Research Article

Characterization and Enhancement of Flight Planning Predictability under Wind Uncertainty

Javier García-Heras , Manuel Soler , and Daniel González-Arribas

Bioengineering and Aerospace Engineering, Universidad Carlos III, Leganés, Madrid, Spain

Correspondence should be addressed to Javier García-Heras; javier.garcia-heras@uc3m.es

Received 16 November 2018; Revised 22 March 2019; Accepted 11 April 2019; Published 17 June 2019

Academic Editor: Kenneth M. Sobel

Copyright © 2019 Javier García-Heras et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel study is presented aiming at characterizing and illustrating potential enhancements in flight planning predictability due to the effects of wind uncertainty. A robust optimal control methodology is employed to calculate robust flight plans. Wind uncertainty is retrieved out of Ensemble Probabilistic Forecasts. Different wind approximation functions are compared, typifying errors, and illustrating its importance for accurate solving of the robust optimal control problem. A set of key performance indicators is defined for the quantification of uncertainty in terms of flight time and fuel consumption. Two different case studies are presented and discussed. The first one is based on a representative sample of the whole 2016 year for a single origin-destination and a forecast time step of 6 hours. As for the second, we select the most uncertain day together with a multiorigin-destination set of flights with forecast time steps up to 2 days.

1. Introduction

The Air Traffic Management (ATM) System is undergoing a paradigm shift aiming at enhancing its environmental impact, capacity, safety, and efficiency. Different worldwide programmes are fostering this transformation, mainly the Single Sky ATM Research (SESAR) in Europe and Next Generation Air Transportation System (NextGen) in USA, yet also in the Asia/Pacific region, the Collaborative Actions for Renovation of Air Traffic Systems (CARATS) in Japan and the New Generation ATM System (CNAS) in China. Improvement in the ATM system predictability has been identified as paramount to achieve the above-mentioned high-level goals (see, e.g., SESAR-JU. European ATM Master Plan Edition 2015). All ATM system actors, including pilots, air traffic controllers, airlines, dispatchers, air navigation service providers, meteorological offices, and the network manager, are daily facing the effects of uncertainty. Should the capacity of the system be increased while maintaining high safety standards and improving the overall performance,

uncertainty levels in ATM must be reduced and, when possible, find new strategies to deal with and eventually reduce it. For instance, it is well-known that the sector's capacity is under-declared due to uncertain entry/occupancy counts, thus limiting the capacity of the system. Also, adding unnecessary fuel leads to a nonneglectable loss of efficiency, e.g., Hao et al. in [1] conclude that 1 min in reducing flight time dispersion could save between \$120 and \$452 million per year only considering US domestic airlines.

Uncertainty in ATM is nevertheless a heavily involved, multilayered, and interrelated phenomena. The analysis of uncertainty in ATM should take into the different scales, yet also the different sources that introduce uncertainty into the system. For deeper insight on this taxonomy, refer to [2] (Chapter 4). From a scale perspective, ATM uncertainty can be grouped into macroscale (the air transport network), mesoscale (traffic scale), and microscale (single flight). Previous work analyzing uncertainty effects at both macro- and mesoscales include, e.g., Goyens et al. [3], who studied delay propagation across the European network; Valenzuela et al.

in [4], who studied probabilistic sector loading considering wind uncertainties; and Cook et al. [5], who applied complex network theory to the ATM system. Needless to say, the focus of the present paper is on flight uncertainty (microlevel).

Regarding flight uncertainty, [2] (Chapter 4) distinguishes four time frames, namely, strategic (flow management planning level, from months up to two/three hours before the off-block time), predeparture (flight dispatching planning level, from two/three hours up to off-block time), gate-to-gate (execution both airborne and taxi times), and postarrival (from on-block time). Source-wise, ATM uncertainty falls into the following five categories: data uncertainty, operational uncertainty, equipment uncertainty, and meteorological uncertainty. Indeed, meteorological uncertainty has been recognized as one of the most (if not the most) important sources introducing uncertainty into the ATM system (see also [2]). All five uncertainty sources obviously affect flight predictability at its different time frames, which is mainly due to [6] not perfect knowledge of the initial conditions (e.g., off-block time, departure time, and initial mass), inaccurate aircraft performance model (e.g., weight, aerodynamics, Earth models, and atmospheric variables), uncertain meteorological conditions (e.g., convection and wind), navigation and equipment errors (e.g., flight management system, avionics, and flight technical errors), and ATM-related operational uncertainty (regulations, ATC advisories). Some previous work on initial condition error estimation include, e.g., Sun et al. [7], who estimate aircraft initial mass using ADS-B data. Also, Vazquez and Rivas [8] studied the propagation of initial mass uncertainty. In [6], the authors use Polynomial Chaos Expansion (PCE) to propagate uncertainties along a given flight intent considering disturbances due to initial conditions, the aircraft performance model, the Earth and atmospheric model, the aircraft intent, and wind.

We are herein interested in both the strategic and predeparture time frames, the ones compatible with flight planning. Yet, analyzing meteorological uncertainty and its effects on flight planning, to the best of the authors' knowledge, has received not enough attention so far. Indeed, wind uncertainty modelling in [6] considers uniform wind field distributions based on a deterministic forecast. More advanced features are demanded, such as the use of Ensemble Prediction System (EPS) forecasts, which provide probabilistic forecasts.

One of the first (if not the first) works on EPS-based wind uncertainty modelling and flight trajectories is due to Cheung et al. [9]. They used EPS information to study the impact of wind uncertainty in flight duration on a given, fixed route over the North Atlantic. A great impact in flight duration was observed, with no seasonal variation, and an important decrease when the time frame gets closer to the used forecast. Moreover, Vazquez et al. [10] analyzed the influence of along-track wind uncertainty (also based on EPS) in fuel mass distribution based on a probabilistic transformation method. Both studies can be casted as trajectory prediction problems since the flight path is known a priori, and thus, they are only capable of characterizing uncertainty. The ambition in this paper is not only to characterize but also to reduce uncertainty. For that purpose, flight planning algorithms that take into consideration predictability should be

derived. Research to this end is rather recent and still needs more insight.

For flight planning purposes, a rather straightforward approach is to extend some of the probabilistic trajectory prediction approaches (e.g., [9, 10]) to consider a higher algorithmic level in which the flight path is found using the discrete optimization method. This was followed by Cheung et al. in [11], where they used a Dijkstra-based trajectory predictor and also evaluated the quality of different EPSs, and by Franco et al. [12], who addressed the optimization of a North Atlantic crossing also using a Dijkstra algorithm together with their probabilistic trajectory predictor based on the probabilistic transformation method. A different approach is due to Legrand et al. [13], who presented a robust wind (obtained from EPS) optimal trajectory design algorithm based on dynamic programming and minimum-energy trajectory clustering. It should be noted that all these three approaches for flight planning under wind uncertainty are somehow based on discrete optimization. In [14], Gonzalez-Arribas et al. presented a robust optimal control approach to flight planning under wind uncertainty (obtained from EPS) in which the flight plan can be optimized considering efficiency and predictability. The present paper is based on the latter methodology, with the aim at generalizing it to different scenarios and modelling features, and thus advance in the understanding of the effects of meteorological uncertainty.

All in all, the contribution of this paper is threefold. First, we extend the methodology in [14] to consider spline wind approximations. It is shown that the errors committed in wind approximation (using, e.g., polynomial approximation such in [14]) have a tremendous impact, leading to ill-conditioned solutions. Second, instead on analyzing a single flight (as in previous work), we generalize for different days (and thus meteorological conditions) and for different origin-destinations. We also analyze the effects of wind uncertainty at different flight planning horizons, i.e., predeparture (3 hours before departure), which could be useful for flight dispatchers, and strategic (1 day before departure), which could be useful for the network manager for probabilistic capacity-demand balancing. Third, we characterize the quality of the flight plan prediction (and its uncertainty) against a hypothetical realization of winds based on reanalysis information. We note (as will be exposed) that further insights in the latter are demanded in future research and this third contribution remains as a preliminary step.

The methodology has been sketched in Figure 1. Firstly, wind data is collected from the EPS forecast. Secondly, wind forecasts are approximated by Lipschitz-continuous functions in order to model wind in a form compatible with the robust optimal planning methodology in [14]. Thirdly, the so-termed Robust Optimal Control Problem (ROCP) is solved, obtaining a Pareto frontier with different trade-offs between fuel consumption and predictability. Then, the flight path (obtained as a solution to the ROCP and defined by a sequence of waypoints) is introduced into a standard trajectory predictor (TP) ([15], Chapter 3), which *flies* the precalculated route using wind information retrieved from the original source (raw, tabular), either EPS or reanalysis. Finally, flight times, fuel burns, and its associated dispersions

FIGURE 1: Methodological diagram.

are analyzed. Different Key Performance Indicators (KPIs) are defined to that end.

Two different case studies are presented. The first focuses on a single prediction horizon (referred to as step), choosing a single origin-destination flight (over the North Atlantic), and a systematic date analysis along year 2016 (choosing the 1st, 5th, 10th, 15th, 20th, and 25th of each month as representative days). The second focuses on a single day (5th October 2016, the one that resulted to be the most unpredictable one in the former analysis), multiple origin-destination flights (based on the schedule of an important airline), and with forecasting steps ranging from 0 to 48 hours (by 6 hours). The first study aims at characterizing the seasonal influence at the predeparture time horizon, i.e., compatible with flight dispatching time frames; the second aims at the geographical influence, yet on a more

strategic vision, i.e., compatible with the network manager's time frames.

2. Robust Flight Planning

2.1. Robust Optimal Control Methodology. Our approach is based on the robust optimal control methodology for aircraft trajectory optimization problems presented in [14]. We will here summarize the method for the sake of completeness.

2.1.1. Robust Optimal Control. Let us consider a dynamical system with a randomly parametrized differential-algebraic equation and constraints. Uncertainty is described using a standard Kolmogorov probability space $(\Omega, \mathcal{F}, P)$, composed of a sample space of possible outcomes Ω , a σ -algebra of events $\mathcal{F}$ containing sets of outcomes, and the probability

function P assigning probabilities to each of these events. The uncertain parameters of the system will be modelled as a no time-dependent random variable $\xi : \Omega \rightarrow \mathbb{R}^{n_\xi}$. For each possible outcome $\omega \in \Omega$, the random variables take a different value $\xi(\omega)$.

Let us denote the state vector $\mathbf{x} \in \mathbb{R}^{n_x}$, the control vector $\mathbf{u} \in \mathbb{R}^{n_u}$, the algebraic variables $\mathbf{z} \in \mathbb{R}^{n_z}$, and $t \in \mathbb{R}$ as the independent variable (typically time). For each outcome $\omega_0 \in \Omega$, there exist a unique trajectory path $t \rightarrow (\mathbf{x}(\omega_0, t), \mathbf{z}(\omega_0, t), \mathbf{u}(\omega_0, t))$ that corresponds to the realization of the random variables $\xi(\omega_0)$. The dynamics of the system are given by the functions $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \rightarrow \mathbb{R}^{n_x}$, $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \rightarrow \mathbb{R}^{n_h}$, and $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \rightarrow \mathbb{R}^{n_g}$, such that valid trajectories fulfill the conditions almost surely (i.e., with probability 1) (the $\leq$ sign applies in an element-wise fashion in equation (3) and analogous equations.):

$$\frac{d}{dt}\mathbf{x}(\omega, t) = f(\mathbf{x}(\omega, t), \mathbf{z}(\omega, t), \mathbf{u}(\omega, t), \xi(\omega), t), \quad (1)$$

$$h(\mathbf{x}(\omega, t), \mathbf{z}(\omega, t), \mathbf{u}(\omega, t), \xi(\omega), t) = 0, \quad (2)$$

$$g_L \leq g(\mathbf{x}(\omega, t), \mathbf{z}(\omega, t), \mathbf{u}(\omega, t), \xi(\omega), t) \leq g_U, \quad (3)$$

where $\omega \in \Omega$ is the sample point on the underlying abstract probability space. Therefore, for each possible scenario or realization of the random parameters $\xi(\omega)$, the trajectory will follow the deterministic differential equation (1) for the corresponding fixed value of ξ . We employ the notation $\mathbf{x}(\omega, t)$ and $\mathbf{u}(\omega, t)$ in order to emphasize the fact that the trajectory depends on the realization of the random parameters. Note that, because the random parameters are constant in time and thus the system is not described by a full stochastic differential equation (as in, for example, [16]), there is no need to consider filtered σ -algebras to deal with potential measurability issues.

We follow the optimal *guidance* scheme of [14] that relies on the notion of *tracked* states.

Definition 1. A state is said to be *tracked* if its trajectory is assumed to be independent of the realization of the random variables almost surely, i.e., with probability one.

In this concept, the controls are specific to each scenario in order to guarantee that the tracked states follow the unique computed trajectory of those states that are tracked. Naturally, the practical consideration of this framework requires a dynamical system where the controls can be computed online to ensure that the realization of the trajectory of the tracked states stays close to its precomputed value. This is the case of the aircraft, since the flight control systems of an aircraft can follow a given route and vertical profile. For an in-depth explanation of this concept and why it is adequate for a flight planning problem, we refer the reader to [14]. We now proceed to illustrate the main concepts.

Definition 2. The amount *control degrees of freedom* of the dynamical system is defined, in this work, as the number of

controls and free algebraic variables minus the number of algebraic restrictions: $d = n_z + n_u - n_h$.

Let $q_x \leq \min \{n_x, d\}$ be the number of tracked states; without loss of generality, we can assume that the tracked states are the first q_x states (rearrange the state vector otherwise), i.e.,

$$\mathbf{x} = \begin{bmatrix} x_1 \cdots x_{q_x} x_{q_x+1} \cdots x_{n_x} \end{bmatrix}^T = \begin{bmatrix} \mathbf{x}_q \\ \mathbf{x}_r \end{bmatrix}, \quad (4)$$

where $\mathbf{x}_q$ is the tracked part of the state vector and $\mathbf{x}_r$ is the untracked part. Let $\mathbb{I}_n$ be the identity matrix of shape $n \times n$ and $0_{n_1, n_2}$ be the zero matrix (i.e., a matrix with zeroes in all its entries) of shape $n_1 \times n_2$. We define the matrix $E_x \in \mathbb{R}^{q_x \times n_x}$ as

$$E_x = \begin{bmatrix} \mathbb{I}_{q_x} & 0_{q_x, n_x - q_x} \end{bmatrix}. \quad (5)$$

This matrix transforms the state vector into the tracked states vector $\mathbf{x}_q = E_x \mathbf{x}$ that contains only the states whose evolution is equal in all scenarios. In this work, $q_x = d$ will hold so that there are enough tracked states to consume all available control degrees of freedom. This is not necessary in general. With the aid of the tracking matrices, we can now define the tracking conditions (which, again, apply almost surely):

$$\begin{aligned} E_x(\mathbf{x}(\omega_1, t) - \mathbf{x}(\omega_2, t)) &= 0, \forall t, \forall \omega_1, \omega_2 \in \Omega, \\ E_z(\mathbf{z}(\omega_1, t) - \mathbf{z}(\omega_2, t)) &= 0, \forall t, \forall \omega_1, \omega_2 \in \Omega, \\ E_u(\mathbf{u}(\omega_1, t) - \mathbf{u}(\omega_2, t)) &= 0, \forall t, \forall \omega_1, \omega_2 \in \Omega. \end{aligned} \quad (6)$$

The tracking conditions enforce equality in the tracked variables between realizations: note that, $E_x(\mathbf{x}(\omega_1, t) - \mathbf{x}(\omega_2, t))$ is the vector of differences between the tracked states in outcome ω_1 and the tracked states in outcome ω_2 . The other two conditions are analogous tracking conditions for the dependent variables and the controls.

Let us define a Bolza-type functional in optimal control problems:

$$\hat{J} = \Phi(t_0, t_f, \mathbf{x}(t_0), \mathbf{x}(t_f)) + \int_{t_0}^{t_f} \mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{u}, \xi, t) dt. \quad (7)$$

For each possible scenario or realization of the random parameters $\xi(\omega)$, we have an objective value of $\hat{J}$. In order to build a scalar objective function, we rely on both expectations and dispersions. Without generality, the latter can be characterized using the central moments, the range, or the interquartile range (IQR), just to mention a few. In our case, and in order to account for worst-case scenarios, we choose the range.

Let us then define the range R of $\hat{J}$ as

$$R(\hat{J}) = \left(\max_{\omega \in \Omega} \hat{J} - \min_{\omega \in \Omega} \hat{J} \right) \quad (8)$$

The robust optimal control problem with tracking is now formulated as follows (an alternative formulation could, for instance, rely on the n^{th} central moment of a random variable $\hat{J}$ as $\mu_n = \mathbb{E}[(J - \mathbb{E}[J])^n]$ and then substituting the range in $R(\hat{J})$ by μ_n) [14]:

$$\min J = \mathbb{E}[\hat{J}] + \text{dp} \cdot \left(\max_{\omega \in \Omega} \hat{J} - \min_{\omega \in \Omega} \hat{J} \right) \quad (9)$$

subject to: initial conditions

$$\mathbf{x}(t_0) = \mathbf{x}_0; \quad (10)$$

final conditions

$$\mathbb{E}[\Psi(t_f, \mathbf{x}(t_f))] = 0; \quad (11)$$

differential-algebraic equations and constraints (1)–(3) grouped as

$$\begin{aligned} \frac{d}{dt} \mathbf{x} &= f(\mathbf{x}, \mathbf{z}, \mathbf{u}, \xi, t), \\ h(\mathbf{x}, \mathbf{z}, \mathbf{u}, \xi, t) &= 0, \\ g_L &\leq g(\mathbf{x}, \mathbf{z}, \mathbf{u}, \xi, t) \leq g_U, \end{aligned} \quad (12)$$

and the tracking conditions (6).

Where in the above, $\mathbb{E}[\cdot]$ is the expectation operator associated with the probability space $(\Omega, \mathcal{F}, P)$; the terminal cost or the Mayer term is $\Phi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}$; the running cost or the Lagrange term is $\mathcal{L} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \rightarrow \mathbb{R}$; dp is a weighting parameter (dispersion penalty); and function $\Psi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}$ denotes the final conditions.

2.1.2. Probabilistic Discretization. In order to deal with potentially continuous random variables, we make use of discretization techniques. These methods will be referred to as stochastic quadrature rules (SQRs).

Definition 3. Given a random variable with probability measure P , a *stochastic quadrature rule* (SQR) is defined as a function or algorithm that produces a finite set of quadrature points $\{\xi_k\}$, $k \in \{1, \dots, N\}$ and weights $\{w_k\}$, $k \in \{1, \dots, N\}$ that approximates the integral $I = \mathbb{E}[G(\xi)] = \int G(\xi) dP$ with the sum:

$$QG = \sum_{k=1}^N w_k G(\xi_k), \quad (13)$$

with the approximation converging to the true value as $N \rightarrow \infty$.

In the present work, we employ a trivial quadrature rule (each ensemble member is a scenario with $w_k = 1/N$), as uncertainty is already characterized by discrete scenarios

from EPS forecasts. However, the integration of additional sources of uncertainty in the future may require the usage of a nontrivial SQR.

Let $\mathbf{x}_q(t) : \mathbb{R} \rightarrow \mathbb{R}^{q_x}$ define a trajectory for the tracked states and $\mathbf{z}_q(t)$ and $\mathbf{u}_q(t)$ analogously. Suppose a SQR has been chosen, with a number of points N (in our case, corresponding to the N possible scenarios). For each one of these points ξ_k , the tracking trajectory $(\mathbf{x}_q, \mathbf{z}_q, \mathbf{u}_q)(t)$ defines a unique trajectory given a full set of initial conditions; we will now collect each one of these N trajectories in a *trajectory ensemble*. We define the trajectory ensemble associated with a tracking trajectory $(\mathbf{x}_q, \mathbf{z}_q, \mathbf{u}_q)(t)$ as the set of trajectories $\{(\mathbf{x}_k, \mathbf{z}_k, \mathbf{u}_k)(t)\}$ with $k \in \{1, \dots, N\}$ such that the trajectory k is generated by the initial conditions $\mathbf{x}_k(t_0) = \mathbf{x}_0$ and the tracking trajectory with $\xi = \xi_k$, i.e.,

$$\begin{aligned} \frac{d}{dt} \mathbf{x}_k &= f(\mathbf{x}_k, \mathbf{z}_k, \mathbf{u}_k, \xi_k, t), \\ h(\mathbf{x}_k, \mathbf{z}_k, \mathbf{u}_k, \xi_k, t) &= 0, \\ g_L &\leq g(\mathbf{x}_k, \mathbf{z}_k, \mathbf{u}_k, \xi_k, t) \leq g_U, \\ E_x \mathbf{x}_k(t) &= \mathbf{x}_q(t), \\ E_z \mathbf{z}_k(t) &= \mathbf{z}_q(t), \\ E_u \mathbf{u}_k(t) &= \mathbf{u}_q(t). \end{aligned} \quad (14)$$

We will now build a virtual dynamical system where the state vectors of all the trajectories in the trajectory ensemble are merged into a single large state vector, which represents the collective trajectory. Its state vector $\mathbf{x}_E \in \mathbb{R}^{n_x N}$ contains the state vector of all the trajectories in the ensemble (the control $\mathbf{u}_E \in \mathbb{R}^{n_u N}$ and algebraic $\mathbf{z}_E \in \mathbb{R}^{n_z N}$ vectors are analogous), and each individual trajectory follows the dynamics associated with each scenario (with $\xi = \xi_k$).

$$\begin{aligned} \mathbf{x}_E &= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{bmatrix}; \\ \mathbf{z}_E &= \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_N \end{bmatrix}; \\ \mathbf{u}_E &= \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_N \end{bmatrix}. \end{aligned} \quad (15)$$

We define the differential equation, algebraic equations, and inequality constraints of this augmented dynamical system as

$$f_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) = \begin{bmatrix} f(\mathbf{x}_1, \mathbf{z}_1, \mathbf{u}_1, \xi_1, t) \\ \vdots \\ f(\mathbf{x}_N, \mathbf{z}_N, \mathbf{u}_N, \xi_N, t) \end{bmatrix}, \quad (16)$$

$$h_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) = \begin{bmatrix} h(\mathbf{x}_1, \mathbf{z}_1, \mathbf{u}_1, \xi_1, t) \\ \vdots \\ h(\mathbf{x}_N, \mathbf{z}_N, \mathbf{u}_N, \xi_N, t) \end{bmatrix}, \quad (17)$$

$$g_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) = \begin{bmatrix} g(\mathbf{x}_1, \mathbf{z}_1, \mathbf{u}_1, \xi_1, t) \\ \vdots \\ g(\mathbf{x}_N, \mathbf{z}_N, \mathbf{u}_N, \xi_N, t) \end{bmatrix}. \quad (18)$$

Let us define the approximate cost functional, divided into Mayer and Lagrange terms, using the trajectory ensemble:

$$\hat{J}_E = \Phi_E(\mathbf{x}_E(t_0), \mathbf{x}_E(t_f)) + \int_{t_0}^{t_f} \mathcal{L}_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) dt, \quad (19)$$

$$\Phi_E(\mathbf{x}_E(t_0), \mathbf{x}_E(t_f)) = \sum_{k=1}^N w_k \phi(\mathbf{x}_k(t_0), \mathbf{x}_k(t_f)), \quad (20)$$

$$\mathcal{L}_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) = \sum_{k=1}^N w_k \mathcal{L}(\mathbf{x}_k, \mathbf{z}_k, \mathbf{u}_k, t), \quad (21)$$

and discretize the boundary conditions as

$$\Psi_E(t_f, \mathbf{x}_E(\omega, t_f)) = \sum_{k=1}^N w_k \Psi(t_f, \mathbf{x}_k(t_f)). \quad (22)$$

For concise writing of the discretization of the tracking conditions (6), we will define the matrix $E_x^N \in \mathbb{R}^{q_x(N-1) \times n_x N}$ as

$$E_x^N = \begin{bmatrix} E_x & & \\ & \ddots & \\ & & E_x \end{bmatrix} \begin{bmatrix} \mathbb{I}_{n_x} & -\mathbb{I}_{n_x} & & \\ & \mathbb{I}_{n_x} & -\mathbb{I}_{n_x} & \\ & & \ddots & \ddots \\ & & & \mathbb{I}_{n_x} & -\mathbb{I}_{n_x} \end{bmatrix}. \quad (23)$$

$E_z^N \in \mathbb{R}^{q_z(N-1) \times n_z N}$ and $E_u^N \in \mathbb{R}^{q_u(N-1) \times n_u N}$ can be defined in analogous fashion. These matrices map the ensemble state vector to the differences in the tracked states between trajectories.

Making use of equations (14), (15), (16), (17), and (18) as well as equations (19), (20), (21), (22), and (23), we can complete now the formulation of the deterministic approximant:

$$\left. \begin{array}{ll} \text{minimize} & J_E = \mathbb{E}[\hat{J}_E] + \text{dp} \cdot R(\hat{J}_E) \\ \text{subject to} & \dot{\mathbf{x}}_E = f_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) \\ & h_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) = 0 \\ & I_G g_L \leq g_E(\mathbf{x}_E, \mathbf{z}_E, \mathbf{u}_E, t) \leq I_G g_U \\ & E_x^N \mathbf{x}_E = 0 \\ & E_z^N \mathbf{z}_E = 0 \\ & E_u^N \mathbf{u}_E = 0 \\ & \psi_E(t_0, t_f, \mathbf{x}_E(\omega, t_0), \mathbf{x}_E(\omega, t_f)) = 0 \end{array} \right\}, \quad (24)$$

where $I_G = [\mathbb{I}_{n_g} \cdots \mathbb{I}_{n_g}]^T \in \mathbb{R}^{n_g N \times n_g}$.

2.2. Application to Flight Planning

2.2.1. Modelling Assumptions. We consider a free-routing airspace and 3-DoF point-mass model based on BADA 3 [17]. We restrict ourselves to the analysis of the cruise phase for the sake of illustration. In addition, we assume the constant flight level and airspeed profiles. Note however that our methodology could be extended to full 4D problems. We consider an ellipsoidal Earth as in the WGS84 model (which offers slightly more precision than a spheroid model), with radii of curvature of ellipsoid meridian and prime vertical denoted by R_M and R_N , respectively. We take the wind and temperature fields from an EPS forecast and compute density with the ideal gas law (as the pressure is determined by the flight level). We assume the heading as a control variable instead of the bank angle, thus allowing it to change instantaneously.

All in all, aircraft dynamics can be described by the following set of differential equations expressed in the Geographic Reference System (GRS) coordinate system (f in equation (12)):

$$\begin{aligned} \dot{\mathbf{x}} &= \frac{d}{dt} \begin{bmatrix} \phi \\ \lambda \end{bmatrix} \\ &= \begin{bmatrix} (R_N(\phi) + h)^{-1} (v \cos(\chi) + w_x(\phi, \lambda, t)) \\ (R_M(\phi) + h)^{-1} \cos^{-1}(\phi) (v \sin(\chi) + w_y(\phi, \lambda, t)) \end{bmatrix}, \end{aligned} \quad (25)$$

where ϕ is the latitude, λ is the longitude, v is the true airspeed, h is the geodetic altitude, χ is the heading, and w_x and w_y are the zonal and meridional components of the wind. The control vector is composed by the heading χ .

2.2.2. System Reformulation. We reformulate the dynamical system above as a differential-algebraic system (DAE) with the addition of the ground speed v_G as an algebraic variable and the course ψ as a control variable, linked to the remaining variables by two new equality constraints. The reformulated system is given by the dynamical function:

$$\frac{d}{dt} \begin{bmatrix} \phi \\ \lambda \end{bmatrix} = \begin{bmatrix} (R_N + h)^{-1} v_G \cos \psi \\ (R_M + h)^{-1} \cos^{-1}(\phi) v_G \sin \psi \end{bmatrix}, \quad (26)$$

and the equality constraints (h in equation (12)):

$$\begin{aligned} v_G \cos \psi &= v \cos(\chi) + w_x(\phi, \lambda, t), \\ v_G \sin \psi &= v \sin(\chi) + w_y(\phi, \lambda, t). \end{aligned} \quad (27)$$

We add the inequality constraint:

$$v_G \geq 0 \quad (28)$$

to ensure uniqueness of v_G and ψ (otherwise, $(-v_G^*, \psi^* + \pi/2)$ would produce the same left-hand side of equation (27) as (v_G^*, ψ^*)).

2.2.3. Time to Distance Transformation. We apply a coordinate transformation to make the distance flown along the route (s) be the new independent variable. This allows us to apply our methodology in a manner that is consistent with existing planning and flight procedures (again, see the discussion in [14]). As a consequence, the time t becomes a state variable, and the new dynamical function can be obtained by dividing the time derivatives by $ds/dt = v_G$:

$$\frac{d}{ds} \begin{bmatrix} \phi \\ \lambda \\ t \end{bmatrix} = \begin{bmatrix} (R_N + h)^{-1} \cos \psi \\ (R_M + h)^{-1} \cos^{-1}(\phi) \sin \psi \\ v_G^{-1} \end{bmatrix}. \quad (29)$$

All constraints remain the same as in the untransformed system of differential-algebraic equations.

2.2.4. Problem Statement. An ensemble forecast contains a set of ensemble members, each one defining a different wind forecast (and, therefore, different functions w_x and w_y). Consider the ensemble contains N members, then we define N scenarios, each one having weight $w_k = 1/N$. We choose to track the course ψ , i.e., the function $\psi(s)$ is the same in every scenario (thus, we do not need to implement scenario-specific versions). As a consequence of (29), this implies that the evolution of the latitude ϕ is unique (as it only depends on the evolution of the unique variable ψ) and λ is also unique (as it only depends on ϕ and ψ). Therefore, the position variables also act like tracked variables, which is a desired goal (since we want to obtain a unique route).

We can define the dynamical system associated with the trajectory ensemble with the dynamical function:

$$\frac{d}{ds} \begin{bmatrix} \phi \\ \lambda \\ t_1 \\ \vdots \\ t_N \end{bmatrix} = \begin{bmatrix} (R_N + h)^{-1} \cos \psi \\ (R_M + h)^{-1} \cos^{-1}(\phi) \sin \psi \\ 1/v_{G,1} \\ \vdots \\ 1/v_{G,N} \end{bmatrix},$$

$$\begin{bmatrix} v_{G,1} \cos(\psi) \\ \vdots \\ v_{G,N} \cos(\psi) \\ v_{G,1} \sin(\psi) \\ \vdots \\ v_{G,N} \sin(\psi) \end{bmatrix} = \begin{bmatrix} v \cos(\chi_i) + w_{y,1}(\phi, \lambda) \\ \vdots \\ v \cos(\chi_i) + w_{y,N}(\phi, \lambda) \\ v \sin(\chi_i) + w_{x,1}(\phi, \lambda) \\ \vdots \\ v \sin(\chi_i) + w_{x,N}(\phi, \lambda) \end{bmatrix}. \quad (30)$$

The boundary conditions are

$$\begin{aligned} (\phi(0), \lambda(0)) &= (\phi_0, \lambda_0), \\ (\phi(s_f), \lambda(s_f)) &= (\phi_f, \lambda_f), \\ t_i(0) &= 0 \forall i \in \{1, \dots, N\}, \\ t_{f,\min} &\leq t_i(r_f) \leq t_{f,\max} \forall i \in \{1, \dots, N\}, \end{aligned} \quad (31)$$

where $t_{f,\min}$ and $t_{f,\max}$ are scalar decision variables.

We complete the definition of the discretized robust optimal control problem by adding the cost function. We apply the robust optimal control methodology to find routes that minimize a weighted sum of average fuel consumption and flight time dispersion (weighted with the dispersion penalty parameter dp). The cost functional to minimize is then

$$J = \mathbb{E}[t_i(r_f)] + dp \cdot \underbrace{(\max(t_i(r_f)) - \min(t_i(r_f)))}_{\text{uncertainty in } t_f}. \quad (32)$$

Note that $(\max(t_i(r_f)) - \min(t_i(r_f)))$ is the difference between the earliest and the latest arrival time.

With this cost functional, the parameter dp regulates the solution flight plan depending on the preferences of the flight planner. High values of dp will produce more predictable trajectories by avoiding regions where the wind is more unpredictable. Low values of the parameter will produce flight plans that are more efficient in average but less predictable.

3. Trajectory Predictor

The trajectory predictor determines the aircraft trajectory for a weather scenario. The objective of the TP is to provide the aircraft variable values, i.e., flight times, aircraft mass, and ground speed, along the previously defined route (sequence of nodes/waypoints) provided by the ROC problem.

Mass at discrete node/waypoint $i + 1$ is computed following equation (33) (note that this equation can be applied because the aircraft is flying at constant speed and altitude):

$$m_{i+1} = \sqrt{\frac{A_{\text{aux}}}{B_{\text{aux}}}} \tan(E + F), \quad (33)$$

where

$$\begin{aligned}
 A_{\text{aux}} &= \eta \cdot A, \\
 B_{\text{aux}} &= \eta \cdot B, \\
 \eta &= \frac{C_{f1}}{(1000 \cdot 60)} \cdot \left(1 + \frac{v \cdot kt2ms}{C_{f2}} \right), \\
 A &= 0.5 \cdot \rho \cdot S \cdot v^2 \cdot CD_{0_{CR}}, \\
 B &= \frac{2 \cdot CD_{2_{CR}} \cdot g}{(0.5 \cdot \rho \cdot S \cdot v^2)}, \\
 E &= \arctan \left(\sqrt{\frac{B_{\text{aux}}}{A_{\text{aux}}}} \cdot m_i \right), \\
 F &= \sqrt{B_{\text{aux}} \cdot A_{\text{aux}}} \cdot (t_i - t_{i+1}),
 \end{aligned} \tag{34}$$

where C_{f1} and C_{f2} are first and second thrust-specific fuel consumption coefficient, respectively, v is the true airspeed (TAS), $kt2ms$ is a constant to convert m/s to knots, S is the reference wing surface area, ρ is the air density, $CD_{0_{CR}}$ is the parasitic drag coefficient (cruise), $CD_{2_{CR}}$ is the induced drag coefficient (cruise), g is the gravity, t_i and t_{i+1} are the flight time at discrete node/waypoint i and node/waypoint $i+1$, respectively, and m_i is the aircraft mass at node/waypoint i .

Ground speed is computed as

$$v_{G_{i+1}} = \sqrt{(v^2 - w_{yPr}^2)} + w_{xPr}, \tag{35}$$

where

$$\begin{aligned}
 w_{xPr} &= \cos(\psi_{(i) \rightarrow (i+1)}) \cdot w_y + \sin(\psi_{(i) \rightarrow (i+1)}) \cdot w_x, \\
 w_{yPr} &= \cos(\psi_{(i) \rightarrow (i+1)}) \cdot w_x - \sin(\psi_{(i) \rightarrow (i+1)}) \cdot w_y.
 \end{aligned} \tag{36}$$

Note that $\psi_{(i) \rightarrow (i+1)}$ is the aircraft course between node/waypoint i and node/waypoint $i+1$ calculated as shown in equation (37) to (38).

$$\psi_{(i) \rightarrow (i+1)} = \text{sign}(\lambda_{i+1} - \lambda_i) \cdot \frac{\pi}{2}, \quad \forall \phi_i = \phi_{i+1}, \tag{37}$$

$$\begin{aligned}
 \psi_{(i) \rightarrow (i+1)} &= \arctan 2 \left(\lambda_{i+1} - \lambda_i, \ln \left[\frac{\tan(\pi/4 - \phi_i/2)}{\tan(\pi/4 - \phi_{i+1}/2)} \right] \right), \\
 &\quad \forall \phi_i \neq \phi_{i+1}.
 \end{aligned} \tag{38}$$

Flight time at node/waypoint $i+1$ is computed as the distance (d) between nodes/waypoints i and $i+1$ over v_G . It is assumed that v_G is constant between nodes/waypoints. Distance between nodes/waypoint is the orthodromic distance.

$$t_{i+1} = t_i + \frac{d_{(i) \rightarrow (i+1)}}{v_{G_{i+1}}}. \tag{39}$$

4. Wind Data and Modelling

4.1. Wind Data. In order to incorporate wind into the models, we make use of two meteorological products, namely, EPS, which provide ensemble-wise uncertainty information and reanalysis, considered to be a sufficiently good approximation to weather's realization.

4.1.1. EPS. Uncertainty of wind fields will be derived from EPS. Ensemble forecasting is a prediction technique that generates a representative sample of the possible future states of the atmosphere. An ensemble forecast is a collection of typically 10 to 50 weather forecasts (referred to as members) with a common valid time, which can be obtained using different Numerical Weather Prediction (NWP) models with varying initial conditions. Each forecast is based on a model which is close, but not identical, to the best estimate of the model equations, thus representing also the influence of model uncertainties on forecast error. Thus, the spread of solutions can be used as a measure of uncertainty. In this paper, we focus on the output data of the global ensemble forecast system ECMWF EPS. Data can be accessed (among others) at the TIGGE dataset by the European Center for Medium-Range Weather Forecasts (ECMWF) [18] (<http://apps.ecmwf.int/datasets/http://apps.ecmwf.int/datasets/>).

(1) *ECMWF EPS.* The European Centre for Medium-Range Weather Forecasts EPS (ECMWF EPS) is based on 51 ensemble members with approximately 32 km resolution up to forecast day 10 and 65 km resolution thereafter, with 62 vertical levels.

4.1.2. Reanalysis. Meteorological reanalysis is based on meteorological data assimilation. The aim is at assimilating historical observational data spanning an extended period, using a given model. Reanalysis is considered to be the best estimate on many variables (such as winds and temperature).

ECMWF ERA-5 reanalysis will be used in the course of the validation activities. The data assimilation system used to produce ERA-5 includes a 4-dimensional variational analysis with a 12-hour analysis window. The spatial resolution of the dataset is approximately 31 km (in the high resolution version) on 32 vertical levels from the surface up to 0.01 hPa. ERA-5 products are normally updated once daily, with a delay of two months to allow for quality assurance and for correcting technical problems with the production, if any. ERA-5 data can be downloaded from the ECMWF Public Datasets.

4.2. Wind Approximation. Note that in order to incorporate winds in Problem (24), functions $w_x(\phi, \lambda, t): \mathbb{R}^3 \rightarrow \mathbb{R}$ and $w_y(\phi, \lambda, t): \mathbb{R}^3 \rightarrow \mathbb{R}$ in equation (27) must ideally be class C^2 functions, i.e., continuous and twice derivable functions (in lack of this, they should be at least Lipschitz continuous functions). This in practice translates into the need of approximating raw, tabular data from either EPS or reanalysis into smooth functions with the abovementioned properties.

FIGURE 2: Polynomial approximation accuracy in L^2 (RMS) norm (a) and L^∞ (maximum) norm (b) [14].

We describe two possible approximations: a multiregion polynomial regression and a spline interpolation method. Approximation errors cannot be obviated; they end up distorting results as we will show in the sequel.

4.2.1. Multiregion Polynomial Regression. We seek to compute an approximation of the wind field in the area of interest. We subdivide this area into regions and consider an approximating polynomial for each region:

$$p_{a,s}(\phi, \lambda) = \sum_{\substack{i+j \leq n \\ i,j \geq 0}} a_{s,ij} \phi^i \lambda^j, \quad (40)$$

where s denotes the index of the region.

We compute the coefficients $a_{s,ij}$ by solving the following least squares problem:

$$\min_{a_{s,ij} \in \mathbb{R}} \sum_{s=1}^{N_{\text{regions}}} \sum_{k=1}^N \sum_{l=1}^M (p_{a,s}(\phi_{kl,s}, \lambda_{kl,s}) - u_{kl,s})^2, \quad (41)$$

where the points $(\phi_{kl,s}, \lambda_{kl,s})$ are generated by applying to a rectangular, equispaced grid on $[0, 1] \times [0, 1]$ the affine transformation $T_s : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $T_s([0, 1] \times [0, 1]) = R_s$, where $R_s \subset \mathbb{R}^2$ denotes the region number s . The values $u_{kl,s}$ are generated by sampling a cubic spline interpolant of the original grid values at this new grid.

We also add continuity and differentiability requirements at the boundaries between regions. Consider the boundary between region s and region $s+1$ and parametrize it by $\xi_s \in [0, 1]$ by building the affine transformations $(\phi_s(\xi_s), \lambda_s(\xi_s))$. By composition, we can build the polynomials:

$$\begin{aligned} b_{a,s}(\xi_s) &= p_{a,s}(\phi_s(\xi_s), \lambda_s(\xi_s)), \\ b_{a,s+1}(\xi_s) &= p_{a,s+1}(\phi_s(\xi_s), \lambda_s(\xi_s)), \end{aligned} \quad (42)$$

that represent the value of the field approximation at the boundaries at each one of the adjacent regions. By construction, both b_s and b_{s+1} are polynomials of degree n in ξ_s ; therefore, by equating them at $n+1$ points, we can guarantee that they are equal. We choose $n+1$ equispaced points along the boundary, $\{\xi_{s,m}\} (m \in \{0, \dots, n\})$ and augment problem (41) with the equality constraints $b_{a,s}(\xi_{s,m}) = b_{a,s+1}(\xi_{s,m})$. Note that these are constraints on the value of the coefficients a , which are the decision variables for this problem. This constraints guarantee continuity between regions.

In order to deal with differentiability constraints, we similarly define the polynomials:

$$\begin{aligned} b_{a,s}^{q,r}(\xi_s) &= \frac{\partial^q}{\partial \phi^q} \frac{\partial^r}{\partial \lambda^r} p_{a,s}(\phi_s(\xi_s), \lambda_s(\xi_s)), \\ b_{a,s+1}^{q,r}(\xi_s) &= \frac{\partial^q}{\partial \phi^q} \frac{\partial^r}{\partial \lambda^r} p_{a,s+1}(\phi_s(\xi_s), \lambda_s(\xi_s)), \end{aligned} \quad (43)$$

where $q, r \geq 0$ are the orders of the differentiability constraints. The degree of this polynomial on ξ is $n - q - r$, so we only need to equate $b_{a,s}^{q,r}$ and $b_{a,s+1}^{q,r}$ in $n + 1 - q - r$ points in order to enforce differentiability across boundaries.

Figure 2 presents the approximation error of the global and regional regression approaches, measured in a band of ± 3 deg around the reference trajectory as computed in [14]. A significant improvement can be observed in both number of regions and degree of the polynomial, thus justifying the advantages of this approach. Nonetheless, a residual error always persists and that is why we resort to spline interpolation.

4.2.2. Spline Interpolation. An alternative method to do wind field spline interpolation is a well-established field of knowledge. Amodei and Benbourhim in [19] introduced a new family of vector field splines with application to wind fields. More recently, Le Guyader et al. [20] studied the use of a spline-bases approximation considering conservative vector

fields as it happens with winds (winds derive from temperature potentials). We herein use the b-spline interpolation functionality implemented in CasADi [21], a symbolic framework for algorithmic differentiation and numeric optimization. The spline S takes values in the region of interest $[\lambda, \phi]$ and maps them to $\mathbb{R}$.

4.3. Wind Errors. To illustrate errors with both wind approximation models and its effects in the results to Problem (24), we bring herein an example. It corresponds to case 5th October 2016 in the single O-D flight case study, so please refer therein for more details.

Needless to say, it can be readily observed comparing Figures 3(a) and 3(b) that the wind uncertainty (computed as the standard deviation of wind norm across all members in the EPS, see equation (44)) in both multiregion polynomial regression and spline interpolation approaches substantially differ. One could qualitatively say that there is a *displacement of the wind uncertainty*. This is mainly due to the fact that wind approximation errors (see definition below) committed using the multiregion polynomial regression itself are of the same order of magnitude as the uncertainty. This fact can be observed in Figure 3(d); indeed, the higher the uncertainty, the higher the error. This is of course expected, since regression is a smoothing technique. On the contrary (see Figure 3(c)), spline interpolation errors are neglectable with orders of magnitude of $1 \cdot e - 12$.

$$\sigma_w = \sqrt{\sigma_{w_x}^2 + \sigma_{w_y}^2}, \quad (44)$$

where σ_w is the wind uncertainty, and σ_{w_x} and σ_{w_y} represent the standard deviation of wind east-west and north-south components across all members in the EPS, respectively.

$$\begin{aligned} \sigma_{w_x} &= \sqrt{\frac{\sum_{n=1}^N (w_{x_n} - \bar{w}_x)^2}{N}}, \\ \sigma_{w_y} &= \sqrt{\frac{\sum_{n=1}^N (w_{y_n} - \bar{w}_y)^2}{N}}, \end{aligned} \quad (45)$$

where w_{x_n} and w_{y_n} are the wind east-west and north-south components of member n in the EPS and $\bar{w}_x$ and $\bar{w}_y$ represent the average with east-west and north-south components across all members in the EPS, respectively, being N is the number of members in the EPS, i.e.,

$$\begin{aligned} \bar{w}_x &= \frac{\sum_{n=1}^N w_{x_n}}{N}, \\ \bar{w}_y &= \frac{\sum_{n=1}^N w_{y_n}}{N}. \end{aligned} \quad (46)$$

Let us define wind uncertainty error as $\varepsilon_{\sigma_w} = \sigma_{w_f} - \sigma_{w_a}$, where σ_{w_f} is the standard deviation of wind across all members in the EPS forecast (considering raw, tabular wind data)

and σ_{w_a} is the standard deviation of wind across all members in the EPS (considering the polynomial/n-spline approximation functions.)

Figure 4 provides evidence on errors' effects in the calculated trajectories (solving Problem (24) with both wind approximations). While darker trajectories (representing the problem computed with low dp values, thus weighting less the effects of uncertainty) are very similar, brighter ones behave in rather dissimilar manner. Indeed, in Figure 4(b), trajectories computed with the multiregion polynomial regression approximation and high dp values (bright colors) seek to deviate southwards to avoid a region of high uncertainty (according to the approximated wind). This region is indeed biased by approximation errors. If one looks at Figure 4(a) (approximated with splines and thus with no error), the high-uncertainty region has moved slightly south and thus trajectories do not go south anymore, rather go straight.

We bring now readers' attention to Figure 5. Flight times and flight dispersions are represented for different dp values in boxplots, which provide qualitative information on quartiles (upper and lower box) and median (red line). The whiskers indicate variability outside the upper and lower quartiles. Outliers are plotted (should they exist) as individual points. All in all, one can readily have an idea of median (or mean) flight times and dispersions. The solution to Problem (24) should provide increasing average flight times and reduced time dispersions as dp values grow. This is as expected in Figure 5(a), where results are represented with the direct output of the optimization problem (in which wind reads as in the approximated functions). However, if one takes the obtained discrete path (λ, θ, h) and computes flight times using a trajectory predictor (TP) over the original raw, tabular wind, results do not show this behaviour anymore. This can be noticed in Figure 5(b), e.g., comparing $dp = 0$ and $dp = 1$. Also, looking now at Figure 5(c), which includes the same simulation with the spline interpolation (note that herein wind errors are zero), flight times are substantially different; yet, dispersions are consistent.

Please refer to Table 1 for quantitative data, including also data for fuel consumptions and fuel dispersions with spline wind approximation. Remark also that this analysis has been made for the whole scenario in Case 1 and this reasoning holds for the whole set of flights. These findings motivate the use of precise wind approximation methodology, though reader should advert there is always a trade-off between accurate modelling and solvability. All in all, we resorted to spline interpolation to circumvent these problems.

5. Key Performance Indicators (KPIs)

The following variables will be analyzed:

- (i) Total flight time (FT)
- (ii) Total fuel consumption (FC)

For the sake of simplicity, henceforth, variable x will be sued (denoting either FT or FC interchangeably). Note the reader that the solution to the robust flight planning provides

FIGURE 3: Wind approximations and associated errors (in m/s).

a discrete number N (equal to the number of members in the ensemble forecast) of possible values for both flight time and fuel consumption. Thus, in order to compare the robust flight

planning results with hypothetical realizations (considered to be that of the reanalysis), we should statistically characterize flight times and fuel consumptions. Figure 6 sketches the

FIGURE 4: Optimal paths (for different dp values).

FIGURE 5: Boxplot diagrams showing flight dispersion results for different dp values.

TABLE 1: Top five days in terms of time and fuel dispersion. Tol = 30 sec. (FT) and 70 kg (FC).

Variable	Date	FT (sec.)					FC (kg)				
		dp0	dp5	dp10	dp15	dp20	dp0	dp5	dp10	dp15	dp20
$\mathbb{R}(x)^{R-Trj}$	2016-Oct-05	337.0	227.6	215.4	211.1	207.1	487.2	328.8	311.0	304.7	298.8
$\bar{x}^{R-Trj}$		20073.5	20203.0	20296.9	20365.0	20454.4	30525.4	30712.5	30848.2	30946.4	31075.5
x^{Rea}		20205.7	20295.1	20383.3	20450.0	20538.4	30716.5	30845.6	30972.9	31069.2	31196.7
Δx^{Mean}		-132.2	-92.1	-86.4	-85.1	-84.0	-191.1	-133.0	-124.7	-122.8	-121.3
$\mathbb{B}(x)^{IQR}$		False	False	False	False	False	False	False	False	False	False
$\mathbb{B}(x)^{Exp\pm Tol}$		False	False	False	False	False	False	False	False	False	False
$\%x^{Rea\pm Tol}$		16.0	20.0	24.0	24.0	24.0	20.0	28.0	30.0	34.0	34.0
$\mathbb{R}(x)^{R-Trj}$	2016-Sep-05	314.2	247.0	225.7	203.4	182.6	451.8	354.9	324.2	291.9	261.6
$\bar{x}^{R-Trj}$		21185.9	21327.5	21477.4	21605.9	22001.7	32129.1	32332.7	32548.1	32732.5	33300.0
x^{Rea}		21217.1	21358.7	21509.5	21650.4	22025.4	32174.1	32377.6	32594.1	32796.4	33333.9
Δx^{Mean}		-31.2	-31.2	-32.1	-44.5	-23.6	-45.0	-44.9	-46.0	-63.9	-33.9
$\mathbb{B}(x)^{IQR}$		True	True	True	True	True	True	True	True	True	True
$\mathbb{B}(x)^{Exp\pm Tol}$		False	False	False	False	True	True	True	True	True	True
$\%x^{Rea\pm Tol}$		36.0	32.0	30.0	30.0	26.0	56.0	48.0	48.0	46.0	50.0
$\mathbb{R}(x)^{R-Trj}$	2016-May-05	300.7	176.1	129.5	107.4	100.0	435.7	254.9	187.3	155.1	144.3
$\bar{x}^{R-Trj}$		19584.8	19751.8	20046.4	20256.5	20447.0	29818.2	30060.0	30486.2	30789.8	31064.8
x^{Rea}		19578.0	19750.7	20054.3	20249.2	20443.3	29808.4	30058.4	30497.6	30779.2	31059.5
Δx^{Mean}		6.8	1.1	-7.9	7.4	3.7	9.8	1.5	-11.4	10.6	5.3
$\mathbb{B}(x)^{IQR}$		True	True	True	True	True	True	True	True	True	True
$\mathbb{B}(x)^{Exp\pm Tol}$		True	True	True	True	True	True	True	True	True	True
$\%x^{Rea\pm Tol}$		62.0	66.0	74.0	66.0	62.0	84.0	82.0	86.0	94.0	94.0
$\mathbb{R}(x)^{R-Trj}$	2016-Jan-25	261.7	180.2	173.2	168.1	149.0	380.2	261.7	251.5	244.1	215.8
$\bar{x}^{R-Trj}$		18981.2	19025.9	19083.4	19153.7	19638.3	28942.4	29007.3	29090.8	29192.9	29895.7
x^{Rea}		19011.9	19070.8	19128.9	19199.9	19700.5	28987.0	29072.6	29157.0	29260.0	29985.7
Δx^{Mean}		-30.7	-44.9	-45.5	-46.2	-62.2	-44.6	-65.3	-66.1	-67.1	-90.0
$\mathbb{B}(x)^{IQR}$		False	False	False	False	False	False	False	False	False	False
$\mathbb{B}(x)^{Exp\pm Tol}$		False	False	False	False	False	True	True	True	True	False
$\%x^{Rea\pm Tol}$		34.0	32.0	26.0	24.0	22.0	62.0	52.0	50.0	50.0	38.0
$\mathbb{R}(x)^{R-Trj}$	2016-May-20	234.8	189.5	166.6	151.9	121.4	341.3	275.4	241.9	220.5	175.7
$\bar{x}^{R-Trj}$		18873.7	18976.7	19140.0	19260.1	19778.1	28786.1	28935.8	29173.1	29347.3	30098.0
x^{Rea}		18895.8	18996.3	19145.6	19253.6	19768.9	28818.2	28964.3	29181.1	29337.9	30084.9
Δx^{Mean}		-22.1	-19.6	-5.5	6.5	9.1	-32.2	-28.5	-8.0	9.4	13.2
$\mathbb{B}(x)^{IQR}$		True	True	True	True	True	True	True	True	True	True
$\mathbb{B}(x)^{Exp\pm Tol}$		True	True	True	True	True	True	True	True	True	True
$\%x^{Rea\pm Tol}$		42.0	46.0	64.0	64.0	70.0	62.0	68.0	82.0	86.0	90.0

solutions provided by the robust trajectory predictor (including its statistical characterization in quartiles and mean values) and the realization (based on reanalysis).

Let us consider the following sets:

- (i) $\mathcal{F} \rightarrow$ (set of flights)

- (ii) $\mathcal{DP} \rightarrow$ (set of dp values)
- (iii) $\mathcal{S} \rightarrow$ (set of forecast steps in the EPS)
- (iv) $\mathcal{N} \rightarrow$ (set of ensemble members in the EPS)

Let us consider the following variables:

$$\begin{aligned} x_{i,j,k,n}^{R-Trj}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}, \forall n \in \mathcal{N}; \\ x_{i,j,k}^{Rea}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \end{aligned} \quad (47)$$

where super-index R-Trj refers to variables obtained from the robust trajectory predictor and computed using EPS forecasts, and super-index Rea refers to the trajectory flown under a hypothetical realization of wind (considered to be the reanalysis).

Taking mean values across all members in the EPS:

$$\bar{x}_{i,j,k}^{R-Trj} = \sum_{n=1}^N \frac{x_{i,j,k,n}^{R-Trj}}{N}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}. \quad (48)$$

Consider also the following conditions:

$$x_{i,j,k}^{Rea} \in \text{IQR}_{i,j,k}^{R-Trj}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \quad (49)$$

$$x_{i,j,k}^{Rea} \in [\bar{x}_{i,j,k}^{R-Trj} \pm \text{tol}], \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \quad (50)$$

$$x_{i,j,k,n}^{R-Trj} \in [x_{i,j,k}^{Rea} \pm \text{tol}], \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}, \forall n \in \mathcal{N}; \quad (51)$$

where tol is a user-specified tolerance.

Let us define the indicator functions that takes value one if the condition is true and zero elsewhere:

$$\begin{aligned} \mathbb{1}_{\text{cond.}(49),(i,j,k)} \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \mathbb{1}_{\text{cond.}(50),(i,j,k)} \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \mathbb{1}_{\text{cond.}(51),(i,j,k,n)} \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}, \forall n \in \mathcal{N}. \end{aligned} \quad (52)$$

The flight-wise Key Validation Indicators are defined as follows:

$$\begin{aligned} \mathbb{R}(x)_{i,j,k}^{R-Trj} &= \max_{n \in \mathcal{N}} x_{i,j,k,n}^{R-Trj} - \min_{n \in \mathcal{N}} x_{i,j,k,n}^{R-Trj}, \\ &\quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \Delta x_{i,j,k}^{\text{Mean}} &= \bar{x}_{i,j,k}^{R-Trj} - x_{i,j,k}^{Rea}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \mathbb{B}(x)_{i,j,k}^{\text{IQR}} &= \mathbb{1}_{\text{cond.}(49),(i,j,k)}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \mathbb{B}(x)_{i,j,k}^{\text{Exp}\pm\text{Tol}} &= \mathbb{1}_{\text{cond.}(50),(i,j,k)}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \%x_{i,j,k}^{\text{Rea}\pm\text{Tol}} &= \sum_{n \in \mathcal{N}} \frac{\mathbb{1}_{\text{cond.}(51),(i,j,k,n)}}{N}, \quad \forall i \in \mathcal{F}, \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \end{aligned} \quad (53)$$

FIGURE 6: Schematic boxplot including the N robust prediction and the realization of a hypothetical variable.

where $\mathbb{R}(x)_{i,j,k}^{R-Trj}$ represents the predicted flight time or fuel burnt dispersions using the robust flight planing methodology; $\Delta x_{i,j,k}^{\text{Mean}}$ represents the deviations of the expected value (the mean) with respect to the observation (assumed herein to be the calculated using reanalysis); $\mathbb{B}(\cdot)$ denotes boolean, and thus, the associated indicator can take value true or false, depending on whether the realization fits into the defined domains (IQR; expected prediction $\pm$ tolerances) for both flight time and fuel consumption; last but not the least, $\%(\cdot)$ denotes the percentage of predictions that fin within the realization $\pm$ a given tolerance. Notice that these indicators are calculated for all flights, all dp values and all forecast steps.

Let us also define the following aggregated KPIs:

$$\begin{aligned} \bar{\mathbb{R}}(x)_{j,k} &= \sum_{i \in \mathcal{F}} \frac{\mathbb{R}(x)_{i,j,k}^{R-Trj}}{N_{\text{flights}}}, \quad \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \bar{\Delta x}_{j,k} &= \sum_{i \in \mathcal{F}} \frac{\Delta x_{i,j,k}^{\text{Mean}}}{N_{\text{flights}}}, \quad \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \% \mathbb{B}(x)_{j,k}^{\text{IQR}} &= \sum_{i \in \mathcal{F}} \frac{\mathbb{B}(x)_{i,j,k}^{\text{IQR}}}{N_{\text{flights}}}, \quad \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \% \mathbb{B}(x)_{j,k}^{\text{Exp}\pm\text{Tol}} &= \sum_{i \in \mathcal{F}} \frac{\mathbb{B}(x)_{i,j,k}^{\text{Exp}\pm\text{Tol}}}{N_{\text{flights}}}, \quad \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \\ \% \bar{x}_{j,k}^{\text{Rea}\pm\text{Tol}} &= \sum_{i \in \mathcal{F}} \frac{\% x_{i,j,k}^{\text{Rea}\pm\text{Tol}}}{N_{\text{flights}}}, \quad \forall j \in \mathcal{DP}, \forall k \in \mathcal{S}; \end{aligned} \quad (54)$$

where they provide an aggregated vision over all flights in the different scenarios of the same variables as above, namely, flight time and fuel burnt dispersions, and deviations of the predicted mean flight time and fuel burnt with respect to the observation.

6. Case Studies

Two case studies have been selected in order to characterize and illustrate potential reductions in both flight time and fuel burnt uncertainty, namely,

- (1) A single origin-destination flight (over the North Atlantic) along a set of characteristic days along year

FIGURE 7: 2016-Oct-05 results: boxplots denote predicted uncertainties (based on EPS data and computed using the robust flight planning algorithm); red stars denote realized values (based on reanalysis data).

FIGURE 8: 2016-Oct-05: trade-off solutions of the problem.

2016. The flights are computed considering a single step of 6 hours, i.e., a characteristic times at flight dispatching level. Please refer to Section 6.1 for additional details

- (2) A multiple origin-destination set of flights in a representative day, i.e., day with more time dispersion in the single origin-destination case study. The flights are computed considering multiple steps, from 0 to 48 hours, i.e., characteristic times at flow management level. Please refer to Section 6.2 for additional details

All robust flight planning have been simulated using the robust flight planning approach in Section 2.1, considering EPS forecasts as input data and approximated using the spline interpolation in Section 4. The realization of the trajectory is computed using the route provided by the robust flight planning (a sequence of (lat, long) coordinates) and considering reanalysis wind data approximated by spline interpolation.

A trajectory predictor is used for that purpose. KPIs introduced in Section 5 are used to quantify results.

6.1. Single O-D Flight, Single Step, Multiple Days. We choose a flight between KJFK and LPPT, flying at constant barometric FL380, and constant Mach of 0.82. A free routing airspace is considered. Aircraft is modelled as a A330 BADA3 model with an initial mass of 200000 kg. We select a set of characteristic days along year 2016, all together 72 days (1st, 5th, 10th, 15th, 20th, and 25th of each month). The corresponding EPS forecasts produced at 00.00 of the day and with forecast of 06 step are considered. We present in Table 1 the top five days in terms of time and fuel dispersions, respectively. We proceed on discussing the results for the day with the greatest dispersion (2016-Oct-05, herein termed representative day).

6.1.1. Representative Day “2016-Oct-05.” Figure 7 includes boxplot information on flight times, fuel consumptions, time dispersions, and fuel burnt dispersions for this representative

FIGURE 9: Aggregated dispersion boxplots for the single origin-destination case study.

day and different dp values ranging 0 to 50. Furthermore, the realization computed using reanalysis is represented (as a star). It can be seen that flight time dispersions and fuel burnt consumptions systematically reduce as dp increases. Similarly, expected flight times and fuel burns increase. It is also interesting to remark that ΔFT^{Mean} and ΔFC^{Mean} decrease as dp increases; in other words, the predicted values get closer to the realization. Indeed, it is remarkable that the greater the dp value is, the more likely the realization to fit within the predicted bounds.

Quantitative information is presented in Table 1 (row 1). For the sake of illustration, it can be observed that maximum dispersions $\mathbb{R}(FT)^{\text{R-Trj}}$ can reach up to roughly 337 sec. This

dispersion could be reduced to 207 sec. (130 sec. or 38% reduction) by setting $dp = 20$. This would come at a cost of roughly 381 sec. of extra flight time in average $\bar{FT}^{\text{R-Trj}}$. Similarly, in terms of consumption, dispersion can be reduced from 487 kg to 298 kg (189 kg or 38%), all in all with a cost of roughly 550 kg of extra fuel burnt. The Pareto frontier of the problem weighting dispersion and time/consumption is presented in Figure 8. The reader is referred therein and to Table 1 for a quantitative insight on other intermediate trade-offs and additional days.

Moreover, the deviations of the expected time (as predicted) with respect to the realization ΔFT^{Mean} is decreasing as dp increases, from 132 sec. to 84 sec. (negative sign means

the trajectory flown under reanalysis has greater flight time than under the EPS forecast), i.e., roughly 48.2 sec. or, in other words, predicted trajectories with $dp = 20$ are 47.6% more predictable. Similarly with fuel consumption ΔFC^{Mean} , we reduce the deviations with respect to the expected consumption from 191 kg to 121 kg, i.e., 70 kg of reduction.

Last but not the least, it holds (as expected) that the higher dp , the greater the percentage of prediction fitting within the realization $\pm$ tolerances $\%x^{\text{Rea}\pm\text{Tol}}$. On the contrary, it should be noted that the realization is not fitting within the predicted interquartile range $\mathbb{B}(\text{FT})^{\text{IQR}}$ (the same reasoning holds *mutatis mutandis* for FC), neither within the expected flight time ± 30 sec. range $\mathbb{B}(\text{FT})^{\text{Exp}\pm\text{Tol}}$. Notice however that the realization hits closer to the IQR as dp grows (see Figure 7). This is partially due to the high uncertainty. In other words, we are able to substantially improve predictability; however, the predictability is still rather poor. The reader is referred to Section 6.3 for a discussion on accuracy of realizations (as considered herein).

6.1.2. 2016 Aggregated Results. Figure 9 shows aggregated values for dispersion and for all 72 days considered. Table 2 presents all aggregated KPIs. On average, flight time dispersion can be reduced roughly 1 min (from 151 sec. to 92 sec.) by computing flight plans with dp equal to 20. Similarly, fuel consumption dispersion could be reduced roughly 85 kg (from 219 to 133 kg). This represents 39% dispersion reduction in both times and fuel consumption. This nevertheless would come at an average cost of extra 300 sec. and 450 kg of fuel.

As for the indicators that take into consideration the realizations, we can observe that deviations from the expected times and fuels decrease slightly (from 29.5 to 27.3 sec. and 42.7 to 39.4 kg, respectively). It is however remarkable to look at the percentage of times the realization is within the IQR (around 40%), yet also at the percentage of time flights are within the expected value $\pm$ a given tolerance (30 sec. and 70 kg in this case), over 60% and almost 90% for times and consumptions, respectively. Again, readers are referred to Section 6.3 for a discussion on accuracy of realizations (as considered herein).

6.2. Multiple O-D Flights, Multiple Steps, Single Day. We keep the focus on the same day “2016-Oct-05.” We choose now a set of 164 flights flying to different destinations (according to British Airways’ schedule for that day). Figure 10 provides a snapshot of the flight set. We assume all of them fly at constant barometric altitude FL380 and constant Mach M0.82. A free routing airspace is also assumed. Different aircraft types are considered, all modeled as BADA3, and considering their corresponding reference masses as initial mass. As for wind, we take EPS forecasts produced on 2016-Oct-05 at 00.00 with forecast steps ranging from 00 to 48 hours.

For the sake of representativeness, we have clustered them into 64 short-, 46 medium-, and 54 long-haul flights. We present the top three long-haul flights in Table 3. Top three medium- and short-haul flights can be seen in Table 4 and Table 5, respectively. All in all, results show a consistent behaviour; dispersions are systematically reduced

TABLE 2: Aggregated KPIs. Tol = 30 sec. (FT) and 70 kg (FC).

Variable	dp0	dp5	dp10	dp15	dp20
$\overline{\mathbb{R}}(\text{FT})$	151.6	113.3	102.6	96.7	92.6
$\overline{\mathbb{R}}(\text{FC})$	219.4	163.9	148.4	139.7	133.7
$\overline{\Delta\text{FT}}^{\text{Mean}}$	29.5	25.8	26.3	27.4	27.3
$\overline{\Delta\text{FC}}^{\text{Mean}}$	42.7	37.3	38.0	39.6	39.4
$\overline{\text{FT}}^{\text{R-Trj}}$	19918.2	19980.9	20061.5	20136.1	20218.2
$\overline{\text{FC}}^{\text{R-Trj}}$	30298.6	30389.2	30505.7	30613.5	30732.0
$\%\mathbb{B}(\text{FT})^{\text{IQR}}$	38.9	45.8	44.4	38.9	40.3
$\%\mathbb{B}(\text{FC})^{\text{IQR}}$	38.9	45.8	44.4	38.9	40.3
$\%\mathbb{B}(\text{FT})^{\text{Exp}\pm\text{Tol}}$	58.3	65.3	66.7	63.9	62.5
$\%\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	87.5	93.1	87.5	88.9	87.5
$\%\text{FT}^{\text{Rea}\pm\text{Tol}}$	47.6	52.2	52.8	51.9	52.6
$\%\text{FC}^{\text{Rea}\pm\text{Tol}}$	69.1	73.8	74.7	74.2	74.9

as we increase dp values; yet, they increase as we increase the forecast step. Needless to say, this corresponds to the expected behaviour. We discuss in what follows flight EGLL to SBGL (first row in Table 3).

6.2.1. Representative Flight EGLL to SBGL. First, it is interesting to look at how dispersions grow as prediction horizon (or step) does. Indeed, at one day look-ahead time, they can be more than 10 minutes; yet, at two days look-ahead time, they can go up 20 minutes. Second, it is also very important to observe how we are able to substantially reduce flight time uncertainty as we increase the dp parameter. For instance, at one day prediction step, we are able to cut it by almost two (459 to 249 sec., i.e., 46% reduction) by moving from $dp = 0$ to $dp = 10$. All in all, this could be of course very relevant for the ATFM units at pretactical level, where they do the capacity-demand balancing the day before operation. Nonetheless, this increase of predictability would come at the cost of increasing flight time and fuel consumption (roughly 800 sec. or 1000 kg in this case). Other trade-off values (also for the other top 2 flights) can be checked both in Table 3 and Figure 11, where the Pareto frontiers of the problem are shown for this specific flight. Moreover, Figure 12 shows boxplot information, including the obtained realization. It can be seen that dispersions reduce as dp increases. On the contrary, expected flight times and fuel burns increase. It is also interesting to remark that in general, $\Delta\text{FT}^{\text{Mean}}$ and $\Delta\text{FC}^{\text{Mean}}$ decrease as dp increases; in other words, the predicted values get closer to the realization. Nonetheless, no conclusive results can be extracted in this sense as forecast step increases. Similar analysis holds for other flights. Of course, medium- and short-haul flights present lower dispersion values. Detailed information on them can be checked in Tables 3, 6–9.

6.2.2. Aggregated Results. An aggregated analysis (for all 164 flights considered) is presented in both Table 5 and

TABLE 3: Top three long-haul flights in terms of flight time dispersion. Tol = 30 sec.

Variable	Airp. Orig-Dest	Step 0			Step 12			Step 24			Step 36			Step 48		
		dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGLL-SBGL	253.3	142.6	122.0	358.3	237.2	219.4	459.2	249.8	213.9	495.5	388.8	340.8	1126.4	492.8	465.2
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGLL-SBGL	38013.3	38370.5	38766.7	37873.9	38205.9	38517.7	37997.2	38689.3	39371.4	38419.6	38958.8	39787.6	39329.7	39810.0	40357.3
FI^{Rea}	EGLL-SBGL	38093.9	38470.9	38854.3	37809.4	38170.0	38469.1	37780.1	38474.2	39158.0	38289.4	38804.2	39644.4	39122.9	39776.1	40306.1
$\Delta \text{FT}^{\text{Mean}}$	EGLL-SBGL	-80.6	-100.4	-87.6	64.5	35.9	48.6	217.1	215.1	213.3	130.2	154.6	143.2	206.9	33.9	51.3
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGLL-SBGL	False	False	False	True	True	True	False	False	False	False	False	False	False	True	True
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGLL-SBGL	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGLL-SBGL	12.0	2.0	4.0	26.0	24.0	26.0	6.0	0.0	0.0	8.0	4.0	10.0	6.0	8.0	8.0
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	TAPA-EGKK	211.6	84.3	78.9	451.0	219.2	212.3	537.0	332.7	328.3	765.1	429.4	389.6	1112.7	567.9	459.9
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	TAPA-EGKK	25561.4	25752.3	25934.8	25786.3	26191.7	26480.6	26385.1	27416.1	27954.4	26873.5	27678.2	28396.5	26496.8	27666.7	29487.5
FI^{Rea}	TAPA-EGKK	25450.9	25647.3	25834.2	25792.5	26189.1	26469.5	26411.1	27464.6	27973.4	26961.0	27707.6	28401.7	26612.5	27552.8	29427.6
$\Delta \text{FT}^{\text{Mean}}$	TAPA-EGKK	110.5	105.0	100.7	-6.2	2.6	11.0	-26.0	-48.5	-19.0	-87.4	-29.4	-5.1	-115.7	113.9	59.9
$\mathbb{B}(\text{FT})^{\text{IQR}}$	TAPA-EGKK	False	False	False	True	True	True	True	True	True	True	True	True	False	False	True
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	TAPA-EGKK	False	False	False	True	True	True	True	False	True	False	True	True	False	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	TAPA-EGKK	2.0	0.0	0.0	24.0	36.0	34.0	24.0	18.0	18.0	20.0	8.0	6.0	2.0	14.0	12.0
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGKK-MMUN	254.4	76.9	70.5	370.3	246.9	220.0	639.8	321.0	273.3	899.3	460.4	436.5	971.6	574.9	545.2
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGKK-MMUN	35990.1	36365.6	36598.8	34405.0	34821.4	35212.8	34633.1	36014.6	36426.2	35048.7	36491.1	37286.9	35639.6	36281.0	37070.0
FI^{Rea}	EGKK-MMUN	36124.7	36428.3	36662.4	34309.6	34768.9	35146.8	34691.2	36117.0	36612.5	35033.4	36511.5	37317.4	35718.1	36458.5	37230.3
$\Delta \text{FT}^{\text{Mean}}$	EGKK-MMUN	-134.5	-62.7	-63.6	95.5	52.5	66.0	-58.1	-102.4	-186.2	15.3	-20.4	-30.5	-78.5	-177.6	-160.3
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGKK-MMUN	False	False	False	False	False	False	True	False	False	True	True	True	True	False	False
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGKK-MMUN	False	False	False	False	False	False	False	False	False	True	True	False	False	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGKK-MMUN	6.0	16.0	4.0	20.0	32.0	20.0	16.0	10.0	0.0	4.0	10.0	12.0	16.0	8.0	8.0

TABLE 4: Top three medium-haul flights in terms of flight time dispersion. Tol = 30 sec.

Variable	Airp. Orig-Dest	Step 0			Step 12			Step 24			Step 36			Step 48		
		dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGLL-KSAN	71.3	35.7	33.4	156.3	118.6	86.8	181.0	126.6	121.7	537.8	349.2	249.8	896.0	331.9	256.7
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGLL-KSAN	9758.4	9881.2	9911.8	9594.4	9746.3	9958.0	9882.8	10006.9	10090.9	10284.7	11149.2	12503.2	10959.2	11772.1	12744.0
FI^{Rea}	EGLL-KSAN	9753.3	9896.4	9926.9	9574.7	9725.5	9940.2	9842.9	9967.6	10054.3	10235.0	11078.0	12461.5	10932.5	11958.0	12917.7
$\Delta \text{FT}^{\text{Mean}}$	EGLL-KSAN	5.1	-15.2	-15.1	19.7	20.8	17.8	39.8	39.3	36.6	49.7	71.1	41.7	26.7	-185.9	-173.6
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGLL-KSAN	True	False	False	True	True	False	False	False	False	True	False	True	True	False	False
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGLL-KSAN	True	True	True	True	True	True	False	False	False	False	False	False	True	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGLL-KSAN	96.0	90.0	88.0	52.0	64.0	66.0	38.0	48.0	48.0	20.0	22.0	34.0	6.0	0.0	0.0
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGKK-LPMA	70.9	35.6	34.0	150.2	130.5	109.5	169.6	120.2	117.3	508.8	334.2	212.7	867.3	337.2	267.7
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGKK-LPMA	9721.7	9822.8	9845.3	9504.8	9558.5	9833.1	9740.0	9883.2	9936.2	10122.9	10897.7	11698.1	10772.7	11758.9	12555.2
FI^{Rea}	EGKK-LPMA	9722.8	9841.3	9863.4	9484.7	9536.2	9806.3	9698.2	9841.7	9896.1	10063.8	10831.1	11690.9	10723.3	11947.4	12734.8
$\Delta \text{FT}^{\text{Mean}}$	EGKK-LPMA	-1.1	-18.5	-18.1	20.1	22.2	26.7	41.8	41.5	40.1	59.1	66.6	7.2	49.5	-188.5	-179.6
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGKK-LPMA	True	False	False	True	True	False	False	False	False	True	False	True	True	False	False
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGKK-LPMA	True	True	True	True	True	True	False	False	False	False	False	True	False	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGKK-LPMA	96.0	78.0	82.0	52.0	54.0	52.0	36.0	42.0	42.0	20.0	22.0	32.0	8.0	0.0	0.0
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGLL-KSJC	65.8	46.7	35.5	143.3	132.2	128.3	286.0	224.1	205.9	555.8	360.3	262.4	610.9	234.5	232.2
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGLL-KSJC	9084.1	9173.2	9349.9	9402.7	9446.7	9519.9	9879.4	10090.3	10354.7	10365.6	10443.8	11113.9	9842.0	10702.4	10810.3
FI^{Rea}	EGLL-KSJC	9045.3	9127.0	9301.8	9377.9	9423.4	9496.8	9871.3	10095.7	10356.0	10235.9	10298.1	10985.4	9901.9	10725.1	10828.5
$\Delta \text{FT}^{\text{Mean}}$	EGLL-KSJC	38.9	46.2	48.1	24.8	23.3	23.1	8.1	-5.4	-1.3	129.7	145.7	128.5	-59.9	-22.7	-18.2
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGLL-KSJC	False	False	False	False	False	False	True	True	True	False	False	False	False	True	True
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGLL-KSJC	False	False	False	True	True	True	True	True	True	False	False	False	False	True	True
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGLL-KSJC	24.0	16.0	0.0	56.0	58.0	60.0	38.0	36.0	40.0	14.0	12.0	20.0	14.0	24.0	22.0

TABLE 5: Top three short-haul flights in terms of flight time dispersion. Tol = 30 sec.

Variable	Airp. Orig-Dest	Step 0			Step 12			Step 24			Step 36			Step 48		
		dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGKK-LPPR	39.2	36.8	36.1	95.8	86.2	83.2	151.2	84.3	75.1	309.3	291.6	134.3	493.8	272.3	244.8
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGKK-LPPR	5409.5	5415.3	5425.8	5097.5	5139.4	5186.8	5049.8	5280.3	5405.8	5171.1	5227.0	7044.2	5571.7	6514.7	6882.0
FI^{Rea}	EGKK-LPPR	5409.4	5414.9	5425.5	5085.2	5128.9	5176.8	5044.7	5270.2	5395.4	5130.7	5183.4	7021.9	5564.1	6617.9	6980.3
$\Delta \text{FT}^{\text{Mean}}$	EGKK-LPPR	0.1	0.4	0.4	12.3	10.5	10.0	5.0	10.1	10.4	40.4	43.6	22.3	7.5	-103.2	-98.2
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGKK-LPPR	True	True	True	True	True	True	True	False	False	True	False	True	True	False	False
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGKK-LPPR	True	True	True	True	True	True	True	True	True	False	False	True	True	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGKK-LPPR	100.0	100.0	100.0	72.0	78.0	80.0	76.0	86.0	88.0	30.0	30.0	44.0	22.0	20.0	18.0
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGLL-KDEN	31.8	30.5	29.4	120.0	94.0	84.4	118.0	84.7	80.4	254.0	120.7	116.2	451.1	371.1	346.6
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGLL-KDEN	5703.1	5709.9	5727.2	5450.9	5622.9	5746.9	5153.0	5284.7	5347.3	5194.3	5887.1	5951.4	5490.1	5742.3	6093.2
FI^{Rea}	EGLL-KDEN	5689.1	5697.5	5716.4	5437.7	5603.3	5727.4	5141.8	5272.5	5334.0	5140.9	5865.5	5931.2	5401.5	5628.3	5988.2
$\Delta \text{FT}^{\text{Mean}}$	EGLL-KDEN	14.0	12.3	10.8	13.2	19.5	19.5	11.2	12.1	13.3	53.4	21.6	20.2	88.6	114.0	105.0
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGLL-KDEN	False	False	False	True	False	False	True	False	False	False	True	True	True	False	False
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGLL-KDEN	True	True	True	True	True	True	True	True	True	False	True	True	False	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGLL-KDEN	100.0	100.0	100.0	76.0	72.0	74.0	84.0	84.0	84.0	32.0	56.0	58.0	12.0	18.0	18.0
$\mathbb{R}(\text{FT})^{\text{R-} \text{Trj}}$	EGLL-LEMD	31.1	29.9	29.3	105.0	69.8	67.8	115.1	85.4	80.8	221.5	108.4	106.4	402.4	344.8	316.3
$\overline{\text{FI}}^{\text{R-} \text{Trj}}$	EGLL-LEMD	5329.8	5334.9	5345.3	5153.9	5261.4	5290.6	4826.6	4902.1	4969.3	4828.0	5359.0	5389.2	5078.5	5268.6	5685.1
FI^{Rea}	EGLL-LEMD	5312.5	5319.4	5329.4	5142.1	5257.0	5285.5	4823.3	4898.2	4964.1	4784.7	5339.1	5369.7	4982.3	5157.5	5582.8
$\Delta \text{FT}^{\text{Mean}}$	EGLL-LEMD	17.3	15.5	15.9	11.8	4.4	5.2	3.3	3.9	5.1	43.3	19.9	19.5	96.2	111.1	102.3
$\mathbb{B}(\text{FT})^{\text{IQR}}$	EGLL-LEMD	False	False	False	True	True	True	True	True	True	False	True	True	False	False	False
$\mathbb{B}(\text{FT})^{\text{Exp} \pm \text{Tol}}$	EGLL-LEMD	True	True	True	True	True	True	True	True	True	False	True	True	False	False	False
$\% \text{FT}^{\text{Rea} \pm \text{Tol}}$	EGLL-LEMD	92.0	96.0	94.0	78.0	94.0	94.0	90.0	90.0	90.0	40.0	66.0	66.0	16.0	18.0	18.0

(a) Pareto front time dispersion vs. average flight time

(b) Pareto front fuel dispersion vs. average fuel consumption

FIGURE 11: EGLL to SBGL results trade-off solutions of the problem: dispersion vs. dp values; and flight times/fuel consumption vs. dispersion.

(a) Flight times (spline)

(b) Fuel consumptions (spline)

FIGURE 12: EGLL to SBGL results. Boxplots denote predicted uncertainties (based on EPS data and computed using the robust flight planning algorithm); red stars denote realized values (based on reanalysis data).

The single origin-destination flight (KJFK-LPPT), over a set of days along year 2016, was computed considering a forecast step used that represents the characteristic flight dispatching level. In average, the cost of increasing 1 sec. of predictability is 5 seconds of extra flight time. Also in average, the cost of reducing fuel burnt uncertainty 1 kg is 5 kg of extra fuel burnt. For some extreme days, increasing 1 sec. predictability (reducing also fuel burnt uncertainty by 1 kg) would cost 2.5 sec. (2.5 kg) of extra fuel burnt. This improvement in flight predictability has some direct relationship with airline savings via fuel reserves, yet also has important indirect effects such as less ATC tactical intervention, less delays, and thus potential miss connections. Further studies on

quantifying both these direct and indirect effects should be addressed. As for the multiple origin-destination analysis, we covered a global study considering 164 flights with different origin-destinations. The forecast step used (up to two days) represents the characteristic air traffic flow management pretactical and tactical time horizons. Similarly, for 1 to 2 days forecast steps, the cost of increasing 1 sec. of predictability is between 5 and 5.5 seconds of extra flight time (also, 5 to 5.5 kg of extra fuel burnt for reducing 1 kg of fuel burnt uncertainty). Yet, we can observe that the higher the forecasting step, the higher the uncertainty: uncertainty grows between 50 and 70 secs. in flight times every 12 hours when considering minimum time trajectories. We are

TABLE 6: Top three long-haul flights in terms of fuel consumption dispersion. Tol = 70 kg.

Variable	Airp. Orig-Dest	Step 0			Step 12			Step 24			Step 36			Step 48		
		dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGLL-SBGL	346.1	194.5	166.2	489.7	323.8	299.1	627.4	340.4	290.6	675.8	529.0	462.1	1529.8	668.2	629.4
$\overline{\text{FC}}^{\text{R-Tri}}$	EGLL-SBGL	56916.7	57404.2	57944.3	56726.1	57179.7	57605.0	56894.6	57838.8	58766.7	57471.2	58205.8	59331.6	58710.0	59361.9	60103.3
FC^{Rea}	EGLL-SBGL	57026.8	57541.2	58063.5	56638.0	57130.7	57538.8	56597.9	57545.7	58476.8	57293.6	57995.4	59137.4	58429.0	59316.0	60034.0
$\Delta \text{FC}^{\text{Mean}}$	EGLL-SBGL	-110.1	-137.0	-119.3	88.2	49.1	66.2	296.8	293.1	289.9	177.6	210.4	194.3	281.0	46.0	69.3
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGLL-SBGL	False	False	False	True	True	True	False	False	False	False	False	False	False	True	True
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGLL-SBGL	False	False	False	False	True	True	False	False	False	False	False	False	False	True	True
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGLL-SBGL	28.0	22.0	30.0	36.0	44.0	46.0	8.0	0.0	0.0	18.0	22.0	26.0	14.0	22.0	14.0
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	TAPA-EGKK	305.7	121.7	113.7	650.8	315.7	305.3	772.6	476.4	468.9	1098.5	614.2	555.3	1600.1	812.4	652.5
$\overline{\text{FC}}^{\text{R-Tri}}$	TAPA-EGKK	39430.5	39706.2	39969.5	39755.2	40339.7	40755.5	40618.1	42098.0	42868.0	41320.1	42473.1	43498.9	40778.7	42456.7	45050.3
FC^{Rea}	TAPA-EGKK	39271.0	39554.6	39824.3	39764.2	40336.0	40739.6	40655.6	42167.5	42895.1	41445.6	42515.3	43506.3	40945.2	42293.8	44965.3
$\Delta \text{FC}^{\text{Mean}}$	TAPA-EGKK	159.6	151.6	145.2	-9.0	3.7	15.9	-37.4	-69.5	-27.1	-125.6	-42.2	-7.4	-166.5	162.9	85.0
$\mathbb{B}(\text{FC})^{\text{IQR}}$	TAPA-EGKK	False	False	False	True	True	True	True	True	True	True	True	True	False	False	True
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	TAPA-EGKK	False	False	False	True	True	True	True	True	True	False	True	True	False	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	TAPA-EGKK	2.0	0.0	0.0	32.0	50.0	54.0	30.0	30.0	22.0	26.0	20.0	16.0	4.0	18.0	14.0
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGKK-MMUN	350.6	105.8	96.9	513.8	342.0	304.3	886.6	442.3	375.9	1244.3	633.0	598.1	1340.9	791.2	747.7
$\overline{\text{FC}}^{\text{R-Tri}}$	EGKK-MMUN	54140.7	54657.7	54978.4	51948.7	52525.9	53067.6	52265.0	54174.3	54741.1	52840.5	54830.2	55922.6	53657.0	54541.1	55625.3
FC^{Rea}	EGKK-MMUN	54326.0	54743.9	55065.8	51816.3	52453.2	52976.4	52345.5	54315.5	54997.1	52819.5	54858.3	55964.5	53765.5	54785.5	55845.1
$\Delta \text{FC}^{\text{Mean}}$	EGKK-MMUN	-185.4	-86.2	-87.4	132.4	72.7	91.2	-80.6	-141.1	-256.0	21.0	-28.1	-41.8	-108.5	-244.4	-219.8
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGKK-MMUN	False	False	False	False	False	False	True	False	False	True	True	True	True	False	False
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGKK-MMUN	False	False	False	False	False	False	False	False	False	True	True	True	False	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGKK-MMUN	10.0	32.0	36.0	32.0	42.0	42.0	26.0	24.0	0.0	14.0	18.0	22.0	20.0	14.0	12.0

TABLE 7: Aggregated KPI multiple origin-destination. Tol = 30 sec. (FT) and 70 kg (FC).

Variable	Step 0			Step 12			Step 24			Step 36			Step 48		
	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\overline{R(FT)}$	90.2	53.9	49.1	137.2	99.7	90.8	198.9	135.4	122.6	274.8	187.1	162.3	344.1	232.5	211.2
$\overline{R(FC)}$	130.0	77.2	70.2	198.9	143.1	129.9	284.9	192.9	174.8	396.3	269.9	234.5	480.2	321.3	288.2
ΔFC^{Mean}	23.6	24.1	23.1	17.5	19.0	19.4	21.1	22.2	21.9	39.8	39.1	40.1	56.4	50.3	49.9
ΔFC^{Mean}	33.3	34.1	32.8	24.1	26.8	27.3	30.4	32.6	31.8	56.3	55.2	57.1	79.1	68.2	69.4
$\overline{FT}^{R-Tri}$	13296.2	13394.3	13477.5	12762.9	12891.1	13032.8	12779.8	13007.1	13192.0	12618.0	12918.4	13246.0	11654.9	12021.9	12319.4
$\overline{FC}^{R-Tri}$	20564.1	20706.8	20827.7	19699.5	19888.6	20094.8	19643.2	19971.6	20232.1	19537.2	19972.1	20433.9	17678.7	18188.0	18639.7
$\%B(FT)^{IQR}$	23.8	25.6	25.6	61.2	56.2	53.1	70.6	60.0	63.8	54.5	42.3	39.7	64.6	51.4	48.6
$\%B(FC)^{IQR}$	23.8	25.6	25.6	61.2	56.2	53.1	70.6	60.0	63.8	54.5	42.3	39.7	64.6	51.4	48.6
$\%B(FT)^{Exp\pm Tol}$	74.4	70.7	72.0	85.6	81.9	83.8	78.8	77.5	78.1	54.5	54.5	55.1	52.1	54.9	54.2
$\%B(FC)^{Exp\pm Tol}$	89.0	86.6	87.8	91.2	88.8	89.4	89.4	90.0	88.8	75.0	74.4	73.7	61.1	65.3	66.7
$\%FT^{Real\pm Tol}$	69.8	69.3	70.4	68.0	70.7	71.4	58.2	61.7	63.1	41.0	43.6	45.3	37.0	41.6	42.1
$\%FC^{Real\pm Tol}$	83.4	83.9	85.1	80.7	82.0	83.1	74.8	77.3	78.0	63.3	65.4	67.0	54.8	59.6	60.2

TABLE 8: Top three medium-haul flights in terms of fuel consumption dispersion. Tol = 70 kg.

Variable	Airp. Orig-Dest	Step 0			Step 12			Step 24			Step 36			Step 48		
		dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGLL-KSAN	144.4	72.3	67.6	317.0	240.3	175.6	366.3	256.1	245.9	1085.2	700.3	496.1	1799.1	662.7	509.0
$\overline{\text{FC}}^{\text{R-Tri}}$	EGLL-KSAN	20521.3	20770.0	20831.9	20188.8	20496.8	20925.4	20773.1	21024.3	21194.1	21585.3	23324.6	26026.5	22943.1	24570.9	26504.3
FC^{Rea}	EGLL-KSAN	20511.0	20800.8	20862.4	20148.9	20454.6	20889.4	20692.4	20944.8	21120.1	21485.1	23181.9	25943.7	22889.8	24941.9	26848.3
$\Delta \text{FC}^{\text{Mean}}$	EGLL-KSAN	10.2	-30.8	-30.5	40.0	42.2	36.0	80.7	79.4	74.0	100.2	142.6	82.8	53.3	-371.0	-344.0
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGLL-KSAN	True	False	False	True	True	False	False	False	False	True	False	True	True	False	False
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGLL-KSAN	True	True	True	True	True	True	False	False	False	False	False	False	False	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGLL-KSAN	96.0	76.0	78.0	42.0	52.0	58.0	32.0	36.0	38.0	20.0	18.0	26.0	6.0	0.0	0.0
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGKK-LPMA	76.9	38.6	36.9	163.0	141.6	118.7	183.9	130.3	127.1	551.4	361.8	230.0	939.0	364.6	289.1
$\overline{\text{FC}}^{\text{R-Tri}}$	EGKK-LPMA	10624.5	10734.2	10758.5	10389.2	10447.5	10745.3	10644.4	10799.6	10857.1	11059.4	11898.7	12764.7	11763.4	12830.4	13690.9
FC^{Rea}	EGKK-LPMA	10625.7	10754.2	10778.1	10367.4	10423.4	10716.3	10599.0	10754.7	10813.6	10995.4	11826.6	12756.9	11709.8	13034.2	13884.9
$\Delta \text{FC}^{\text{Mean}}$	EGKK-LPMA	-1.2	-20.0	-19.6	21.8	24.1	29.0	45.3	45.0	43.5	64.0	72.1	7.8	53.5	-203.8	-194.0
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGKK-LPMA	True	False	False	True	True	False	False	False	False	True	False	True	True	False	False
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGKK-LPMA	True	True	True	True	True	True	True	True	True	False	False	True	False	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGKK-LPMA	100.0	100.0	100.0	74.0	72.0	72.0	60.0	58.0	60.0	30.0	34.0	50.0	12.0	4.0	0.0
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGLL-KSJC	97.0	68.8	52.3	211.0	194.6	188.7	419.9	328.8	301.6	814.0	527.6	383.0	897.1	342.9	339.5
$\overline{\text{FC}}^{\text{R-Tri}}$	EGLL-KSJC	13700.8	13832.1	14092.4	14170.0	14234.8	14342.5	14870.8	15180.3	15567.9	15583.8	15698.4	16678.0	14815.9	16076.9	16234.6
FC^{Rea}	EGLL-KSJC	13643.5	13764.1	14021.6	14133.5	14200.5	14308.5	14858.9	15188.2	15569.8	15393.9	15485.0	16490.4	14903.9	16110.0	16261.3
$\Delta \text{FC}^{\text{Mean}}$	EGLL-KSJC	57.3	68.1	70.8	36.5	34.3	34.0	11.8	-7.9	-1.9	189.9	213.4	187.6	-88.0	-33.2	-26.7
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGLL-KSJC	False	False	False	False	False	False	True	True	True	False	False	False	False	True	True
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGLL-KSJC	False	False	False	True	True	True	True	True	True	False	False	False	False	True	True
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGLL-KSJC	36.0	18.0	16.0	66.0	64.0	62.0	40.0	44.0	52.0	14.0	12.0	20.0	22.0	26.0	24.0

TABLE 9: Top three short-haul flights in terms of fuel consumption dispersion. Tol is 30 kg.

Variable	Airp. Orig-Dest	Step 0			Step 12			Step 24			Step 36			Step 48		
		dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20	dp0	dp10	dp20
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGKK-LPPR	42.8	40.2	39.4	104.6	94.2	90.9	165.1	92.0	82.0	337.8	318.4	146.3	539.0	296.7	266.6
$\overline{\text{FC}}^{\text{R-Tri}}$	EGKK-LPPR	5932.7	5938.9	5950.5	5592.0	5637.7	5689.5	5539.8	5791.5	5928.6	5672.3	5733.3	7715.0	6109.7	7138.2	7538.4
FC^{Rea}	EGKK-LPPR	5932.6	5938.5	5950.1	5578.5	5626.2	5678.6	5534.3	5780.5	5917.3	5628.2	5685.8	7690.7	6101.5	7250.7	7645.4
$\Delta \text{FC}^{\text{Mean}}$	EGKK-LPPR	0.1	0.4	0.4	13.4	11.5	10.9	5.5	11.0	11.3	44.1	47.6	24.3	8.2	-112.5	-107.0
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGKK-LPPR	True	True	True	True	True	True	True	False	False	True	False	True	True	False	False
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGKK-LPPR	True	True	True	True	True	True	True	True	True	False	False	True	True	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGKK-LPPR	100.0	100.0	100.0	70.0	70.0	72.0	74.0	86.0	86.0	28.0	26.0	38.0	20.0	16.0	16.0
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGLL-KDEN	66.3	63.8	61.4	251.1	196.4	176.3	247.6	177.5	168.4	532.7	251.7	242.3	943.9	774.9	721.8
$\overline{\text{FC}}^{\text{R-Tri}}$	EGLL-KDEN	12179.2	12193.4	12229.7	11652.1	12011.6	12270.7	11027.8	11303.9	11435.1	11114.5	12563.3	12697.3	11733.9	12261.2	12992.9
FC^{Rea}	EGLL-KDEN	12150.1	12167.7	12207.1	11624.4	11970.8	12230.0	11004.4	11278.5	11407.3	11002.5	12518.4	12655.2	11548.6	12023.1	12774.2
$\Delta \text{FC}^{\text{Mean}}$	EGLL-KDEN	29.2	25.7	22.6	27.6	40.8	40.7	23.5	25.4	27.9	112.0	45.0	42.1	185.4	238.1	218.7
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGLL-KDEN	False	False	False	True	False	False	True	False	False	False	True	True	True	False	False
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGLL-KDEN	True	True	True	True	False	False	True	True	True	False	False	False	False	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGLL-KDEN	52.0	58.0	62.0	32.0	34.0	36.0	52.0	48.0	48.0	10.0	40.0	38.0	2.0	10.0	8.0
$\mathbb{R}(\text{FC})^{\text{R-Tri}}$	EGLL-LEMD	33.7	32.5	31.8	114.1	75.8	73.6	125.1	92.8	87.8	240.7	117.7	115.5	437.0	374.3	343.2
$\overline{\text{FC}}^{\text{R-Tri}}$	EGLL-LEMD	5810.6	5816.1	5827.4	5619.7	5736.4	5768.1	5264.1	5346.2	5419.1	5265.7	5842.3	5875.1	5537.7	5744.1	6196.3
FC^{Rea}	EGLL-LEMD	5791.9	5799.3	5810.1	5606.8	5731.6	5762.5	5260.5	5341.9	5413.5	5218.6	5820.7	5853.9	5433.2	5623.5	6085.3
$\Delta \text{FC}^{\text{Mean}}$	EGLL-LEMD	18.7	16.8	17.3	12.9	4.8	5.6	3.6	4.3	5.6	47.1	21.6	21.2	104.5	120.6	111.0
$\mathbb{B}(\text{FC})^{\text{IQR}}$	EGLL-LEMD	False	False	False	True	True	True	True	True	True	False	True	True	False	False	False
$\mathbb{B}(\text{FC})^{\text{Exp}\pm\text{Tol}}$	EGLL-LEMD	True	True	True	True	True	True	True	True	True	False	True	True	False	False	False
$\% \text{FC}^{\text{Rea}\pm\text{Tol}}$	EGLL-LEMD	88.0	92.0	92.0	74.0	94.0	94.0	90.0	90.0	88.0	36.0	60.0	60.0	16.0	18.0	16.0

FIGURE 13: Aggregated dispersion boxplots multiple origin-destination (long haul).

FIGURE 14: Aggregated dispersion boxplots multiple origin-destination case study.

FIGURE 15: Aggregated dispersion boxplot multiple origin-destination (medium haul).

FIGURE 16: Aggregated dispersion boxplot multiple origin-destination (short haul).

however able to reduce flight time uncertainty by roughly 40% in average. The direct effect of this would be a more accurate sector loading forecasting, leading to increase the capacity of the system. Indirectly, airlines would benefit from less regulations.

Data Availability

This research is using weather data from the ECMWF MARS public databases (<https://www.ecmwf.int>); therefore, the work can be replicated. Also, BADA data was used and it is accessible from the Eurocontrol website (<https://simulations.eurocontrol.int/solutions/bada-aircraft-performance-model/>).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is partially supported by the Spanish Government through the project entitled “Analysis and Optimization of Aircraft Trajectories under the Effects of Meteorological Uncertainty-OptMet” (TRA2014-58413-C2-2-R) (<https://optmet.wordpress.com/>). The very first ideas of the present paper were introduced in a talk in the Workshop on Uncertainty and Air Traffic Management organized by the OptMet project on Wednesday 25th, October 2017 in Madrid (Spain).

References

- [1] L. Hao, M. Hansen, M. S. Ryerson, and M. S. Ryerson, “Fueling for contingencies: the hidden cost of unpredictability in the air transportation system,” *Transportation Research Part D: Transport and Environment*, vol. 44, pp. 199–210, 2016.
- [2] A. Cook and D. Rivas, *Complexity Science in Air Traffic Management*, Routledge, London and New York, 2016.
- [3] F. Goyens, F. Gonze, A. Simonetto, E. Huens, J. Boucquoy, and R. M. Jungers, “A probabilistic model for precedence rules and reactionary delay in air traffic management,” in *ICRAT 2018*, pp. 1–8, Castelldefels, Catalonia, Spain, 2018.
- [4] A. Valenzuela, A. Franco, D. Rivas, J. García-Heras, and M. Soler, “Effects of reducing wind-induced trajectory uncertainty on sector demand,” in *Seventh SESAR Innovation Days*, pp. 1–8, Belgrade, Serbia Republic, 2017.
- [5] A. Cook, H. A. P. Blom, F. Lillo et al., “Applying complexity science to air traffic management,” *Journal of Air Transport Management*, vol. 42, pp. 149–158, 2015.
- [6] E. Casado, M. L. Civita, M. Vilaplana, and E. W. McGookin, “Quantification of aircraft trajectory prediction uncertainty using polynomial chaos expansions,” in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, pp. 1–11, St. Petersburg, FL, USA, 2017.
- [7] J. Sun, J. Ellerbroek, and J. M. Hoekstra, “Aircraft initial mass estimation using Bayesian inference method,” *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 59–73, 2018.
- [8] R. Vazquez and D. Rivas, “Propagation of initial mass uncertainty in aircraft cruise flight,” *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 2, pp. 415–429, 2013.
- [9] J. Cheung, J. L. Brenguier, J. Heijstek, A. Marsman, and H. Wells, “Sensitivity of flight durations to uncertainties in numerical weather prediction,” in *Fourth SESAR Innovation Days*, UPM, Madrid, 2014.
- [10] R. Vazquez, D. Rivas, and A. Franco, “Stochastic analysis of fuel consumption in aircraft cruise subject to along-track wind uncertainty,” *Aerospace Science and Technology*, vol. 66, pp. 304–314, 2017.
- [11] J. Cheung, A. Hally, J. Heijstek, A. Marsman, and J. L. Brenguier, “Recommendations on trajectory selection in flight planning based on weather uncertainty,” in *Fifth SESAR Innovation Days*, Bologna, Italy, 2015.
- [12] A. Franco, D. Rivas, and A. Valenzuela, “Probabilistic aircraft trajectory prediction in cruise flight considering ensemble wind forecasts,” *Aerospace Science and Technology*, vol. 82–83, pp. 350–362, 2018, In Press.
- [13] K. Legrand, S. Puechmorel, D. Delahaye, and Y. Zhu, “Robust aircraft optimal trajectory in the presence of wind,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 33, no. 11, pp. 30–38, 2018, In press.

- [14] D. Gonzalez-Arribas, M. Soler, and M. Sanjurjo-Rivo, "Robust aircraft trajectory planning under wind uncertainty using optimal control," *Journal of Guidance, Control, and Dynamics*, vol. 41, no. 3, pp. 673–688, 2018.
- [15] J. García-Heras, *Analysis of the geometric altimetry to support aircraft optimal trajectories within the future 4D trajectory management*, Universidad Politécnica de Madrid. School of Aeronautical Engineering, Madrid, 2014.
- [16] B. Øksendal, *Stochastic Differential Equations*, Springer, Berlin, Heidelberg, 1998.
- [17] A. Nuic, "User Manual for the Base of Aircraft Data (BADA)," rev 3.13, 2015.
- [18] Y. Y. Park, R. Buizza, and M. Leutbecher, "TIGGE: preliminary results on comparing and combining ensembles," *Quarterly Journal of the Royal Meteorological Society.*, vol. 134, no. 637, pp. 2029–2050, 2008.
- [19] L. Amodèi and M. N. Benbourhim, "A vector spline approximation," *Journal of Approximation Theory.*, vol. 67, no. 1, pp. 51–79, 1991.
- [20] C. Le Guyader, D. Apprato, and C. Gout, "Spline approximation of gradient fields: applications to wind velocity fields," *Mathematics and Computers in Simulation*, vol. 97, pp. 260–279, 2014.
- [21] J. Andersson, J. Åkesson, and M. Diehl, "CasADi: a symbolic package for automatic differentiation and optimal control," in *Recent Advances in Algorithmic Differentiation*, Springer, Berlin, Heidelberg, 2012.