Noname manuscript No.
(will be inserted by the editor)

Efficient and Fair Traffic Flow Management for On-demand Air Mobility

Christopher Chin · Karthik Gopalakrishnan · Hamsa Balakrishnan · Maxim Egorov · Antony Evans ·

Received: date / Accepted: date

Abstract The increased use of drones and air-taxis is expected to make airspace resources more congested, necessitating the use of Unmanned Aircraft Systems Traffic Management (UTM) initiatives to ensure safe and efficient operations. Typically, strategic UTM involves solving an optimization problem that ensures that proposed flight schedules do not exceed airspace and vertiport capacities. However, the dynamic nature and low lead-time of applications such as on-demand delivery and urban air mobility traffic may reduce the efficiency and fairness of strategic UTM. We first discuss the adaptation of three fairness metrics into a traffic flow management problem (TFMP). Then, with computational simulations of a drone package delivery scenario in Toulouse, we evaluate trade-offs in the TFMP between efficiency and fairness, as well as between different fairness metrics. We show that system fairness can be improved with little loss in efficiency. We also consider two approaches to the integrated scheduling of both high lead-time flights (i.e., flights with a schedule known in advance) and low lead-time flights in a rolling horizon optimization framework. We compare the performance of both approaches for different horizon lengths and under varying proportions of high and low lead-time flights.

This work was supported in part by A^3 by Airbus under contract number 40008574, and NASA under grant number 80NSSC19K1607.

Christopher Chin, Karthik Gopalakrishnan, Hamsa Balakrishnan
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA, USA
E-mail: {chychin, karthikg, hamsa}@mit.edu

Maxim Egorov, Tony Evans
Airbus UTM
Sunnyvale, CA, USA
E-mail: {maxim.egorov, tony.evans}@airbus-sv.com

Keywords Fairness · Equity · Efficiency · Air Traffic Flow Management · UAS Traffic Management

1 Introduction

The increasing demand for Unmanned Aircraft Systems (UAS) and Urban Air Mobility (UAM) applications, such as package delivery and air taxis, is expected to transform the skies. Recent studies estimate a demand for over 170,000 package-delivery drone flights/hour over Paris by the year 2035, with some urban areas projected to see a 200× increase in UAS operations, and a 30× increase in UAM operations [1, 2]. With increasing traffic, there is a need for congestion management algorithms to ensure the safe operation of these vehicles. Further, due to the dynamic and unpredictable nature of flight traffic, UAS Traffic Management (also referred to as UTM) will require both tactical (i.e., near real-time) and strategic (i.e., minutes or hours in advance) management.

In this paper, we focus on the strategic aspects of UTM. The starting point of our research is the classic air traffic flow management (ATFM) problem. The key idea behind ATFM is to proactively manage congestion by anticipating traffic demand and predicting the usage of various airspace and airport resources, with respect to their capacities. Delays are then assigned to aircraft, either before departure (on the ground) or in the air (through airborne holds or speed changes), to meet resource capacity constraints. An integer program (IP) formulation called the traffic flow management problem (TFMP) is used to assign delays, based on resource capacity constraints, aircraft performance characteristics (e.g., maximum hold times), and desired flight trajectories. The overarching objective of ATFM, and the corresponding IP, is to improve system efficiency by absorbing unavoidable delays on the ground, where they are less costly, rather than in the air.

In the quest for efficiency, fairness can often be sacrificed. Fairness has been recognized as a key consideration by industry stakeholders and regulators [3]. Incorporating fairness can minimize strategic behavior like exaggerating vehicle limitations or filing overly conservative flight plans, and can improve competitiveness by preventing the monopolization of the airspace. In this paper, we consider three well-established fairness measures from ATFM, namely reversals, overtaking, and time-ordered deviation. There are, however, several unique characteristics of UTM that necessitate the development of new algorithms beyond those used in present-day commercial aviation. Contributing factors include a higher density of operations, low-endurance of flights, autonomous vehicles, and dynamic flight operations.

Of these factors, the dynamic and unpredictable nature of the flight traffic is most significant. Because of the need to sell tickets to passengers in advance, present-day commercial air traffic is highly predictable with schedules filed at least a few weeks before departure. This enables the implementation of strategic congestion mitigation strategies such as re-routes, and speed or altitude changes for known flight schedules, hours in advance of their departure times. However, on-demand aerial services where flight destinations and routes are decided minutes, rather than months, in advance make strategic planning challenging. Our paper quantitatively explores the effect of dynamic demand on the efficiency and fairness of the TFMP solution in greater detail. More formally, we investigate the effect of increasing the proportion of flights that are “pop-ups” (i.e., flights that have a low file-ahead time such that they were not included in earlier horizons). Finally, we propose two candidate approaches to incorporate flights with low lead-times in our traffic flow management framework, and identify conditions under which they will be beneficial.

1.1 Problem description

Our paper focuses on the design of strategic air traffic flow management frameworks for highly dynamic flight traffic. The standard TFMP formulation has been shown to find efficient solutions [4]. However, efficient solutions may sacrifice fairness between flights and operators. For example, an efficient solution may require that a flight f_1 arrives at a resource before another flight f_2, even though f_2 was originally scheduled to arrive at that resource before f_1. This “schedule reversal” is one example notion of fairness. Other notions of fairness are discussed in Section 2. Balancing between efficiency (low system-wide delay costs) and fairness (equitable distribution of delays) can be achieved for TFMP by appropriately modifying the objective of the optimization problem [5, 6].

However, this optimization problem relies on the advance knowledge of proposed flight trajectories along with airspace and airport capacities for the planning time horizon. When there is a lack of certainty on proposed flight plans due to the on-demand nature of the system, our experiments show that this approach is not as effective (Section 5.2). This can be easily explained by reasoning about the scope of the optimization problem. When flight schedules are known well in advance, the TFMP optimization can be run over longer time-horizons with more flights to find a globally optimal solution. In contrast, when flight schedules are known with low lead times, the planning horizon is smaller and it is not possible to proactively plan as many trajectories at once to minimize delay costs and maximize fairness. We expect, though, that there is still a benefit to strategic traffic management over relying solely on a tactical system, like First-Come-First-Serve (FCFS).

1.2 Prior works

The efficient allocation of constrained airspace and airport resources has been studied extensively. The problem of managing congestion by assigning delays on the ground, through the Ground Delay Program, is a simplified version of the traffic flow management problem. Initial studies focused on the optimization formulation for the case of a single airport [7] or groups of airports [8]. The ATFM problem, which considers both airspace and ground resources, is more challenging to address; however, significant progress has been made in solving this problem over the past two decades [9, 4, 10]. Other works have studied fairness in air traffic flow management. Fairness and equity have been considered in the context of arrivals at a capacity-constrained airport [11, 12]. In the context of trajectory-based operations, max-min fairness [13], cost-based penalization for fairness and equity [14], and accrued delay [15] have been considered. Within the TFMP, fairness has previously been defined in terms of three popular metrics: reversals (change in arrival order among flights at resources relative to schedule) and overtakings (magnitude of reversals) [5], and time-ordered deviation [6] (excess delay beyond maximum expected delay for each flight). These will be described in-depth in Section 2. In addition, the trade-off between efficiency and notions of flight/passenger fairness was explored in [16]. Our previous work in [17] looked at efficiency and fairness when operators have different preferences and market shares.

Efficiency and fairness of the TFMP in a UTM context is an area of growing interest [18]. Proposed ideas for congestion management include airspace auctions [19] or decentralized algorithms [20]. While a large focus has been on developing tactical conflict resolution algorithms, very limited work attempts to solve a strategic global congestion management problem [10]. The key difference between our focus and prior works is in the extension of the classic traffic flow management optimization into a framework best suited for dynamic traffic demands. Past work suggests that, in a strategic decentralized FCFS setting, flights with a lower file-ahead time may experience significantly more delay than flights with higher file-ahead times. [20].

Lastly, stochastic optimization [21, 22, 23], robust optimization [24, 25], and chance-constrained programming [26] are other popular frameworks for decision making under uncertainty. However, these works only consider uncertainty in travel times or airport capacity and do not consider uncertain traffic demand. Additionally, these formulations only focus on efficiency and ignore fairness considerations.

1.3 Contributions and findings

The contributions of this paper are threefold. First, we identify the nuances of the UTM context that prevent the direct application of ATFM solutions. Second, our work emphasizes that there is no all encompassing metric of fairness, and that the choice of metric may be critical in determining the optimal allocation of resources. Finally, we use realistic simulations, including trajectory data from an Airbus simulator, battery-life-based flight time constraints, dynamic demand with low file-ahead times, and a rolling horizon implementation to evaluate the trade-offs between efficient and fair solutions in a practical UTM setting. Our major findings are as follows:

1. For the three fairness metrics studied (reversals, overtaking, and time-order deviation), a significant improvement in fairness can be obtained in exchange for little to no decrease in system efficiency.
2. Incorporating reversals can increase time-order deviation (improving one fairness metric, but deteriorating another), whereas incorporating time-order deviation can decrease reversals/overtaking (improving two metrics).
3. Demand with low file-ahead times ("pop-ups") can be incorporated in the TFMP when using a rolling horizon. Pop-ups can either be inserted or forced to wait until the next horizon. With a high pop-up fraction, using shorter horizons and inserting pop-ups into the schedule is best. With a low pop-up fraction, using longer horizons is acceptable as long as pop-ups are inserted, and shorter horizons perform well when either inserting pop-ups or delaying them until the next planning horizon.

1.4 Outline

In Section 2 we present the traffic flow management problem (TFMP) in its baseline form, as well as with fairness metrics of reversals, overtaking, or time-order deviation incorporated. In Section 3 we describe the two alternative approaches to handling dynamic demand: inserting pop-ups and delaying pop-ups until the next horizon. We then describe the Toulouse package delivery scenario that we used and the experiments we performed to evaluate the trade-offs between fairness and efficiency, and the impact of dynamic demand. In Section 5 we present our experimental results before summarizing this paper in Section 6.

2 The Traffic Flow Management Problem (TFMP)

In this section, we present the main formulation for the traffic flow management problem. We describe three metrics to measure fairness and show how they can be incorporated into the optimization. We first build off of the classical traffic flow management problem (TFMP) formulation [9].

2.1 Setup and Notations

We consider a discrete-time traffic flow management problem, which uses the following notation described below.

$\mathcal{T}$:	Set of time periods $\{1, \ldots, T\}$
ΔT :	Length of each time period
$\mathcal{A}$:	Set of all vertiports
$\mathcal{S}$:	Set of all airspace sectors
$\mathcal{F}$:	Set of all flights
$C(s,t)$:	Capacity of sector $s \in \mathcal{S}$ at time t
$A(a,t)$:	Arrival capacity of vertiport a at time t
$D(a,t)$:	Departure capacity of vertiport a at time t
a_f :	Scheduled arrival time for flight $f \in \mathcal{F}$
d_f :	Scheduled departure time for flight $f \in \mathcal{F}$
$\mathcal{S}^f$:	Sequence of sectors in flight f's trajectory
$\mathcal{S}^f_j$:	Next sector after j in flight f's trajectory
$\mathcal{P}^f_j$:	Sector preceding j in flight f's trajectory
orig_f :	Origin vertiport for flight f
dest_f :	Destination vertiport for flight f
$l_{f,s}$:	Minimum time spent by flight f in sector s
M :	Maximum delay for each flight
T^f_j :	Set of feasible time periods for flight f to arrive at resource $j \in \mathcal{S} \cup \mathcal{A}$ (vertiport or sector)
$\overline{T}^f_j$:	Latest time in the set T^f_j
$\underline{T}^f_j$:	Earliest time in the set T^f_j
$w^f_{j,t}$:	A binary variable that is 1 when flight $f \in \mathcal{F}$ has arrived at resource $j \in \mathcal{A} \cup \mathcal{S}$ at or before time t

2.2 Baseline TFMP

We impose a maximum delay across all flights (M) such that no single flight is overly penalized. M is used to construct T^f_j, the feasible arrival times at each resource. The objective function minimizes the total delay cost (TDC). The expression for (TDC) is assumed to be of the form $\text{TDC} = \beta(\text{GD}^{1+\epsilon} + \alpha\text{AD}^{1+\epsilon})$, where GD is ground delay, AD is airborne delay, β is delay to cost scale factor, and $\alpha \geq 1$ is the ratio of airborne delay cost to ground delay cost. Note that ϵ makes these costs super-linear in the delay duration, as we prefer even distribution of delays across flights over skewed delay distributions. For example, setting ϵ to be a small positive number (≤ 0.05) guides the optimization solver to allocate 2 minutes of delay each for two flights rather than 4 minutes of delay to a single flight, even though the total delay would be the same for both solutions. In other words, this super-linear cost structure breaks ties between solutions that result in the same total delay cost. Without loss of generality, we set $\beta = 1$. Since TD = AD + GD, we have:

$$\text{TDC} = \alpha\text{TD}^{1+\epsilon} + (1-\alpha)\text{GD}^{1+\epsilon} \tag{1}$$

If a flight departs at time t, then its ground delay (GD) is $(t - d_f)$. Also, if this flight lands at time t, the total delay (TD) is $(t - a_f)$. Thus, TDC can be re-written as below.

$$\begin{aligned}\text{TDC} = \sum_{f\in\mathcal{F}}\Big(&\sum_{t\in T^f_{dest_f}} \alpha(t-a_f)^{1+\epsilon}(w^f_{dest_f,t} - w^f_{dest_f,t-1}) \\ &- \sum_{t\in T^f_{orig_f}} (\alpha-1)(t-d_f)^{1+\epsilon}(w^f_{orig_f,t} - w^f_{orig_f,t-1})\Big)\end{aligned} \tag{2}$$

The key aspect of the formulation that lends computational tractability to larger-scale problems is the choice of the decision variable $w^f_{j,t}$, which is a binary variable that is non-decreasing in time (Constraints (3g) and (3h)). Flight f is said to enter a resource i (which could be a vertiport or a sector) *at* time t if $(w^f_{i,t} - w^f_{i,t-1}) = 1$. The following constraints must be satisfied:

$$\sum_{f\in\mathcal{F}:\, orig_f=k} (w^f_{k,t} - w^f_{k,t-1}) \leq D(k,t),\ \forall k \in \mathcal{A}, t \in \mathcal{T} \tag{3a}$$

$$\sum_{f\in\mathcal{F}:\, dest_f=k} (w^f_{k,t} - w^f_{k,t-1}) \leq A(k,t),\ \forall k \in \mathcal{A}, t \in \mathcal{T} \tag{3b}$$

$$\sum_{f\in\mathcal{F}:\, i\in\mathcal{S}^f, j=\mathcal{S}^f_i} (w^f_{i,t} - w^f_{j,t-1}) \leq C(j,t),\ \forall t \in \mathcal{T} \tag{3c}$$

$$w^f_{i,t} = 0,\ \forall f \in \mathcal{F}, t = \underline{\mathrm{T}}^f_{j-1}, i = \mathcal{S}\cup\mathcal{A} \tag{3d}$$

$$w^f_{i,t} = 1,\ \forall f \in \mathcal{F}, t = \overline{T}^f_j, i = \mathcal{S}\cup\mathcal{A} \tag{3e}$$

$$\begin{aligned} w^f_{i,t} - w^f_{j,t-l_{f,j}} \leq 0,\ &\forall f \in \mathcal{F}, t \in T^f_i, \\ &i \in \mathcal{S}^f : i \neq \text{orig}_f, j = \mathcal{P}^f_j \end{aligned} \tag{3f}$$

$$w^f_{i,t-1} - w^f_{i,t} \leq 0,\ \forall f \in \mathcal{F}, i \in \mathcal{S}^f, t \in T^f_i \tag{3g}$$

$$w^f_{i,t} \in \{0,1\},\ \forall f \in \mathcal{F}, i \in \mathcal{S}^f, t \in T^f_i \tag{3h}$$

Constraint (3a) enforces departure capacity for each vertiport at each timestep by summing across all flights that departed. Constraint (3b) does the same for arrivals, and constraint (3c) does the same for enroute sectors by tracking the current location of each flight f based on values of $w^f_{s,t}$ for consecutive s in S_f, the list of sectors that f traverses. Constraint (3d) ensures that a flight does not reach a sector before the earliest feasible time. Analogously, constraint (3e) enforces that a flight must arrive at a sector before the latest feasible time. The minimum time to be spent in each sector is described in Constraint (3f).

2.3 Fairness Metrics

We focus on three candidate notions of fairness, which we describe qualitatively below. We then incorporate them into the baseline TFMP formulation.

2.3.1 Reversals and overtaking [5]

According to this notion, a fair solution is one in which the relative ordering of arrivals at any resource is preserved according to published schedules. More precisely, a flip in the ordering of flight arrivals at a sector or a vertiport with respect to the original schedule is called a *reversal*, and the magnitude of the reversal, in terms of the difference in arrival times is referred to as *overtaking*. Two additional sets for reversals and overtaking are defined below.

- R^j : Pairs of reversible flights
- $T^r_{f,f',j}$: Set of time periods common for flights f and f' where a reversal could occur at resource j
- λ_r : Penalty factor for reversals
- λ_o : Penalty factor for overtaking

For reversals, we define a new variable $s_{f,f',j}$ which is 1 if flight f' arrives before flight f at resource j, where f was scheduled to arrive before f', and 0 otherwise. In the objective function, we sum the previously defined TDC with the total number of reversals multiplied by a weight λ_r.

$$\min \quad \text{TDC} + \lambda_r \sum_{j\in S,(f,f')\in R^j} s_{f,f',j} \tag{4}$$

The following constraint must be satisfied:

$$s_{f,f',j} = \max(0, w^{f'}_{j,t} - w^f_{j,t})\ \forall t \in T^r_{f,f',j} \tag{5}$$

For overtaking, we define a new variable $s^i_{f,f',j}$ which is 1 if flight f' arrives but flight f does not arrive by time $\underline{\mathrm{T}}^f_j + i$ in resource j, where f was scheduled to arrive before f', and 0 otherwise. The objective function looks similar to incorporating reversals, but note that $s^i_{f,f',j}$ is summed over the cardinality of $T^r_{f,f',j}$.

$$\min \quad \text{TDC} \; + \; \lambda_o \sum_{j \in S, (f,f') \in R^j}^{|T^r_{f,f',j}|} s^i_{f,f',j} \tag{6}$$

The following constraint must be satisfied:

$$s^i_{f,f',j} = \max\left(0, w^{f'}_{j,\underline{\mathrm{T}}^f_j + i} - w^f_{j,\underline{\mathrm{T}}^f_j + i}\right) \tag{7}$$

2.3.2 Time-order deviation [6]

In this section, we describe the time-order deviation metric used to quantify fairness. We calculate the first-come first-serve (FCFS) arrival time FCFS^f_i for each flight f at resource i that it goes through, assuming that i was the only constrained resource. With first-come first-serve, arrival slots are assigned to flights according to the original schedule ordering. For each flight, we then calculate the maximum FCFS delay d^{FCFS}_f.

- FCFS^f_i : First-come first-serve arrival time for flight f at resource i assuming that i was the only constrained resource ($i \in \mathcal{S} \cup \mathcal{A}$)
- d^{FCFS}_f : Maximum FCFS delay for flight f
- $c^f_{\text{TODA}}(t)$: Additional delay cost when flight f is delayed for time t
- λ_t : Penalty factor for time-order deviation

The intuition behind time-order deviation is as follows. When there are multiple constrained resources, each flight should not expect to incur any less delay than it would incur as a result of only the most constrained resource along its route. In other words, there is a notion of *expected delay*, that every flight is inherently entitled to be assigned, and only delays beyond this expected delay should be equalized among the multiple flights. Thus, for every flight $f \in \mathcal{F}$, the maximum delay that it would have been assigned as a result of only the most constraining resource is

$$d^{\text{FCFS}}_f \triangleq \max_{i \in \mathcal{S} \cup \mathcal{A}} \text{FCFS}^f_i \tag{8}$$

Thus, the modified optimization problem is

$$\min \quad \text{TDC} \; + \; \lambda_t \sum_f \sum_{t=a_f}^{T} c^f_{\text{TODA}}(t)(w^f_{\text{dest}_f,t} - w^f_{\text{dest}_f,t-1}),$$

$$\text{where} \qquad c^f_{\text{TODA}}(t) = (\max\{0, t - a_f - d^{\text{FCFS}}_f\})^{1+\epsilon}. \tag{9}$$

3 Method: Incorporating dynamic demand

Typically, the input demand to the standard TFMP formulation is not only deterministic, but also known well in advance such that the TFMP is only solved once. Given the on-demand nature of many UTM applications, this is not a realistic assumption. One way around this challenge is to implement a rolling horizon version of the TFMP. With a rolling horizon of length H (in time-periods, as defined in Section 2.1), we intend to solve the TFMP once for every horizon (i.e., every H time-periods). If we solve a horizon at time t, we solve the TFMP for flights with scheduled departure times in the range $[t, t + H - 1]$. Once a flight is assigned a revised schedule, it is fixed and acts as a constraint for flights in the next horizon. The time that a flight files its trajectory is denoted by r_f, and its scheduled departure time is d_f. The difference between the scheduled departure time and the filing time is called the *file-ahead time*, and is expressed as $\Delta_f = d_f - r_f$. Note that $\Delta_f \geq 0$, since $d_f > r_f$. The starting time of the horizon that contains the scheduled departure time of flight f is denoted as h_f. The rolling horizon implementation thus far works if the following conditions hold for every flight f:

$$h_f = \max(H * i) \mid H * i < d_f, \text{ where } i = 1, 2, 3, ... \tag{10}$$

$$\Delta_f > d_f - h_f \tag{11}$$

Thus, if a flight wants to depart toward the end of a horizon (i.e, $d_f - h_f$ is large), then its Δ_f must be high.

In some settings, it may be reasonable to require that all flights have sufficient file-ahead times. However, it may be unfair or unreasonable to require flights to file their trajectories before the start of a horizon, particularly for long horizons that require a high Δ_f. We expect that while some flights will have sufficiently high file-ahead times, others will not. This raises the question of how to incorporate flights with a low file-ahead time in the TFMP framework. We call a flight a *pop-up* if equation (10) and (11) do not hold. A pop-up flight files its trajectory after the start of the horizon that it wants to depart in. There are many ways to account for this dynamic demand. An extreme way is to abandon the strategic TFMP framework and tactically schedule flights in the order that they file, but this would remove the benefits of scheduling multiple flights at once. Instead, we prefer to incorporate the dynamic demand in the existing framework. We consider two such options for handling pop-ups below:

1. **Insert pop-ups:** Consider a pop-up flight f with a departure time d_f such that $h_f < d_f < h_f + H$. By the time flight f files its trajectory, non pop-up flights scheduled to depart between h_f and $h_f + H$ will have already been scheduled. When a pop-up files its flight plan, we run a one-flight TFMP with vertiport and sector capacities reduced to account for previously scheduled flights.

Note that we are not modifying previously scheduled flights, so the TFMP will have to find available capacity for the pop-up. If the maximum delay M for flight f is not high enough and demand is very high, the TFMP may be infeasible. We do not consider such cases.

2. **Delay pop-ups:** Consider the same situation above. Rather than insert the pop-up into the schedule, we delay the pop-up until the next horizon. The scheduled departure time at the origin and the scheduled arrival times at the enroute sectors and destination remain the same. However, the feasible times (e.g., earliest departure time) for the pop-up are updated to reflect its initial delay because of its shift to the next horizon. The initial pop-up delay is equal to $h_f + H - r_f$ with the following bounds: $0 < h_f + H - r_f < H$. In the worst case, the pop-up will file right after the start of the horizon ($r_f - h_f \approx 0$) and incur a delay close to H.

There are qualitative trade-offs between the two pop-up options. Option 1 (inserting pop-ups) may lead to an inefficient schedule for the current horizon that exacerbates delay in the next horizon. Option 2 (delaying pop-ups) allows the pop-up to be batched with the flights of the next horizon, but forces pop-ups to incur an initial delay before being solved by the TFMP. We will quantify the differences between these two options in experiments described in the next section.

4 Experimental Setup

4.1 Scenario generation

We use a package delivery scenario created by Airbus where four operators in Toulouse, France have warehouses on the outskirts of the city and make deliveries in locations randomly distributed around the city [27]. The vertiport traffic is determined through a Poisson process. Each flight has a desired 4D trajectory (three spatial dimensions with time as the fourth dimension). For simplicity, only the outbound flight segments, from the warehouse to the delivery site, are considered. We used two demand scenarios: 50 flights/hour and 25 flights/hour per vertiport. Fig. 1 shows the scenario with 50 flights/hour.

One of the key requirements of the TFMP formulation is that time is discretized into timesteps. We rounded sector entry and exit times to the nearest 60 s, while ensuring that each flight spent at least one timestep in each sector. We set a sector time discretization threshold of 3 s, and omit a sector from a flight's trajectory if it spent less than 3 s in it. Also, the TFMP formulation requires that a flight may only traverse through a sector once. We smoothed the trajectory in eight instances where a flight entered a sector multiple times. For example, a flight that entered sector A, briefly left to sector B, then reentered sector A would be modified to stay in sector A.

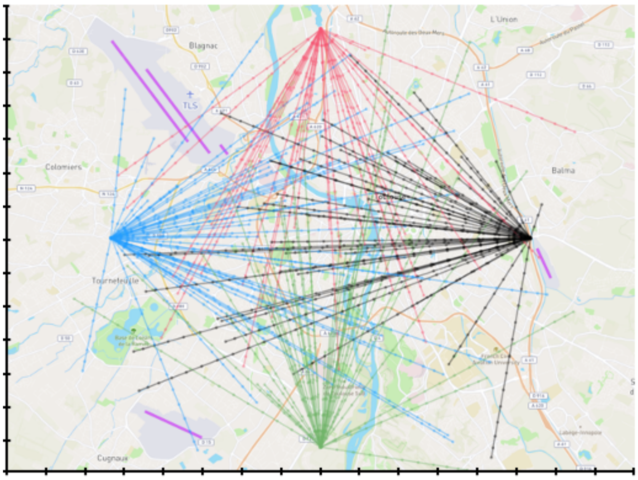

Fig. 1 Flight trajectories shown from 4 vertiports in a 16 km × 14 km region, with axis ticks along the border denoting 1 km sector boundaries.

An additional factor that we accounted for was maximum battery life, which we assumed to be 20 min. We used the remaining battery life and the unimpeded time-to-destination to calculate an upper-bound on airborne delay for each flight at each sector. Table 1 lists some additional parameters used for the experiment. One important parameter is the ratio of airborne delay cost to ground delay cost, represented by α. With higher values of α, airborne delay will be more penalized relative to ground delay. Thus, to minimize airborne delay, the total delay (ground + airborne delay) may increase. In contrast, with lower values of α, the TFMP solution will approach the minimum total delay solution. We took this ratio to be three based on [28], although different values of α could easily be justified. For an analysis of the effect of different α values on fairness, see [17].

Table 1 List of Parameters

Parameter	Value
Timestep Size	60 s
Sector X-Y Dimensions	1 km × 1 km
Sector Z Dimension (Height)	65 m
Sector Capacity	1 per sector
Departure Capacity	2 per timestep
Sector Discretization Threshold	3 s
Maximum Battery Life	20 min.
Airborne Delay Cost to Ground Delay Cost Ratio	$\alpha = 3$

4.2 Fairness-Efficiency Tradeoff

We seek to evaluate the fairness-efficiency tradeoff when incorporating one of three fairness metrics: reversals, overtaking, or time-order deviation. Recall that the weight that a fairness metric is given is represented by λ_r, λ_o, or λ_t. We vary these values to generate fairness-efficiency curves. There are four main variations of the TFMP objective function that we use: total delay cost only ("baseline") and total delay cost with penalization of reversals, overtaking, or time-order deviation. We use total delay cost as a measure of efficiency (refer to equation (2)). Note that total delay cost is distinct from total delay, as it penalizes airborne delay three times more than ground delay. In this initial round of experiments, we assume that the TFMP is only solved once. This means that there are no pop-ups, as all flights have sufficiently large file-ahead times such that the system is aware of them at the beginning of the experiment.

4.3 Experiments with Pop-ups

In these sets of experiments, we now relax the assumption that there are no pop-ups. We define the proportion of all flights that are pop-ups as the *pop-up fraction*, denoted by p. Given p, we randomly select the appropriate number of flights to be pop-ups. For bookkeeping purposes, flights with departure times that are equal to the start of a horizon are eligible to be pop-ups, but their departure times are shifted to 1 s later. Each pop-up is randomly assigned an integer file time r_f with a discrete uniform distribution on the interval $[h_f \mathinner{..} d_f]$, Recall that h_f is the start time of the horizon containing flight f, and d_f is the scheduled departure time of f. Besides pop-up fraction, there are three other parameters we vary. We have two options for handling pop-ups: Option 1 (inserting pop-ups) and Option 2 (delaying pop-ups). We also vary the horizon length, with larger horizon lengths meaning that fewer planning horizons—with several flights in them—are solved. Finally, we still have the choice of which TFMP objective function to use. Since the selection of pop-up flights is random, we can test scenarios with different sets of pop-up flights. However, when making direct comparisons between horizon lengths, pop-up options, or TFMP objective function, we use the same random seed so that the same set of flights are pop-ups.

5 Results

5.1 Fairness-Efficiency Tradeoffs

Incorporating fairness metrics in the objective function results in an inherent tradeoff between fairness and efficiency, measured in terms of the total delay cost. In the baseline formulation, the objective function is simply the total delay cost, without any fairness considerations. Thus, when incorporating fairness metrics in the objective function, the total delay cost either remains the same or increases as the additional terms drive the solution away from the optimal delay cost. In return, we expect fairness to increase. Additionally, we want to evaluate the effect of incorporating one fairness metric in the objective on other fairness metrics.

Fig. 2 shows the average number of reversals per flight and the total delay cost when minimizing total delay cost for various scenarios. The "Baseline" case minimizes the total delay cost, and the other three cases ("Reversal", "Overtaking", "TODA") incorporate one of the three fairness metrics. Results are shown for a high demand scenario (vertiport demand of 50 flights/hour) and a low demand scenario (25 flights/hour). For each scenario, there is one data point for the baseline case, but several data points for reversals, overtaking, and TODA, corresponding to different λ_r, λ_o, and λ_t values, respectively.

We first look at the results of incorporating reversals as a fairness metric in the low demand scenario (shown as blue hexagon points). As λ_r increases, the number of reversals decreases and the total delay cost increases relative to baseline (shown as a black square). For small λ_r values, it is possible to reduce the number of reversals with no increase in total delay cost. For example, when λ_r = 0.4 the number of reversals per flight decreases to 0.23 (compared to 0.54 in the baseline) with no increase in total delay cost. With further increases in λ_r, decreases in reversals are smaller and become increasingly expensive in terms of the total delay cost. At λ_r = 10 the optimal solution has only 3 reversals (equivalent to an average of 0.03 reversals per flight) but an average delay cost per flight of 1.86 (a 19% increase compared to 1.56 in the baseline). Overall, the average number of reversals decays exponentially with increasing total delay cost. This is because to prevent a pair of flights from being reversed, it may be necessary for one flight to incur excess delay. In the absence of limitations on the maximum delay a flight can endure, the number of reversals could be driven to zero at the cost of very high total delay.

In the high demand scenario, the new baseline (shown as a black circle) has a higher average number of reversals and average total delay cost than the previous baseline corresponding to a demand of 25 flights/hour. This is expected, as more congestion leads to more flight interactions and potential for reversals. Incorporating reversals in the objective has a similar effect as doing so with lower demand. The trade-off curve has a similar shape, and for very high λ_r, the average number of reversals approaches zero while the average total delay cost increases substantially.

Incorporating overtaking produces nearly identical results as when incorporating reversals. In many cases, they have identical optimal solutions, not only with regard to fair-

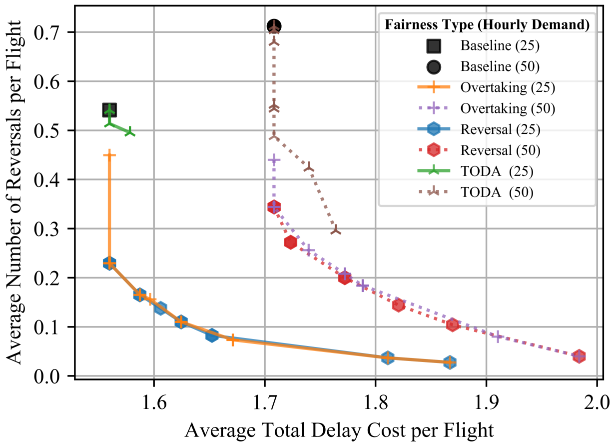

Fig. 2 Reversals vs. Total Delay Cost (TDC) when incorporating different fairness metrics. The hourly demand level is shown in parentheses.

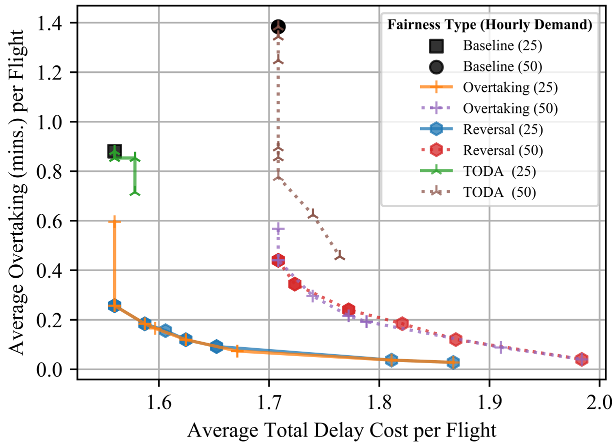

Fig. 3 Overtaking vs. Total Delay Cost (TDC).

ness and efficiency, but also concerning schedule and delay allocation. This is expected since the two fairness metrics are intertwined, with overtaking measuring the magnitude of time duration that a given pair of flights was reversed. Whereas reversals and overtaking are nearly in lockstep, time-order deviation behaves differently from reversals or overtaking. For small λ_t, incorporating time-order deviation can lead to a decrease in the average number of reversals with little to no increase in the total delay cost, especially for the high demand scenario. However, incorporating time-order deviation does not decrease the average number of reversals as much as explicitly incorporating reversals. For larger λ_t, the optimal solution does not change and no further reductions in reversals are apparent.

Fig. 3 is similar to Fig. 2 but displays average overtaking (in minutes) instead of the number of reversals on the y-axis. Since reversals and overtaking are closely related, it comes as no surprise that the efficiency-fairness tradeoff of both are similar. Average overtaking decreases exponentially in relation to the total delay cost, and for very large λ_r or λ_o, it is possible to reduce overtaking to zero, albeit at a great expense to the total delay cost. Incorporating time-order deviation impacts overtaking similarly to the way it impacted reversals.

Fig. 4 shows the average time-order deviation (in minutes) on the y-axis. We first consider how incorporating time-order deviation in the objective affects the average time-order deviation per flight. As λ_t increases, the average time-order deviation decreases and the average total delay cost increases. The decreases in time-order deviation are modest, but more pronounced in the high demand scenario, for which the tradeoff between the average time-order deviation and the average total delay cost is linear. At $\lambda_t = 2$, the average time-order deviation decreases by 4.5% and the total delay cost increases by 3%. The increase in total delay cost happens despite a reduction in total delay (from 208 min in the baseline to 201 min)—this is because the airborne delay (which is 3x more costly than ground delay) increases.

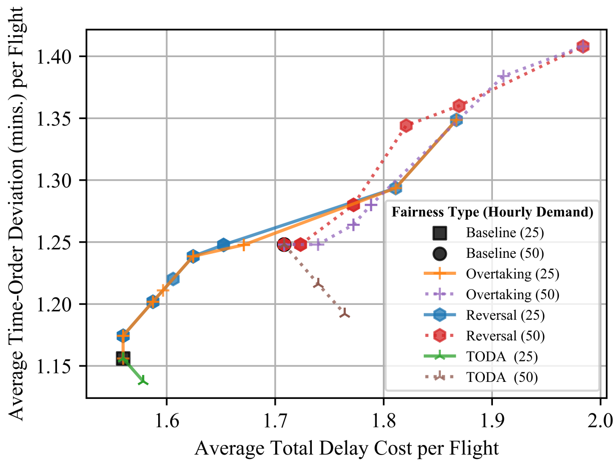

Fig. 4 Time-Order Deviation vs. Total Delay Cost (TDC).

While penalizing reversals or overtaking can drive its value to zero, it is not possible to drive the average time-order deviation to zero, no matter how large λ_t gets. This is inherent to the way time-order deviation is defined (9). If all flights have delay assigned greater than or equal to their maximum expected delay, time-order deviation cannot be reduced by reallocating delay to flights that have delay assigned less than their maximum expected delay. Instead, time-order deviation can only be decreased by also decreasing total delay. Thus, when all flights have delay assigned that is greater than or equal to their maximum expected delay and the total delay has been minimized, then the time-order deviation is also minimized. This is the case with the high demand scenario. No flight was assigned delay less than its maximum expected delay, and minimizing total delay rather than total delay cost leads to an optimal solution with the same 201 min of total delay seen with λ_t = 2. Incorporating reversals or overtaking results in a 17% increase in average time-order deviation in the low demand scenario and a 13% increase in the high demand scenario. In contrast, incorporating time-order deviation can slightly decrease reversals or overtaking.

While the improvement in the average time-order deviation when penalizing time-order deviation may appear modest, there is another benefit. Since the cost coefficient for time-order deviation is a super-linear function, evenly distributed time-order deviation is preferred over lopsided distributions. As such, incorporating time-order deviation also reduces the standard deviation of time-order deviation across flights. As λ_t increases, the standard deviation decreases; $\lambda_t = 2$ results in a 27% decrease in the standard deviation of time-order deviation relative to the baseline. Further, incorporating time-order deviation bounds the loss in efficiency while remaining robust to the choice of λ_t. These observations suggest that time-order deviation may be a suitable fairness metric in practice.

5.2 Handling Dynamic Demand

We now turn to the case when we have dynamic demand in the form of pop-up flights. We experimented with several different pop-up fractions, horizon lengths, and TFMP objective functions. Note that in the absence of pop-ups, it is most efficient for the horizon to be as large as possible so that the TFMP has knowledge on as many flights as possible. However, with a larger horizon, flights are forced to file their flight plans earlier to avoid being pop-ups.

Recognizing that pop-up flights are difficult to eliminate, we tested two approaches to handling pop-ups: Option 1 (insert pop-ups) and Option 2 (delay pop-ups). We are interested in the trade-offs between these different parameters. The following results are with the baseline TFMP objective. As expected, we find that as the pop-up fraction increases, efficiency and fairness decrease. Fig. 5 shows the average total delay cost per flight across the two pop-up options, two horizon lengths (5 and 30 min) in scenarios with pop-up fraction 0.1 and 0.5. Each bar is an average of 100 runs. We start with a pop-up fraction of 0.1. We see that Option 1 with a 30-minute horizon performed the best overall (had the lowest average total delay cost per flight). Option 2 with a 30-minute horizon performed poorly because of the high delay assigned to pop-ups that need to wait until the next horizon. With a 5-minute horizon, Option 1 and Option 2 performed similarly to each other, with Option 2 having a very slight edge. Fig. 6 is arranged similarly to Fig. 5 but shows average reversals per flight rather than average total delay cost per flight. While inserting pop-ups with a large horizon (Option 1 with 30-minute horizon) had the best efficiency, delaying pop-ups with a small horizon (Option 2 with 5-minute horizon) had the fewest reversals per flight.

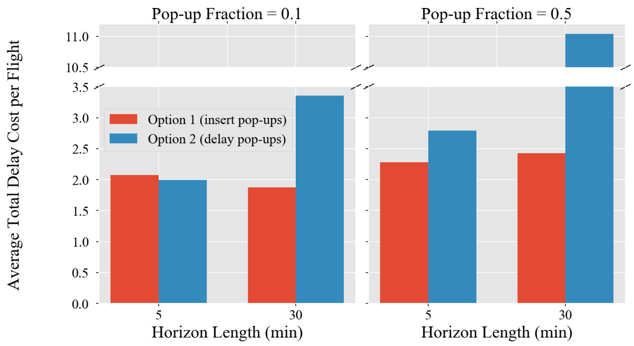

Fig. 5 Average Total Delay Cost per Flight by Horizon Length and Pop-up Option.

Moving on to pop-up fraction 0.5, Option 1 with the 5-minute horizon had the lowest total delay cost and average reversals. In contrast to with pop-up fraction 0.1 and a 5-minute horizon, Option 2 had a higher delay cost than Option 1. In addition, Option 1 with the 30-minute horizon performed worse than with the 5-minute horizon. The best combination was Option 1 with a 5-minute horizon. With more pop-ups, the TFMP horizon needs to be smaller to reduce the number of non pop-ups that are scheduled before (and thus block) each pop-up. Consider a pop-up with desired departure time of 9:02 and a horizon with a start time h_f of 9:00. With a 30-minute horizon, non pop-ups scheduled to depart between 9:00 and 9:30 will be scheduled before the pop-up, but with a 5-minute horizon, only non pop-ups scheduled to depart between 9:00 and 9:05 will block the pop-up. There is a break in the y-axis of Fig. 5 and Fig. 6 because Option 2 with a 30-minute horizon has such high delay. We saw with pop-up fraction 0.1 that Option 2 does not pair well with a large horizon because of the delay that pop-ups are forced to incur before they are scheduled. This trend is further highlighted with pop-up fraction 0.5.

We saw similar trends when incorporating reversals, overtaking, or time-order deviation in the objective. Thus, we do not show versions of Fig. 5 and Fig. 6 for these objectives. Instead, Table 5.2 shows the effect of incorporating reversals on total delay cost and reversals. Each row corresponds to a combination of pop-up fraction, horizon length, and pop-up option. Note that the two combinations that were shown to be impractical are not included (30-minute horizon with Option 2 for pop-up fraction 0.1; 5 or 30-minute horizon with Option 2 for pop-up fraction 0.5). Table 5.2 is similar to Table 5.2 but shows the effect of incorporating time-order deviation in the objective.

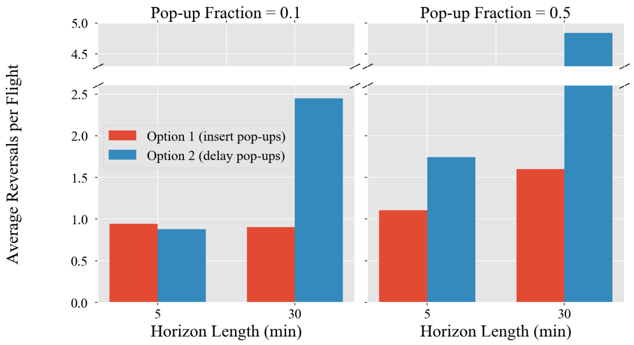

Fig. 6 Average Reversals per Flight by Horizon Length and Pop-up Option.

Parameters			**Penalizing Reversals**	
Horizon Length (min)	**Pop-up Option**	**Pop-up Fraction**	**% Change Total Delay Cost**	**% Change Reversals**
5	1	0.1	0.17	-26.8
5	2	0.1	4.54	-26.6
30	1	0.1	6.96	-11.8
5	1	0.5	0.12	-14.9
30	1	0.5	0.57	-2.34

Table 2 Change in total delay cost and reversals when incorporating reversals in the TFMP. Option 1 is inserting pop-ups; option 2 is delaying pop-ups.

The first takeaway is that even in the rolling horizon setting, incorporating reversals/time-order deviation is effective in improving fairness. Next, we observe that fairness improves by a larger relative amount with 0.1 pop-up fraction than with 0.5 pop-up fraction. All else equal, from an efficiency and fairness perspective, the system benefits from having fewer pop-ups. We also note that we get larger improvements in fairness with a 5-minute horizon rather than a 30-minute horizon. With a longer horizon, pop-ups are scheduled after more non pop-ups (in Option 1 and Option 2), and forced to delay longer to wait for the next horizon (in Option 2 only).

Parameters			**Penalizing Time-order Deviation**	
Horizon Length (min)	**Pop-up Option**	**Pop-up Fraction**	**% Change Total Delay Cost**	**% Change Time-order Deviation**
5	1	0.1	3.56	-7.18
5	2	0.1	4.09	-14.83
30	1	0.1	4.60	-1.71
5	1	0.5	3.36	-11.73
30	1	0.5	2.24	-5.56

Table 3 Change in total delay cost and TOD when incorporating TOD in the TFMP

6 Conclusions

This paper explores incorporating fairness metrics in UTM. Before deciding the extent to which fairness is implemented, it is important to choose a metric that defines fairness. From our analysis, time-order deviation appears to be a promising metric for fairness as it is robust to the specific choice of the λ penalty, strives for equality in a relative sense rather than on an absolute scale, and does not significantly compromise efficiency. However, it is also worth remembering that a significant fraction of the improvements in fairness and reversals can be obtained for a small penalty in delays if the appropriate λ is chosen. The UTFM framework can be used to evaluate the centralized efficiency and fairness of any trajectory set (e.g., trajectories with different demand profiles or "geofenced" airspace restrictions).

We also tested our formulations in a rolling horizon framework where not every flight files their flight plan sufficiently in advance to be considered in the time horizon that contains their scheduled departure time. We considered two options for handling pop-ups: inserting them into the schedule, and delaying them until the next horizon. We found that with a low occurrence of pop-ups, either using longer horizons and inserting pop-ups, or using shorter horizons with either pop-up option, are acceptable. However, with high occurrences of pop-ups, it is best to use a short horizon and insert pop-ups. In our experiments found that it was beneficial to incorporate fairness into the rolling horizon framework.

Several promising directions for further research remain. This paper did not consider rerouting, which is common in commercial aviation and will likely occur in UTM as well. Airborne flights could choose to reroute around congested areas, or flights on the ground could choose to alter their route from the start to reduce the incurred delay. Approaches like trajectory option sets (TOS) exist for commercial aviation wherein each flight submits a set of unique trajectories along with acceptable delays, but such research for UTM is less explored. In addition, the inputs to a UTM optimization may not be specific trajectories but may be volumes of airspace that operators reserve. There are several related research questions on how to manage airspace reservations, requirements, and constraints on requests for airspace, and fairness between different-sized operators. Finally, with the rolling horizon framework, pop-ups are at a disadvantage because flights with longer file-ahead times are allocated resources before them. In previous work, researchers have suggested that limiting the time in advance that resources can be allocated/reserved would improve fairness between early-filers and late-filers [20]. The design of such a system, and the evaluation of its efficiency and fairness, remain open challenges.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. MM Doole, J Ellerbroek, and JM Hoekstra. Drone delivery: Urban airspace traffic density estimation. In *8th SESAR Innovation Days*, 2018.
2. K. Balakrishnan, J. Polastre, J. Mooberry, R. Golding, and P. Sachs. Blueprint for the Sky: The roadmap for the safe integration of autonomous aircraft. Technical report, Airbus UTM, 2018.
3. J. Rios. Strategic deconfliction: System requirements. *NASA UAS Traffic Management (UTM) Project*, 2018.
4. D. Bertsimas, G. Lulli, and A. Odoni. An integer optimization approach to large-scale air traffic flow management. *Operations research*, 59:211–227, 2011.
5. D. Bertsimas and S. Gupta. Fairness and collaboration in network air traffic flow management: an optimization approach. *Transportation Science*, 50(1), 2015.
6. C. Barnhart, D. Bertsimas, C. Caramanis, and D. Fearing. Equitable and efficient coordination in traffic flow management. *Transportation Science*, 46(2), 2012.
7. O. Richetta and A. Odoni. Dynamic solution to the ground-holding problem in air traffic control. *Transportation Research Part A: Policy and Practice*, 28(3): 167–185, 1994.
8. P. Vranas, D. Bertsimas, and A. Odoni. The multi-airport ground-holding problem in air traffic control. *Operations Research*, 42(2):249–261, 1994.
9. D. Bertsimas and S. Patterson. The air traffic flow management problem with enroute capacities. *Operations research*, 46(3):406–422, 1998.
10. H. Balakrishnan and B. Chandran. A distributed framework for traffic flow management in the presence of unmanned aircraft. In *USA/Europe Air Traffic Management R&D Seminar*, June 2017.
11. B. Kotnyek and O. Richetta. Equitable models for the stochastic ground-holding problem under collaborative decision making. *Transportation Science*, 40(2), 2006.
12. T. Vossen, M. Ball, R. Hoffman, and M. Wambsganss. A general approach to equity in traffic flow management and its application to mitigating exemption bias in ground delay programs. *Air Traffic Control Quarterly*, 11(4):277–292, 2003.
13. O. Rodionova, H. Arneson, B. Sridhar, and A. Evans. Efficient trajectory options allocation for the collaborative trajectory options program. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference*. IEEE, 2017.
14. I. del Pozo de Poza, M. Vilaplana Ruiz, and C. Goodchild. Assessing fairness and equity in trajectory-based operations. In *9th AIAA ATIO Conference*, 2009.
15. H. Idris, C. Chin, and A. Evans. Accrued delay application in trajectory-based operations. In *USA/Europe Air Traffic Management R&D Seminar*, 2019.
16. A. Montlaur and L. Delgado. Flight and passenger efficiency-fairness trade-off for ATFM delay assignment. *Journal of Air Transport Management*, 83: 101758, 2020.
17. C. Chin, K. Gopalakrishnan, M. Egorov, A. Evans, and H. Balakrishnan. Efficiency and fairness in unmanned air traffic flow management. *IEEE Transactions on Intelligent Transportation Systems*, 20(9), 2021.
18. P. Kopardekar, J. Rios, T. Prevot, M. Johnson, J. Jung, and J. Robinson. Unmanned aircraft system traffic management (UTM) concept of operations. In *16th AIAA ATIO Conference*, 2016.
19. B. Skorup. Auctioning airspace. *Mercatus Center*, 2018.
20. A. Evans, M. Egorov, and S. Munn. Fairness in Decentralized Strategic Deconfliction in UTM. In *AIAA Scitech 2020 Forum*, page 2203, 2020.
21. A. Nilim and L. El Ghaoui. Algorithms for air traffic flow management under stochastic environments. In *Proceedings of the 2004 American Control Conference*, volume 4, pages 3429–3434. IEEE, 2004.
22. G. Andreatta, P. Dell'Olmo, and G. Lulli. An aggregate stochastic programming model for air traffic flow management. *European Journal of Operational Research*, 215(3):697–704, 2011.
23. A. Agustı, A. Alonso-Ayuso, L. Escudero, and C. Pizarro. On air traffic flow management with rerouting. part ii: Stochastic case. *European Journal of Operational Research*, 219(1):167–177, 2012.
24. KKH Ng, CKM Lee, F. Chan, and Y. Qin. Robust aircraft sequencing and scheduling problem with arrival/departure delay using the min-max regret approach. *Transportation Research Part E: Logistics and Transportation Review*, 106:115–136, 2017.
25. KKH Ng, CH Chen, and CKM Lee. Mathematical programming formulations for robust airside terminal traffic flow optimisation problem. *Computers & Industrial Engineering*, 154:107119, 2021.
26. J. Chen, L. Chen, and D. Sun. Air traffic flow management under uncertainty using chance-constrained optimization. *Transportation Research Part B: Methodological*, 102:124–141, 2017.
27. M. Egorov, V. Kuroda, and P. Sachs. Encounter aware flight planning in the unmanned airspace. In *Integrated Communications, Navigation & Surveillance Conf.*, 2019.
28. Y. Liu and M. Hansen. Evaluation of the Performance of Ground Delay Programs. *Transportation Research Record*, 2400(1):54–64, January 2014.