uc3m | Universidad Carlos III de Madrid

e-Archivo

This is a postprint version of the published document at:

Soler, M., Zou, B. y Hansen, M. (2014). Flight Trajectory Design in the Presence of Contrails: Application of a Multiphase Mixed-Integer Optimal Control Approach. *Transportation Research Part C: Emerging Technologies*, 48, pp.172-194.

DOI: https://doi.org/10.1016/j.trc.2014.08.009

© 2014 Elsevier Ltd. All rights reserved.

This article is licensed under a under a Creative Commons Attribution Non-Commercial No Derivatives License 4.0 International License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Flight Trajectory Design in the Presence of Contrails: Application of a Multiphase Mixed-Integer Optimal Control Approach

Manuel Soler, Bo Zou, Mark Hansen[a,b,c]

[a]*Department of Bioengineering and Aerospace Engineering, Universidad Carlos III, Madrid, Spain. e-mail: masolera@ing.uc3m.es*
[b]*Department of Civil and Materials Engineering, University of Illinois at Chicago, Illinois, USA. e-mail: bzou@uic.edu*
[c]*Department of Civil and Environmental Engineering, University California at Berkeley, California, USA. e-mail: mhansen@ce.berkeley.edu*

Abstract

In this paper we study the 4D trajectory planning problem in a contrail sensitive environment. We identify the control inputs that steer the aircraft from the initial fix to the final fix following a horizontal route of waypoints while performing step climbs and descents, in order to minimize the overall flying cost of fuel consumption, CO_2 emissions, passenger travel time, and persistent contrail formation. The optimal trajectory design problem is formulated as a multiphase mixed integer optimal control problem, which is converted into a mixed integer non-linear program by first making the unknown switching times part of the state, then applying a Hermite-Simpson direct collocation method, and finally introducing binary variables to model the decision making. We solve the mixed-integer nonlinear program using a branch-and-bound algorithm. The numerical results show the effectiveness of the approach.

Keywords: Flight 4D trajectory design; Persistent contrails; Climate impact; Mixed-integer optimal control; Mixed-integer non-linear program.

1. Introduction

Worldwide aviation and the associated greenhouse gas emissions have witnessed significant growth over the past decades. This growing trend is likely to continue in the foreseeable future [21, 33]. [29] project that the greenhouse gas emission from the aviation sector will increase by 60% and 300% by 2030 and 2050, respectively. The share of emitted CO_2 in the global total is also expected to become more important, from 2% in 1999 to 3-5% in 2050 [35]. In terms of anthropogenic radiative forcing, an estimate from the United Kingdom (UK) Royal Commission of Environmental Pollution (RCEP) suggests that the aviation sector will be responsible for 6% of the global total by 2050 [36].

The climate change impact of aircraft operations comes from multiple constituents, with carbon dioxide (CO_2) the most known one. Aviation induced NOx also tends to increase tropospheric ozone and reduce methane. However, the increase in radiative forcing associated with ozone is largely offset by the methane reduction, resulting in a relatively small net positive NOx impact compared to the CO_2 impact in aviation operations [58]. Another important source of aviation-induced climate change that has garnered growing attention is the formation of contrails, which are line-shaped clouds composed of ice particles and formed in the wake of jet aircraft at high altitude where the ambient temperature is very low. The physics of contrail formation is well documented and known as the Schmidt-Appleman criterion [39, 2]. A more recent review of the conditions for contrail formation from aircraft exhausts can be found in [40].

Contrails evaporate quickly if the ambient air is dry, but can persist if the ambient air is humid enough. Like natural high clouds, persistent contrails modify the radiation budget of the earth-atmosphere system by reducing the outgoing terrestrial radiation more than they reflect solar radiation, resulting in warming of the earth's surface [57]. Quantifying the climate impact of persistent contrails has attracted considerable research interests over the past, although consensus has yet to be achieved. The general conclusion is that the magnitude of contrail climate impact is non-negligible compared to that of CO_2 [35, 41, 28]. At the high end among the existing estimates, the greenhouse effect from aviation induced contrails is approximately 10 times higher than from CO_2 emitted by aircraft [28]. Accounting for the formation of persistent contrails, therefore, is indispensable to mitigating the overall aviation induced climate impact.

Persistent contrail can only be formed when aircraft fly into parts of the airspace in which both the Schmidt-Appleman criterion is met and the atmosphere is sufficient humid.[1] In this paper, we term such airspace as Persistent Contrail Formation Areas, or PCFAs. Mitigating aviation induced contrail formation therefore involves adjustment of flight profiles -both vertically and horizontally- in order to avoid PCFAs. Prevailing approaches for modeling the spatial adjustment of flight trajectories include mathematical programming, simulation, and optimal control. [54] develops a linear program with binary decision variables for flight level allocation subject to operational feasibility constraints. Also, a linear program with binary decision variables is used by [15] to consider a tradeoff between contrail formation and emissions in the US airspace. Mixed integer programming techniques are employed in [12, 13], which minimizes aircraft fuel cost while avoiding the formation of persistent contrails. A more recent effort by [60] formulates a binary integer program that allows for both altitude and heading modifications while minimizing the total flying cost in a dynamic, contrail sensitive environment. Using simulation tools, [20], [14], [49], and [46] model efficient trajectories that minimizes the combined

[1] in the remainder of the paper, *contrail formation* always refers to *persistent contrail formation*.

climate impacts of aircraft CO_2 emissions and contrails (and oxides of nitrogen in the latter study), and the tradeoff between the climate impact and aircraft operating cost. Several other attempts have been made to examine vertical displacement of flight paths in order to avoid PCFAs [57, 58, 17, 15]. These studies find altitude adjustment to be an effective strategy to significantly reduce contrail production, but with resulting increase in flight separation minima violations, and add workload for air traffic controllers to resolve the trajectory conflicts [57] and [56].

While using mathematical programming and simulation techniques in aircraft trajectory design allows many constraints (e.g., flight separation minima, maximum workload for air traffic controllers) to be considered, the two approaches have limited capability to model aircraft dynamics. This deficiency can be largely overcome by employing the optimal control approach, which provides control inputs as part of the solution to steer aircraft in the airspace. Previous attempts in using optimal control focus on horizontal design of flight paths with the trade-off between persistent contrail formation and aircraft fuel consumption, under the assumption of constant airspeed and for a range of separate flying altitudes [45] and [47]. However, for a given flying altitude, it has been shown that variable airspeed profiles are more efficient than constant airspeed profiles [34, 18]. Allowing aircraft to alter flight altitude is also important in order to permit aircraft maneuvering and thus avoid PCFAs. Joint consideration of variable flying speed and allocation of flight levels remain largely absent in the PCFA-constrained aircraft trajectory literature.

This study contributes to the existing literature by adopting a multiphase mixed-integer optimal control approach, which incorporates both integer and continuous variables into an optimal control problem, to determine the optimal 4-D (time plus 3-D space) aircraft trajectory which allows for varying flying speed and altitude, in the presence of PCFAs. The multiphase mixed-integer optimal control approach has been recently used, but only for horizontal flight paths design with waypoint allocation [44, 8, 42]. Altitude change and contrail formations are not considered in those studies. We formalize the research question in the present paper as follows: given an aircraft point mass dynamical model, a route composed of a sequence of horizontal waypoints, and a vertical structure of airspace with multiple permitted flight levels, find the control inputs that steer the aircraft from the initial fix to the final fix following the horizontal waypoints while performing appropriate step climbs/descents, in order to minimize the overall cost of fuel consumption, passenger travel time, and climate impact from CO_2 emission and persistent contrail formation. Binary variables model decision making processes, which herein involve the optimal allocation of flight levels. The times at which both the waypoints and the different flight levels are reached are also obtained within the solution. As part of the modeling efforts, this research presents an approach to estimate the unit costs of fuel consumption, CO_2 emission, and persistent contrail formation.

We solve the multiphase mixed-integer optimal control problem through a multi-step process. First, we reduce the multiphase mixed-integer optimal control problem (MIOCP) to a conventional mixed integer optimal control problem

by making the unknown switching times part of the state [43]. We then apply a collocation method based on the Hermite-Simpson Gauss-Lobatto quadrature rules [25] to convert the conventional mixed integer optimal control problem to a mixed integer non-linear program (MINLP). The MINLP is solved using a branch and bound algorithm to obtain the control inputs and optimal flight trajectory.

The paper continues with a general description of a multiphase MIOCP formulation in Section 2, which is ensued by a proposed solution procedure in Section 3. We devote Section 4 to exposing aircraft dynamics. A detailed case study is presented in Sections 5-7, including the problem setup (Section 5), the determination of various cost factors (Section 6), and the presentation and discussions of the results (Section 7). Section 8 summarizes and offers future research directions.

2. Multiphase MIOCP problem formulation

The multiphase motion of an aircraft can be modeled by a switched dynamical system. A switched dynamical system is composed of a set of differential-algebraic dynamic subsystems:

$$\Sigma_q = \begin{bmatrix} \dot{x}_q = f_q(x_q(t), u_q(t), z, t) \\ 0 = g_q(x_q(t), u_q(t), z, t) \end{bmatrix}, q \in \{0, 1, \ldots, n\}.$$

where f_q describes the right-hand side of the differential equation $\dot{x}_q(t) = f_q(x_q(t), u_q(t), z, t)$, g_q describes the algebraic constraints $0 = g_q(x_q(t), u_q(t), t)$, and $q = \{0, 1, \ldots, n\}$ denotes the set of indexes for different dynamical subsystems (also referred herein as phases). $t \in [t_I, t_F] \subset \mathbb{R}$ is the time variable, $x_q(t) \in \mathbb{R}^{n_{x_q}}$ is the state vector in phase q and at time t, $\dot{x}_q(t) \in \mathbb{R}^{n_{x_q}}$ its time derivative, and $u_q(t) \in \mathbb{R}^{n_{u_q}}$ the vector of control inputs in phase q and at time t, which is assumed to be measurable. $z \in \mathbb{R}^{n_z}$ denotes a vector of parameters.

In order to control a switched dynamical system, both a control input, $u_q(t)$, $q = \{0, 1, \ldots, n\}$, and a switching sequence, σ, need to be specified. We assume that the set of admissible control inputs is a set of piecewise continuous functions in $t \in [t_I, t_F]$.

A switching sequence σ in $[t_I, t_F]$ is defined as the timed sequence of $n + 1$ active dynamical systems

$$\sigma = [(\tilde{t}_0, \Sigma_0), (\tilde{t}_1, \Sigma_1), \ldots, (\tilde{t}_n, \Sigma_n)], \tag{1}$$

where $0 \leq n < \infty$, and Σ_q represent the dynamical subsystem for $q \in \{0, 1, \ldots, n\}$. Let

$$\tilde{t}_1 \leq \tilde{t}_2 \leq \cdots \leq \tilde{t}_{n-1} \leq \tilde{t}_n \tag{2}$$

be the switching times between phases. Let also $t_I = \tilde{t}_0$ and $t_F = \tilde{t}_{n+1}$ be the initial and final time, respectively. Notice that $t_I = \tilde{t}_0 \leq \tilde{t}_1$ and $\tilde{t}_n \leq \tilde{t}_{n+1} = t_F$. In Equation (1), the pair $(\tilde{t}_q, \Sigma_q)$ indicates that at time $\tilde{t}_q$ the dynamic equation of the switched system changes from $q - 1$ to q, with $\tilde{t}_0 = t_I$ and $\tilde{t}_{n+1} = t_F$. As a consequence, in the time interval $[\tilde{t}_q, \tilde{t}_{q+1})$ the system evolution is governed

by the dynamic subsystem Σ_q. In the interval $[\tilde{t}_n, \tilde{t}_{n+1}]$ the active dynamic subsystem is Σ_n. The phenomena that trigger the switching instants can be either due to capture conditions, in which case the switch is referred to as internally forced or autonomous switch, or in response to the control law, in which case the switch is referred to as externally forced or controlled switch [9] and [59].

Let us define $n+1$ vectors of binary variables $v_q \in \{0,1\}^{n_{v_q}}$ to model decision making processes within phase q, $q = \{0,\ldots,n\}$. A decision making process consists in the selection of one option among a set of alternatives, e.g., the selection of one flight level among a set of possible flight levels during one phase of the flight.

The multiphase MIOCP can be stated as follows:

$$\min J(x_q, u_q, v_q, z, t) = \sum_{q=0}^{n} \Big[\int_{\tilde{t}_q}^{\tilde{t}_{q+1}} L_q[x_q(t), u_q(t), z, t]dt + E_q[x_q(\tilde{t}_q), v_q, z], \Big] \quad (3)$$

subject to

$$\dot{x}_q(t) = f_q[x_q(t), u_q(t), z, t], t \in [\tilde{t}_q, \tilde{t}_{q+1}], \quad (4)$$
$$g_q[x_q(t), u_q(t), z, t] = 0, t \in [\tilde{t}_q, \tilde{t}_{q+1}], \quad (5)$$
$$c_q[x_q(t), u_q(t), z, t] \leq 0, t \in [\tilde{t}_q, \tilde{t}_{q+1}], \quad (6)$$
$$x(t_I) = x_I, \quad (7)$$
$$\psi(x(t_F)) = 0, \quad (8)$$
$$r^{\text{eq}}[x_q(\tilde{t}_q), v_q, z] = 0, \quad (9)$$
$$r^{\text{ineq}}[x_q(\tilde{t}_q), v_q, z] \leq 0, \quad (10)$$
$$x_{q+1}(\tilde{t}_{q+1}) = G_q[x_q(\tilde{t}_{q+1}), u_q(\tilde{t}_{q+1}), z], \quad (11)$$

for $q \in \{0,\ldots,n\}$. Notice that in this particular form of multiphase MIOCP binary decision variables are only present in the objective function (3), (9) and (10). A more general form can be consulted in [37] and [38].

The objective function (3) is in Bolza form and contains a Lagrange term

$$\int_{\tilde{t}_q}^{\tilde{t}_{q+1}} L_q[x_q(t), u_q(t), z, t]dt$$

and a Mayer term $E_q[x_q(\tilde{t}_q), v_q, z]$. Both L_q and E_q are assumed to be Lipschitz-continuous and their derivatives are also Lipschitz-continuous in their arguments. The Lagrange term represents a running cost, whereas the Mayer terms represent a terminal cost. As an illustration, a classical Lagrange objective function is to minimize the total amount of energy during a process; a classical Mayer objective function is to minimize the final time. The presence of binary variables in the Mayer terms denotes that the decision-making process involves different terminal costs for the different alternatives.

(4) and (5) with $f_q \in \mathbb{R}^{n_{f_q}}$ and $g_q \in \mathbb{R}^{n_{g_q}}$ represent the differential-algebraic equations for the switched system in phase q. (6) with $c_q \in \mathbb{R}^{n_{c_q}}$ are the inequality constraints in phase q. (7) and (8) denote the initial and final boundary conditions of the problem. (9) with $r^{eq} \in \mathbb{R}^{n_{r^{eq}}}$ and (10) with $r^{ineq} \in \mathbb{R}^{n_{r^{ineq}}}$ represent the equality and inequality interior point constraints respectively, which are assumed to be twice differentiable. The presence of binary variables in (9) and (10) implies that there might be a set of alternative interior point equality and inequality constraints per phase. The decision making process would be to select one of the alternatives in each phase. (11) describes the transition conditions between phases to ensure continuity, which are usually of the form $x_{q+1}(\tilde{t}_{q+1}) = x_q(\tilde{t}_{q+1})$. It should be noted that the dimensions $n_{x_q}, n_{u_q}, n_{v_q}, n_{c_q}, n_{r^{ineq}}, n_{r^{eq}}, n_{f_q}, n_{g_q}$ are not necessarily identical for each phase. The solution to this problem is composed of $x_q(t), u_q(t), t \in [t_q, t_{q+1}]$, the switching times $\tilde{t}_1, \ldots, \tilde{t}_n$, and v_q.

As an illustration: (4) and (5) represent the differential-algebraic equation system that governs the motion of the dynamical system, e.g., an aircraft; (6) models the physical limits of performance of the dynamical system, typically expressed as upper and lower bounds on both states and control variables; (7) and (8) denote the boundary (initial and final, respectively) conditions of the process in which the system is involved; (9) models equality constraints to be enforced during the process (not at the boundaries), e.g., to reach a mid-position, and (10) models inequality constraints to be enforced also during the process, e.g., to enforce that the above mentioned mid-position is reached within a prescribed time window; (11) is only to ensure that the state variables that describe the evolution of the system evolve continuously. The presence of binary variables in (9) and (10) denotes that the definition of interior point equality and equality constraints involves a decision-making process, i.e., there might be a set of alternative mid-positions to be reached with their corresponding time windows. The particular expressions for the elements of the MIOCP in the case study herein exposed are stated in Sections 4-6.

3. Multiphase MIOCP solution approach

The MIOCP solution approach is a multi-step process as illustrated in Figure 1.

3.1. Problem reformulation

We tackle the multiphase MIOCP problem by first making the unknown switching times part of the state and then introducing a new independent variable τ which determines the switching times to be fixed. The new formulation results in a conventional MIOCP problem. A linear relation exists between the new variable τ and time t, but the slope of this linear relation changes on each interval between two switches. These slopes, which are part of the solution to the multiphase MIOCP problem, are time scaling factors that determine the optimal switching times. The mathematical details of this formulation can be found in [43]. Appendix A briefly summarizes the fundamentals of this reformulation.

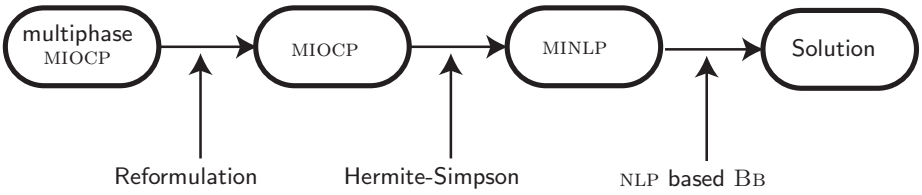


Figure 1: Solution approach flow chart.

3.2. Hermite-Simpson Gauss-Lobatto collocation method

A collocation approach is used to deal with the differential constraints and the Lagrange term in the objective function. After applying the technique described in the previous section, (3) and (4) are modified, and the switching times are included in the state vector.

Collocation methods enforce the dynamic equations through quadrature rules or interpolation [25]. A suitable interpolating function, or interpolant, is chosen such that it passes through the state values and maintains the state derivatives at the nodes spanning one interval, or subinterval, of time. The interpolant is then evaluated at points between nodes, called collocation points. At each collocation point, a constraint equating the interpolant derivative to the state derivative function is introduced to ensure that the equations of motion are approximately satisfied across the entire interval of time [55].

In particular, the Hermite-Simpson Gauss-Lobatto integration scheme is used in the present study [25]. The motivation behind the use of this integration scheme is related to its order of accuracy which is a measure of the effectiveness of a numerical method to approximate a solution. The Hermite-Simpson Gauss-Lobatto integration scheme has an order of accuracy 4 whereas, for instance, the trapezoid (the 2^{nd} degree Gauss-Lobatto integration scheme) has an order of accuracy 2. The greater is the order of accuracy, the greater is the reduction in error if the stepsize is made smaller. Therefore, as the order of accuracy increases, a specified accuracy may be achieved with larger stepsizes, i.e., with a smaller number of variables. Nevertheless, increasing the order of accuracy results in more complicated quadrature rules, and thus a compromise between accuracy and complexity must be found. Even though there exist higher degree methods, e.g., the fourth and fifth degree Gauss-Lobatto collocation methods [26], their application to multiphase problems with high number of phases is computationally very intense. Overall, using the Hermite-Simpson collocation method will reduce the number of variables drastically (at relatively low complexity), thus improving the computation times in solving the resulting MINLP problem.

Summing up, by using this class of numerical method, the MIOCP problem is converted into a MINLP problem. Interested readers can refer to [25] for mathematical details. Appendix B briefly summarizes the fundamentals of this method.

3.3. MINLP *solution approach*

In theory, MINLP problems are nondeterministic polynomial time hard (NP-hard) to solve, and represent one of the most challenging problems in computational optimization. In addition, due to aircraft dynamics and constraints, the feasible set of the MINLP problem is not convex and highly nonlinear. A simple algorithm for determining the mode sequence is to enumerate all possible values for v_q, solve the associated optimal control problems, and pick the best solution. However, this approach becomes impractical for more than a few binary variables. A common, alternative method to solve larger problems is to perform an implicit enumeration using the branch and bound algorithm [5]. We provide a sketch of this implicit enumeration method below and highlight the particularities that arise when applying the method to solve the MIOCP.

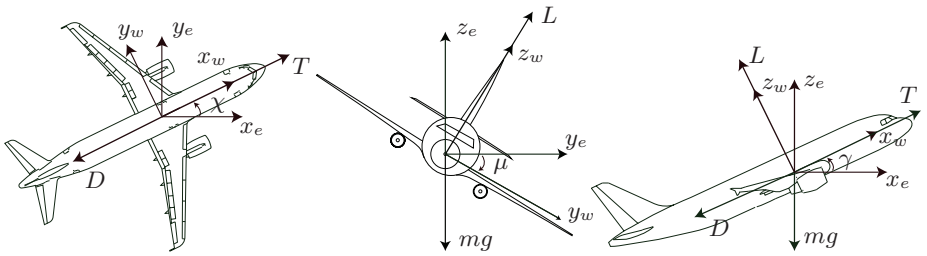
The branch and bound algorithm is a divide-and-conquer method. The problem is divided by partitioning the set of feasible solutions into smaller and smaller subsets. The conquering is done by computing bounds on the cost of the best feasible solution in each subset and discarding subsets whose lower bound exceeds a known feasible solution. The branch and bound algorithm produces an exact optimal solution when the bound used in each subproblem is a valid lower bound. In our case, we relax the binary variables to find a lower bound for the multiphase MIOCP, i.e. we let $v_q \in [0, 1]$ and solve the associated NLP program. However, obtaining a true lower bound on the value of the multiphase MIOCP is a difficult task due to presence of nonconvex dynamics and constraints in the NLP. As a consequence, the approach does not rely on a true lower bound but rather uses approximate solutions. In that case the procedure is heuristic (i.e. does not return the exact optimal solution). The quality of the final solution depends on the quality of the approximation. To the best of our knowledge, there is no theoretical guarantee on the quality of the approximation. More details on the heuristic algorithm are given in Appendix C.

We use the open-source solver Bonmin [6] which implements several different algorithms to solve MINLP programs. We call Bonmin through the AMPL modeling language. Reference [8] provides further details about the NLP branch and bound algorithm employed in this paper. Appendix C briefly summarizes the fundamentals of the algorithm.

4. Aircraft Dynamics

In designing optimal aircraft trajectories, it is common to consider a 3 degree-of-freedom dynamic model that describes the point variable-mass motion of the aircraft over a spherical flat-earth model. We consider a symmetric flight, i.e., there is no sideslip and all forces lie in the plane of symmetry of the aircraft. Wind is included in the model due to its considerable effects on fuel consumption and flight time. The equations of motion of the aircraft are:

$$\frac{d}{dt}\begin{bmatrix} V \\ \chi \\ \gamma \\ \lambda_e \\ \theta_e \\ h_e \\ m \end{bmatrix} = \begin{bmatrix} \frac{T(t)-D(h_e(t),V(t),C_L(t))-m(t)\cdot g\cdot \sin\gamma(t)}{m(t)} \\ \frac{L(h_e(t),V(t),C_L(t))\cdot \sin\mu(t)}{m(t)\cdot V(t)\cdot \cos\gamma(t)} \\ \frac{L(h_e(t),V(t),C_L(t))\cdot \cos\mu(t)-m(t)\cdot g\cdot \cos\gamma(t)}{m(t)\cdot V(t)} \\ \frac{V(t)\cdot \cos\gamma(t)\cdot \cos\chi(t)}{R\cdot \cos\theta_e(t)} + W_x(\lambda_e(t),\theta_e(t),h_e(t)) \\ \frac{V(t)\cdot \cos\gamma(t)\cdot \sin\chi(t)}{R} + W_y(\lambda_e(t),\theta_e(t),h_e(t)) \\ V(t)\cdot \sin\gamma(t) + W_z(\lambda_e(t),\theta_e(t),h_e(t)) \\ -T(t)\cdot \eta(V(t)) \end{bmatrix}. \tag{12}$$



(a) Top view (b) Front view (c) Lateral view

Figure 2: Aircraft state and forces

In the above, the three kinematic equations are expressed in a ground based reference frame (x_e, y_e, z_e) and the three dynamic equations are expressed in an aircraft-attached reference frame (x_w, y_w, z_w), as shown in Figure 2. The states are: λ_e, θ_e, h_e referring to the aircraft 3D position (longitude, latitude, altitude); V, χ, γ, m referring to the true airspeed, heading angle, flight path angle, and the mass of the aircraft, respectively. R is the radius of earth and η is the speed dependent fuel efficiency coefficient. T is the thrust and μ is the bank angle. Lift $L = C_L S\hat{q}$ and drag $D = C_D S\hat{q}$ are the components of the aerodynamic force, where S is the reference wing surface area; $\hat{q} = \frac{1}{2}\rho V^2$ is the dynamic pressure; C_L and C_D are the lift and drag coefficients. A parabolic drag polar $C_D = C_{D0} + KC_L^2$, where C_{D0} is the parasite coefficient of drag and K is the induced coefficient of drag, and a standard atmosphere are assumed. C_L is a known function of the angle of attack, α, and the Mach number, M. W_x, W_y, and W_y denote the components of the wind vector, $W = (W_x, W_y, W_z)$, expressed in a ground based reference frame (x_e, y_e, z_e). The aircraft position in 2D is approximated as $x_e = \lambda_e \cdot (R + h_e) \cdot \cos\theta_e$ and $y_e = (R + h_e) \cdot \theta_e$. The bank angle μ, the engine thrust T, and the coefficient of lift C_L are the inputs, i.e., $u(t) = (T(t), \mu(t), C_L(t))$. Notice that the specific fuel flow η is a speed dependent function as illustrated in (27).

4.1. Wind Model

The wind velocity is often considered as the sum of two terms: a deterministic component which represents the meteorological predictions available to the ATC, and a stochastic component which represents the uncertainty in

these predictions. In the scope of this framework, the wind is considered as deterministic.

The meteorological predictions are obtained from the Rapid Update Circle (RUC), a numerical model developed at the National Oceanic and Atmospheric Administration (NOAA) forecasts system laboratory [4] and [3]. Forecasts are provided via GRIded Binary (GRIB) files. Data are provided four times a day into a $1^\circ \times 1.25^\circ$ grid and 14 different barometric altitudes. The wind field is assumed to be stationary, i.e., its evolution over time has not been considered.

In particular, GRIB files provide wind forecasts giving the three components (west, north and vertical) of the wind velocity vector, $W = (W_x, W_y, W_z)$, at each node of the grid. GRIB data are given in spherical coordinates, i.e., longitude (λ_e), latitude (θ_e) and altitude (h_e). An analytical representation of wind data is needed, for which multiple polynomial regression analysis is used within this framework.

The general form of a polynomial regression model for m independent variables is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \varepsilon, \tag{13}$$

where $\beta_0, \beta_1, \ldots, \beta_m$ are the regression coefficients that need to be estimated and ε is the error component reflecting the difference between an individual's observed response Y and the true average response $\mu_{Y|X_1,X_2,\ldots,X_m}$.

In order to estimate the coefficients of the multiple polynomial regression equation, least-squares are considered. Thus, the estimated regression equation can be expressed as:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_m X_m,$$

In order to estimate the wind velocity vector, in principle a three-dimensional multiple variable polynomial regression should be used. We assume that the vertical component of the wind vector is negligible (which is essentially true according to GRIB data), i.e., $W_z \approx 0$, and that the two horizontal components of the wind velocity vector, i.e., W_x, W_y, are not correlated. Under these assumptions, two independent multivariete polynomial regressions as described previously are performed to estimate the east-west and south-north wind components.

In such a form, the wind velocity east-west component can be expressed as:

$$\begin{aligned} W_x = p^x_m(r) + \varepsilon_x = {} & \beta^x_{000} + \beta^x_{100} x_e + \beta^x_{010} y_e + \beta^x_{001} z_e + \beta^x_{200} x_e^2 + \beta^x_{110} x_e y_e \\ & + \beta^x_{020} y_e^2 + \beta^x_{101} x_e z_e + \beta^x_{011} y_e z_e + \cdots + \beta^x_{00m} z_e^m + \varepsilon_x, \end{aligned} \tag{14}$$

The wind velocity south-north component can be expressed as:

$$\begin{aligned} W_y = p^y_m(r) + \varepsilon_y = {} & \beta^y_{000} + \beta^y_{100} x_e + \beta^y_{010} y_e + \beta^y_{001} z_e + \beta^y_{200} x_e^2 + \beta^y_{110} x_e y_e \\ & + \beta^y_{020} y_e^2 + \beta^y_{101} x_e z_e + \beta^y_{011} y_e z_e + \cdots + \beta^y_{00m} z_e^m + \varepsilon_y, \end{aligned} \tag{15}$$

We refer interested reader to [42, Chapter 4] for a deeper insight into this method, including justification of the assumptions, discussion about other approaches, e.g., splines interpolation, and in-depth analysis of the goodness-of-fit tests.

4.2. Flight Modes

In the en-route portion of a flight, unless there is an Air Traffic Control (ATC) requirement to deviate the flight from its nominal paths in order to avoid a potential conflict, aircraft typically fly horizontal segments connecting waypoints and perform step climbs/descents to modify their flight levels. We consider different flight dynamics as the modes of the switched system. A 3-D flight plan can be subdivided into a sequence of modes pertaining to different flights segments, i.e., cruising at constant altitude between waypoints or ascending/descending to change flight level. We characterize these maneuvers by two modes: *control speed and heading* (mode $\mathbb{H}$) for the cruising at constant altitude, and *control altitude* (mode $\mathbb{V}$) for the ascent/descent to change flight level.

Under mode $\mathbb{H}$ (control speed and heading), the aircraft is considered as flying into a horizontal plane with variable speed and variable heading. γ, $\dot{\gamma}$, and $\dot{h}$ are set to zero. Thus, the following algebraic constraint applies: $L\cos\mu = mg$. The engine thrust and the bank angle are the control inputs, i.e., $u(t) = (\mu(t), T(t))$.

Under mode $\mathbb{V}$ (Control altitude), the aircraft is considered as flying into vertical plane, performing a leveled-wing climb/descent. Therefore, the bank angle μ is set to zero. Without loss of generality, we let $\chi = 0$. The thrust and the lift coefficients are the control inputs, i.e., $u(t) = (T(t), C_L(t))$.

4.3. Flight Envelope Constraints

Flight envelope constraints refer to constraints due to the geometry of the aircraft, structural limitations, engine power, and aerodynamic characteristics. They are derived from aircraft performance limitations. Typically, these constraints can be expressed as upper and lower bounds on both states and controls. We use the BADA performance limitations model and parameters [32]:

$$
\begin{array}{ll}
0 \le h(t) \le \min[h_{M0}, h_u(t)], & \gamma_{min} \le \gamma(t) \le \gamma_{max}, \\
M(t) \le M_{M0}, & m_{min} \le m(t) \le m_{max}, \\
\dot{V}(t) \le \bar{a}_l, & C_v V_s(t) \le V(t) \le V_{Mo}, \\
\dot{\gamma}(t)V(t) \le \bar{a}_n, & 0 \le C_L(t) \le C_{L_{max}}, \\
T_{min}(t) \le T(t) \le T_{max}(t), & \mu(t) \le \bar{\mu},
\end{array}
$$

where h_{M0} is the maximum reachable altitude; $h_u(t)$ is the maximum operative altitude at a given mass (it increases as fuel is burned); γ_{min} and γ_{max} are the minimum and maximum recommended flight path angles, respectively; m_{min} and m_{max} are the minimum and maximum aircraft mass, respectively; $M(t)$ is the Mach number and M_{M_0} is the maximum operating Mach number; C_v is the minimum speed coefficient, $V_s(t)$ is the stall speed, and V_{M_0} is the maximum operating calibrated airspeed; $\bar{a}_n$ and $\bar{a}_l$ are respectively the maximum normal and longitudinal accelerations for civilian aircraft, respectively; $C_{L_{max}}$ is the maximum coefficient of lift; T_{min} and T_{max} correspond to the minimum and maximum available thrust, respectively; $\bar{\mu}$ corresponds to the maximum bank angle due to structural limitations.

Additionally, the following operational constraints apply: $\dot{h} \geq 0$ $[ft/min]$, during step climbs; $\dot{h} \leq 0$ $[ft/min]$, during step descents; climbs/descent descents are limited to a maximum duration of 4 minutes.

Note that several flight envelope constraints are nonconvex.

5. Case Study: Problem set up

We consider a B757-200 BADA 3.6 [32] model aircraft performing the en-route part of a flight San Francisco (SFO) - New York (JFK) between the waypoint[2] PEONS as initial fix and the waypoint MAGIO as final fix. The route is composed of a series of waypoints, as shown in Table 1. All pairs formed by two consecutive waypoints are connected by bi-directional airways. Along the air route, multiple flight levels are available for flying with vertical separation of 1000 feet between two neighboring ones. In air traffic practice, flights flying opposite directions are assigned two non-overlapping set of flight levels. In this study, we assume that eastbound flights are assigned odd flight levels separated 2000 feet. The aircraft can fly at any of the following flight levels:

$$\{FL270, FL290, FL310, FL330, FL350, FL370, FL390, FL410\} \tag{16}$$

Notice that according to the path constraints stated in Section 4.3, in practical terms it results in a limitation of performing step climbs up/down to three levels at a time.

5.1. Wind data

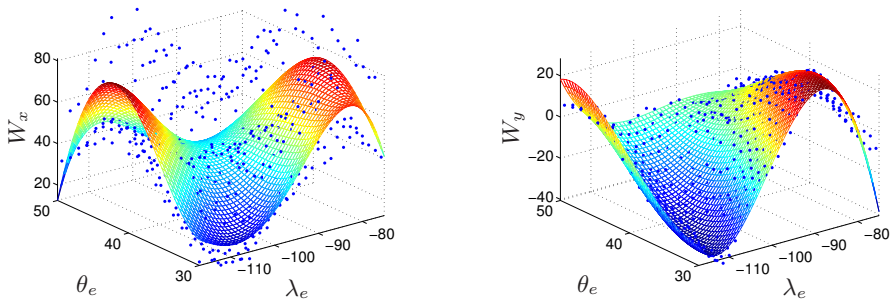
The forecast wind of March 17, 2014 in the Northern America region has been considered. Using nonlinear regression, wind tabular data have been converted to analytic functions, which are valid within a domain covering continental US, i.e., $\lambda_e \in$ [-120°,-70°] and $\theta_e \in [30°, 50°]$. Then, the resulting polynomials P_m^x and P_m^y [3] in functions (14)-(15) have been included in the set of equations (12). The goodness-of-fit, measured by R^2, yielded 0.84 for W_x and 0.95 for W_y. Figure 3 shows both forecast tabular data (blue dots) and analytic functions (surfaces) for W_x and W_y at 200 [Hpa] (h=38612 [ft]). Note that the regression analysis provides satisfactory results in terms of R^2 only in the horizontal plane (at a given altitude), polynomials fitting the whole 3D space (longitude, latitude, and altitude) provide less satisfactory model fit with R^2 values below 0.5. In this study, we make the approximation that wind is equal for all flight levels.

5.2. PCFA regions

As already discussed in the introduction, persistent contrail can only be formed when aircraft fly into parts of the airspace in which both the Schmidt-Appleman criterion is met and the atmosphere is sufficient humid.

[2] Waypoints may be a simple named point in space or may be associated with existing navigational aids, intersections, or fixes.

[3] The estimates β are available upon request.



(a) $W_x(\lambda_e, \theta_e)$ (b) $W_y(\lambda_e, \theta_e)$

Figure 3: Wind forecast tabular data (blue dots) and analytic functions (surfaces) for W_x and W_y at 200 [Hpa] (h=38612 [ft]).

Table 1: Route's waypoints, navaids and fixes.

Name	Type	Longitude	Latitude
PEONS	WAYPOINT (RNAV)	-119.1674°	38.5035°
INSLO	WAYPOINT (RNAV)	-117.2981°	38.6791°
DTA	VOR-TAC (NAVAID)	-112.5055°	39.6791°
MTU	VOR-DME (NAVAID)	-110.1270°	40.1490°
CHE	VOR-DME (NAVAID)	-107.3049°	40.5200°
HANKI	REPORTING POINT	-102.9301°	41.6319°
KATES	REPORTING POINT	-96.7746°	42.5525°
FOD	VOR-TAC (NAVAID)	-942947°	42.6111°
KG75M	NRS-WAYPOINT	-88°	42.5°
DAFLU	REPORTING POINT	-82.7055°	42.3791°
JHW	VOR-DME (NAVAID)	-79.1213°	42.1886°
MAGIO	REPORTING POINT	-76.5964°	41.5373°

Following the Schmidt-Appleman criterion, contrails form in the regions of airspace that have ambient relative humidity with respect to water (RH_w) greater than a critical value r_{contr}. Regions with RH_w greater or equal than 100% are excluded because clouds are already present [23]. Contrails will persist when the environmental relative humidity with respect to ice (RH_i) is greater than 100% [16]. Thus, persistent contrail formation areas are defined as the regions of airspace that have: $r_{contr} \leq RH_w < 100\%$ and $RH_i \geq 100\%$.

Following [15], the estimated critical relative humidity for contrail formation at a given temperature T (in degrees Celsius) can be calculated as:

$$r_{contr} = \frac{G(T - T_{contr}) + e_{sat}^{liq}(T_{contr})}{e_{sat}^{liq}(T)}, \tag{17}$$

where $e_{sat}^{liq}(T)$ is the saturation vapor pressure over water at a given temperature. The estimated threshold temperature (in degrees Celsius) for contrail formation at liquid saturation is:

$$T_{contr} = -46.46 + 9.43 \log(G - 0.053) + 0.72 \log^2(G - 0.053), \tag{18}$$

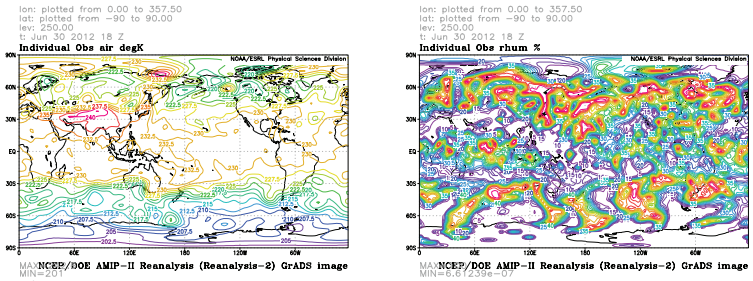


Figure 4: Air temperature and relative humidity June the 30th, 2012 at time 18.00 Z and for an barometric altitude of 250Hpa. Image provided by Physical Sciences Division, Earth System Research Laboratory, NOAA, Boulder, Colorado, from their Web site at http://www.esrl.noaa.gov/psd/.

where

$$G = \frac{EI_{H_2O} C_p P}{\epsilon Q (1-\eta)}. \tag{19}$$

In equation (19), EI_{H_2O} is the emission index of water vapor, C_p is the isobaric heat capacity of air, P is the ambient air pressure, ϵ is the ratio of molecular masses of water and dry air, Q is the specific heat combustion, and η is the average propulsion efficiency of the jet engine.

RH_i is calculated by the following formula [1]:

$$RH_i = RH_w \frac{6.0612 \exp^{\frac{18.102T}{249.52+T}}}{6.1162 \exp \frac{22.577T}{237.78+T}}, \tag{20}$$

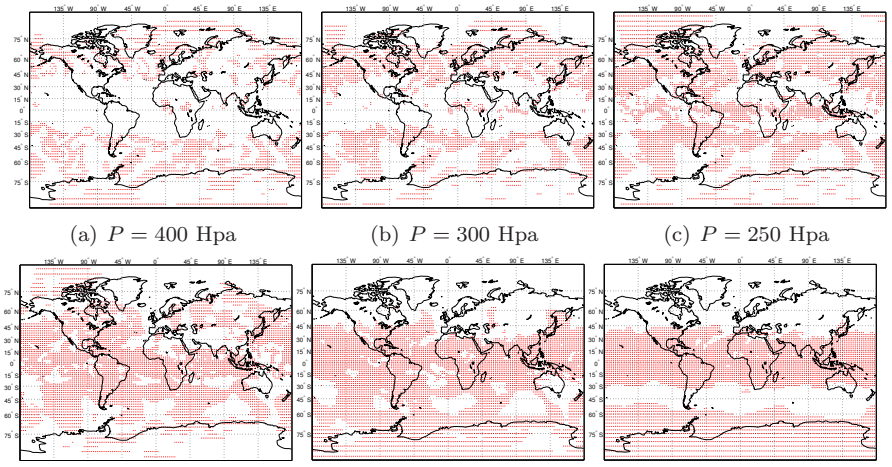
where T is the temperature in degrees Celsius.

The construction of PCFA regions is based on atmospheric data of air temperature and relative humidity on June 30, 2012, which are retrieved from the NCEP/DOE AMIP-II Reanalysis data provided by the System Research Laboratory at the National Oceanic & Atmospheric Administration (NOAA)[4]. The data have a global longitude-latitude grid resolution of 2.5°×2.5°. Regarding the vertical resolution, the data are provided in 17 pressure levels (hPa): 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10. Figure 4 shows the values of air temperature and relative humidity for June 30, 2012 at time 18.00 Z[5] (10.00 a.m. PDT).

To identify the PCFA regions, we first compute the latitude-longitude grid points that are favorable to the formation of persistent contrails at different barometric altitudes (defines the pressure). This is done using the data on air temperature and relative humidity and formulas (17-20) with the following parameter values: $EI_{H_2O} = 1.25$; $C_p = 1004$ $[J/kgK]$; $\epsilon = 0.6222$;

[4]The data have been downloaded from NOAA website @ http://www.esrl.noaa.gov/psd/

[5]Z-hour corresponds to Universal Time Coordinates (UTC). The Pacific Daylight Time (PDT) is given by UTC - 8 hours.



(a) $P = 400$ Hpa (b) $P = 300$ Hpa (c) $P = 250$ Hpa

(d) $P = 200$ Hpa (e) $P = 150$ Hpa (f) $P = 100$ Hpa

Figure 5: Longitude-latitude grid points that present favorable conditions for persistent contrail formation for different barometric altitudes.

$Q = 43x10^6$ $[J/kg]$; and $\eta = 0.15$. The longitude-latitude grid points with favorable conditions for persistent contrail formation are represented as red dots in Figure 5 for different barometric altitudes.

The PCFA information based on the barometric altitudes above cannot be directly used in our trajectory optimization model, which uses flight levels. What we need is air temperature and relative humidity values for the flight levels defined in set (16). To this end, we use the International Standard Atmosphere (ISA) equations to convert altitude into barometric altitude, and then run a linear interpolation using the data of air temperature and relative humidity corresponding to the 17 pressure levels for the desired flight levels (already converted into barometric altitude). We then apply formulas (17-20) to derive the PCFA regions. The PCFA coverage over the continental US at the different flight levels appears later on in Figure 8.

5.3. Problem Statement

A flight plan is prepared on the ground before take-off and provides information on the route composed of waypoints, flight levels, speeds, times, and fuel for different flight phases, alternative airports, and other relevant data for the flight, so that the aircraft properly receives support from ATC and executes safe operations. In the trajectory optimization problem under study in this paper, the solution to the problem would result in the intended flight plan.

Consider a switched system as described in Section 2 with the switching sequence (1). Following Table 1, the aircraft is constrained to start at the initial waypoint PEONS, pass through the subsequent n_{wp} waypoints, and eventually arrive at the final waypoint MAGIO. For simplicity, we assume that the aircraft can only perform step climbs/descents right after reaching a waypoint, and

that a maximum of four step climbs/descents are allowed during the flight, i.e., $n_{sc} = 4$. Given that $n_{wp} + n_{sc} = n$, $n + 1$ phases will be identified during the motion of the aircraft.

Each phase starts either when the aircraft position coincides with a waypoint on the route given in Table 1, or when the aircraft reaches one of the n_{fl} flight levels specified in set (16) ($n_{fl} = 8$) when a step climb/descent is permitted. Because the waypoints are prescribed, the decision making is over the set of flight levels that represents alternative capture conditions for switching, i.e., all switchings herein considered are internally forced (autonomous) switches.

As is already described in Section 3.1, the switchings are considered fixed (time invariant) with respect to the new independent variable τ. Thus, time invariant binary variables can be used to model the decision making processes. Let v be a $n_{sc} \times n_{fl}$ dimensional binary variable vector. Each of the binary variables $v_{i,j},\ i = 1, \ldots, n_{sc},\ j = 1, \ldots, n_{fl}$, is set to one if the flight reaches flight level FL_j after performing step climb/descent i. The constraints on the flight levels can be expressed as follows:

$$h(\tilde{t}_i) = \sum_{j=1}^{n_{fl}} v_{i,j} \cdot\ FL_{i,j}, i = 1, \ldots, n_{sc}. \tag{21}$$

In addition,

$$\sum_{j=1}^{n_{fl}} v_{i,j} = 1,\ i = 1, \ldots, n_{sc}. \tag{22}$$

(21) means that, if $v_{i,j} = 1$, then $h(\tilde{t}_i) = FL_{i,j}$, i.e., the aircraft will reach flight level j after performing step climb/descent i at time $\tilde{t}_i$. (22) stipulates that the aircraft must select a single flight level, $FL_{i,j}$, after every step climb/descent.

Let $\bar{X} = \{\bar{x}_1, \ldots, \bar{x}_{n_{wp}}\}$, with $\bar{x}_k = (\lambda_k, \theta_k)$, be the set of waypoints

$$\{\text{Inslo, Dta, Mtu, Che, Hanki, Kates, Fod, Kg75m, Daflu, Jhw, Magio}\}, \tag{23}$$

whose locations were given in Table 1.

We express the constraints on the waypoints as follows:

$$\bar{x}(\tilde{t}_k) = \bar{x}_k, k = 1, \ldots, n_{wp}, \tag{24}$$

which states that the aircraft must overfly position $\bar{x}$ at time $\tilde{t}_k$.

Table 2 presents the elements assumed to be known in the flight plan with $n_{sc} = 4$, $n_{fl} = 8$, and $n_{wp} = 10$.[6] The first column denotes the phases, which indicates the ordered sequence of step climbs/descents and passing through

[6]Note that leg in aeronautical terminology refers to a segment of the flight, typically a track joining two waypoints.

Table 2: Flight plan model.

Phase	Name	Σ_q	Trigger
1	leg 1	Mode $\mathbb{H}$	INSLO
2	Step Climb/Descent 1	Mode $\mathbb{V}$	$\sum_{j=1}^{nfl} v_{1,j} FL_{1,j}$
3	leg 2	Mode $\mathbb{H}$	DTA
4	leg 3	Mode $\mathbb{H}$	MTU
5	Step Climb/Descent 2	Mode $\mathbb{V}$	$\sum_{j=1}^{nfl} v_{2,j} FL_{2,j}$
6	leg 4	Mode $\mathbb{H}$	CHE
7	leg 5	Mode $\mathbb{H}$	HANKI
8	Step Climb/Descent 3	Mode $\mathbb{V}$	$\sum_{j=1}^{nfl} v_{3,j} FL_{3,j}$
9	leg 6	Mode $\mathbb{H}$	KATES
10	leg 7	Mode $\mathbb{H}$	FOD
11	Step Climb/Descent 4	Mode $\mathbb{V}$	$\sum_{j=1}^{nfl} v_{4,j} FL_{4,j}$
12	leg 8	Mode $\mathbb{H}$	KG75M
13	leg 9	Mode $\mathbb{H}$	DAFLU
14	leg 10	Mode $\mathbb{H}$	JHW
15	leg 11	Mode $\mathbb{H}$	MAGIO

different waypoints, as interpreted by column 2. The aircraft flies each phase under either Mode $\mathbb{H}$ or Mode $\mathbb{V}$ as defined in Subsection 4.2. The specific mode, displayed in column 3, is determined by the aircraft dynamics equations and constraints presented in Section 4. The forth column shows the trigger conditions for switching. Because these trigger conditions correspond to capture conditions, the switchings are internally forced (autonomous) and modeled as interior point constraints as stated in equations (21) and (24). Notice that the decision making is performed over the set of flight levels that represent alternative capture conditions for switching.

The objective function to be minimized, stated in its more general form in equation (3), is given in the Bolza form as follows:

$$J = C_t \cdot t_f + (C_F + C_{CO_2}) \cdot \left[\sum_{q=0}^{n} \int_{\tilde{t}_q}^{\tilde{t}_{q+1}} \dot{m}_q(t)\, dt \right] + \mathcal{D}, \tag{25}$$

where C_t represents in-flight travel time cost over all passengers per unit of time in [\$/s] and t_f represents the total flight time; $\dot{m}_q(t)$ is the fuel flow of the aircraft during phase q and C_F and C_{CO_2} correspond to the unit fuel and CO_2 emission cost, both measured in dollars per kilogram of fuel consumed. Therefore, the first two terms on the right hand side of (25) represent time, fuel and CO_2 emission cost, which are continuous. The second term $\mathcal{D}$ denotes the persistent contrail formation cost, which is discrete because it can only be generated when the aircraft flies into a PCFA region. The explicit form of $\mathcal{D}$ can be expressed by:

$$\mathcal{D} = \delta_I \cdot C_C(FL_I) \cdot (\tilde{t_1} - t_I) + \sum_{k=1}^{2}\sum_{j=1}^{n_{fl}} \delta_{kj} \cdot C_C(FL_{1,j}) \cdot v_{1,j} \cdot (\tilde{t}_{k+1} - \tilde{t}_k) +$$
$$\sum_{k=3}^{4}\sum_{j=1}^{n_{fl}} \delta_{kj} \cdot C_C(FL_{2,j}) \cdot v_{2,j} \cdot (\tilde{t}_{k+1} - \tilde{t}_k) + \sum_{k=5}^{6}\sum_{j=1}^{n_{fl}} \delta_{kj} \cdot C_C(FL_{3,j}) \cdot v_{3,j} \cdot (\tilde{t}_{k+1} - \tilde{t}_k) +$$
$$\sum_{k=7}^{n_{wp}}\sum_{j=1}^{n_{fl}} \delta_{kj} \cdot C_C(FL_{4,j}) \cdot v_{4,j} \cdot (\tilde{t}_{k+1} - \tilde{t}_k). \quad (26)$$

Two points are worth mentioning for (26). First, we neglect the effects of contrail formation cost during step climbs/descents, and only consider contrail formation cost while the aircraft flies between waypoints. That is also why we use time index k in (26). Second, a factor $\delta \in [0, 1]$ is introduced when applying the contrail cost numbers that depend on the flight level, i.e. $C_C(FL)$, and appear later in Table 4, to account for the extent of PCFA coverage during a given flight phase. Specifically, if a leg k at a certain flight level j is entirely covered by a PCFA, $\delta_{kj} = 1$; if a leg k at a certain flight level j does not overlap at all with a PCFA region, $\delta_{kj} = 0$; in other cases, we consider the great circle distance between waypoints and set δ_{kj} as the proportion of track that is covered by PCFAs. Notice that the unit cost is given in $\$/s$ and therefore it must be multiplied by the overflying time.

The initial boundary conditions of the problem are set as follows: $V(t_I) = 220$ [m/s], $\gamma(t_I) = 0°$, $\chi(t_I) = 0.115$ [rad], $m(t_I) = 76160$ [Kg], and the initial waypoint PEONS, i.e,, $\bar{x}(t_I) = \{-119.1674°, 38.5035°\}$. The initial time t_I, without any loss of generality, is set to zero. The final boundary conditions are specified by reaching the waypoint MAGIO, i.e., $\bar{x}(\tilde{t}_F) = \{-76.5964°, 41.5373°\}$. Notice that we do not impose constraints on when the final waypoint MAGIO is reached.

The selection of the number of samples for the collocation method explained in Section 3.2 is done by comparing solutions to MINLP problem computed with increasing number of subintervals in each phase until a negligible change in the objective function is achieved. Also, smooth functions for both states and controls are sought. Based on these criteria, the Hermite-Simpson collocation method is applied with a discretization using a total of $n = 180$, with $n_q = 12$, $q = 1, \ldots, 15$.

6. Case Study: Cost factor determination

Whereas there is little doubt about the importance of appropriate estimates for different climatic and cost components in identifying optimal flight strategies, research in this area remains limited. [12] assumes infinite cost when a flight flies into PCFAs. This is certainly inconsistent with reality, in which airlines care about cost associated with fuel consumption and CO_2 emissions. [58] argue that

changing altitudes to avoid PFCAs is justified if the radiative forcings from CO_2 and contrails are of comparable magnitude. This justification could be fine-tuned when more accurate climate impact estimates from CO_2 and contrails become available.

The Williams et al. study further gives rise to the issue of appropriate metric(s) to use for comparing the climate impact of CO_2 and persistent contrails. As admitted by [57] and [50], using radiative forcing as the measure for climate impact must be done with caution since radiative forcing includes the impact of all historic flights and does not account for the timescales of emissions. Given the distinct physical characteristics of different greenhouse gas agents, Global Warming Potential (GWP) and Global Temperature Potential (GTP) represent better metrics to quantify the true climate impact of different gases. While contrail is not categorized as a gas agent per se, we follow [60] by choosing GWP to unify the climate impact of CO_2 and contrail formation.[7] We incorporate it together with the cost of fuel consumption into the flight planning design. In what follows, the unit cost impacts from fuel consumption, CO_2 emission, and contrail formation are estimated.

6.1. Fuel cost

Fuel consumption is determined based upon the differential equation $\dot{m} = -T \cdot \eta$ in set (12), where η is the specific fuel consumption in $\frac{kg}{N \cdot s}$ and can be expressed following BADA 3.6 [32] as:

$$\eta = \frac{C_{f1}}{60000} \cdot (1 + \frac{3600 \cdot V}{1852 \cdot C_{f2}}), \tag{27}$$

where C_{f1} and C_{f2} are BADA parameters, and V represents the true airspeed.

Clearly, fuel consumption rate depends upon the true airspeed. In order to associate fuel consumption with cost, one simply needs to consider the amount of fuel consumed and the cost of jet fuel. Using the average kerosene price and its typical density, unit fuel cost is estimated to be:

$$C_F = \$1.30048 \text{ /(kg of fuel consumed)}. \tag{28}$$

6.2. CO_2 emissions cost

The unit CO_2 cost is obtained by multiplying the fixed ratio of CO_2 emissions and jet fuel consumption (in kg of CO_2 per kg of fuel consumed), and the social cost of CO_2. We follow [52] and [24], and use \$35 as the mean social cost for one ton of CO_2 emitted. For each kilogram of fuel consumed, the associated CO_2 emission cost is:

$$C_{CO_2} = \$0.11 \text{ /(kg of fuel consumed)} \tag{29}$$

[7] [31] use GWP to unify the effect of CO_2, H_2O, and NO_x emissions, but not contrail formation. Independent of our study, [48] employ average GTP to assess the climate impact of contrails and CO_2 in air traffic operations.

6.3. Passenger travel time

The travel cost per unit time for all passengers onboard is the product of passengers' value of time (VOT), aircraft seat capacity, and load factor. Following the U.S. Department of Transportation guidance on the economic value of passenger travel time [51], we use wage rate as a proxy for passenger VOT. The mean wage rate of \$21.35/hr in the U.S. in 2010 is used [10]. Assuming 200 seats in a typical B757 configuration with a factor of occupancy of 80%, the unit travel time cost C_t for all passengers onboard amounts to \$ 0.9488/sec.

6.4. Contrail formation cost

As already mentioned in the beginning of this section, we use GWP to quantify the climate impact of different greenhouse gas agents. Under its general definition, GWP for trace gas i is defined as the time-integrated commitment to radiative forcing from the instantaneous release of one kg of gas i relative to that of one kg of CO_2, i.e.:

$$GWP_i(H) = \frac{\int_0^H RF_i(t)dt}{\int_0^H RF_{CO_2}(t)dt} = \frac{\int_0^H a_i c_i(t)dt}{\int_0^H a_{CO_2} c_{CO_2}(t)dt}, \tag{30}$$

where H represents the time horizon and t is the time. $RF_{CO_2}(t)$ and $RF_i(t)$ denote the radiative forcing of CO_2 and gas i (in $[\frac{W}{m^2 Kg}]$) at time t respectively. a_{CO_2} and a_i denote the instantaneous radiative forcing from one kg emission of CO_2 and gas i respectively. $c_{CO_2}(t)$ and $c_i(t)$ denote the remaining abundance of CO_2 and gas i (due to decay) from their respective impulse of one [Kg] of CO_2 and gas i at time t. Clearly, GPW_i is a dimensionless measure.

To be consistent with the International Panel on Climate Change (IPCC) recommendations, we consider three time horizons for H: 20, 100, and 500 years [19]. The use of different time horizons can be interpreted as a discount rate, and reflect different value judgements about the importance of emission impacts that may occur in the near and far future.

The GWP metric was initially developed to quantify the climate impacts between different gas agents with CO_2 as the reference gas. However, when it comes to contrails the use of GWP becomes less straightforward since contrail is not a type of physical gas. To address this issue, [19] relates the total forcing due to contrails to the total emissions of CO_2 by the aviation fleet and calculates the average GWP for the present-day fleet on a per kg of CO_2 basis, assuming a 5-hour lifetime of contrails. The $GWP_{contrail}(H)$ is given by:

$$GWP_{contrail}(H) = \frac{\int_0^H RF^T_{contrail}(t)dt}{M_{CO_2} \int_0^H RF_{CO_2}(t)dt}, \tag{31}$$

where $RF^T_{contrail}(t)$ is the total RF in a year from all aviation-induced contrails at time period t; M_{CO_2} represents the total CO_2 (in kg) emitted in the same year. The first row in Table 3 reports the estimated $GWP_{contrail}$ values from [19], for three time horizons: 20, 100, and 500 years.

In the subsequent analysis, we use the above aggregated $GWP_{contrail}(H)$ as representative values given the limited body of literature on contrail formation at the individual flight level. On the other hand, persistent contrails can only be produced within PCFA regions, whereas CO_2 is emitted as long as the aircraft flies in the air. To compare the unit climate impact between contrail and CO_2 when the aircraft flies in a PCFA region, the $GWP_{contrail}(H)$ in (31) needs to be scaled up to reflect the actual amount of time flights spent in PCFAs. To this end, we introduce an adjusted $GWP^*_{contrail}(H)$ as below:

$$GWP^*_{contrail}(H) = \frac{GWP_{contrail}(H)}{\nu}, \tag{32}$$

where ν is the aggregate portion of time the global fleet spends in PCFAs. Previous estimates report that airlines cruise 10-20% of the time in air masses that are humid enough for the formation of persistent contrails [30, 41, 22]. In light of these estimates, we choose $\nu = 15\%$. The resulting $GWP^*_{contrail}(H)$ values are shown in the second row of Table 3.

Table 3: GWP for contrail with different time horizon choices

	H=20 Years	H=100 Years	H=500 Years
$GWP_{contrail}(H)$	0.74	0.21	0.064
$GWP^*_{contrail}(H)$	4.93	1.40	0.423

To further convert the climate impact of contrails to dollar values, we assume that $GWP^*_{contrail}(H)$ equals the ratio of the social cost of formed contrails over that of emitted CO_2 per unit of time when the aircraft flies inside PCFAs. The unit contrail formation cost [\$/s] is the unit CO_2 emission cost [\$/Kg] in equation (29) multiplied by the $GWP^*_{contrail}(H)$ values and the nominal fuel flow [Kg/s] at a given altitude, the latter of which is the fuel flow at the nominal speed for a given altitude and obtained from BADA 3.6 [32]. Because the nominal fuel flow varies by altitude, the contrail formation cost is a function of the flying altitude, in addition to the choice of time horizons (Table 4). The total contrail cost when the flight flies through a PCFA is the product of the unit contrail cost of the flight and the time spent in the PCFA region.

Table 4: Contrail cost [\$/s] with different time horizon choices.

FL	Fuel Flow [Kg/min]	$C_C(FL)$ [\$/s] H=20 Years	H=100 Years	H=500 Years
270	60.7	0.883	0.2337	0.0706
290	61.7	0.8976	0.2375	0.0718
310	62.7	0.9121	0.2414.	0.0729
330	62.2	0.9049	0.2395	0.0724
350	60.1	0.8743	0.2314	0.0699
370	58.7	0.8539	0.2260	0.0683
390	58.1	0.8452	0.2237	0.0676
410	58.0	0.8438	0.2233	0.0675

Table 5: Computational times.

	H=20 Years	H=100 Years	H=500 Years	Min. Operational cost
Computational time [s]	235.87	120.72	686.14	730.04

7. Case study: Solution procedure and results

The resulting problem is a non-convex MINLP program. The existing off-the-shelf MINLP solvers can solve problems to optimality but are typically limited to a small number of variables.[8] In this study we use a two-step heuristic approach implemented in the solver Bonmin. In the first step, the integer values are relaxed to continuous domain [0,1], and the resulting NLP subproblems are solved to local optimality using the non-linear solver Ipopt. Then, using a branch-and-bound framework, we seek a solution that satisfies the integer requirements among the different subproblems. The branching rules are heuristics to reduce the size of the tree search: we do not explore the whole tree of solutions, only those branches in which *good enough*[9] solutions are expected.

We consider four different experiments. The first three experiments consider contrail cost and use time horizons of 20, 100, and 500 years, respectively, and the fourth one does not take into account the cost of persistent contrail formation, but only the cost of passenger travel, fuel and CO_2 emissions. The last case provides a minimum operational cost baseline.

With regard to computational issues, the problem has 2976 variables (32 integer), 2352 equality constraints, and 1252 inequality constraints. Table 5 shows computational times for the different experiments. Computations were performed on a Mac OS X 2.56 GHz laptop with 4 GB RAM. In order to improve computational time we have properly scaled the differential equations. We have also ordered them in the proper way so that the resulting Jacobian and Hessian matrixes are sparse. In addition, we have fed the MINLP program with an improved initial guess that can be calculated, for instance, based on standard aircraft performances. As a result, computational time is relatively low, always within less than 15 minutes. Typically, the more important cost related to binary variables is, i.e., contrail cost in this case, the less computational time the algorithm employs. This is due to the fact that during the branching process, if the algorithm detects groups (branches) of solutions that differ a lot from others (in this case, due to the existence of different discrete cost associated with contrail formation) it is able to rapidly discard the least efficient ones. For more information on this point, the reader is invited to the discussion on the sensitivity of the algorithm to discrete costs in [8].

Figure 6 displays the horizontal profile of the obtained optimal path, which connects all predetermined waypoints in sequence, and it is identical for all four cases by construction.

[8] For example, we have tried to use Couenne solver which did not yield any solution.
[9] Please, refer to Appendix C for a more rigorous explanation.

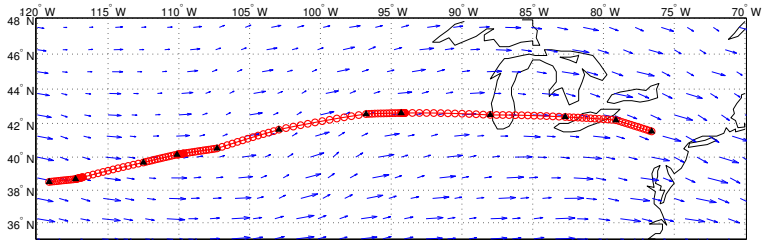


Figure 6: Horizontal profile.

Table 6: Flight level allocation.

	H=20 Years	H=100 Years	H=500 Years	Min. Operational cost
Step Climb 1	$FL350$	$FL350$	$FL370$	$FL370$
Step Climb 2	$FL370$	$FL370$	$FL410$	$FL410$
Step Climb 3	$FL350$	$FL350$	$FL410$	$FL410$
Step Climb 4	$FL290$	$FL290$	$FL410$	$FL410$

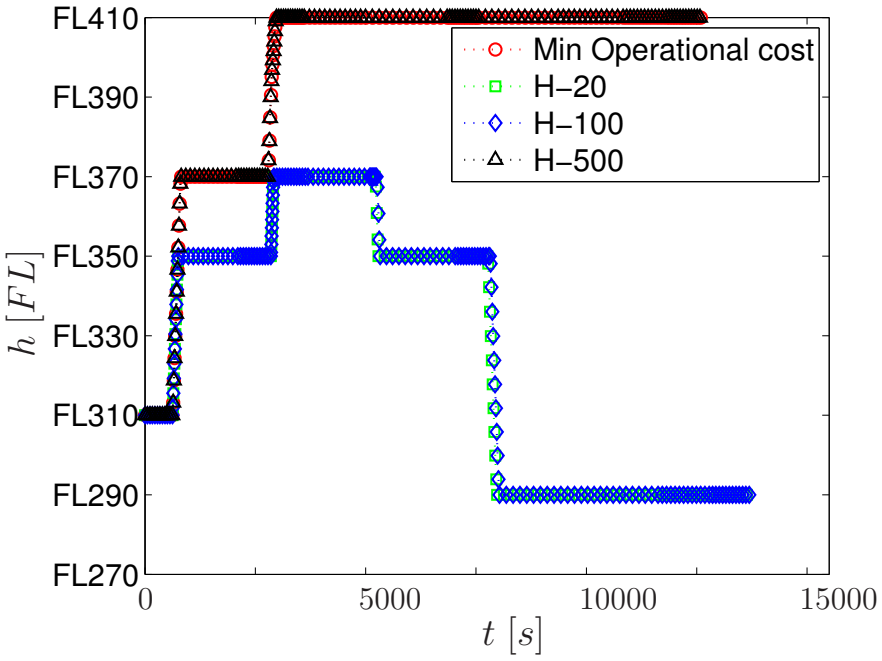


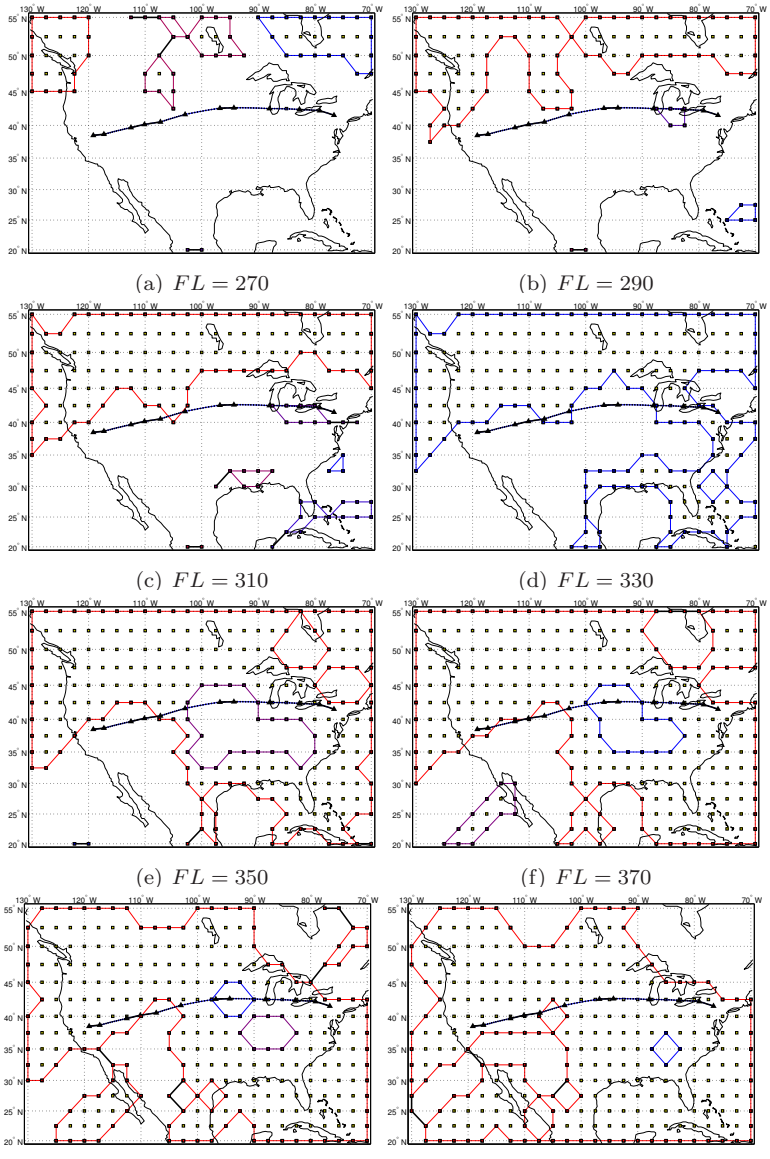
Figure 7: Vertical profiles: square-green line corresponds to H-20; diamond-blue line corresponds to H-100 (H-20 and H-100 share the same vertical profile); triangle-black corresponds to H-500; circle-red line corresponds to minimum operational cost (H-500 and minimum operational cost share the same vertical profile). The markers correspond to the computed samples.

Table 6 reports, for the different experiments, the flight levels that are achieved by the aircraft after each of the four permitted step climbs/descents. The corresponding vertical flight profiles are illustrated in Figure 7. The minimum operational cost (which includes passenger travel time, fuel, and CO_2 costs) seeks to fly as high as possible (subject to altitude constraints[10]) in order to consume less fuel and be able to fly faster as the speed of sound decreases with altitude. The scenario H-500 present the same vertical profiles. In this case, the cost due to contrail is small enough to permit the aircraft to fly in PFCA regions while saving fuel and time. On the contrary, H-20 and H-100 scenarios present a different vertical profile. In these cases, the flight would spend less time at higher altitude, avoiding the most significative PCFA regions.

Looking at Figure 8, one can observe that most of the route goes into PCFA regions if one flies at FL390 or FL410. This is what the minimum operational cost and H-500 trajectory would do, precisely because cost due to contrails is either not considered or very little. However, when one considers cost due to contrails for time horizons of 20 and 100 years, the first step climb is performed up to FL350 in both cases. This is the highest flight level at which the aircraft would not fly into a PCFA region before reaching the waypoint MTU. Thereafter, the aircraft would follow vertical profiles that incur some contrail impact in exchange for reduced CO_2 emissions, fuel burnt, and reduced flight time (notice that FL370 and FL350 during tracks joining MTU and FOD go into PCFA regions). The formation of persistent contrails, which depends on the time horizon, clearly has a substantial impact on the optimal vertical profile.

Values of the objective function, fuel consumption, flight time, and total contrail cost can be consulted in Table 7. For more detailed breakdown information, the switching times (labeled *Swt. times*), cumulative fuel consumption (labeled *con.*) by the end of each leg, and contrail formation cost (labeled *CC*) for each leg are documented in Table 8. For instance, by comparing H-20 and H-100 scenarios with H-500 and minimum operational cost examples, we can quantify that the contrail mitigation strategy costs in approximately 300-350 kg of extra consumed fuel (together with its correspondent extra CO_2 emissions) and 1100 sec more flight time. Notice that, even though we have two pairs (H-20 and H-100 on the one hand; H-500 and min. operational cost on the other) of experiments, each resulting in almost the same solution, slightly different fuel consumptions and flight times are observed. This difference is attributed to the different contrail costs incurred when flying into PCFA regions. Since the unit cost is expressed in dollars per time unit, with higher contrails costs it makes sense to accelerate the aircraft. This can be particularly observed in phase 7 when comparing H-20 and H-100 flight times.

[10]Notice that, due to operational constraints, aircraft have an operational ceiling that increases as fuel is burnt. Moreover, the aircraft is only permitted to ascent/descent three flight levels at a time. In this sense an aircraft might not be able to reach a determined FL at the beginning of the flight, but however, reach it later on when the weight of the aircraft is smaller due to fuel burnt.



(a) $FL = 270$ (b) $FL = 290$

(c) $FL = 310$ (d) $FL = 330$

(e) $FL = 350$ (f) $FL = 370$

(g) $FL = 390$ (h) $FL = 410$

Figure 8: Favorable regions of contrail formation over USA at different flight levels. The horizontal route is depicted to illustrate how the same horizontal route under different flight levels might increase/reduce potential persistent contrail generation.

Table 7: Objective, consumption, flight time, and contrail cost values.

	H=20 Years	H=100 Years	H=500 Years	Min. Operational cost
Objective [$]	28283	27900	26997	26358
Consumption [kg]	10513	10474	10229	10215
Flight time [s]	13636	13689.4	12574	12598
Contrail Cost [$]	516.5	138.9	638.7	-

Table 8: Switching times [s], Fuel Consumed [kg], and Contrail Costs [$].

H-20			H-100			H-500			Min. Operational cost	
Swt. times	Con.	CC	Swt. times	Con.	CC	Swt. times	Con.	CC	Swt. times	Con.
$\tilde{t}_1 = 622.1$	464.2	0	$\tilde{t}_1 = 622$	464.3	0	$\tilde{t}_1 = 620.8$	471.4	0	$\tilde{t}_1 = 620.8$	470.59
$\tilde{t}_2 = 756.6$	660.5	–	$\tilde{t}_2 = 756.5$	660.6	–	$\tilde{t}_2 = 819.6$	760	–	$\tilde{t}_2 = 819.9$	759.2
$\tilde{t}_3 = 2104.7$	1775.1	0	$\tilde{t}_3 = 2104.4$	1775.3	0	$\tilde{t}_3 = 2061.4$	1819.6	22.7	$\tilde{t}_3 = 2064.4$	1816.3
$\tilde{t}_4 = 2850.6$	2420.4	0	$\tilde{t}_4 = 2850.6$	2417.8	0	$\tilde{t}_4 = 2783.7$	2454.9	16.3	$\tilde{t}_4 = 2788.9$	2451
$\tilde{t}_5 = 2921$	2515.6	–	$\tilde{t}_5 = 2921.3$	2510.1	–	$\tilde{t}_5 = 2978.7$	2683.6	–	$\tilde{t}_5 = 2985.5$	2680.3
$\tilde{t}_6 = 3685.7$	3185.2	94	$\tilde{t}_6 = 3695.5$	3171.1	25.2	$\tilde{t}_6 = 3611.2$	3206.3	55.8	$\tilde{t}_6 = 3616.9$	3202.2
$\tilde{t}_7 = 5049.7$	4426.3	294.7	$\tilde{t}_7 = 5088.6$	4388.7	79.6	$\tilde{t}_7 = 4997.9$	4344.2	32.4	$\tilde{t}_7 = 5003.6$	4340
$\tilde{t}_8 = 5289.7$	4574.1	–	$\tilde{t}_8 = 5328.6$	4537.5	–	$\tilde{t}_8 = 4997.9$	4344.2	–	$\tilde{t}_8 = 5003.6$	4340
$\tilde{t}_9 = 6973.5$	5974.4	127.3	$\tilde{t}_9 = 7027.3$	5932.1	33.9	$\tilde{t}_9 = 6844.3$	5842.1	124.6	$\tilde{t}_9 = 6850$	5838.1
$\tilde{t}_{10} = 7736.7$	6579.7	0	$\tilde{t}_{10} = 7790.4$	6538.8	0	$\tilde{t}_{10} = 7560.1$	6417.9	48.3	$\tilde{t}_{10} = 7565.9$	6413.8
$\tilde{t}_{11} = 7976.7$	6635.5	–	$\tilde{t}_{11} = 8030.4$	6584.7	–	$\tilde{t}_{11} = 7560.1$	6417.9	–	$\tilde{t}_{11} = 7565.9$	6413.8
$\tilde{t}_{12} = 9831.2$	7986.5	0	$\tilde{t}_{12} = 9884.7$	7946.2	0	$\tilde{t}_{12} = 9358.8$	7852	121.4	$\tilde{t}_{12} = 9364.5$	7848.1
$\tilde{t}_{13} = 11571.5$	9210.7	0	$\tilde{t}_{13} = 11624.7$	9171	0	$\tilde{t}_{13} = 10831.5$	9013.7	99.4	$\tilde{t}_{13} = 10838.7$	9006.6
$\tilde{t}_{14} = 12731.4$	9999.8	0	$\tilde{t}_{14} = 12784.5$	9960.4	0	$\tilde{t}_{14} = 11806.2$	9772.7	65.8	$\tilde{t}_{14} = 11821.8$	9758.6
$\tilde{t}_{15} = 13636.2$	10513	0	$\tilde{t}_{15} = 13689.2$	10474	0	$\tilde{t}_{15} = 12574.8$	10229	51.9	$\tilde{t}_{15} = 12594.2$	10216

Analyzing the velocity profiles, in Figure 9 we observe how velocity slightly decreases as fuel is burnt. This is because the optimal cruising speed (in terms of minimum fuel consumption) is reduced as the aircraft loses weight due to fuel consumption. Notice that the optimal cruising speed depends on the state of the aircraft, mainly on altitude and mass. Indeed, for both H-500 and min. operational cost, the state of the aircraft is such that the optimal cruising speed exceeds the maximum Mach number, and this explains the constant velocity profile at FL410. It is worth mentioning that this behavior is anomalous, since BADA 3.6 aerodynamic model does not consider compressibility effects.[11] This modeling issue leads to lower than real drag at high Mach numbers, resulting in higher than real optimum speeds.

The evolution of the remaining state and control variables, i.e., heading angle, flight path angle, mass, bank angle, thrust, and coefficient of lift, is provided in Appendix D. In general, these variables follow similar patterns across all four cases.

We have shown that solving a rather complete aircraft dynamic model such as Eq. (12) is feasible and requires low computation time in our study. Nonetheless, one might be interested in comparing the performance of such a dynamic model with that of a simplified version. To this end, we also consider the fol-

[11] Compressibility effects refer to any of the effects that result from changes in the flow field about an airplane when the velocity at some point in the field reaches the local speed of sound. In particular, these effects modify dramatically the distribution of pressures over the airfoil/wing resulting in an important increase in drag force.

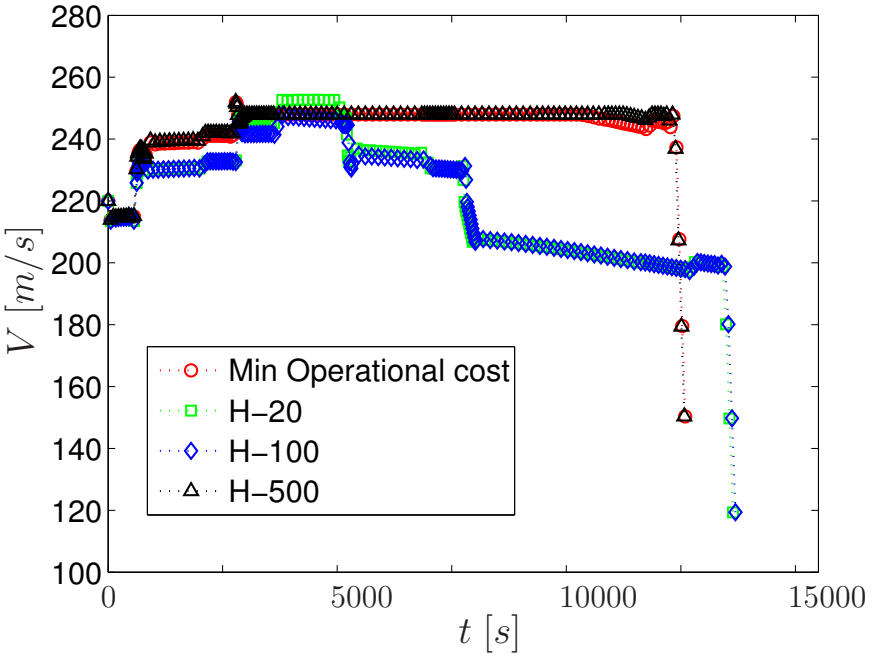


Figure 9: velocity: square-green line corresponds to H-20; diamond-blue line corresponds to H-100; triangle-black corresponds to H-500; circle-red line corresponds to minimum operational cost . The markers correspond to the computed samples.

lowing simplified dynamic model for aircraft performance

$$\frac{d}{dt}\begin{bmatrix} V \\ \gamma \\ s \\ h_e \\ m \end{bmatrix} = \begin{bmatrix} \frac{T(t)-D(h_e(t),V(t),C_L(t))-m(t)\cdot g\cdot \sin\gamma(t)}{m(t)} \\ \frac{L(h_e(t),V(t),C_L(t))-m(t)\cdot g\cdot \cos\gamma(t)}{m(t)\cdot V(t)} \\ V(t)\cdot\cos\gamma(t)+W(s(t)) \\ V(t)\cdot\sin\gamma(t) \\ -T(t)\cdot\eta(V(t)) \end{bmatrix}. \tag{33}$$

In this model, the lateral navigation is enforced to follow the nominal path, i.e., the minimum distance path between waypoints. This distance is denoted s. Therefore, the bank angle does not act as a control anymore, and one can assume χ and θ_e to be known. Indeed, this is a 2DOF aircraft dynamic model.

If one does not take into consideration the effects of wind, it is certainly true that the simplified model in Eq. (33) might be more appropriate. (it provides essentially the same solutions with computational times always below 120 sec.) This is in part due to the fact that the optimal path coincides with the minimum distance path (Great Circle Distance (GCD) path).

However, when wind effects are taken into consideration, the optimal path does not necessarily coincide with the GCD path. In the case that the trajectory

is constrained through the overfly of several waypoints (as it is the case herein), both paths can still slightly differ. In this case, the solutions to the problem using the simplified model in Eq. (33) are similar to the ones obtained using the complete dynamic model in Eq. (12), with computational times always below 160 sec. Therefore, when the effects of wind substantially modify the optimal path, e.g., in free routing airspaces, a complete 3DOF model is preferred. In areas of highly constrained airspace, one might think of using the simplified model to reduce computational efforts. The simplified model may also be more appropriate for modeling trajectories of multiple aircraft. On the other hand, given the operational trends of allowing more areas of free routing, the advantage of using the complete 3DOF model would become more manifested.

8. Conclusions

This study presents a promising beginning towards understanding the optimal flight trajectories in the presence of contrail formations. The multiphase MIOCP approach provides a solution to the problem combining both discrete and continuous elements, which was an unsolved issue in aircraft trajectory optimization. By analyzing trade-offs between passenger travel time, fuel, $CO2$ emissions, and contrail-based costs, we obtain the discrete sequence of flight levels, the continuous optimal control law, the complete continuous state of the aircraft over time domain, including the 4D optimal trajectory, and the overflying times over waypoints and after performing step climbs/descents. Therefore, the solution to the problem provides a powerful tool for designing more efficient trajectories within a future, environmentally friendlier air navigation system.

Results show identical vertical profiles for min. operational cost and H-500 scenarios, which may suggest that the contrail formation cost under H-500 becomes so small that trajectories will not be affected. On the contrary, by comparing the H-20 and H.100 scenarios with the minimum operational cost example, we conclude that the contrail mitigation strategy costs approximately 300-350 kg of extra consumed fuel (together with its correspondent extra CO_2 emissions) and 1100 sec of extra flight time. Therefore, this study is a step toward developing flight planning policies that include climate-based (contrails and emission sources) cost within operational costs.

Future efforts are headed towards extending the application to multi-aircraft scenarios. This would be of interest in order to assess the viability of this approach in terms of flow management, sector capacity or controllers workload.

Acknowledgments

This research was carried out at NEXTOR, University of California at Berkeley, where Manuel Soler was a visiting scholar hosted by Prof. Mark Hansen. The authors would like to thank all members of NEXTOR at UC Berkeley for fruitful and interesting discussions. This work was partially financed by the URJC, encouraging mobility in PhD candidates with the following program: *Programa propio de fomento y desarrollo de la investigación.*

Appendix A. Reformulation

The multiphase optimal control problem is converted into a conventional optimal control problem by making the unknown switching times part of the state and then introducing a new independent variable with respect to which the switching times are fixed. In this reformulated problem, there is a linear relation between the new variable and time, but the slope of this linear relation changes on each interval between two switches. These slopes, which are part of the solution to the optimal control problem, are actually time scaling factors that determine the optimal switching times. See Figure A.10.

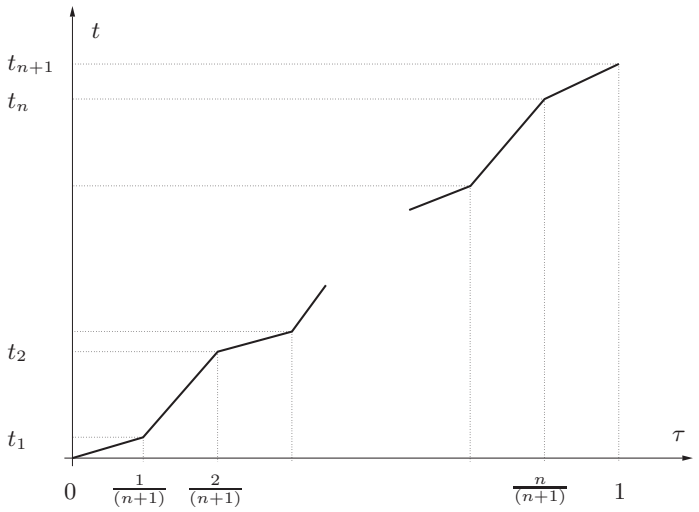


Figure A.10: Relation between scaled time, τ, and real (unscaled) time t.

Without loss of generality, we can assume that $t_I = t_0 = 0$ and $t_F = t_{n+1} = 1$. The first step is to introduce the new state variables, $x_{n_x+1}, \ldots, x_{n_x+n}$, which correspond to the switching times, t_q, $q \in \{1, 2, \ldots, n\}$, i.e., $x_{n_x+q} = t_q$, with $\dot{x}_{n_x+q} = 0$.

We then introduce the new independent variable, τ. The relation between τ and t changes on each interval $[t_q, t_{q+1}]$. We establish piecewise linear correspondence between time, t, and the new independent variable, τ, so that for every chosen fixed point, τ_q, $q = 1, \ldots, n$, t equals t_q. Any monotonically increasing sequence of n numbers on interval $[0, 1]$ could be used. We set $\tau_q = q/(n+1)$. As a result we obtain the following expression

$$t = \begin{cases} (n+1)x_{n_x+1}\tau, & 0 \le \tau \le \frac{1}{n+1} \\ \dots \\ (n+1)(x_{n_x+q+1} - x_{n_x+q})\tau + (q+1)x_{n_x+q} - qx_{n_x+q+1}, & \frac{q}{n+1} < \tau \le \frac{q+1}{n+1} \\ \dots \\ (n+1)(1 - x_{n_x+n})\tau + (n+1)x_{n_x+n} - n, & \frac{n}{n+1} < \tau \le 1 \end{cases}$$

By introducing the new independent variable, τ, the differential equation in (2) on the interval $[t_q, t_{q+1}]$ becomes:

$$x' = (n+1)(x_{n_x+q+1} - x_{n_x+q})\hat{f}_q(x, u, \tau), \tag{A.1}$$

where $(\cdot)'$ denotes the derivative of $(\cdot)$ with respect to the new independent variable, τ, and

$$\hat{f}_q(x, u, \tau) = f_q(x, u, t(\tau)).$$

Let $\hat{x}$ be the extended state vector

$$\hat{x} = [x_1, \dots, x_n, x_{n_x+1}, \dots, x_{n_x+n}]^T,$$

Then, define on each interval $\frac{q}{n+1} < \tau \le \frac{q+1}{n+1}$

$$\hat{L}(\hat{x}, u, \tau) = (n+1)(x_{n_x+q+1} - x_{n_x+q})L(x, u, t(\tau)).$$

We can rewrite the objective functional (considering the Mayer term only applies at the final state value) as follows:

$$J = E\left(\hat{x}(1)\right) + \int_0^{\frac{1}{n+1}} \hat{L}(\hat{x}, u, \tau)d\tau + \dots + \int_{\frac{n}{n+1}}^1 \hat{L}(\hat{x}, u, \tau)d\tau = E\left(\hat{x}(1)\right) + \int_0^1 \hat{L}(\hat{x}, u, \tau)d\tau, \tag{A.2}$$

and the task is to minimize J in the extended state space, subject to the parameterized system given in (A.1), and to the corresponding path constraints. The new equivalent problem is a conventional optimal control problem. The last n components of the optimal solution of this problem, $\hat{x}^*$, will be the optimal switching times $t_q, q = 1, \dots, n$.

Appendix B. Hermite-Simpson Gauss-Lobatto collocation method

For the Hermite-Simpson Gauss-Lobatto integration rule the root of the corresponding Jacobi polynomial (the collocation point) is 0, yielding the following approximate integration rule:

$$\int_{t_i}^{t_{i+1}} f(t)dt \approx \frac{h_i}{6}[f(t_i) + 4f(t_{i,C}) + f(t_{i+1})], \tag{B.1}$$

which has an order of accuracy of 4.

For the sake of clarity in the exposition, consider the differential equation $\frac{dx}{dt} = f(x)$. Simpson's rule in equation (B.1) is formulated considering a quadratic approximation of the integrand, and thus the state as a function of time $x(t)$ must be approximated by a cubic polynomial. Moreover, the polynomial used to interpolate $f(x)$ at the endpoints and center points of the subinterval is obtained as an integration of the above mentioned cubic polynomial. In this case, parameters representing the state at the endpoints x_i and x_{i+1} are used to formulate a constraint. Knowing x_i, x_{i+1}, $f_i = f(x_i)$ and $f_{i+1} = f(x_{i+1})$, a Hermite-cubic polynomial representing the state $x(t)$ between the endpoint times t_i and t_{i+1} can be constructed such as both the values and first derivatives of the interpolant polynomial coincide with the values and first derivatives of function $f(x)$ at the extremes of the subinterval. Figure B.11 illustrates it. Such polynomial is used to generate an internal collocation point $x_{i,C}$ per subinterval, whose numerical expression is

$$x_{i,C} = \frac{1}{2}(x_i + x_{i+1}) + \frac{h_i}{8}(f(x_i) - f(x_{i+1})), \tag{B.2}$$

where $x_{i,C}$ is a discrete approximation for $x(t)$ at $t_{i,C} = \frac{t_i + t_{i+1}}{2}$ and $i = 0, \ldots, N_d - 1$. The Simpson's system constraint is then formulated using $x_{i,C}$ to evaluate the system equation resulting in a discrete value at center point of the subinterval $f_{i,C} = f(x_{i,C})$. Then, by enforcing $f_{i,C}$ to be equal to the first time derivative of the Hermite-cubic interpolant polynomial at the center point of the subinterval, i.e., $\dot{x}_{i,C} = f_{i,C}$, one defect equation per subinterval is generated:

$$c_i^S(x_i, x_{i+1}) = x_i - x_{i+1} + \frac{h_i}{6}(f(x_i) + 4f(x_{i,C}) + f(x_{i+1})) = 0, \tag{B.3}$$

with $i = 0, \ldots, N_d - 1$. These constraints are known as Hermite-Simpson defect constraints [25].

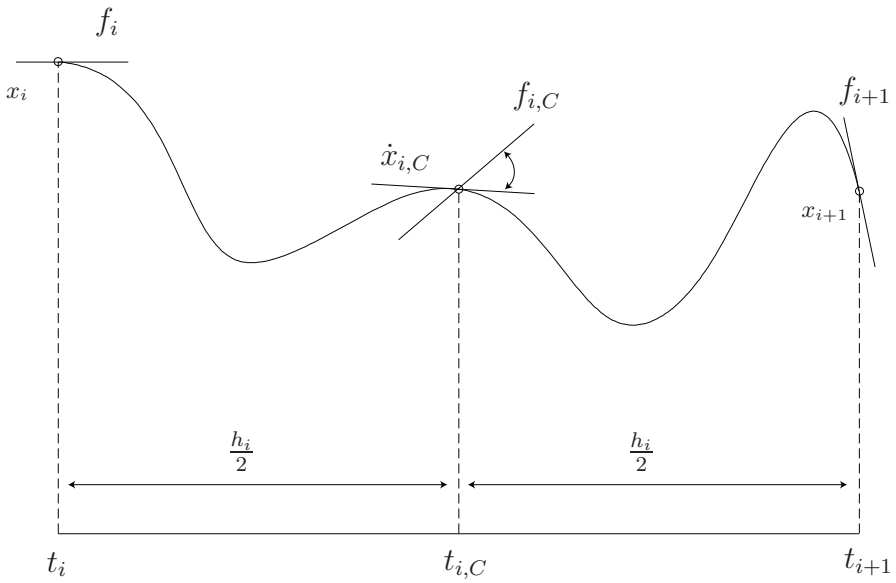


Figure B.11: Hermite-Simson collocation scheme.

Appendix C. MINLP approach

The heuristic algorithm employed to solve the resulting non-convex MINLP can be described as follows:

Approximations are computed by using the relaxed integer optimal control problem (RIOCP), where the constraints $v^q \in \{0,1\}$ are relaxed to $v^q \in [0,1]$, for $q = 0,\ldots,n$. A locally optimal solution to the RIOCP can be computed with a nonlinear programming algorithm, for instance the interior point algorithm implemented by IPOPT [53]. The branch-and-bound framework is then used to find a solution that satisfies also the integrity requirements $v^q \in \{0,1\}$, $q = 0,\ldots,n$. This variant of branch-and-bound is usually called NLP based branch-and-bound or NLP BB for short [7].

The first step of the branch-and-bound algorithm is to solve the RIOCP. If the solution obtained by solving the RIOCP is integer feasible (all variables v^q take value 0 or 1) it specifies a sequence of points and the algorithm stops. If no solution to the RIOCP is found the algorithm stops. If an upper bound $\beta_{\mathcal{U}}$ on the value of the optimal solution is known and the value of the solution of the RIOCP is above $\beta_{\mathcal{U}}$, the algorithm also stops (fixing the infeasibility of the solution should increase the objective value of the solution). Otherwise, the algorithm divides the feasible region in two by fixing one of the variables v^q such that $\hat{v}^q \notin \{0,1\}$ to 0 and to 1 successively.

Applying the above steps recursively, yields to a tree $\mathcal{T}$ of partial assignment for the binary variables. At each node of this tree, a subset $\mathcal{L}$ of the variables v^q are fixed to 0 and a subset $\mathcal{U}$ of the variables v^q are fixed to 1, and a local

optimum of the restriction of the RIOCP where the variables in $\mathcal{L}$ and $\mathcal{U}$ are fixed is to be sought. This restricted relaxed optimal control program is referred to as RIOCP($\mathcal{L}$,$\mathcal{U}$). The value of the upper bound $\beta_{\mathcal{U}}$ is initially $+\infty$ and is updated whenever a new integer feasible solution is found such that the cost is improved.

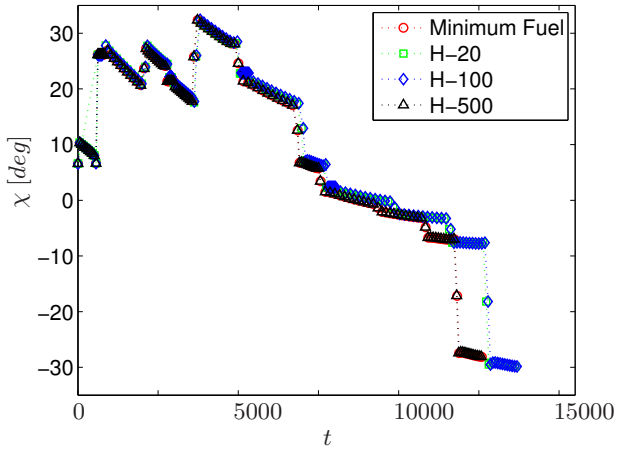
The pseudo-code of the NLP BB is given in Algorithm 1. Several solvers implement this algorithm for example MINLP BB [27] and SBB [11]. In this paper, the solver BONMIN is used [6]. BONMIN is an open-source MINLP solver implementing several different algorithms for solving mixed integer nonlinear optimization problems. Source code and binaries of BONMIN are available from COIN-OR (http://www.coin-or.org). BONMIN is called through the AMPL modeling language.

Two critical steps for the practical efficiency of algorithm 1 which have not been explicated are the selection of the next subproblem to evaluate (step 2), and the choice of the variable to divide the feasible region (step 5). For these two steps standard rules implemented in BONMIN are used. The subproblem selected in step 2 is always the one with lowest β^q (best-bound rule). Whereas for choice of the variable $\hat{v}^q$, a default strategy in BONMIN is used which is a combination of strong-branching and pseudo-costs [7].

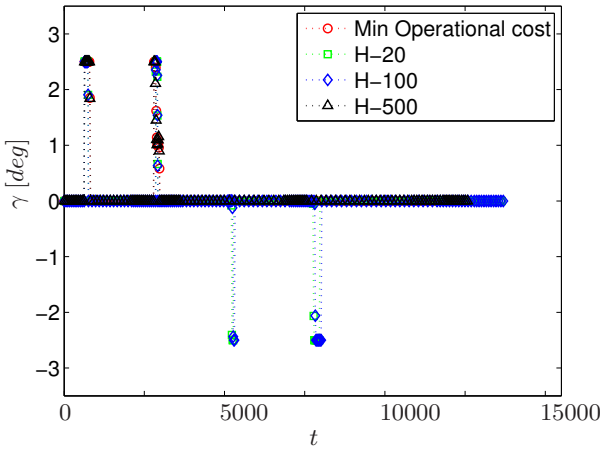
Algorithm 1 NLP BB

0. **Initialize.**
   $\mathcal{T} \leftarrow \{(\emptyset,\emptyset)\}$. $\beta_{\mathcal{U}} = \infty$. $v^{q*} \leftarrow$ NONE.
1. **Terminate?**
   Is $\mathcal{T} = \emptyset$? If so, stop and return the sequence described by v^{q*}.
2. **Select.**
   Choose and delete a problem $\mathcal{N}^l = (\mathcal{L}^l,\mathcal{U}^l)$ from $\mathcal{T}$.
3. **Evaluate.**
   Solve the RIOCP($\mathcal{L}^l$, $\mathcal{U}^l$). If no solution can be found go to step 1, else let $\beta^R_{\mathcal{L}^l,\mathcal{U}^l}$ be its objective function value and $\hat{v}^q$ be the values for the relaxed binary variables.
4. **Prune.**
   If $\beta^R_{\mathcal{L}^l,\mathcal{U}^l} \geq \beta_{\mathcal{U}}$ go to step 1. If $\hat{v}^q \notin \{0,1\}^{n+1} \times \{0,1\}^{n_{vq}}$ go to step 5, else let $\beta_{\mathcal{U}} \leftarrow \beta^R_{\mathcal{L},\mathcal{U}}$, $v^{q*} \leftarrow \hat{v}^q$, and delete from $\mathcal{T}$ all problems with $\beta^q_{\mathcal{L}} \geq \beta_{\mathcal{U}}$. Go to step 1.
5. **Divide.**
   Create two new nodes $\mathcal{N}^{|\mathcal{T}|}$, and $\mathcal{N}^{|\mathcal{T}|+1}$. Choose $\hat{q}$ such that $\hat{v}^{\hat{q}} \notin \{0,1\}$. Let $\beta^{|\mathcal{T}|} \leftarrow \beta^{|\mathcal{T}|+1} \leftarrow \beta^R_{\mathcal{L},\mathcal{U}}$ and add the problem $\mathcal{N}^{|\mathcal{T}|} = (\mathcal{L}^l \cup v^{\hat{q}},\mathcal{U}^l)$ and $\mathcal{N}^{|\mathcal{T}|+1} = (\mathcal{L}^l,\mathcal{U}^l \cup v^{\hat{q}})$. Go to 1.

Appendix D. State and Control Variables

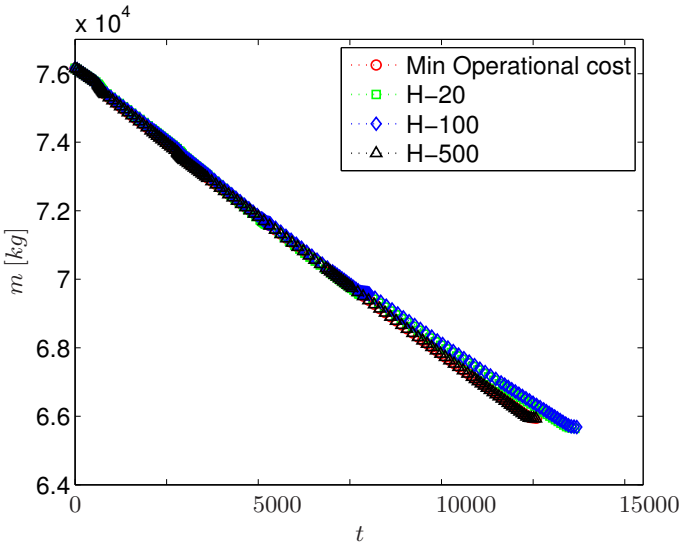


(a) χ

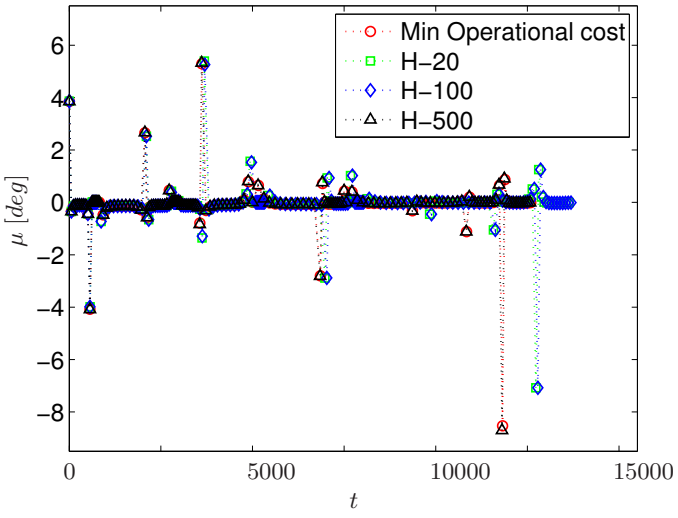


(b) γ

Figure D.12: State and control variables: square-green line corresponds to H-20; diamond-blue line corresponds to H-100; triangle-black corresponds to H-500; circle-red line corresponds to minimum operational cost . The markers correspond to the computed samples.

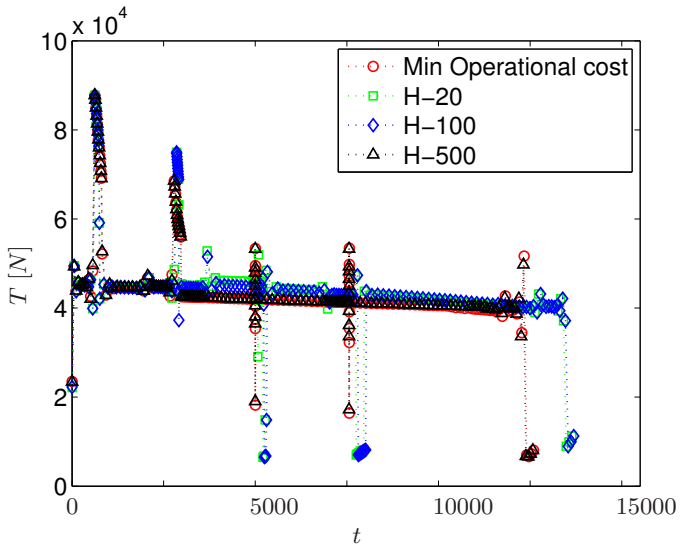


(a) m

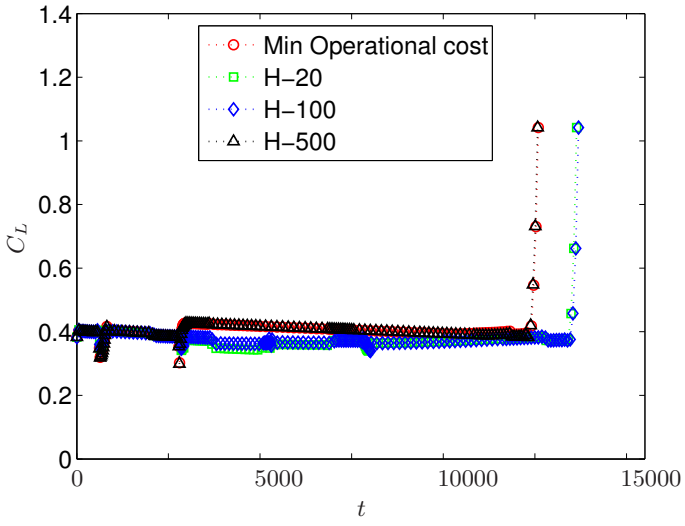


(b) μ

Figure D.13: State and control variables: square-green line corresponds to H-20; diamond-blue line corresponds to H-100; triangle-black corresponds to H-500; circle-red line corresponds to minimum operational cost . The markers correspond to the computed samples.



(a) T



(b) C_L

Figure D.14: State and control variables: square-green line corresponds to H-20; diamond-blue line corresponds to H-100; triangle-black corresponds to H-500; circle-red line corresponds to minimum operational cost . The markers correspond to the computed samples.

References

[1] Alduchov, O., Eskridge, R., 1996. Improved magnus form approximation of saturation vapor pressure. Journal of Applied Meteorology 35, 601–609.

[2] Appleman, H., 1953. The formation of exhaust condensation trails by jet aircraft. Bull. Amer. Meteor. Soc 34, 14–20.

[3] Benjamin, S., Brundage, K., Miller, P., Smith, T., Grell, G., Kim, D., Brown, J., Schlatter, T., Morone, L., 1994. The rapid update cycle at NMC, in: Preprints, 10th Conference on Numerical Weather Prediction, Portland, Oregon, USA, American Meteorological Society, pp. 566–568.

[4] Benjamin, S., Dévényi, D., Weygandt, S., Brundage, K., Brown, J., Grell, G., Kim, D., Schwartz, B., Smirnova, T., Smith, T., et al., 2004. An hourly assimilation–forecast cycle: the RUC. Monthly Weather Review 132, 495–518.

[5] Bixby, R., Fenelon, M., Gu, Z., Rothberg, E., Wunderling, R., 2004. The Sharpest Cut. SIAM. chapter Mixed-Integer Programming: A Progress Report. MPS-SIAM Series on Optimization, pp. 309–326.

[6] Bonami, P., Biegler, L.T., Conn, A.R., Cornuéjols, G., Grossmann, I.E., Laird, C.D., Lee, J., Lodi, A., Margot, F., Sawaya, N., Wächter, A., 2008. An algorithmic framework for convex mixed integer nonlinear programs. Discrete Optimization 5, 186–204.

[7] Bonami, P., Kilinç, M., Linderoth, J., 2012. Algorithms and software for convex mixed integer nonlinear programs, in: Lee, J., Leyffer, S. (Eds.), Mixed Integer Nonlinear Programming. Springer New York. volume 154 of *The IMA Volumes in Mathematics and its Applications*, pp. 1–39.

[8] Bonami, P., Olivares, A., Soler, M., Staffetti, E., 2013. Multiphase mixed-integer optimal control approach to aircraft trajectory optimization. Journal of Guidance, Control, and Dynamics 36, 1267–1277.

[9] Branicky, M.S., Borkar, V.S., Mitter, S.K., 1998. A unified framework for hybrid control: Model and optimal control theory. IEEE Transactions on Automatic Control 43, 31–45.

[10] Bureau of Labor Statistics, 2010. Occupational employment and wages. Technical Report. BLS Economics News Release Web. http://www.bls.gov/news.release/ocwage.htm. Accessed October 19 2011.

[11] Bussieck, M., Drud, A., 2001. SBB: a new solver for mixed integer nonlinear programming. GAMS Development Corp .

[12] Campbell, S., Neogi, N., Bragg, M., 2008. An optimal strategy for persistent contrail avoidance, in: AIAA Guidance, Navigation and Control Conference.

[13] Campbell, S.E., Bragg, M.B., Neogi, N.A., 2013. Fuel-optimal trajectory generation for persistent contrail mitigation. Journal of Guidance, Control, and Dynamics 36, 1741–1750.

[14] Chen, N.Y., Sridhar, B., Ng, H.K., Li, J., 2014. Evaluating tradeoffs between environmental impact and operational costs for enroute air traffic, in: Guidance, Navigation, and Control Conference.

[15] Chen, N.Y., Sridhar, B., Ng, H.N., 2012. Tradeoff between contrail reduction and emission in United States national airspace. Journal of Aircraft 49, 1367–1375.

[16] Duda, D., Minnis, P., Palikonda, R., 2005. Estimated contrail frequency and coverage over the contiguous united states from numerical weather prediction analyses and flight track data. Meteorologische Zeitschrift 14, 537–548.

[17] Fichter, C., Marquart, S., Sausen, R., Lee, D., 2005. The impact of cruise altitude on contrails and related radiative forcing. Meteorologische Zeitschrift 14, 563–572.

[18] Franco, A., Rivas, D., Valenzuela, A., 2010. Minimum-fuel cruise at constant altitude with fixed arrival time. Journal of Guidance, Control and Dynamics 33, 280–285.

[19] Fuglestvedt, J., Shine, K., Berntsen, T., Cook, J., Lee, D., Stenke, A., Skeie, R., Velders, G., Waitz, I., 2010. Transport impacts on atmosphere and climate: Metrics. Atmospheric Environment 44, 4648–4677.

[20] Gao, H., 2013. Aircraft cruise phase altitude optimization considering contrail avoidance. Master's thesis. Massachusetts Institute of Technology.

[21] (GAO), U., 2009. Aviation and climate change: aircraft emissions expected to grow, but technological and operational improvements and government policies can help control emissions. Technical Report. Government Accountability Office (GAO). Report to the Congressional Committees.

[22] Gierens, K., Schumann, U., Helten, M., Smit, H., Marenco, A., 1999. A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of mozaic measurements, in: Annales Geophysicae, Springer. pp. 1218–1226.

[23] Gierens, K., Schumann, U., Smit, H., Helten, M., Zangl, G., 1997. Determination of humidity and temperature fluctuations based on mozaic data and parametrisation of persistent contrail coverage for general circulation models, in: Annales Geophysicae, Springer. pp. 1057–1066.

[24] Greenstone, M., Kopits, E., Wolverton, A., 2011. Estimating the social cost of carbon for use in us federal rulemakings: A summary and interpretation. Technical Report. National Bureau of Economic Research.

[25] Hargraves, C.R., Paris, S.W., 1987. Direct trajectory optimization using nonlinear programming and collocation. Journal of Guidance, Control and Dynamics 10, 338–342.

[26] Herman, A.L., Conway, B.A., 1996. Direct optimization using collocation based on igh-order gauss-lobatto quadrature rules. Journal of Guidance, Control, and Dynamics 19, 592–599.

[27] Leyffer, S., 1998. User manual for MINLP-BB. University of Dundee.

[28] Mannstein, H., Schumann, U., 2005. Aircraft induced contrail cirrus over europe. Meteorologische Zeitschrift 14, 549–554.

[29] McCollum, D., Gould, G., Greene, D., 2009. Greenhouse gas emissions from aviation and marine transportation: mitigation potential and policies. Technical Report. Report prepared for the Pew Center on Global Climate Change.

[30] Minnis, P., Schumann, U., Doelling, D., Gierens, K., Fahey, D., 1999. Global distribution of contrail radiative forcing. Geophysical Research Letters 26, 1853–1856.

[31] Ng, H., Sridhar, B., Grabbe, S., Chen, N., 2011. Cross-polar aircraft trajectory optimization and the potential climate impact. Technical Report. NASA Web. http://www.aviationsystemsdivision.arc.nasa.gov/publications/2011/DASC2011_Ng.pdf. Accessed May 2 2013.

[32] Nuic, A., 2005. User Manual for the base of Aircraft Data (BADA) Revision 3.6. Eurocontrol Experimental Center.

[33] Owen, B., Lee, D., 2006. Allocation of international aviation emissions from scheduled air traffic – Future cases, 2005 to 2050 (Report 3 of 3), study on the allocation of emissions from international aviation to the UK inventory – CPEG 7. Final Report. Technical Report. DEFRA Global Atmosphere Global Atmosphere Division, Manchester Metropolitan University, UK.

[34] Pargett, D.M., Ardema, M.D., 2007. Flight path optimization at constant altitude. Journal of Guidance, Control, and Dynamics 30, 1197.

[35] Penner, J.E., Lister, D., Griggs, D.J., Dokken, D.J., McFarland, M., 1999. Aviation and the global atmosphere. Aviation and the Global Atmosphere, by Joyce E. Penner and David Lister and David J. Griggs and David J. Dokken and Mack McFarland, pp. 384. ISBN 0521663008. Cambridge, UK: Cambridge University Press, June 1999. 1.

[36] Royal Commission, 2002. The environmental effects of civil aircraft in flight. Technical Report. Royal Commission of Environmental Pollution, TR, London, England, UK.

[37] Sager, S., 2006. Numerical methods for mixed-integer optimal control problems. Ph.D. thesis. Universität Heidelberg, Germany.

[38] Sager, S., Reinelt, G., Bock, H., 2009. Direct methods with maximal lower bound for mixed-integer optimal control problems. Mathematical Programming 118, 109–149.

[39] Schimdt, E., 1941. Die Entstehung von Eisnevel aus den Auspuffgasen von Flugmotoren. Schriften der Deutschen Akademie fur Luftfahrtforschung 14, 1–15.

[40] Schumann, U., 1996. On conditions for contrail formation from aircraft exhausts. Meteorologische Zeitschrift-Berlin- 5, 4–23.

[41] Schumann, U., 2005. Formation, properties and climatic effects of contrails. Comptes Rendus Physique 6, 549–565.

[42] Soler, M., 2013. Commercial Aircraft Trajectory Planning based on Multiphase Mixed-Integer Optimal Control. Ph.D. thesis. Department of Signal Theory and Communications, Universidad Rey Juan Carlos.

[43] Soler, M., Olivares, A., Staffetti, E., 2010. Hybrid optimal control approach to commercial aircraft trajectory optimization. Journal of Guidance, Control, and Dynamics 33, 985–991.

[44] Soler, M., Olivares, A., Staffetti, E., Bonami, P., 2011. En-route optimal flight planning constrained to pass through waypoints using minlp., in: Proceedings of 9th USA/Europe Air Traffic Management Research and Development Seminar (ATM 2011), Berlin, Germany.

[45] Sridhar, B., Chen, N., Ng, H., Linke, F., Center, D., 2011a. Design of aircraft trajectories based on trade-offs between emission sources, Ninth USA/Europe Air Traffic Management Research and Development Seminar.

[46] Sridhar, B., Chen, N.Y., Ng, H.K., 2014. Aircraft trajectory design based on reducing the combined effects of carbon-dioxide, oxides of nitrogen and contrails. AIAA Modeling and Simulation Technologies Conference .

[47] Sridhar, B., Ng, H., Chen, N., 2011b. Aircraft trajectory optimization and contrails avoidance in the presence of winds. Journal of Guidance, Control, and Dynamics 34, 1577–1583.

[48] Sridhar, B., Ng, H.K., Chen, N.Y., 2012. Integration of linear dynamic emission and climate models with air traffic simulations, in: AIAA Guidance, Navigation and Control Conference, Minneapolis, MN.

[49] Sridhar, B., Ng, H.K., Chen, N.Y., Gao, H., Azzam, M., Hansman, R.J., 2013. Energy efficient trajectory designs for minimizing climate impact of aircraft on various timescales. AIAA Modeling and Simulation Technologies (MST) Conference .

[50] Stuber, N., Forster, P., et al., 2007. The impact of diurnal variations of air traffic on contrail radiative forcing. Atmospheric Chemistry and Physics 7, 3153–3162.

[51] US Department of Transportation, (US DOT), 2011. The value of travel time savings: departmental guidance for conducting economic evaluations. Revision 2. Technical Report. DOT Web. http://www.dot.gov/sites/dot.dev/files/docs/vot_guidance_092811c.pdf. Accessed Feb 23 2012.

[52] U.S. Energy Information Administration, 2012. Voluntary reporting of greenhouse gases program fuel carbon dioxide emission coefficients. http://www.eia.gov/oiaf/1605/coefficients.html. accessed dec 17 2011. Voluntary reporting of greenhouse gases program.

[53] Wächter, A., Biegler, L.T., 2006. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. Math. Programming 106, 25–57.

[54] Wei, P., Sridhar, B., Chen, N., Sun, D., 2012. A linear programming approach to development of contrail reduction strategies satisfying operationally feasible constraints, in: Guidance, Navigation, and Control Conference, Minneapolis, Minnesota.

[55] Williams, P., 2005. Hermite-Legendre-Gauss-Lobatto Direct Transcription Methods in Trajectory Optimization. Advances in the Astronautical Sciences 120.

[56] Williams, V., Noland, R., 2005. Variability of contrail formation conditions and the implications for policies to reduce the climate impacts of aviation. Transportation Research Part D: Transport and Environment 10, 269–280.

[57] Williams, V., Noland, R., Toumi, R., 2002. Reducing the climate change impacts of aviation by restricting cruise altitudes. Transportation Research Part D: Transport and Environment 7, 451–464.

[58] Williams, V., Noland, R., Toumi, R., 2003. Air transport cruise altitude restrictions to minimize contrail formation. Climate Policy 3, 207–219.

[59] Xu, X., 2001. Analysis and design of switched systems. Ph.D. thesis. University of Notre Dame. Notre Dame, Indiana.

[60] Zou, B., Buxi, G.S., Hansen, M., 2013. Optimal 4-d aircraft trajectories in a contrail-sensitive environment. Networks and Spatial Economics , 1–32.