Abstract—This paper introduces a strategy to include linear holding into air traffic flow management (ATFM) initiatives, together with the commonly used ground holding and airborne holding measures. In this way, ATFM performance can be improved when handling delay assignment with uncertainty. Firstly, a trajectory generation method is presented, aiming at computing, per flight, the maximum linear holding realizable using the same fuel as the original nominal flight. This information is assumed to be computed and shared by the different airlines and it is then used to build a network ATFM model to optimally assign ATFM delays, in the scope of trajectory based operations. Hence, the best distribution of delay is optimized at given positions along the flight trajectory (combining the three holding practices together) and taking into account the cost of delay, especially in the fuel consumption. The problem is formulated as a mixed integer linear programming and solved with a commercially-of-the-shelf solver. An illustrative example is given, showing that under the circumstance of capacity recovered ahead of schedule, including linear holding contributes to a notable delay reduction compared to the case where only ground and/or airborne holding apply.

Keywords: linear holding, speed reduction, trajectory optimization, delay assignment, air traffic flow management

I. INTRODUCTION

Air traffic flow management (ATFM) refers to processes of a more strategic nature, involving taking a higher-level view of the overall air traffic rather than controlling specific flights. It detects and resolves demand-capacity imbalances, smoothing aggregate traffic flows and keeping the workload of air traffic control (ATC) under manageable levels. For practical reasons, the institution in charge of ATFM (known as the Network Manager) cannot take care of the specific preferences of one particular flight, since the overall objective of ATFM is typically to reach a global optimum based on some unanimous fairness criteria.

With the paradigm shift for the future air traffic management (ATM) proposed by SESAR and NextGen programs, with a transition from airspace based operations to trajectory based operations (TBO), the airspace users will be expected to increasingly participate in ATM decisions using, in particular, more collaborative decision making (CDM) mechanisms. For example, the SESAR concept of reference business trajectory (RBT), as output of an ATFM negotiation, is the trajectory that the airspace user agrees to fly and the ANSP (air navigation service provider) and airports agree to facilitate.

Under current Ground Delay Programs (GDPs), resources (i.e., arrival slots) are assigned to flights in accordance with a ration-by-schedule (RBS) mechanism (first-scheduled first-served prioritization). It is accompanied with CDM initiatives, such as flight substitution, cancellations, compression, or slot credit substitution, allowing airlines to manage their own flights in line with their specific policies [1]. Following this thought, we could imagine that airlines would be willing to provide specific flight information to the Network Manager in order to improve the performance of the ATFM regulations, especially if potential benefits might exist (e.g., reducing flight delays) along with feasible negotiation mechanisms (e.g., the TBO paradigm).

In this paper, we discuss the applicability of including linear holding (LH), as a specific information provided by airlines to the Network Manager, in the ATFM delay assignment process and aiming at increasing the flexibility when handling delays and improving the performance of ATFM regulations.

The LH strategy was proposed in [2], aimed at partially incurring in the air (by flying slower) the initially assigned ground delays. This strategy was further explored in [3], where aircraft were allowed to cruise at the lowest possible speed at which the fuel consumption remained exactly the same as initially planned. In this situation, if the delays were canceled ahead of schedule, aircraft already airborne and flying slower, could speed up to the initially planned speed and recover part of the delay without extra fuel costs [4]. Recently, an aircraft trajectory optimization technique was adopted in [5], where the whole flight profile was subject of optimization to maximize the achievable LH at no extra fuel cost.

As the core method to perform LH, the speed adjustment has proven successful for several ATM scenarios. For instance, a speed control approach was presented in [6] to transfer delay away from the terminal to the en route, from which significant fuel saving on a per flight basis was also yielded. In [7], a pre-tactical speed control was applied en route to prevent aircraft from performing airborne holding when arriving at a congested airspace. Similar, but more at the tactical level, aircraft in [8] were required to reduce their speed to avoid arriving at the airport before its opening time to reduce unnecessary holdings.
On the other hand, following the pioneering work done in [9], a number of researchers have focused their activity on the development of optimization models for the delay assignment as a short-term measure for traffic flow regulation (see [10] for instance). Further taking into account the capacitate airspace, the problem of controlling release times and speed adjustments of aircraft while airborne for a network of airports (including sectors) has been studied in [11], [12].

The network ATFM model in this paper is based on the above mentioned Bertsimas and Stock-Patterson model [11], which has been widely studied in the last decades. Up to the best of our knowledge, however, among past research less discussion on the operating cost impact of such speed adjustments as delay absorption measure has been done. Given the fact that aircraft speeds are intimately related to fuel consumption, the main contribution of this paper is to include a cost-based LH in this well-known model, and exploit its flexibility to improve ATFM performance.

II. LINEAR HOLDING PRACTICES

In this section the linear holding concept is explained and a methodology, based on trajectory optimization algorithms, is outlined with the purpose to compute the maximum linear holding time without incurring into extra fuel consumption.

A. Linear holding concept

Fig. 1 highlights schematically the main characteristics of the linear holding together with the two commonly used holding practices in current ATM: ground and airborne holding.

Typical airborne holding would consume more fuel due to the extended flight track (assuming no specific speed adjustment), whilst ground holding has no impact in fuel consumption. Due to the increased extra fuel, the airborne holding time is fairly limited, taking account that safety related issues may arise from a reduction of the on-board reserve fuel.

Airlines consider direct operating costs (DOC) when planning their flights [13], which besides fuel consumption also take into account time-related costs. In this context, on-board flight management systems (FMS) allow them to optimize trajectories by means of a Cost Index (CI) input parameter, which expresses the ratio between time-related costs and the cost of fuel. Flying at a CI greater than zero (typically the case) will result into a speed greater than the maximum range speed, since time savings will be also considered.

Notionally, an equivalentairspeed \( V_{eq} \) can be defined by that speed, lower than the speed initially planned for the nominal flight \( V_{nom} \), which produces the same fuel consumption than flying at \( V_{nom} \) (see Fig. 2). Thus, by flying between these two speeds the fuel burnt will not exceed the quantity initially planned, while some LH could be realized.

The amount of LH that can be realized depends on several factors, such as the aircraft type, trip distance, payload, cruise flight level, etc. In previous work [5], this topic has been discussed, exploring the maximum achievable LH for some particular flights on the previous premise of no extra fuel consumption.

From an implementation point of view, ground holding can only be performed at the departure airport, prior to take-off. Airborne holding (including holding patterns or path stretching) can be done at any available airspace, in theory, but practically it is typically performed in designated locations.

The most promising feature of LH is that delay absorption can be flexibly managed through proper speed adjustment along the original planned route, and without incurring extra fuel burnt than initially scheduled if flying between \( V_{eq} \) and \( V_{nom} \). It should be noted that LH could also be performed by burning more fuel, as shown in [3]. Yet, it is out of the scope of this paper to explore the benefits of this strategy.

To see the potential applicability of LH, let us imagine a flight assigned with a certain delay as a result of a GDP, as shown in Fig. 3. In the near future, we could assume that controlled times of arrival (CTA) could be enforced at the destination airport, in order to guarantee the arrival slot allocation computed by the GDP [6]. Our flight could absorb all delay by means of ground holding, as currently done (orange line in Fig. 3); or perform less ground holding but some LH in the air, such that the CTA is still met (blue solid line). If we assume that at some point the GDP is canceled, due to weather improvement for instance (purple line), airborne aircraft could stop the LH, accelerate to the nominal speed, and recover part of its delay (blue dash line). Obviously, if the CTA is not changed, aircraft will finally arrive at the destination airport with the same amount delay than the case where all
delay is served on the ground.

B. Trajectory optimization for linear holding computation

The optimization of an aircraft trajectory requires the definition of a mathematical model representing aircraft dynamics, along with a model for certain atmospheric parameters. In this paper, a point-mass aircraft model, an enhanced performance model using manufacturer certified data, and the International Standard Atmosphere (ISA) have been considered. For more details, the readers may refer to [14].

A generic vertical trajectory can be partitioned into several segments \( i \in \{1, \cdots, N\} \), where different constraints or models may apply. Fig. 4 shows the different segments considered in this paper, where the initial and final points are taken, respectively, at the moment the slats are retracted (after take-off) and extended (before landing).

Before considering LH, the nominal trajectory (i.e., that one initially scheduled by the airline) is generated first. In line with the discussion in Sec. II-A, the optimal trajectory will be that one minimizing a compound cost function \( J \) over the whole time window \( [t^0_i, t^N_i] \), as follows:

\[
\min J = \min \int_{t^0_i}^{t^N_i} (FF(t) + CI) dt,
\]

where \( FF(t) \) is the fuel flow and \( CI \) is the Cost Index.

This optimization problem is subject to several constraints, including the aircraft dynamics, in form of differential equations on the aircraft states; some algebraic event constraints, fixing initial/final states; path constraints, in order to fulfill ATM and operational constraints along the different flight phases (see Fig. 4); and finally, link constraints, to ensure continuity along the flight segments. Then, a non-linear programming (NLP) problem can be formulated and solved with a commercially-of-the-shelf solver. A detailed description of all these constraints and a complete formulation of the problem is given in [14].

Then, based on the nominal trajectory, by changing the optimization objective and constraints, the trajectory can be modified performing some LH. As an example, we will illustrate the case where the maximum LH is realized at no extra fuel cost. The objective function of the problem (1) is replaced by (2), where the total flight time is maximized, whilst subject to an additional constraint on the total fuel consumption, as depicted by (3):

\[
\max J = \max \int_{t^0_i}^{t^N_i} dt,
\]

\[
\text{s.t. } \int_{t^0_i}^{t^N_i} FF(t) dt \leq F_{\text{nom}},
\]

where \( F_{\text{nom}} \) is the fuel consumed for the nominal trajectory.

This makes it clear that the flight as a whole is optimized rather than the climb, cruise or descent phases separately (or even the subdivided flight segments within a particular phase). In this paper, however, the aircraft trajectory is allowed to update tactically (i.e., halfway en route) in response to possible changes of ATFM regulations (see Fig. 3). In order to keep fuel consumption similar to the nominal flight, however, the mass of the aircraft \( M \) is fixed for each trajectory segment \( i \) defined in the vertical flight profile of Fig. 4:

\[
M_{i, LH} = M_{i, nom}, i = CL_1, \cdots, DE_4.
\]

Otherwise, more fuel might be burnt during climb (in the LH trajectory), leaving less fuel available for updating the remaining flight phases (when situation improved).

In addition, the cruise flight level(s) and route should be fixed as well, as pre-tactical re-routings or flight level cappings, as part of a possible ATFM negotiation, are out of the scope of this paper:

\[
H^{CL_4}_{LH} = H^{CL_4}_{nom},\ H^{DE_1}_{LH} = H^{DE_1}_{nom},\ H^{CR_2m}_{LH} = H^{CR_2m}_{nom},
\]

where \( H \) denote the aircraft flight altitude.

Since the angle of climb (descent) varies with speed, the climb (descent) distances will be different at different speeds, meaning that the location of the top of climb (TOC) and top of descent (TOD) will also move. Moreover, the distance at which
each step climb (if any) is performed will not be enforced, considering that possible changes in the TOC and/or TOD could impact on the length of different cruise segments.

Finally, it should be noted that before each cruise flight level, a short cruise segment (with a maximum duration of 1 minute) is added in order to allow speed adjustments (see Fig. 4). A similar segment is added at the end of the last cruise phase in order to adjust to the optimal descent Mach. These segments help to reduce the excessive influences from the link constraints on the flight profile.

III. NETWORK ATFM MODEL WITH LINEAR HOLDING

A network ATFM model is proposed in this section, which assigns delays at designed positions. Ground and airborne holdings remain the default measure for delay assignment, while linear holding is possible for those airlines willing to participate in the ATFM decision making, by sharing certain information with the Network Manager.

A. Problem statement

Existing ATFM network models (see [11] and the references therein), delays could be assigned to flights by means of ground or airborne holding. Airborne holding, however, tends to be less weighted (especially when a long-time delay occurs) because of its higher fuel costs (and potential safety issues) as has been discussed in Sec. II-A.

In the model proposed in this paper, we maintain the above two holding practices, but adding the LH option. Nevertheless, aiming at differentiating the proposed LH with typical airborne holding both appearing in the model, the delay assignment is conducted at specific designed “positions” along the scheduled trajectory, by using the concept of Controlled Time of Arrival (CTA) and Controlled Time of Departure (CTD) at each position. Accordingly, the decision variables of the model are defined as follows:

\[ x_{j,f,t} = \begin{cases} 
1, & \text{if flight } f \text{ departs from the position } j \text{ at time } t \\
0, & \text{otherwise} 
\end{cases} \]

\[ y_{j,f,t} = \begin{cases} 
1, & \text{if flight } f \text{ arrives at the position } j \text{ at time } t \\
0, & \text{otherwise} 
\end{cases} \]

Fig. 5 shows schematically flight time versus distance and the three types of holding strategies are represented: ground holding is performed only at the origin airport; airborne holding can only be performed “at” a given position (the difference between the “departure” and “arrival” time at that position equals to the holding time); and since LH is performed by flying slower, the slope of the lines is increased if compared with the planned schedule.

Recall that we distinguish the typical airborne holding from LH by the fact that when performing the former, the actual flight distance will be extended (either by vectoring or using holding patterns). This flight path “stretching”, however, does not contribute to the execution of trajectory defined by each contiguous point. Thus, the typical airborne holding, on some level, can be seen as a “circling” at a particular position.

It is also worth noting that the “positions” referred here (such as P1, P2 and P3 in Fig. 5) may not correspond to the actual geographical waypoints existing in current airspace. The model in this paper defines entrance and exit positions at each elementary sector that the controlled flight is scheduled to traverse (as well as the two representing origin and destination airports respectively, as shown in Fig. 6).

In addition, the feasible time window shown in Fig. 5 defines a solution space based on the flight schedule, which could largely reduce the number of variables taken into optimization.

The set of contiguous sectors varies for different trajectories (see for instance Fig. 6). From each scheduled trajectory, we could sort all the positions (two for each sector) by scheduled times. Then, if there are duplicate positions (i.e., with the same scheduled time), only one of them is taken, while the others are removed (not for those belonging to the exceptions). As
such, the number of optimization variables could be reduced, given there exist some of sectors sharing the same frontier. Taking into account sector capacities, however, the duplicate positions must be identified by the model as exceptions in this process of variable reduction.

B. Participation of airlines in the ATFM process

As indicated from Sec. II-B, the amount of delay absorption that LH can realize is constrained by the fuel consumption, which again is dependent on aircraft type, take-off mass, flight distance, etc. Thus, from the ATFM perspective, considering all these data would be a daunting work. Moreover, some of the airline’s information is proprietary, such as aircraft mass and fuel consumption figures, which is normally not publicly accessible. From the airline perspective, however, they could have a clear view of all the information of their own flights, and thus have an intimate knowledge of the capability of each particular flight to absorb delays airborne.

Consequently, aiming at including the proposed LH into delay assignment, the ATFM model in this paper requires (only) one more input from airlines than those models found in the literature: the maximum LH bound per flight, along the planned trajectory. In case that no data for this input is provided, the model could still work by setting the default value for LH to zero.

Let us first take a look on what could be found from current flight planning information, as provided by the demand data repository v2 (DDR2) published by EUROCONTROL. Table I presents the detailed information obtained for a specific flight scheduled from Rome Fiumicino Airport (LIRF) to Amsterdam Schiphol Airport (EHAM) shown in Fig. 6. These data include “enter time”, “exit time”, “crossed duration” and “altitude” at each of the sectors (and airports) the aircraft is scheduled to fly, which correspond to the designed “positions” (sector boundaries and airports) used in our model.

Based on the planned vertical profile found in DDR2, the nominal optimal trajectory has been reconstructed, on one hand, using the trajectory optimization methodology described in Sec. II-B (red line in Fig. 7). On the other hand, the blue line in the Figure represents the same flight when performing the maximum amount of LH, while incurring the same fuel consumption during every single flight segment (see Fig. 4). Fig. 7 also shows the true airspeed (TAS) of both trajectories, where the airborne delay generation by means of LH can be easily seen.

Table II summarizes the trajectories of Fig. 7 in form of the current flight planning information (as Table I), but added with a “LH time” bound (see the rightmost column), which equals to the difference of crossed duration between the nominal trajectory and the LH trajectory (negative values appear in climb/descent because of the slightly trajectory differences caused from speed changes), and that is the one that should be provided (by airlines) to the Network Manager.

C. Model formulation

The network ATFM model with the above presented airline-enabled LH is formulated here. As mentioned before, the
TABLE II
Nominal and LH trajectories flight planning information.

<table>
<thead>
<tr>
<th>Position</th>
<th>Nominal optimal trajectory</th>
<th>Linear holding trajectory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enter Position</td>
<td>Exit Position</td>
</tr>
<tr>
<td></td>
<td>Dist (nm)</td>
<td>Time (hh:mm:ss)</td>
</tr>
<tr>
<td>LIRF</td>
<td>2</td>
<td>7:59:00</td>
</tr>
<tr>
<td>LIRRPN1</td>
<td>2</td>
<td>7:59:50</td>
</tr>
<tr>
<td>LIRRWN2</td>
<td>48</td>
<td>8:08:28</td>
</tr>
<tr>
<td>LIRRWN3</td>
<td>80</td>
<td>8:13:06</td>
</tr>
<tr>
<td>LIRRWN5</td>
<td>115</td>
<td>8:17:34</td>
</tr>
<tr>
<td>LIMMWS5</td>
<td>165</td>
<td>8:24:26</td>
</tr>
<tr>
<td>LIMMWN5</td>
<td>229</td>
<td>8:32:43</td>
</tr>
<tr>
<td>LFEEKE</td>
<td>479</td>
<td>9:05:46</td>
</tr>
<tr>
<td>EHAMCR1</td>
<td>796</td>
<td>9:51:44</td>
</tr>
</tbody>
</table>

overall framework of the model is based on the widely-studied Bertsimas and Stock-Patterson model [11].

Notation

- \( f \in F \): set of flights
- \( t \in T \): set of time moments
- \( \tau \in \mathcal{T} \): set of periods for traffic demand
- \( T(\tau) \): \( \tau \)-th time period defined within \( T \)
- \( k \in K \): set of airports
- \( w \in W \): set of sector entrance or exit positions
- \( j \in K \cup W \): set of positions
- \( s \in S \): set of sectors of capacity constrained
- \( P(f,i) \): the departure airport, if \( i = 1 \)
- \( P_f \): \{\( P(f,i) : i \in [1,n_f] \}\}, the positions of \( f \) along the scheduled trajectory in sequence
- \( R(f) \): \{\( R(f,i) : i \in [1,n_f] \}\}, the scheduled time of \( f \) in line with \( P(f) \)
- \( r_{j}^{f} \): the scheduled time of \( f \) at \( j \)
- \( T_{f}^{j} \): the feasible time window for \( f \) at position \( j \)
- \( e_{i,j}^{f} \): the length of feasible time window
- \( a_{i,j}^{f} \): the maximum airborne holding time
- \( z_{i,j}^{f} \): the maximum LH bound of contiguous positions
- \( q_{i,j}^{f} \): the minimum turnaround time
- \( D_{s}^{k}(\tau) \): the departure capacity of airport \( k \) in period \( \tau \)
- \( A_{s}^{k}(\tau) \): the arrival capacity of airport \( k \) in period \( \tau \)

1) Objective function: In this model, the cost of the total delay (TD) is minimized including the costs consequence of ground holding (GH), airborne holding (AH) and linear holding (LH):

\[
\min (cost_{TD}) = \min (GH + \alpha AH + \beta LH),
\]

where \( \alpha \) and \( \beta \) are the cost weighting factors. Since \( TD = GH + AH + LH \), we can substitute LH in (6), yielding to:

\[
\min (cost_{TD}) = \min (GD + (1 - \beta)AH + (1 - \beta)GH)).
\]

Taking into account the fairness of delay assignment, as discussed in [12], the total delay is multiplied by a coefficient \( c_{f} = (t - r_{k}^{f})^{1+\epsilon}, \epsilon > 0 \) in (7). In this way, delays will be assigned moderately across all the flights, instead of unevenly to one particular flight. Accordingly, the objective function can be arranged as:

\[
\min \sum_{f \in F} \left[ \beta c_{f} h_{f} + (\alpha - \beta)a_{f} + (1 - \beta)g_{f} \right],
\]

\[
c_{f} h_{f} = \sum_{t \in T^+_f, P(f,n_f) = k} (t - r_{k}^{f})^{1+\epsilon} y_{f,t},
\]

\[
a_{f} = \sum_{t \in T^+_f, w \in P(f,i): 1 < i < n_f} t(x_{w,f,t} - y_{f,t}),
\]

\[
g_{f} = \sum_{t \in T^+_f, P(f,1) = k} (t - r_{k}^{f})^{1+\epsilon} x_{w,f,t}. \]
2) Flight operations constraints:

\[ \sum_{t \in T_f^j} x^j_{f,t} = 1 \quad \forall f \in F, \forall j \in P_f, \] (9)
\[ \sum_{t \in T_f^j} y^j_{f,t} = 1 \quad \forall f \in F, \forall j \in P_f, \] (10)
\[ x^j_{f,t'} + y^j_{f,t'} \leq 1 \quad \forall f \in F, \forall i \in [1,n_f-1], \forall t \in T_f^j, \]
\[ \forall t' \in T_f^j - [t,t+w^w], \] (11)
\[ x^j_{f,t'} + y^j_{f,t''} \leq 1 \quad \forall f \in F, \forall i \in [1,n_f-1], \forall t \in T_f^j, \]
\[ P(f,i) = j. \] (12)
\[ \forall t'' \in T_f^j - [t+z_{f,j}^j,t+2z_{f,j}^j + v_{f,j}^j], \]
\[ x^k_{f,t''} + y^k_{f,t''} \leq 1 \quad \forall (f,j') \in F, \forall t \in T_f^k, \]
\[ P(f',j) = k, \] (13)
\[ \forall t'' \in T_f^k \cap [t,t+q_{f,j,k}]. \]

Constraints (9) and (10) ensure each flight is assigned with only one slot for departing and arriving, respectively, at position \( j. \)

Constraint (11) imposes a maximum airborne holding time \( w^w \) at each designated position.

Constraint (12) stipulates that the LH performed between two contiguous positions of flight \( f \) should not exceed the maximum LH bound \( v_{f,j}^j \), which is provided by airlines and that is set by 0 as default if no such information is acquired. At the same time it enforces that LH to be non-negative (i.e., flying faster than initially planned is not considered for delay assignment in this model).

Constraint (13) specifies that two connective flights should have at least a turnaround time \( t_{f,j}^j \) with regards to the arrival of flight \( f \) and the departure of \( j' \) at the airport \( k. \)

3) Network capacity constraints:

\[ \sum_{f \in F: P(f,1)=k} x^k_{f,t} \leq D^k(\tau) \quad \forall k \in K, \forall t \in T(\tau), \forall \tau \in \mathcal{T} \] (14)
\[ \sum_{f \in F: P(f,n_f)=k} y^k_{f,t} \leq A^k(\tau) \quad \forall k \in K, \forall t \in T(\tau), \forall \tau \in \mathcal{T} \] (15)
\[ \sum_{f \in F: P(f,i)=w, P(f,i+1)=w'} (x^w_{f,t} - x^w_{f,t}) \leq C^w(\tau) \] (16)
\[ \forall w \in S, \forall t \in T(\tau), \forall \tau \in \mathcal{T}, \forall i \in [1,n_f-1] \]

Constraints (14), (15) and (16) ensure that the traffic demand would not exceed the capacity of departure airport, arrival airport and en route sectors, respectively.

It is worth noting that the flight performing airborne holding in this model is counted within the boundary of its current sector (i.e., before departing the position). Since the capacity values are all defined within a period of time window, they are capable of being modified following the changes of the network environment, such as the improvement of weather conditions or traffic situations.

4) Constraints on decision variables:

\[ x^j_{f,t} \in \{0,1\} \quad \forall f \in F, \forall j \in P_f, \forall t \in T_f^j \] (17)
\[ y^j_{f,t} \in \{0,1\} \quad \forall f \in F, \forall j \in P_f, \forall t \in T_f^j \] (18)

Constraints (17) and (18) state that the decision variables of the model are binary.

Above all, the model can be modified to perform the iterative delay assignment. Assume at the start of the \((\tau + 1)\) th time period, i.e. \( t_\sigma \), the capacity changes from current status of the time period of \( T(\tau) \), and requires for another round of delay assignment. We could simply fix part of the decision variables based on the current results, and optimize the rest of them in the next round of delay assignment.

5) Constraints from updating assignment:

\[ x^j_{f,t}(\tau'+1) = x^j_{f,t}(\tau') \quad \forall f \in F, \forall j \in P_f, \]
\[ \forall t \in T_f^j \cap [t_1,t_\sigma], \tau' \geq 1 \] (19)
\[ y^j_{f,t}(\tau'+1) = y^j_{f,t}(\tau') \quad \forall f \in F, \forall j \in P_f, \]
\[ \forall t \in T_f^j \cap [t_1,t_\sigma], \tau' \geq 1 \] (20)
\[ x^j_{f,t}(\tau'+1) = x^j_{f,t}(\tau') \quad \forall f \in F, \forall t \in T_f^j, \]
\[ t_\sigma \in (t_f^j(\tau'),t_f^j(\tau')) \}, \tau' \geq 1 \] (21)
\[ y^j_{f,t}(\tau'+1) = y^j_{f,t}(\tau') \quad \forall f \in F, \forall t \in T_f^j, \]
\[ t_\sigma \in (t_f^j(\tau'),t_f^j(\tau')) \}, \tau' \geq 1 \] (22)
\[ \sum_{t \in T_f^j: P(f,i)=j, \forall i \in [1,N_f]} tx^j_{f,t}(\tau') = t_f^j(\tau') \] (23)
\[ \sum_{t \in T_f^j: P(f,i+1)=j', \forall i \in [1,N_f]} tx^j_{f,t}(\tau') = t_f^j(\tau') \]

Constraints (19) and (20) enforce that values, prior to time \( t_\sigma \), of the decision variables \((x^j_{f,t}(\tau') \text{ and } y^j_{f,t}(\tau'))\) derived from the first round of optimization should be assigned to those new decision variables \((x^j_{f,t}(\tau'+1) \text{ and } y^j_{f,t}(\tau'+1))\) defined in the same domains \((f,j \text{ and } t)\); \( t_1 \) means the initial time of \( T \); while \( t_\sigma \) represents the initial time defined in the \((\tau + 1)\)th time period \( T(\tau' + 1) \).

Constraints (21) and (22) stipulate that for specifically the flights in the air at time \( t_\sigma \), the new decision variables subject to the second round of optimization must start from the next position after finishing their current flight segment linked by \((j,j')\). It is because the remaining distance within the segment might be not long enough to realize the amount of LH previously provided by airlines, which, however, is based on the calculation by en entire segment. \( t_f^j(\tau') \) and \( t_f^j(\tau') \) are the last assigned departure times for flight \( f \), and can be derived from (23).
IV. ILLUSTRATIVE EXAMPLE

An illustrative example of the methodology introduced in this paper is presented in this section. As stated in [12], network formulations present significant challenges in computational tractability, while some studies have focused exclusively on addressing the computational challenges of the network problem (see for instance [16]). Yet, given the fact that the main objective of this paper is to primarily reveal the effects of including the linear holding practice in ATFM delay assignment, more than improving the model computational performance, we have taken only a small sample into the experiments to analyze the possible results. GAMS has been used as the modeling tool and Xpress v23.01 optimizer bundled into the GAMS suite has been used as the solver.

A. Case of study setup

The sample data chosen involve 156 flights heading towards EHAM airport, as shown in Fig. 8, with their Estimated Times of Arrival and Estimated Times of Off Block both scheduled within the period from 6am to 12am on October 24th 2016. Initial flight schedules and elementary sector crossings have been taken from the DDR2. The departure airports and different sectors modeled are, respectively, 84 and 3555.

For the initial delay assignment, we assume four hot spots: EHAM airport and sectors EDDDLALL1, LFEEEKHRZIU and LFEEEUXE; where the demand exceeds the capacity during the studied period. The pre-regulation and post-regulation traffic demands per hour versus capacity are shown in Fig. 9, for these four spots.

Furthermore, we have also considered a situation where an early capacity recovery occurs at 9am, well before scheduled (12am), for the above four hot spots, which leads to an update of the delay assignment. It is assumed that the updating can be initiated at once while flights can receive and execute immediately the latest delay assignment.

Some other key assumptions have been taken in the computation: 1) the discrete time interval is set to 1min; 2) $\epsilon = 0.05$ is selected as the fairness factor; 3) the cost weights for airborne holding and linear holding are, respectively, $\alpha = 1.2$ and $\beta = 0.8$ with regards to the ground holding; and 4) the LH time bound is approximated as 20% of the planned total trip time [5], if no otherwise specified, and are all shared by airlines to the Network Manager.

B. Results of the delay assignment

Figures 10(a) and 10(b) show how in the initial process of delay assignment, part of the ground holding (and airborne holding) is replaced by LH. Referring to Table III, we can see the total delay has a reduction of 120 minutes after this replacement. This is because, including LH means that more space and periods can be used to absorb delays, rather than only at the departure airports prior to take-off. As a consequence, if multiple node constraints occur at the same time, separating delays at different places and periods would
TABLE III
SUMMARIZED RESULTS FOR THE FOUR CASES OF STUDY.

<table>
<thead>
<tr>
<th>Cases of study</th>
<th>Total delay (min)</th>
<th>Delayed flights (a/c)</th>
<th>Av. total delay (min)</th>
<th>Total AH (min)</th>
<th>Total GH (a/c)</th>
<th>GH flights (a/c)</th>
<th>Av. GH (min)</th>
<th>Total LH (min)</th>
<th>LH flights (a/c)</th>
<th>Av.LH (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial delay assignment with no LH</td>
<td>2421</td>
<td>86</td>
<td>28.15</td>
<td>3</td>
<td>2418</td>
<td>86</td>
<td>28.12</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Initial delay assignment by using LH</td>
<td>2301</td>
<td>97</td>
<td>23.72</td>
<td>1</td>
<td>1681</td>
<td>72</td>
<td>23.35</td>
<td>619</td>
<td>96</td>
<td>6.45</td>
</tr>
<tr>
<td>Updated delay assignment with no LH</td>
<td>1369</td>
<td>66</td>
<td>20.74</td>
<td>2</td>
<td>1367</td>
<td>66</td>
<td>20.71</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Updated delay assignment by using LH</td>
<td>499</td>
<td>38</td>
<td>13.13</td>
<td>0</td>
<td>370</td>
<td>27</td>
<td>13.70</td>
<td>129</td>
<td>24</td>
<td>5.38</td>
</tr>
</tbody>
</table>

Fig. 11. Effects of delay recovery for the flight (LIRF-EHAM) performing LH halfway in the air when encountering the update of delay assignment.

count to reducing the minimum delay required from multiple constraints. Moreover, we can also notice that more flights are included to share the reduced total delay, leading to an even lower average delay for each flight (see Table III).

Figures 10(c) and 10(d) illustrate the case when the early capacity recovery occurs, assuming that the new round of ATFM delay assignment starts immediately after this recovery. Obviously, the flights that have not been serving the ground holding, or have been holding on the ground halfway, can request for an immediate departure, and thus, have their delays (partially) recovered, as revealed by results of Fig. 10(c).

When implementing LH, however, the remaining total delay reduces remarkably once the delay assignment is updated (see Fig. 10(d)). There are two main reasons that could account for these promising results. First, benefiting from the shortening of ground holding, the departure time of one flight can be advanced. Once the delay is updated, less ground holding, and thus less total delay will be realized, as exactly is the case shown in Fig. 10(c). Since most of the flights are observed to substitute part of the ground holding by LH (see Table III), the effects can be enlarged notably. The second reason is because the flexibility of LH compared to the ground holding in terms of delay absorption, as mentioned previously in Sec. II-A. Regarding this scenario, a detailed analysis is given in the next section.

C. Delay recovery for an airborne flight performing LH

In this section, the same flight (LIRF-EHAM) introduced in Sec. III-B is analyzed in detail. During the initial process of delay assignment, this particular flight is allocated with 41mins of delay in total (imposed on the arrival slot), while allowed to wait on the ground for 22mins (i.e., ground holding) but flying slower to absorb the rest of the delay, i.e., 19mins, by means of LH in the air (if no any update occurs later).

After serving 22mins of ground holding followed by encountering the update of delay assignment, at 9am, as shown in Fig. 11(a), the flight starts to recover its nominal trajectory. The process initiates when passing the next designed position.
(229nm) compared to the flight’s current geographical position. Therefore, the updated timeline (the red line) deviates from the actual one performing an amount of LH (19mins, the blue line) which is lower than its maximum LH (25mins, the grey line), and becomes flatter as exactly parallel with the nominal timeline (the black line) for the remaining distance. At the end, there are 14mins of delay saved, reducing the total delay from 41mins to 27mins.

As we can see from Fig. 11(b), the TAS of the actual trajectory performing 19mins of LH lies between the nominal TAS and the one having the maximum LH. Interestingly, since delays are not assigned evenly along the trajectory, we may notice that the actual speed (the blue line) changes progressively during the cruise phase (due to the discrete time step is 1min in this paper), which may result to an increasing of the workload of flight crews. However, given the Required Time of Arrival (RTA) featured in modern on-board FMS, and aimed at autopilot when performing LH, it might not raise too much concerns on the procedures. As for the vertical trajectory illustrated in Fig. 11(c), caused from the changes of climb and descent speeds seen in Fig. 11(b), the geographical positions of TOC and TOD vary from the nominal trajectory.

At last, as shown in Fig. 11(d), recall again that the difference of LH with respect to typical airborne holding locates at whether the extra fuel needs to be consumed. Obviously, without this premise, LH might not be favored by airlines, given a safer and cheaper ground holding is always there. As a consequence, by restricting the fuel along the whole trajectory when optimizing it for LH, the fuel consumption can be constrained to the nominal one. Note, however, that due to limitations of the model (fuel consumption is enforced for each discrete flight segment), the fuel consumption will not be exactly the same as the nominal one and small differences can be appreciated.

V. CONCLUSIONS

In this paper, a cost-based LH practice was merged into the ATFM regulation for delay absorption, together with the commonly seen ground and airborne holdings. In the light of trajectory based operations, airlines’ sharing of maximum LH bounds derived from their own optimal aircraft trajectory generation, could be effectively utilized by the ATFM side as one of the optimization factors considered for delay assignment.

Incorporating the LH means that more space and periods in the network could be used to absorb delays. Provided multiple node constraints occur at the same time, splitting delays at different places and times could contribute to reducing the minimum system delay required from multiple constraints.

Results suggest that benefitting from the replacing of ground holding by LH, the release times can be advanced. Once the delay updated raised from the improvement of network situation, less ground holding, and thus less total delay would be eventually realized. Moreover, if the delays are canceled ahead of schedule, aircraft already airborne and performing LH, could accelerate to the speed as initially planned and recover part of the delay at no extra fuel cost.

ACKNOWLEDGMENT

The authors would like to thank Airbus Industrie for the use of PEP (Performance Engineers Program) suite, which allowed us to undertake realistic aircraft performances simulations. This research is partially supported by grants from the Funds of China Scholarship Council (201506830050) and by a La Caixa-UPC International Mobility Scholarship (Y4262159C).

REFERENCES


AUTHOR BIOGRAPHIES

Yan Xu is currently a Ph.D. candidate in Aerospace Science and Technology at the Technical University of Catalonia. He received his M.Sc. and B.Eng. in Traffic Engineering from Nanjing University of Aeronautics and Astronautics in 2015 and 2012. Email address: yan.xu@student.upc.edu

Dr. Xavier Prats is an Serra i Hunter fellow at the Technical University of Catalonia (UPC). M.Sc. degree in aeronautical engineering from the École Nationale de l’Aviation Civile (ENAC) in Toulouse (France) and Ms. Sc. degree in telecommunications engineering from Telecom Barcelona. Furthermore, he received his Ph.D. in Aerospace Science and Technology from UPC-BarcelonaTech. His main research interests include improving the performance and efficiency of aircraft operations as well as the air traffic management (ATM) system. He co-founded the Icarus research group at UPC and he is currently leading the air transportation research activities within it. Email address: xavier.prats@upc.edu