., 2017). Consequently, the taxiing operations may become less efficient.

To deal with these issues, previous and ongoing SESAR projects considered how to increase the efficiency and predictability of taxiing operations through automation in general (MOTO, 2016; TaCo, 2017). Other projects examined more specifically how to reduce emissions by integrating engine-off taxiing techniques (AEON, 2021), or how to enable human-automation teamwork in the operations through higher levels of automation (ASTAIR, 2023). Moreover, different aspects of airport surface movement operations (ASM Ops) were studied in previous work: for example, (Morris et al., 2016; Roling & Visser, 2008) studied automation within ASM Ops, (Atkin et al., 2010) reviewed optimization approaches, (Weiszer et al., 2020; Zhang et al., 2019) investigated how to create optimal trajectories under multiple objectives, and (Liu et al., 2014) assessed metrics to evaluate predictability when automating airport surface movements.

In practice, the Advanced-Surface Movement Guidance and Control System (A-SMGCS) provides specifications for four services to increase efficiency through the use of automation: the surveillance service to track vehicles, the airport safety support service to alert controllers of potential conflicts, the routing service to determine conflict-free trajectories, and the guidance service to guide the vehicles during taxiing (ICAO, 2004; Lane et al., 2020). Moreover, towards 2050 and beyond, both EASA and SESAR JU define a vision and roadmap to eventually reach autonomous air traffic management, i.e. level 3 in EASA's AI roadmap (AI Roadmap 2.0, 2023), or level 4-5 in the roadmap of SESAR JU (Automation in ATM, 2020). However, the challenges to achieve this long-term vision are manifold. For instance, the role of the human during and beyond the transition as well as the implications of such fully-automated operations remain largely unknown.

In this paper, we explore which operational consequences may result from autonomous surface movement operations at large airports: in terms of efficiency, predictability, delays on taxiways, and runway capacity, while sustaining safety levels. We represent such operations as multi-agent system (MAS) model that plans conflict-free routes for all aircraft on the ground and controls their execution. Multi-agent systems modelling and simulation allow for inherent modularity, flexibility, and expressiveness of a system's structure and dynamics. Both heterogeneous agent properties and different types of interactions between actors as well as randomness can be integrated into an agent-based model (Helbing & Balietti, 2015).

However, to evaluate the operational impact of such a new operational concept, the model must represent the airport surface movement operations realistically. As pushback and engine-start operations considerably effect delays in planning systems (Stergianos et al., 2016), we explicitly model these processes in the MAS control model. Furthermore, we include inbound and outbound holding as these can considerably affect the taxi times of the respective flights: some arriving aircraft must wait at a remote holding location until their stand is free, while some departing aircraft must be delayed to comply with the Calculated Take-Off Time (CTOT) issued by Eurocontrol. Moreover, the taxi times are strongly influenced by the chosen agent kinematics as pilots speed up on straight segments, and slow down in front of curved taxiways. To this end, we

define a maximal and minimal velocity, and use speed limits along curves dependent on the radius of curvature that we calibrate using historic ADS-B data. To the best of our knowledge, this is the first study to use such detailed kinematics to compute trajectories of taxiing aircraft. We thus provide an overview of values found in the literature when describing the model calibration in Section 2.4.

Airport surface movement operations must be safe at all times. To plan conflict-free routes for a set of agents, many different multi-agent path finding algorithms have been developed that model agents as a point (Stern & Sturtevant, 2019). However, to include agent shapes and kinematics in path planning, the domain of multi-agent motion planning (MAMP) offers more suitable algorithmic concepts (Cohen et al., 2019). We combine and extend such state-of-the-art MAMP algorithms to form a routing algorithm that addresses the requirements of airport surface movement operations. Section 2.3 summarises how it computes conflict-free trajectories for all aircraft.

With an implementation of the MAS model in Python, we simulate the autonomous taxiing operations using the flight schedule of two of the busiest days at Amsterdam Airport Schiphol to date. Section 3 outlines the experimental setup. We then analyse and discuss key performance indicators in relation to their historic counterparts in Section 4, list directions for future work in Section 5, and end with concluding remarks in Section 6.

2 Multi-agent system model

The developed multi-agent system (MAS) model for autonomous aircraft taxiing operations has a distributed-hierarchical structure of both centralized and distributed agents, which is illustrated in Fig. 1. The centralized Airport Operations Agent defines and updates the flight schedule and runway configuration, the centralized Routing Agent plans conflict-free trajectories for all Aircraft Agents which are instructed and monitored by distributed Guidance Agents while executing their planned routes. For the autonomous operations considered in this paper, it is assumed that the full control and decision making is done by the agents. Furthermore, we assume that digital means of communication via a datalink such as AeroMACS as well as the surveillance service of the A-SMGCS specification are fully operational.

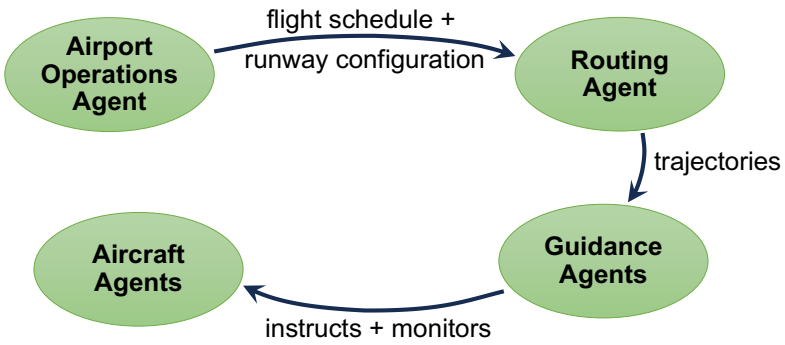

Figure 1. Overview of multi-agent system for autonomous airport surface movement operations

2.1 Model specification

The airport taxiing infrastructure is represented by a graph G = (V, E) comprising vertices V and directional edges E. As example, the layout of Amsterdam Airport Schiphol, which is also used in the simulations presented in this paper, is shown in Fig. 2. Vertices denote aircraft stands (green), taxiway intersections (black), holding points (orange), or stopbars in front of runway entries (red). Each bidirectional taxiway segment between two vertices is constructed from two unidirectional

edges that connect the vertices. Taxiway edges (black) are constructed using Bezier-curves that closely match the taxiway centrelines from a satellite image of Schiphol[1].

The Airport Operations Agent schedules all flights, and updates them whenever new predictions of the underlying A-CDM milestones are available. When the allocated stand of an arriving flight is still occupied by a departing aircraft, or Eurocontrol issued a Calculated Take-Off Time (CTOT) for a departing aircraft, the agent marks the corresponding flight. Such flights are subject to special routes assigned by the Routing Agent to account for the necessary holding, detour, or prioritization during taxiing, as outlined in Section 2.2. Furthermore, the Airport Operations Agent defines the runways in use, i.e. the runway mode of operation (RMO). Active runways and the resulting flight path of arriving or departing flights must not be crossed. Thus, the Airport Operations Agent blocks such taxiway segments by setting layout constraints on them. This mechanism is also applicable for taxiway segments that are temporarily unavailable.

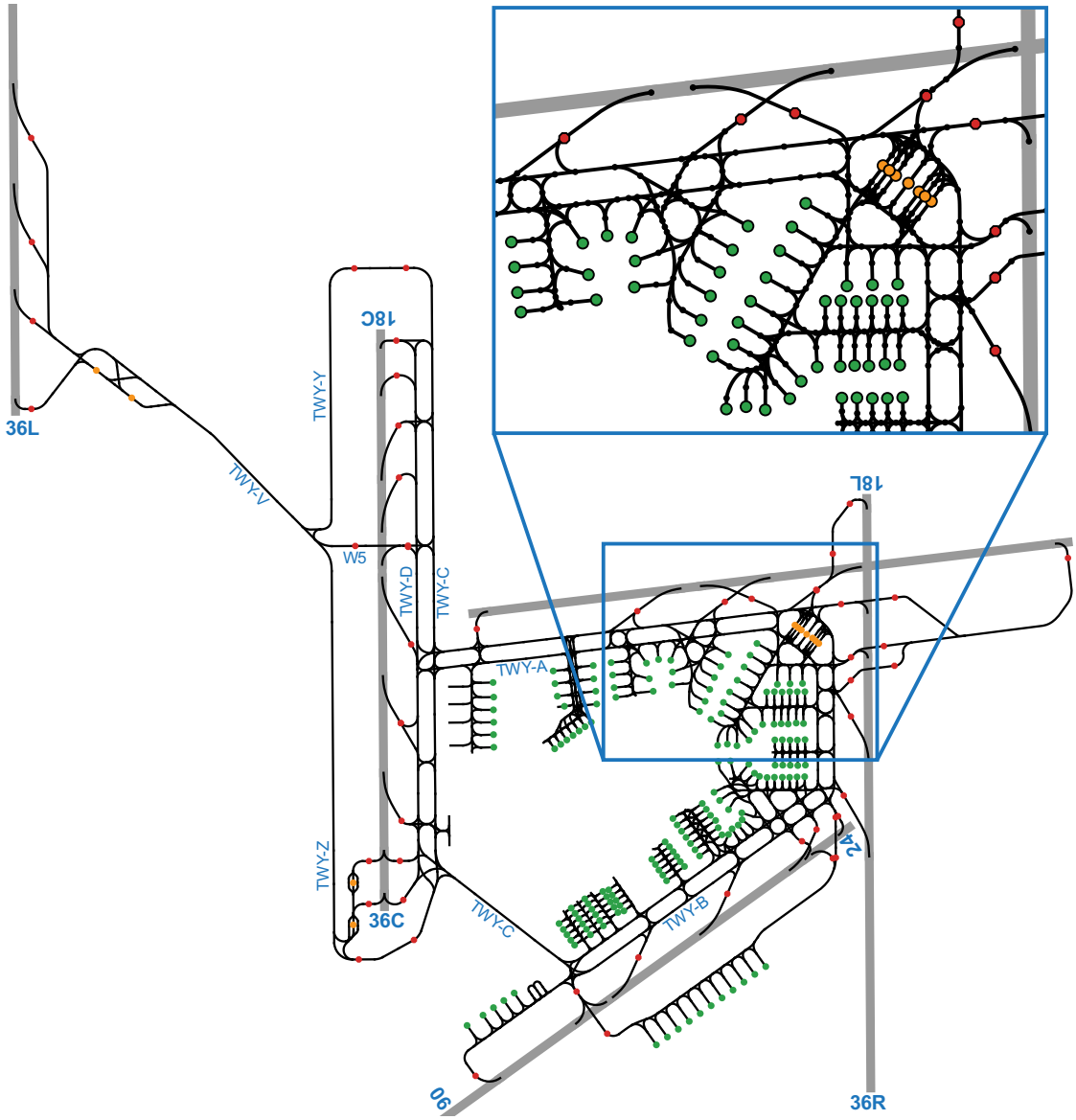

Figure 2. Graph of layout of Amsterdam Airport Schiphol with edges for runways (grey) or taxiways (black), and vertices for taxiway intersections (black), stopbars (red), holding points (orange), or gates (green)

[1] The Bezier-curves are constructed using the X-Plane WorldEditor WED 2.5 (X-Plane, 2024).

Both the flight schedule and constraints are shared with the Routing Agent that computes conflict-free routes for all taxiing aircraft within the upcoming planning window w_{plng}. It re-computes the routing plans when it receives updates from the Airport Operations Agent, or latest after the replanning period[2] h_{plng} has passed. We use motion planning to account for vehicle kinematics and shapes in planning. To ensure conflict-free paths, we deploy a two-level search based on Priority-Based Search (PBS) (Ma et al., 2019) with an augmented version of the Safe Interval Path Planning (SIPP) algorithm (Phillips & Likhachev, 2011). This routing algorithm is presented in Section 2.3, and we describe how we calibrate the model in Section 2.4.

The resulting trajectories are sent to the Guidance Agents which are positioned at every intersection in the taxiway system. Each Guidance Agent controls those Aircraft Agents that are moving towards its location. It instructs them to execute the next part of the planned trajectories, and monitors that the instructions are carried out accordingly. To do so, the Guidance Agents use the airport radar, which reports the position, speed, and heading of all Aircraft Agents while they move over the airport surface. In case the executed movements deviate from the planned routes, the Guidance Agents locally adjust the trajectories to minimize these deviations. However, when the impact becomes too extensive, they request central replanning from the Routing Agent. Once one of the Aircraft Agents has passed the location of a Guidance Agent, it passes the guidance responsibility for that aircraft to the next Guidance Agent along the aircraft's route.

Aircraft Agents represent the aircraft (auto-)pilots and are modelled to be fully cooperative: they thus carry out the instructions as accurately as possible. To account for the different sizes of aircraft, all flights are categorized as one of the 6 aircraft types from the ICAO aerodrome reference codes (ICAO, 2016). They are assumed to have a circular shape with a pre-defined radius according to the type. Table 2b in Section 3 lists these shape-radii.

When planning the trajectories, a safety zone is added around all agents. To this end, we define a general safety distance, as well as a safety distance that an agent has to keep when it is trailing another aircraft. Both safety measures are defined in relation to the shape radii of the corresponding pair of agents. Moreover, two aircraft that consecutively take off from the same runway must have a minimal separation to mitigate the wake turbulence of the preceding aircraft. We use the time-based separation minima from RECAT-EU for that (Rooseleer & Treve, 2018).

2.2 Activity sequence of Aircraft Agents

To take the various surface movement operations into account during path planning, the route of an Aircraft Agent is expressed as a combination of the following three activities:

- Go-to activities have one start vertex and a set of goal vertices. Thus, the routing algorithm gets two degrees of freedom: the path between the vertices, and the time to traverse this path. The regular taxiing between one point to another point at the airport is an exemplary go-to activity.
- Follow activities comprise a predefined ordered list of edges that must be part of the route. Therefore, during routing, time is the only remaining variable as the path cannot be changed. Pushback and push-pull manoeuvres of departing aircraft are examples of such.
- Wait activities define a vertex at which an agent has to wait for a fixed duration. For instance, a wait activity is used to specify the place at which the pushback-truck is decoupled from the aircraft, or the necessary direction-switch of the push-pull manoeuvre within the pushback operations occurs.

Using a combination of these activities, the Routing Agent defines an activity sequence for both departing and arriving aircraft, as depicted in Fig. 3.

[2] The notation of w_{plng} and h_{plng} is commonly used in MAPF literature, see for example (Li et al., 2021).

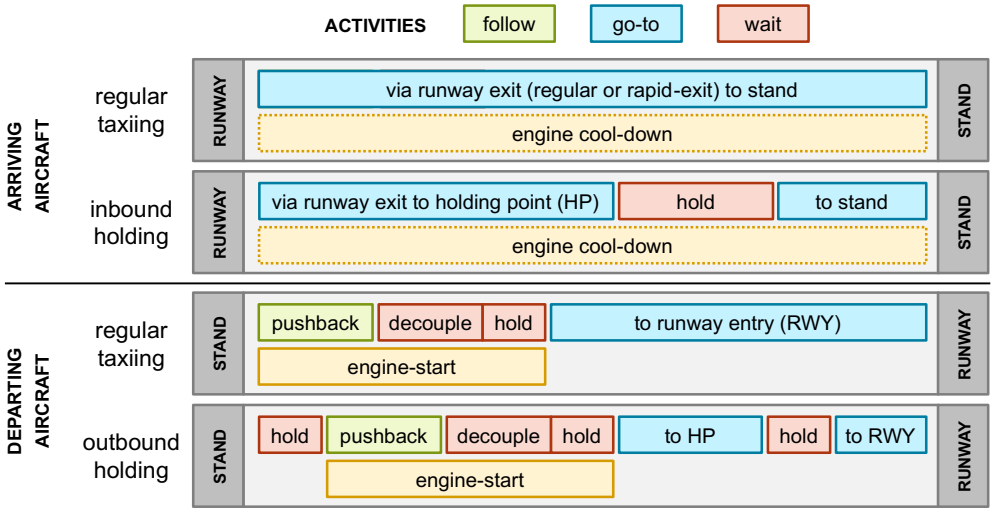

Figure 3. Activity sequence for regular taxiing of arriving and departing aircraft as well as inbound and outbound holding. While engine-start (orange box) of departing aircraft is accounted for, engine cool-down (dotted orange box) is neglected.

1. Engine warmup and cooldown: In the sequence, the warmup and cooldown of the engines represent special cases. The routing algorithm takes the warmup-phase as part of the engine-start manoeuvre and on basis of the aircraft-specific engine-start duration as input value into account. Therefore, if this duration exceeds the time needed till decoupling from the pushback-truck, additional waiting in form of holding is added to the route. We do not model engine cooldown, as it does not have an influence on the routing regarding the kinematics, since the engines are switched off after standstill at the gate.
2. Inbound holding: When an aircraft arrives at the airport, but its stand is still occupied by a departing flight, the Routing Agent has three options to resolve the anticipated stand-conflict: for long conflict durations (case 1), it sends the arriving flight to the remote holding platform (see Fig. 2). Otherwise, it defines a detour along the taxiways (case 2), or reduces the agent's taxi speed for short conflicts (case 3). To this end, the Routing Agent first calculates the single-agent route directly to the stand, i.e. the trajectory without accounting for other aircraft agents, to estimate the severity of the stand-conflict. Then, it computes a single-agent trajectory via the remote holding points. When this detour is insufficient to resolve the stand-conflict, the Routing Agent assigns the remaining time as remote holding duration (case 1), and updates the agent's activity sequence accordingly. In contrast, when the taxi duration now exceeds the time at which the departing aircraft has cleared the stand (case 3), it keeps the original activity sequence of the agent.
3. Outbound holding to comply with CTOT-slots: Similar to inbound holding, the Routing Agent deals with departing flights for which Eurocontrol issued Computed Take-Off Times (CTOT-slots). However, as long as no arriving flight requires the stand, it assigns a holding duration at the agent's stand so that the agent arrives at the runway at the beginning of the CTOT-slot. In case an arriving flight is scheduled for the stand, the Routing Agent sends the departing flight to a remote holding location close to the scheduled runway. It updates the activity sequence of the departing flight accordingly.

2.3 *Routing algorithm*

The Routing Agent carries out multi-agent motion planning for all Aircraft Agents that taxi within the planning window. This two-level routing algorithm uses a low-level search to calculate individual trajectories per aircraft, and coordinates all agents in its high-level search to yield conflict-free trajectories. For the low-level, we extended the Safe Interval Path Planning (SIPP) algorithm (Phillips & Likhachev, 2011), and adapted the Priority-Based Search (PBS) algorithm (Ma et al., 2019) to serve as high-level solver.

PBS carries out multi-agent coordination by assigning a priority order between agents to deconflict their space-time trajectories. To this end, PBS constructs a binary priority tree: first, individual agent trajectories are planned by the low-level search, and are stored as the PBS root-node along with the sum-of-cost of all agent trajectories. We define the cost of a trajectory as sum of the taxiing duration and travelled distance. Then, all conflicts between the trajectories of each agent pair are detected, and one of these conflicting agent pairs is selected (we use the pair of the first conflict in time). To resolve this conflict, a priority-relation between the conflicting pair of agents is established by creating two PBS child-nodes. In each, one of the agents is given priority, and the other agent must give way along its entire trajectory. Per child-node, the low-level solver is invoked as further outlined below, returning a new trajectory with an updated cost. PBS then continues by following a depth-first search: between the two child-nodes, the one with the lowest sum-of-cost is picked. Any remaining conflicts between those agents that do not yet form a priority-relation with each other are detected, and the conflict selection and resolution scheme is repeated. In case no conflicts remain, PBS can terminate as the established priority order results in conflict-free trajectories, which are returned as the path planning solution.

In the low-level search, the route of a deprioritized agent has to be adapted, either by changing its path or altering the speed profile along the path. To this end, we translate all paths into a set of graph reservations: an aircraft temporarily blocks a set of edges during each movement between one vertex and another. The blockage times and set of blocked edges are dependent on the agent's shape, velocity profile, the shapes of other agents, and the safety zone between the shapes.

The SIPP algorithm represents moving obstacles as collision intervals and subsequently defines a set of Safe Intervals (SIs) per graph location, representing time intervals during which an agent can occupy that location. Furthermore, states are defined on vertices and motion profiles with piecewise constant acceleration map the trajectory between states. We augmented SIPP to facilitate the activity sequence of an aircraft as defined by the Routing Agent, and to take the travelling direction as well as the kinematic agent properties into account. Additionally, we use SIs also on edges to deal with the reservations of agents higher in priority.

In the motion generation, we are bound to the agent's kinematic properties for the current activity and the velocity in the current state. A motion that is part of the follow-activity for pushback is for example constrained by a lower maximum speed than regular taxiing in a go-to activity. In addition, vehicles that have maximum velocity in the current state, might not be able to decelerate enough to satisfy a reservation on the next edge or vertex. In this case, it might be required to start decelerating on the edge before the current state. To efficiently account for this, we anticipate based on the agent's current velocity, braking distance, and reservations or velocity restrictions within the braking distance.

2.4 Model calibration

In the MAS model, the agents' motions during route planning are modelled based on constant longitudinal acceleration/deceleration and do not account for slip, i.e. are steady-state motions. Nonetheless, to the best of our knowledge, this is the first study to use such detailed kinematics to compute trajectories of taxiing aircraft. In the following, we thus include an overview of related values found in the literature.

We define a general speed limit of 15 m/s in line with the design taxi speed given in the A-SMGCS manual from ICAO (2004). Except for the dedicated wait-locations, agents must taxi at least with the minimal velocity of 1.5 m/s to avoid stop-and-go during taxiing. For curved taxiway segments, the ICAO manual mentions that speeds up to 10 m/s may occur. Most previous studies on airport surface movement operations define curved segments as turns with a maximal velocity of 5 m/s (Bakowski et al., 2015; Chen et al., 2015; Udluft, 2017). Since we model taxiway curves explicitly through Bezier-curves (see Section 2.1), we define a speed limit v_{curve} per edge by using

$$v_{curve} = \sqrt{a_{lat} * r_{curve}} \quad (1)$$

with the lateral acceleration a_{lat} and the radius of curvature r_{curve} of the respective edge. To obtain r_curve, we use the median value of all curvatures per 1 m-segment of the underlying Bezier-curve. For passenger comfort in public transport, Bae et al. (2022) provide a range for both longitudinal and lateral accelerations of ±0.9 m/s². Furthermore, they claim that a car driver with a normal driving style experiences a lateral acceleration of up to ±4 m/s² and a longitudinal acceleration of -2 m/s² to 1.47 m/s². In contrast, De Winkel et al. (2023) found in empirical studies that the acceptable limits for passenger comfort are 1.23 m/s2 for longitudinal and 0.98 m/s² for lateral acceleration. As noted above, previous studies did not consider lateral accelerations to define turn speeds. For longitudinal acceleration/deceleration, different values are reported: for example, ±0.98 m/s² (Bakowski et al., 2015; Chen et al., 2015), or 0.26 m/s² as acceleration and -0.8 m/s² as deceleration (Udluft, 2017).

To find realistic values for the longitudinal and lateral accelerations of taxiing aircraft, we use historic track data from Schiphol captured by ADS-B receivers that record the aircraft positions during taxiing with a rate of 1 Hz. To this end, we map the positions onto the graph representing the taxiway centrelines, and smooth the resulting trajectories with a Savitzky-Golay filter (window length of 11 s, linear polynomial). This yields the travelled taxi distance along the graph edges as well as the speed and acceleration at each time point of the trajectory. However, we only use the data on the edges that correspond to the main taxiways: while the tracks become too noisy in the bay areas and at aircraft stands, the accelerations on runways for takeoff and landing are not representative of those experienced during taxiing.

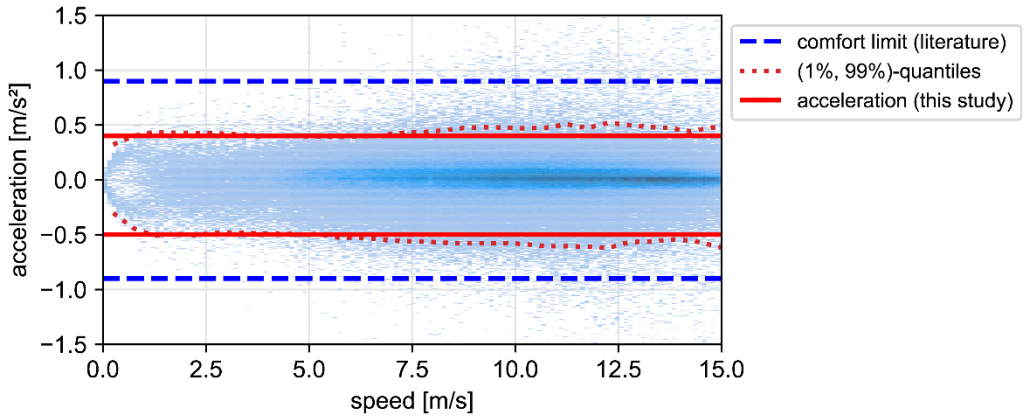

Figure 4. Calibration of longitudinal acceleration and deceleration values with historic track data. Comfort limit obtained from (Bae et al., 2022)

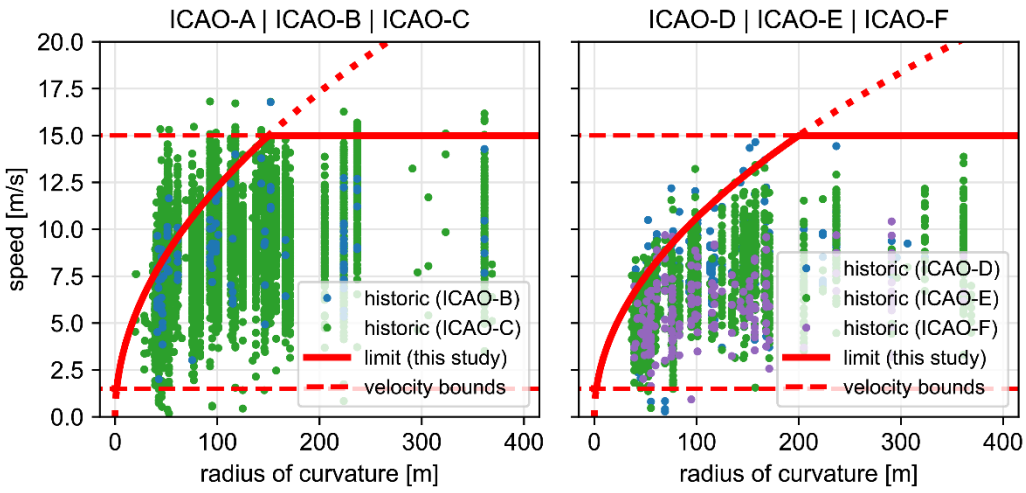

Figure 5. Calibration of curve speed with historic track data

In Fig. 4, the acceleration over velocity of each data point is visualized as 2d-histogram. The 1% and 99% percentile lines of the acceleration values per 0.5 m/s step show that the longitudinal acceleration/deceleration remain similar across different taxi speeds. Therefore, we set the acceleration to 0.4 m/s² and deceleration to -0.5 m/s² independent of an agent's speed. While these values seem low compared to those mentioned in the literature, we argue that using these in planning increases the flexibility during execution: the Guidance Agents have more options to locally adjust the trajectories if necessary (see Section 2.1).

Fig. 5 visualizes the historic curve speeds of different aircraft types as average speed along a curved edge with radius r. The average speed is calculated as $\bar{v} = \Delta d/\Delta t$ with the time difference Δt and travelled distance Δd of the data points per edge along each trajectory. Although higher curve speeds exist, we define the speed limit in curves based on a lateral acceleration of 1.5 m/s² for small aircraft (left plot in Fig. 5) and 1.125 m/s² for large aircraft (right plot). Using Eq. (1), we assign a speed limit per edge within the velocity bounds of 1.5 m/s to 15 m/s.

Table 1 summarizes the kinematic values and lists the main algorithmic parameters used by the routing algorithm. In general, two aircraft agents have to keep a minimal safety distance between them equal to the average of their shape radii. However, when an aircraft is trailing another agent, it has to keep a safety distance of at least 3-times the shape radius of the preceding aircraft, which is in accordance with experts. The planning window w_{plng} and replanning period h_{plng} are provided as ranges with the requirement that $h_{plng} < w_{plng}$.

Table 1. Kinematic and algorithm parameters that are used in the routing algorithm

parameter	value	unit
maximal speed v_{max}	15	m/s
minimal speed v_{min}	1.5	m/s
curve speed v_{curve}	per edge	m/s
acceleration acc	0.4	m/s²
deceleration dec	-0.5	m/s²
safety distance in general	1	averaged shape radius
safety distance trailing	3	shape radii of preceding aircraft
planning window w_{plng}	15 to 60	min
replanning period h_{plng}	3 to 48	min (= 20 % to 80 % of w_{plng})

2.5 Verification and validation

Verification and validation of the simulation model were performed in accordance with validation techniques and tests as described by Sargent (Sargent, 2011). To validate the conceptual model, operational experts from Amsterdam Airport Schiphol were consulted. During implementation, continuous verification was performed. The model was developed in different modules, allowing for the independent testing of the building blocks. In addition, assertion conditions were added to ensure correctness of the internal processes in the code and compiler errors were resolved. With visual animations, we verified that the routes were executed as planned. Furthermore, small test scenarios were created to verify the model's behaviour in the bay areas for pushback, push-pull, and engine start manoeuvres. The activity-based path planning was verified with small test scenarios ensuring correct timings and kinematics. With these scenarios, face validation was performed to ensure that the model performance was as expected. Finally, individual agent behaviour was carefully followed throughout the system to ensure correctness. As safety must not be compromised, we confirmed that indeed no collisions between agents occurred during the execution of the planned conflict-free routes.

As validation of the calibrated agent kinematics, Fig. 6 shows an exemplary historic speed profile over travelled distance (black line) as well as simulated single-agent trajectories along the same path for both the un-calibrated agent kinematics used in (von der Burg & Sharpanskykh, 2023) (green line) and the calibrated ones of this work (orange line). Especially on the long straight

taxiway segments towards runway 36L, the historic trajectories indicate that aircraft often taxi faster than the velocity limit of 15 m/s.

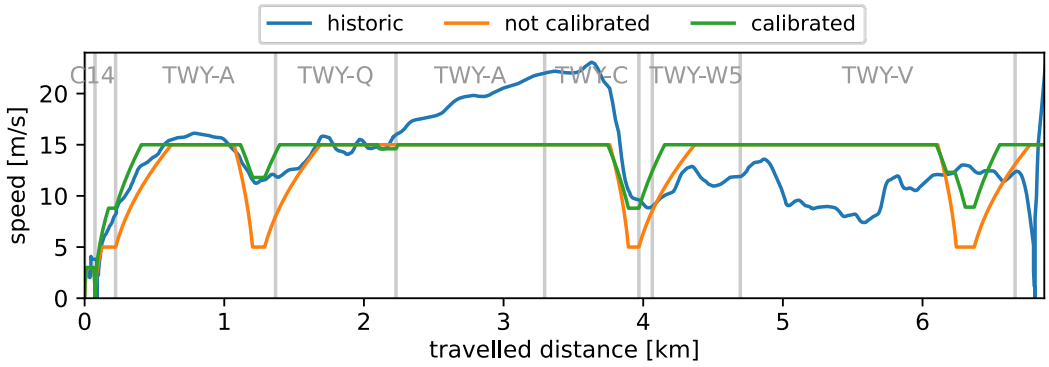

Figure 6. Exemplary historic, un-calibrated, and calibrated speed profiles of a departing flight travelling from stand C14 to runway 36L

3 Experimental setup

In this section, we present the experimental setup to simulate the flight schedules on 17th and 18th July 2019 as two of the most busiest days to date at Amsterdam Airport Schiphol. To this end, we outline additional assumptions for our study and give an overview of the traffic situation on these two days. Furthermore, we describe the analyses and key performance indicators to evaluate the study.

3.1 Study assumptions

We draw the following additional assumptions with respect to the concept of the control architecture outlined above:

- the final flight schedule of the two operational days is used and remains static throughout the simulation
- arriving aircraft are spawned at their historic Actual Landing Time (ALDT); their velocity at rapid-exit taxiways $v_{RET} = v_{max}$, or at regular exits $v_{exit} = v_{curve}$
- departing aircraft are spawned at their historic Actual Off-Block Time (AOBT), but are allowed to hold at the stand; they use the standard pushback path of the stand according to the airport manuals (see (Schiphol - Standaard Pushback per Positie, n.d.))
- the routing algorithm can freely determine the takeoff sequence, but has to comply with the CTOT-slots of the historic data
- all vehicles execute the instructions from the Guidance Agents perfectly, i.e. no deviations to planned routes
- the simulation is executed sequentially, i.e. paused when routes are planned
- the Routing Algorithm must always return a solution for the simulation to continue, i.e. we do not consider any degradation modes for path planning
- the standard taxiway directions at Schiphol are ignored.

These assumptions decrease the complexity within the simulation, and give the routing algorithm more freedom to optimize, limiting the risk of an incomplete solution, i.e. situations in which the routing cannot be done. Moreover, they exclude any uncertainty or noise, making the simulations deterministic. As outlined below, the simulated scenarios are based on the most common runway configurations, and we use flight schedules from two of the busiest days at Schiphol to date. Despite the drawn assumptions, the simulated MAS model represents the airport surface movement operations in detail. We thus deem the conducted simulations suitable to study the key operational consequences when automating ASM Ops. Furthermore, we discuss the main aspects influencing the comparison between historic and simulated operations in Section 4.1.

3.2 *Traffic situation*

At Schiphol, two main runway mode of operations (RMOs) exist: **RMO North** (active on 17th July 2019), and **RMO South** (active on 18th July 2019). During each day, different runway combinations are set. These RMO phases are visualized in Fig. 7.

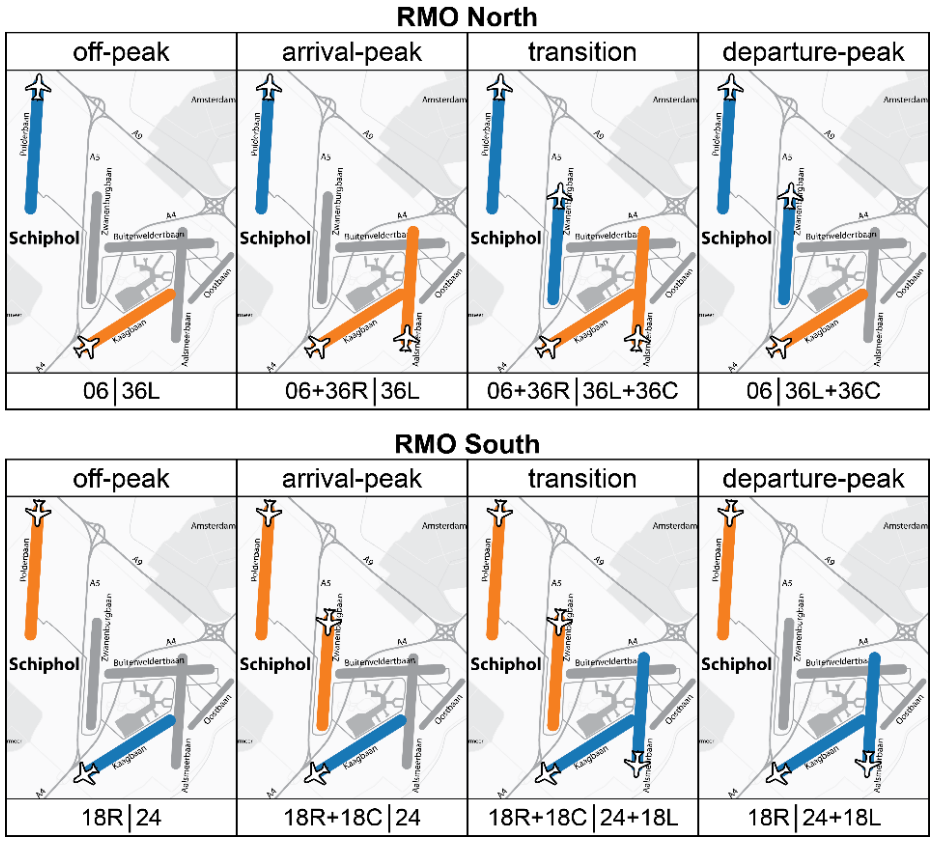

Figure 7. Runway mode of operations (RMO) at Schiphol: different phases of RMO North (top) and RMO South (bottom) with active runways for arrivals (orange) and departures (blue). Background map adapted from (NielsB, 2007).

In the simulation, the same runways are activated according to those that were active in the historic operations. The runway 09/27 was not active during the two days. Aircraft landing on or departing from runway 04/22 are not modelled as these general aviation flights remain foremost within Schiphol East. Table 2a lists the total number of flights, arrivals, departures, the main RMO, and the number of RMO phases for the two simulated days. As mentioned above, each aircraft is categorized as one of the six ICAO-types with an associated shape radius and wake turbulence category (WTC). Table 2b lists these parameters along with the count over the two days.

Table 2. Overview of (a) traffic data and (b) parameters and daily counts of ICAO-types

(a) Traffic data

date	17-07-2019	18-07-2019
flights	1489	1492
arrivals	745	744
departures	744	748
RMO	RMO North	RMO South
RMO phases	19	19

(b) ICAO-types. WTC = wake turbulence category

	parameters		count per day	
	shape [m]	WTC	17-07-2019	18-07-2019
ICAO-A	12	CAT-F	0	0
ICAO-B	25	CAT-E	22	20
ICAO-C	40	CAT-D	1195	1198
ICAO-D	54	CAT-C	37	43
ICAO-E	72	CAT-B	213	206
ICAO-F	80	CAT-A	22	25

3.3 *Overview of analysis and key performance indicators*

We carry out two types of analyses to explore the operational consequences of fully automated airport surface movement operations (ASM Ops), as outlined below. All conducted simulations use the historic flight schedules and RMOs defined above. The taxi time distributions form the basis of our analyses as they are directly linked to the efficiency, predictability, and resulting emissions of airport surface movement operations. Since the taxi times are not necessarily normally distributed, we report the averages, medians, and interquartile ranges. Furthermore, we exclude flights with a hold-type assigned by the Routing Agent due to the potential impact on their taxi times.

Analysis of algorithmic parameters to investigate the influence of the planning window w_{plng} and replanning period h_{plng} on the trade-off between:

1. Efficiency of operations: How do the parameters influence the distribution of taxi times? Intuitively, we expect that longer planning windows combined with shorter replanning periods yield lower and less varying taxi times (Hypothesis H1).
2. Predictability of taxi times: In the simulated operations, how predictable is the remaining taxi time when the planning window changes? To this end, we define the predictability of the taxi time as the actual vs. predicted value with respect to the remaining time till the end of taxiing: recall that per planning round, the MAS plans the routes for all aircraft that are scheduled to start within the planning window and those that are already taxiing. Conflicts with other agents that occur beyond w_{plng} are ignored. This yields a predicted taxi time that is updated in the subsequent planning round, which takes place latest after h_{plng} has passed. We hypothesize that longer planning windows combined with more frequent plan updates (lower h_{plng}) yield more accurate predictions of the remaining taxi time (Hypothesis H2).
3. Computational efficiency of routing algorithm: How long does it take to coordinate the agents per planning round with respect to different sets of w_{plng} and h_{plng}? Intuitively, we expect that longer planning windows require more computational time to reach a conflict-free solution as more agents need to be coordinated (Hypothesis H3).

To this end, we simulate the operations with different values of w_{plng} and h_{plng} within the ranges specified in Table 1. We analyse the first part with the taxi time distributions, the second with the deviations between predicted and actual taxi times, and the third with the computational time in relation to the number of agents per planning round. As outcome of the trade-off, we choose a set of parameters to conduct the following analysis.

Analysis of operational consequences of autonomous airport surface movement operations with respect to:

1. Delay hotspots in the simulated operations: In the routing algorithm, three mechanisms exist to let an agent avoid the reservations of other agents: an alternative path is chosen, its speed is reduced down to the minimal velocity of 1.5 m/s, or the agent has to hold at the stand or engine-start location. The latter two create additional taxi time that we sum per 5 m-segment of the graph underlying the airport layout. This yields the locations where most delays across all flights occur. Per simulated day of operations, we visually examine these congestion areas in the airport layout.
2. Efficiency gain on each of the two days: We compare the distribution of taxi times between historic and simulated operations, and summarize the efficiency gain as change in the key performance indicators related to the taxi time.
3. Predictability of simulated operations compared to the historic operations: we assess the predictability of the taxi times by defining metrics similar to those used by Liu et al. (2014):
    a. The predictability of takeoff times as difference between the Estimated taXi-Out Time (EXOT) and Actual taXi-Out Time (AXOT) of both historic and simulated

operations: to this end, we calculate the root-mean-square prediction error $RMSE$ and the inequality coefficient U as defined below.

b. Likewise, the predictability of in-block times as difference between the Estimated taXi-In Time (EXIT) and the Actual taXi-In Time (AXIT) of both historic and simulated operations.

The prediction error $RMSE$ is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(AXT_i - EXT_i)^2} \quad (2)$$

with the corresponding actual taxi time (AXT) and estimated taxi time (EXT) per flight i as well as the total number of flights n. The inequality coefficient U quantifies the relative prediction error and is calculated as

$$U = \frac{\sqrt{\sum_{i=1}^{n}(AXT_i - EXT_i)^2}}{\sqrt{\sum_{i=1}^{n}(AXT_i)^2}}. \quad (3)$$

When this ratio is zero, all predictions are perfect, while unity denotes that the $RMSE$ equals the error of always predicting a taxi time of zero.

4. Inbound and outbound holding: We assess how many arriving and departing aircraft must hold, and the effect on the taxi times. To this end, we compare the taxi time distributions between historic and simulated operations of all flights that must hold during taxiing.
5. Compliance with CTOT-slots: We analyse the compliance with CTOT-slots and the distribution of takeoff times within the CTOT window between historic and simulated operations.
6. Runway sequencing and usable capacity: We provide an exemplary takeoff order as comparison of the runway sequence between the historic and simulated operations. As indication of the runway use, we provide the maximal hourly throughput and occupancy rate of any departure or arrival runway. The occupancy rate is calculated as the relative time that the runway is blocked due to the minimal wake turbulence separation between two consecutive aircraft. Furthermore, we calculate the remaining takeoff slots per day that comply with the wake turbulence separation of the flights before and after. We compare the historic to the simulated slot count to estimate the surplus runway capacity.

4 Experimental results

To give an overview of the traffic situation, Fig. 8 shows the hourly count of all flights for both the historic (black dotted line) as well as simulated operations (grey line) over the two days of 17th and 18th July 2019. The two curves almost match each other, with the simulated operations showing a slightly lower total count due to the lower taxi times as discussed below. Furthermore, the chart visualizes the count of arriving vs. departing flights: the alternating trend between landings and takeoffs that is characteristic for a hub-and-spoke airport such as Schiphol is clearly visible. This is also reflected in the frequently changing RMO phases over the course of the two days, illustrated by the coloured shades in the figure.

Table 3 lists the mean, median, and interquartile range (IQR) of the taxi times for inbound and outbound flights. The underlying distributions exclude flights that hold explicitly during taxiing, i.e. arriving flights without a free gate and departing flights with an issued CTOT-slot exceeding the required taxi time. The distributions are further discussed in Section 4.2, while the excluded flights are analysed separately in Section 4.6. As predictability metrics, we list the prediction error $RMSE$ and inequality coefficient U, and refer back to these in Section 4.3. Moreover, for any runway in use, we report the maximal throughput and occupancy rate per hour. For arrivals, the indicators are identical between the historic and simulated operations as we used the actual landing time

(ALDT) as spawn-time in the simulation (see the assumptions in Section 3). For departures, the maximal hourly throughput is similar between historic and simulated operations, while the maximal hourly occupancy rate has increased for the simulated operations. Section 4.8 provides more details on the runway sequence and capacity. The listed CTOT-slot violations are further discussed in Section 4.7.

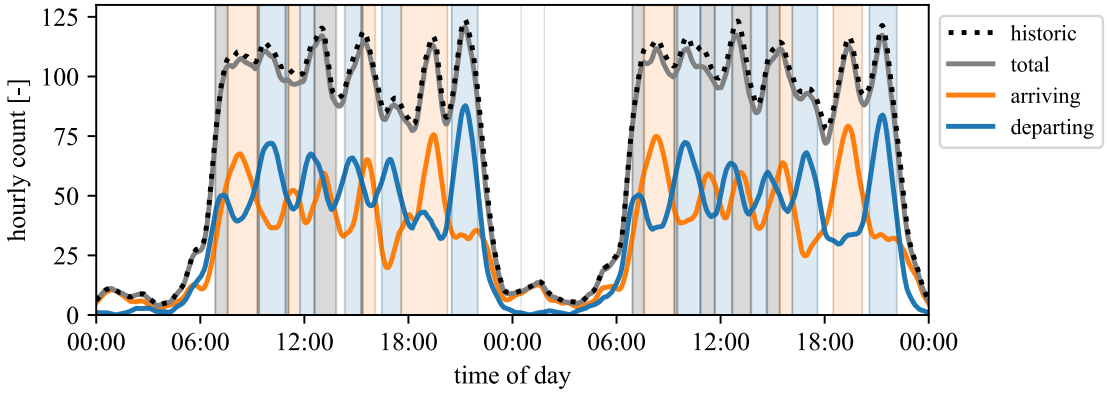

Figure 8. Hourly count of flights over the two days. Shades denote the RMO phase: off-peak (white), arrival-peak (orange), transition (grey), and departure-peak (blue)

Table 3. Comparison of historic and simulated operations with $w_{plng} = 20\text{min}$, $h_{plng} = 10\text{min}$

	date	17-07-2019		18-07-2019	
	operations	historic	simulated	historic	simulated
ARR	mean taxi time	04:38	03:03	10:15	07:40
	median taxi time	04:00	02:53	10:44	08:37
	IQR taxi time	02:51	01:40	05:59	05:13
	RMSE taxi time prediction	01:47	00:02	02:16	00:02
	U taxi time prediction	33.9%	1.1%	20.6%	0.5%
	RWY throughput*	42		38	
	RWY occupancy*	68.7%		69.3%	
DEP	mean taxi time	14:29	11:18	10:20	07:05
	median taxi time	14:32	11:20	10:06	06:34
	IQR taxi time	06:32	04:56	04:20	02:29
	RMSE taxi time prediction	02:45	00:40	02:44	00:35
	U taxi time prediction	18.1%	5.5%	25.2%	7.9%
	RWY throughput*	45	45	43	42
	RWY occupancy*	74.7%	76.1%	73.5%	77.0%
	CTOT-slot violations	5	3	0**	0

*: maximal hourly value for any runway
**: corrected after checking historic A-CDM milestones

4.1 Influencing factors on taxi time comparison

As visualized in Fig. 9, most A-CDM milestones of landings and takeoffs, i.e. Actual Landing Time (ALDT) and Actual Take-Off Time (ATOT), are located on the runway and are thus viewed as sufficiently precise. Likewise, most block-times, i.e. Actual In-Block Time (AIBT) of arriving and Actual Off-Block Time (AOBT) of departing flights, are fairly accurate. However, the tracks of some flights end (1.6 %) or start (4.9 %) further away from the respective stand. Especially for departing flights, an AOBT situated outside the bay area may lead to significantly lower taxi times: the pushback and engine-start procedures require multiple minutes but are likely not captured in the resulting historic taxi time.

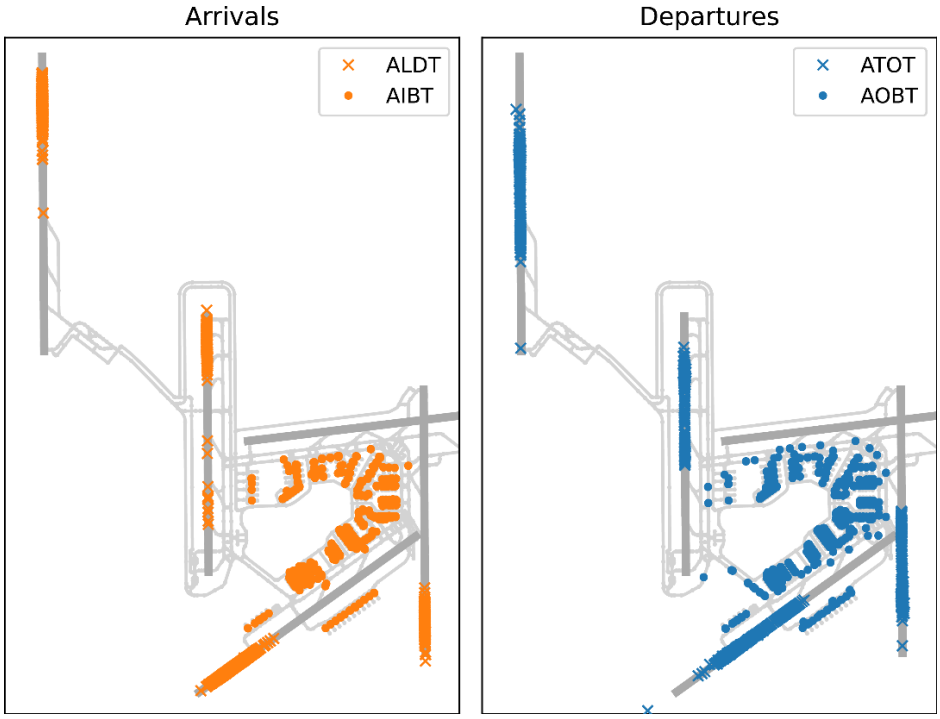

Figure 9. Aircraft positions at historic A-CDM milestones for inbound (left) and outbound flights (right)

Moreover, the historic tracks of departing flights often do not follow the standard pushback paths that the simulated flights must follow. Additionally, the actual engine-start duration cannot be extracted from the historic tracks. In the simulation, aircraft are only allowed to enter the runway when they can immediately take off, i.e. use a rolling takeoff. However, we do not model the takeoff, and instead use the time point of entering the runway as ATOT. While all of these factors introduce potential offsets between the historic and simulated taxi times, we believe that those leading to shorter historic times have a greater impact. Therefore, we did not attempt to correct the historic A-CDM milestones, but use them as such when comparing the taxi times in the following.

4.2 *Comparison of taxi time distributions per runway*

Fig. 10 displays the distributions of taxi times for each arrival and departure runway for the historic and different simulated operations as box-and-whisker plots. These represent the median, first and third quartiles as box, while outliers are marked by whiskers and points. Like Table 3, all distributions exclude flights that hold explicitly during taxiing, which are analysed separately in Section 4.6.

In general, the taxi times from the simulated operations are shorter and vary less. Since the runway 18R/36L is far away from the central part of Schiphol, taxiing to/from this runway takes more time than to any of the other runways. As the departing aircraft have to start their engines to taxi after pushback, their taxi time to any runway is in general longer than for aircraft that land on the same runway. Furthermore, since we used an engine-start time of 6 min for large aircraft (ICAO-D to ICAO-F) in comparison to 3 min for small aircraft, the taxi times of departing aircraft vary more than those of arriving aircraft in the simulated operations.

Of the simulations carried out for the algorithmic analysis, the five shown in Fig. 10 represent different combinations for w_{plng} and h_{plng}. We also tested values for w_{plng} below 15 min, but the routing algorithm did not succeed to find a solution throughout each of the two days. The taxi times do not differ significantly between the simulations, both for arriving and departing flights from all runways. Therefore, we conclude that both the planning window w_{plng} as well as the replanning period h_{plng} do not impact the efficiency of operations within the ranges that we tested, and we reject Hypothesis H1.

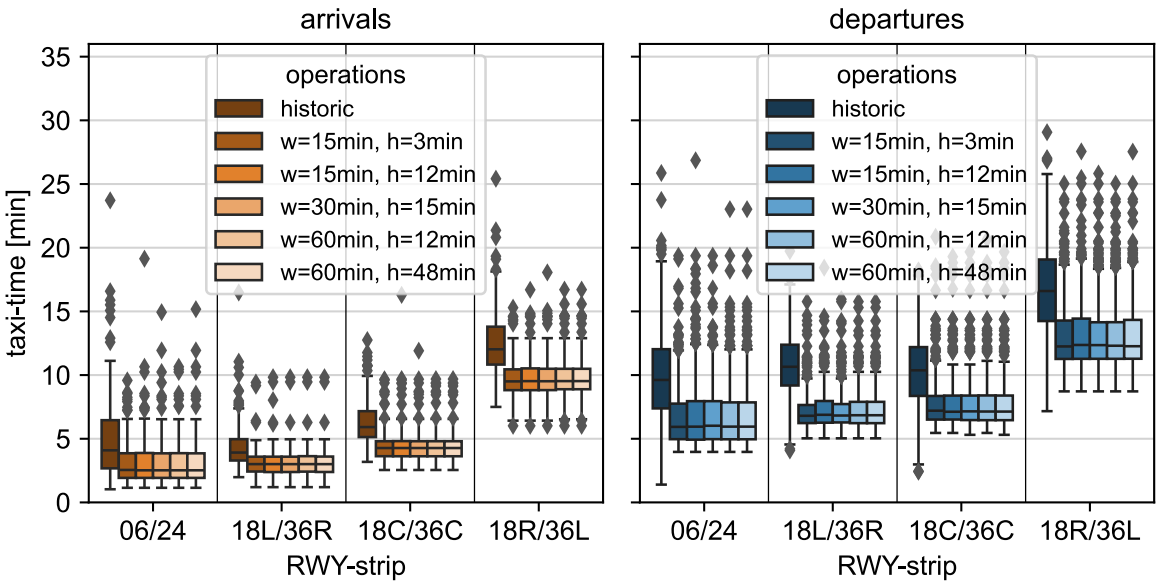

Figure 10. Box-and-whisker plot of historic and simulated taxi times for arrivals and departures per runway-strip

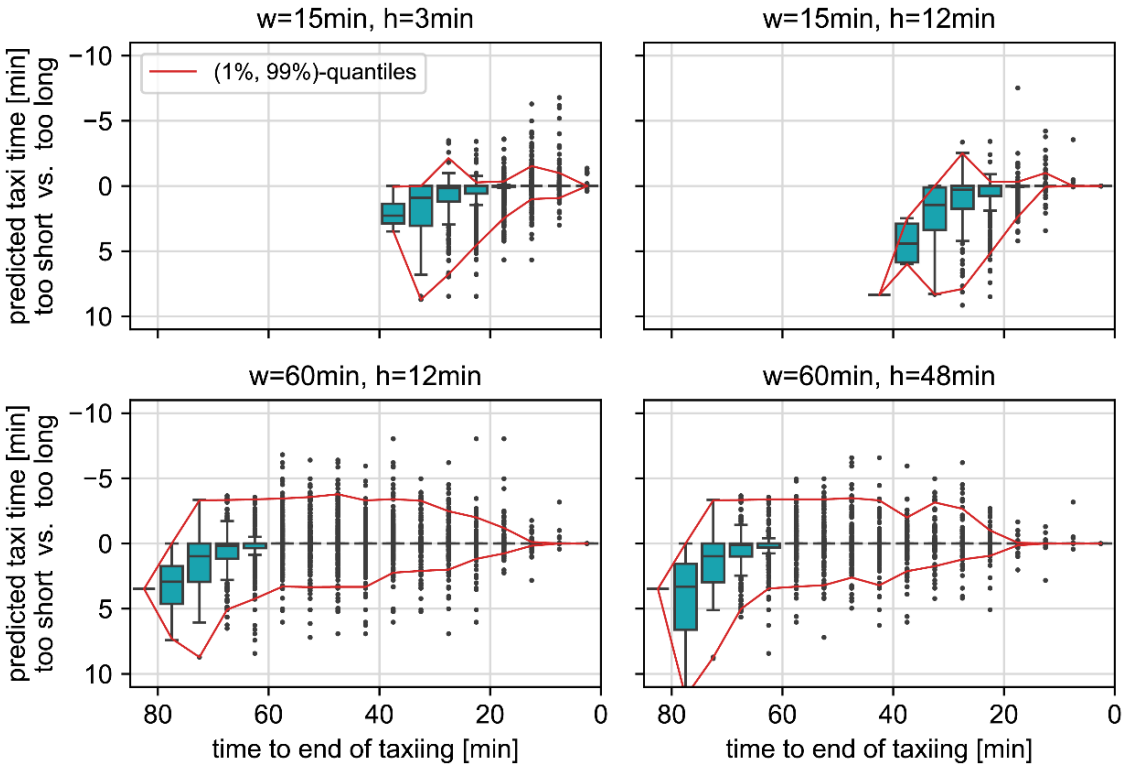

Figure 11. Predictability of taxi time. In each planning round, only the conflicts within the planning window w_{plng} are resolved. The estimated taxi times are then updated in the subsequent planning rounds. Based on the remaining taxi time, the predictions as diff

4.3 *Predictability of taxi times*

Fig. 11 shows the variability of the taxi time predictions with respect to the remaining actual taxi time for four sets of w_{plng} and h_{plng}. The red lines mark the 1 % and 99 % quantiles as indication of the accuracy over the remaining time. For all four simulations, the first predictions underestimate the actual taxi time: the aircraft start taxiing almost at the end of the planning

window and most conflicts are thus not yet resolved. The deviation to the actual taxi time decreases strongly in the following planning rounds. When the remaining taxi time is less than w_{plng}, the difference between predicted and actual taxi time is negligible for more than 50 % of all flights. The accuracy further increases towards the end of taxiing. Longer planning windows yield accurate predictions within a longer duration till the end of taxiing, supporting Hypothesis H2, while h_{plng} has a subordinate effect on the predictability. As listed in Table 3, the $RMSE$ and U values decrease significantly in comparison to the historic operations. Note that this may change when deviations to the planning arise during execution, which we did not model in this work.

4.4 *Computational efficiency of routing algorithm*

The multi-agent system including the routing algorithm are implemented in Python. We ran the simulations on a Windows 10 laptop equipped with a 1.80 GHz Intel Core i7-10610U CPU and 16 GB RAM. In Fig. 12, the runtime of each planning round is plotted over the number of agents that had to be routed for different sets of w_{plng} and h_{plng}. The planning rounds on day 1 (RMO North) are marked by circles, and those of day 2 (RMO South) by crosses. The algorithm scales with approx. quadratic time complexity in relation to the number of agents (supporting Hypothesis H3), while the RMO has a subordinate effect. From the parameter settings we tested, we deem w_{plng} = 20min with h_{plng} = 10min to yield a good trade-off between a fast algorithmic runtime and accurately predicting the remaining time within the last 10 min of taxiing. Thus, we chose these parameters to analyse the operational consequences of the autonomous operations.

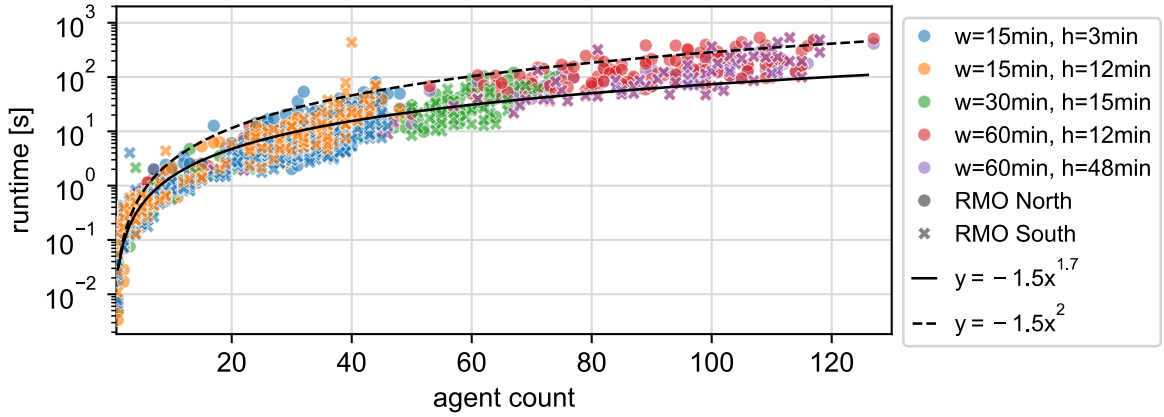

Figure 12. Runtimes of planning rounds over the number of agents

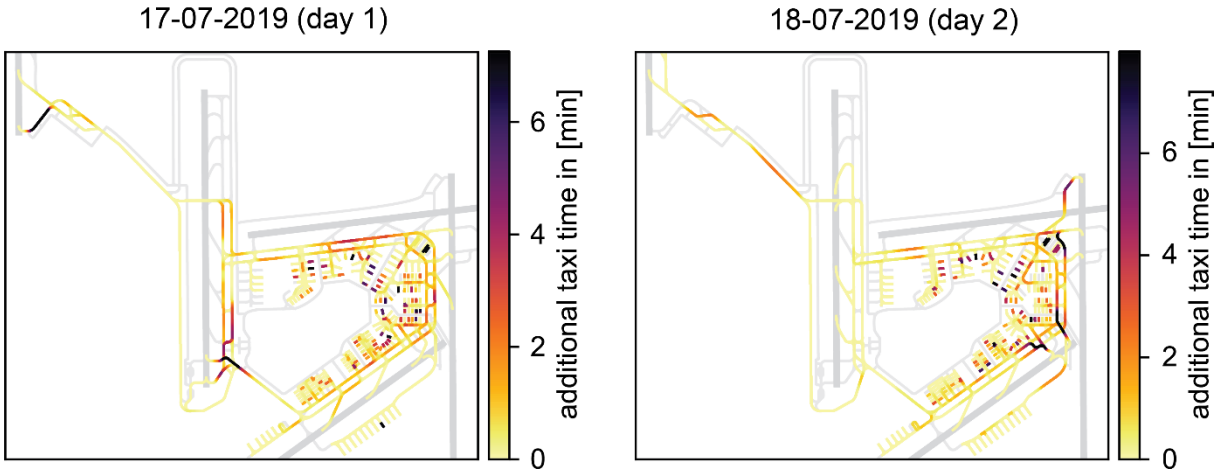

Figure 13. Additional taxi time per layout location on each of the two days, simulated with w_{plng}=20min, h_{plng}=10min.

4.5 *Delay hotspots in the taxiway network*

Fig. 13 visualizes congested areas in the taxiway network on each of the two days. On both days, hotspots form in front of the runway stopbars, the bay areas and stands, as well as on some taxiway segments. On 17th July, one of the latter hotspots is located in front of the crossing of the second departure-runway 36C, which is active only in departure-peaks and transition-phases (see Fig. 7). Aircraft slow down in front of the crossing to await its opening, which the Routing Agent determined to be faster than letting the aircraft go around the Southern end of 18C/36C. In front of the runways, aircraft queue with reduced velocity until the minimal separation time due to wake turbulence of the preceding aircraft has passed. In the bay areas, aircraft mostly hold at their stand, with some more holding after engine-start occurring on the second day.

4.6 *Holding of inbound and outbound flights*

In the taxi time analysis in Section 4.2, we excluded flights with a hold-type assigned by the Routing Agent. Fig. 14 compares the taxi times of these flights between historic operations and the different hold-types of the simulated operations as box-and-whisker plot. In general, not many flights are holding. In comparison to Fig. 10, the taxi times of flights with inbound holding are significantly higher, and are similar between historic and simulated operations, also considering that some of the historic A-CDM milestones end at the holding locations. For most outbound aircraft that must be delayed due to their CTOT-slots, the Routing Agent lets them hold at their stand. The taxi times of the historic operations are slightly longer, which has an influence on the moment that the aircraft take off within the CTOT-slot, as we analyse further below.

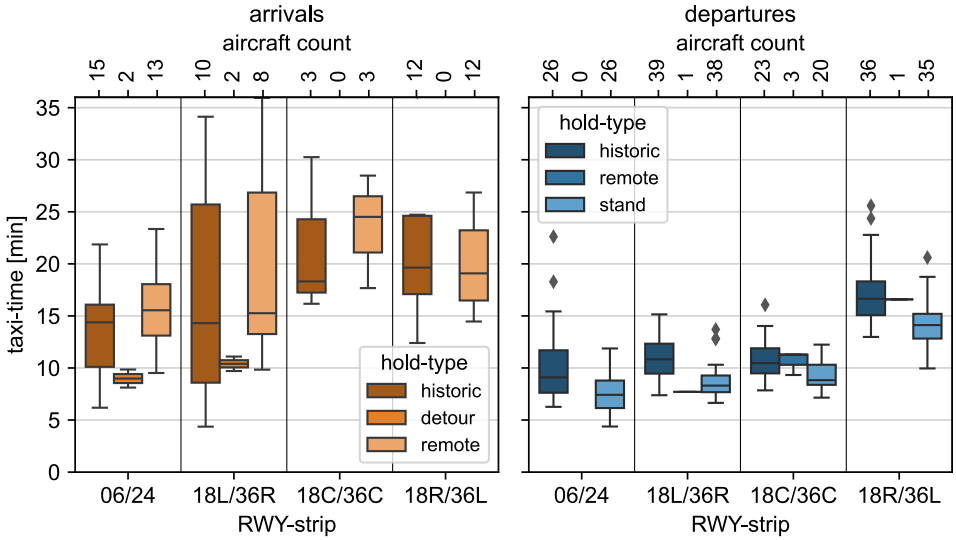

Figure 14. Box-and-whisker plot of historic and simulated taxi times dependent on the hold-type of arrivals and departures per runway-strip; simulation with $w_{plng} = 20min, h_{plng} = 10min$. The hold-type "remote" denotes holding at one of the remote holding locations, "detour" marks trajectories that go via one of these holding points. The hold-type "stand" denotes explicit holding time at the stand, in contrast to implicit holding at the stand as emergent property to resolve conflicts during path planning.

4.7 *CTOT-slots of outbound flights*

Over the two days, a total number of 442 CTOT-slots are assigned by Eurocontrol. Fig. 15 visualizes the compliance between the takeoff times to the CTOT-slots of both historic and simulated operations. While the historic times almost follow a normal distribution centered around the time issued by Eurocontrol, those of the simulated operations are skewed towards the beginning of the CTOT-window. In its current implementation, the routing algorithm optimizes for lowest taxi times and does not attempt to let aircraft take off closer to their calculated takeoff time. From the 442 flights, 8 historic flights (1.8 %) do not comply with their CTOT-slots. In comparison, only a

single simulated flight takes off outside its CTOT-slot. However, we noticed that for this flight, the historic A-CDM milestone occurs multiple minutes after the actual pushback, rendering it a faulty outlier.

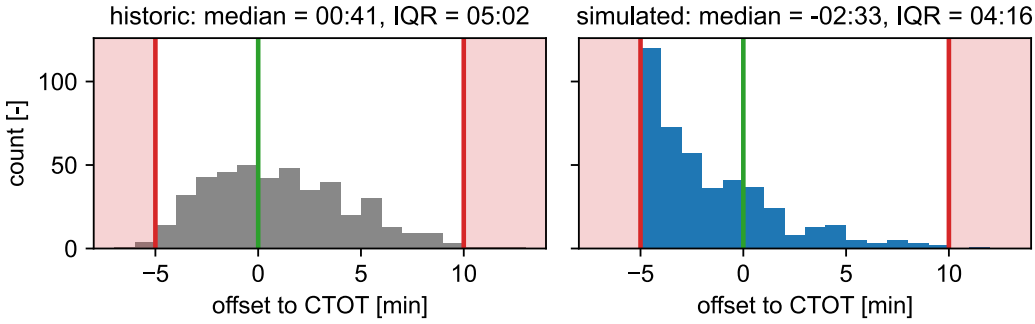

Figure 15. Compliance to CTOT-slot (white area) of historic (left) and simulated operations (right).

4.8 Runway sequence and capacity of outbound flights

As defined in Section 3, we do not use a pre-defined takeoff sequence. Thus, the order of aircraft departing from a runway is different between the historic and simulated operations, as exemplary shown in Fig. 16. The minimal separation between flights according to RECAT-EU is illustrated by the red shades. While the historic order mostly adheres to the RECAT-EU separation, in some cases, two aircraft are separated less than the minimum. We confirmed with the track data that indeed some flights take off around 52 s after the previous one, despite having a larger separation than indicated in the figure due to inaccuracies of the A-CDM milestones.

As emergent property of the MAS, whenever possible, flights are grouped together with minimal separation time between each takeoff. Based on this observation, we count the remaining slots per departure runway for each of the two days, as visualized in Fig. 17. For both historic and simulated operations, takeoff slots remain even during busy departure peaks marked by the blue shades. However, throughout the two days, more slots remain for the simulated operations: in relation to the total number of departing flights, 10.9 % and 12.6 % additional slots are available on each of the two days, respectively. Thus, the MAS better utilizes the potential runway capacity.

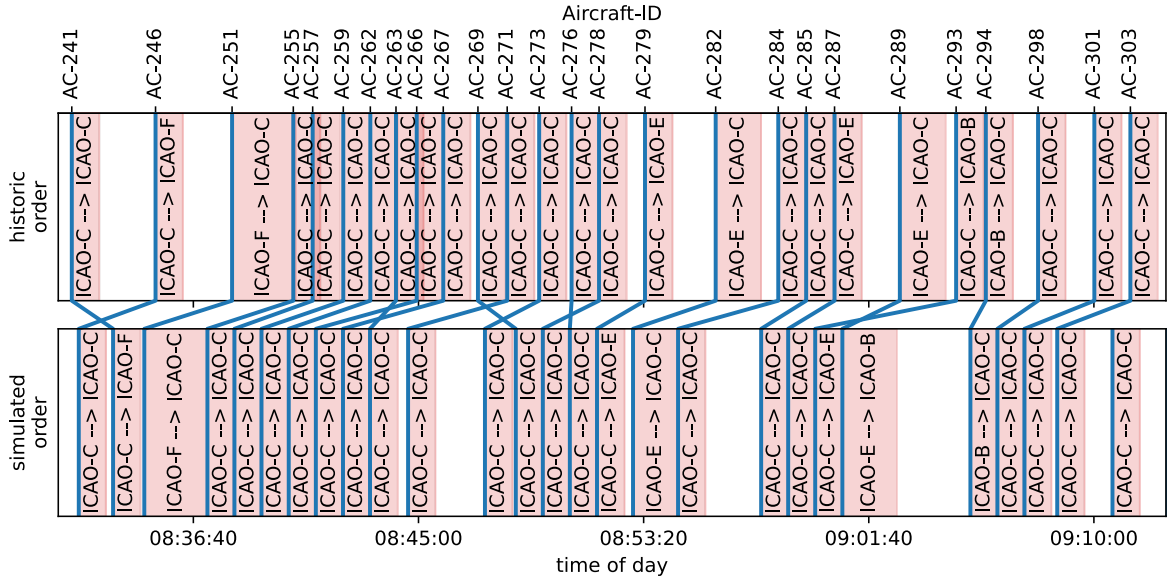

Figure 16. Comparison of exemplary takeoff order at runway 36L between historic (top) and simulated operations (bottom), with actual takeoff time (blue lines) and minimal WTC separation (red shades)

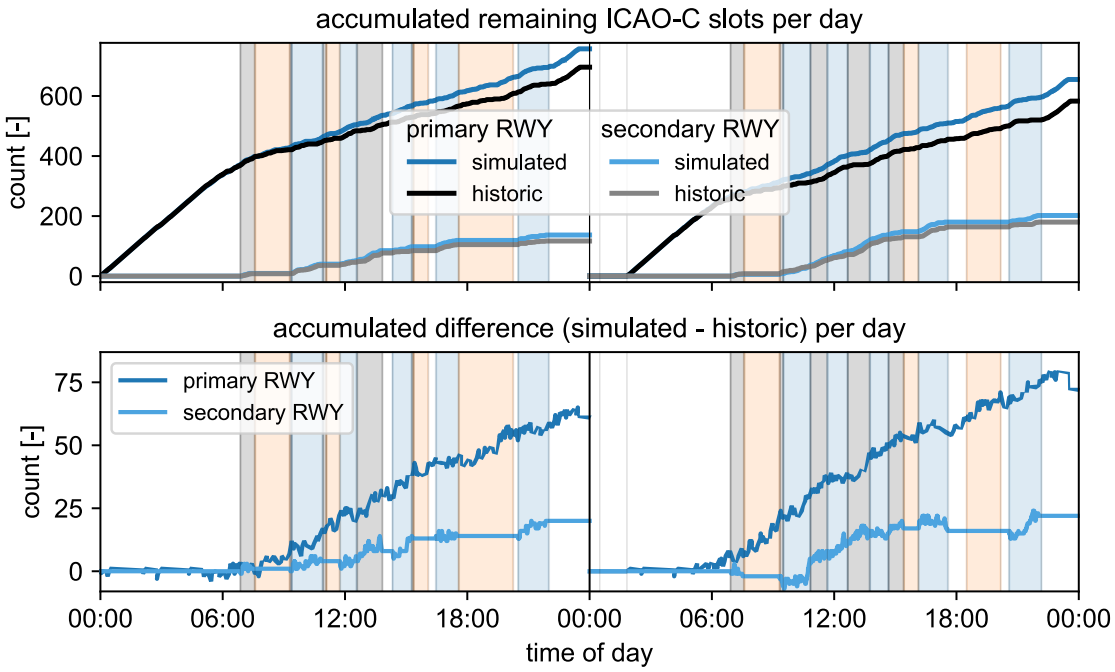

Figure 17. Remaining takeoff slots for ICAO-C aircraft per day for both historic and simulated operations. Shades denote the RMO phase: off-peak (white), arrival-peak (orange), transition (grey), and departure-peak (blue)

5 Future work

In this study, we considered a futuristic scenario of autonomous operations. However, before real-world airport surface movement operations are fully automated, a long transition period will be necessary with many challenges to investigate and resolve. In general, it needs to be explored which tasks could be automated and how controllers could then interact with the automated support tools. Moreover, a suitable architecture of the human-automation teaming that keeps controllers in the loop is needed. The underlying model should also handle automation failures in an effective way, e.g. by defining degradation modes for path planning in case the Routing Algorithm cannot find a conflict-free solution. We will focus on creating a suitable framework and evaluating it in human-in-the-loop simulations in future work.

As continuation of the research presented in this paper, we first plan to analyse how the engine-off taxiing techniques that were explored in AEON (2021) affect the key performance indicators of autonomous taxiing operations. Besides this, aircraft were assumed to execute all commands exactly as instructed. However, many uncertainties arise in real-world operations. Future research should determine how severe different sources of uncertainty are, and how to include these in the model. Moreover, it would be valuable to compare the centralized optimization of the concurrent routes presented in this paper to a baseline-model that represents the current working procedures of ground controllers more closely. As such, we plan to implement a distributed control model with local adjustments in future work.

Furthermore, as part of surface movement operations, various types of ground vehicles may come in close contact to parked and moving aircraft, foremost in the aprons. Their movements must be coordinated with each other and the aircraft. Future versions of the model should include such operations to explore the operational consequences that may result from their automation.

6 Conclusion

In this paper, we presented a hierarchical multi-agent system (MAS) model for autonomous taxiing operations at large airports in which different agents coordinate and control all movements on the

airport surface. As centralized agents, the Airport Operations Agent handles the flight schedule and runway configuration, while the Routing Agent computes conflict-free trajectories for all Aircraft Agents. Their execution is then instructed and monitored by the Guidance Agents. We accounted for aircraft shapes and kinematics during path planning on a high-resolution airport layout. To this end, we calibrated the model using historic ADS-B data. Furthermore, important airport surface movement elements and processes were explicitly included in the model such as pushback, engine-start, inbound holding, complying with CTOT-slots, and adhering to a minimal safety distance during taxiing as well as minimal wake turbulence separation during takeoff. The routing algorithm was evaluated to be well suited for planning conflict-free trajectories of all aircraft.

We analysed the proposed model using the real-world schedules of two of the busiest days at Amsterdam Airport Schiphol, including 19 different runway configurations per day. For the considered simulation conditions, the autonomous operations controlled by the MAS model reduced the average taxi time per flight by around 3 min, or 30 %. Furthermore, the simulated operations decreased the taxi time variability and yielded a high accuracy in predicting the remaining time till the end of taxiing. Counting the remaining takeoff slots at departure runways yielded an increase of approximately 11 % in comparison to the historic operations, meaning that the MAS model better utilizes the available runway capacity. All obtained results demonstrate the potential of the MAS as control model for autonomous airport surface movement operations that are more efficient, predictable, and hence produce less emissions.

Contributor Statement

Conceptualization: Malte von der Burg, Alexei Sharpanskykh
Funding Acquisition: Alexei Sharpanskykh
Investigation: Malte von der Burg
Methodology: Malte von der Burg, Alexei Sharpanskykh
Software: Malte von der Burg
Visualization: Malte von der Burg
Writing – Original Draft: Malte von der Burg
Writing – Review & Editing: Malte von der Burg, Alexei Sharpanskykh

Acknowledgment

This article is part of the ASTAIR - Auto Steer Taxi at AIRport project which has received funding from the SESAR 3 Joint Undertaking under Grant Agreement No 101114684 under European Union's Horizon Europe research and innovation programme.
The authors would like to thank Mr. Jan Post from LVNL for providing the historic track data that made this work possible, and Dr. Barys Shyrokau from the Faculty of Mechanical Engineering, TU Delft for the valuable discussions on acceptable accelerations for passenger comfort.

Conflict Of Interest

There are no conflicts of interest.

References

AEON - Advanced Engine-Off Taxiing Operations. (2021). AEON. Retrieved September 25, 2023, from https://www.aeon-project.eu/

AI roadmap 2.0: Human-centric approach to AI in aviation. (2023, May). *EASA*.

Atkin, J. A. D., Burke, E. K., & Ravizza, S. (2010). The Airport Ground Movement Problem: Past and Current Research and Future Directions. *In 4th International Conference on Research in Air Transportation*.

Automation in air traffic management: Long-term vision and initial research roadmap. (2020). *SESAR JU*.

Bae, I., Moon, J., Jhung, J., Suk, H., Kim, T., Park, H., Cha, J., Kim, J., Kim, D., & Kim, S. (2022, November 18). *Self-Driving like a Human driver instead of a Robocar: Personalized comfortable driving experience for autonomous vehicles*. arXiv: 2001.03908 [cs, eess]. Retrieved March 8, 2024, from http://arxiv.org/abs/2001.03908

Bakowski, D. L., Hooey, B. L., Foyle, D. C., & Wolter, C. A. (2015). NextGen Surface Trajectory-Based Operations (STBO): Evaluating Conformance to a Four-dimensional Trajectory (4DT). *Procedia Manufacturing*, *3*.

Chen, J., Weiszer, M., Stewart, P., & Shabani, M. (2015). Toward a More Realistic, Cost-Effective, and Greener Ground Movement Through Active Routing—Part I: Optimal Speed Profile Generation. *IEEE Transactions on Intelligent Transportation Systems*, *17*.

Chua, Z., Cousy, M., Causse, M., & Lancelot, F. (2017). Initial assessment of the impact of modern taxiing techniques on airport ground control. *Proceedings of HCI-Aero 2016*.

Cohen, L., Uras, T., Kumar, T. K. S., & Koenig, S. (2019). Optimal and bounded-suboptimal multi-agent motion planning. *Proceedings of SoCS-19*.

De Winkel, K. N., Irmak, T., Happee, R., & Shyrokau, B. (2023). Standards for passenger comfort in automated vehicles: Acceleration and jerk. *Applied Ergonomics*, *106*.

Eurocontrol. (2018). European aviation in 2040 - challenges of growth. *Eurocontrol StatisticsForecast Service*.

Eurocontrol. (2021, November 30). Data snapshot #22 on lower summer taxi-out times.

Helbing, D., & Balietti, S. (2015). How to do agent-based simulations in the future: From modeling social mechanisms to emergent phenomena and interactive systems design.

IATA. (2021, October 4). *Net-zero carbon emissions by 2050*. Press Release No. 66. Retrieved April 4, 2023, from https://www.iata.org/en/pressroom/pressroom-archive/2021-releases/2021-10-04-03/

ICAO. (2004). *Advanced Surface Guidance and Control Systems (A-SMGCS) Manual*.

ICAO. (2016). Aerodrome design and operations. In *Annex 14 to the convention on international civil aviation* (7th ed.).

Lane, R., Dubuisson, S., Ellejmi, M., & Cerasi, E. (2020, April 22). *EUROCONTROL specification for A-SMGCS services*.

Li, J., Tinka, A., Kiesel, S., Durham, J. W., Kumar, T. K. S., & Koenig, S. (2021). Lifelong multi-agent path finding in large-scale warehouses. *Proceedings of AAAI-21*.

Liu, Y., Hansen, M., Gupta, G., Malik, W., & Jung, Y. (2014). Predictability impacts of airport surface automation. *Transportation Research Part C: Emerging Technologies*, *44*.

Ma, H., Harabor, D., Stuckey, P. J., Li, J., & Koenig, S. (2019). Searching with consistent prioritization for multi-agent path finding. *Proceedings of AAAI-19*.

Morris, R., Pasareanu, C. S., Luckow, K., Malik, W., Ma, H., Kumar, T. K. S., & Koenig, S. (2016). Planning, Scheduling and Monitoring for Airport Surface Operations. *In AAAI-16 Workshop on Planning for Hybrid Systems*.

NielsB. (2007, August 6). *Schiphol (Amsterdam Airport) overview map*. Wikimedia Commons. Retrieved September 28, 2023, from https://commons.wikimedia.org/wiki/File:Schiphol-overview.png Phillips, M., & Likhachev, M. (2011). SIPP: Safe Interval Path Planning for dynamic environments. *Proceedings of ICRA-11*.

Roling, P. C., & Visser, H. G. (2008). Optimal Airport Surface Traffic Planning Using Mixed-Integer Linear Programming. *International Journal of Aerospace Engineering*, *2008*.

Rooseleer, F., & Treve, V. (2018, February 14). *RECAT-EU: European wake turbulence categorisation and separation minima on approach and departure*.

Sargent, R. G. (2011). Verification and validation of simulation models. *Proceedings of WSC-10*.

Schiphol - standaard pushback per positie. (n.d.). Schiphol. Retrieved August 30, 2022, from https://www.schiphol.nl/en/operations/page/sleep-en-pushbackbewegingen/

SESAR Joint Undertaking | ASTAIR - Auto-Steer Taxi at Airport. (2023). Retrieved September 25, 2023, from https://www.sesarju.eu/projects/astair

SESAR Joint Undertaking | MOTO - the Embodied Remote Tower. (2016). Retrieved September 25, 2023, from https://www.sesarju.eu/projects/moto

SESAR Joint Undertaking | TaCo - Take Control. (2017). Retrieved September 25, 2023, from https ://www.sesarju.eu/projects/taco

Stergianos, C., Atkin, J., Schittekat, P., Nordlander, T., Gerada, C., & Morvan, H. (2016). The importance of considering pushback time and arrivals when routing departures on the ground at airports. *Proceedings of ICAOR-16*.

Stern, R., & Sturtevant, N. R. (2019). Multi-agent pathfinding: Definitions, variants, and benchmarks. *Proceedings of SoCS-19*.

Udluft, H. (2017). *Decentralization in Air Transportation* [Doctoral dissertation, Department of Control and Operations, Aerospace Engineering, TU Delft].

von der Burg, M., & Sharpanskykh, A. (2023). Multi-Agent Planning for Autonomous Airport Surface Movement Operations. *SESAR Innovation Days 2023*.

Weiszer, M., Burke, E. K., & Chen, J. (2020). Multi-objective routing and scheduling for airport ground movement. *Transportation Research Part C: Emerging Technologies*, *119*.

X-Plane. (2024). *WorldEditor (WED)* (Version WED 2.5).

Zhang, M., Huang, Liu, & Li. (2019). Multi-Objective Optimization of Aircraft Taxiing on the Airport Surface with Consideration to Taxiing Conflicts and the Airport Environment. *Sustainability*, *11*.