Trajectory Prediction Sensitivity Analysis Using Monte Carlo Simulations Based on Inputs' Distributions

Rudnyk, Julia; Ellerbroek, Joost; Hoekstra, Jacco

DOI
10.2514/1.D0156

Publication date
2019

Document Version
Accepted author manuscript

Published in
Journal of Air Transportation

Citation (APA)
Rudnyk, J., Ellerbroek, J., & Hoekstra, J. (2019). Trajectory Prediction Sensitivity Analysis Using Monte Carlo Simulations Based on Inputs' Distributions. *Journal of Air Transportation*, *27*(4), 181–198. https://doi.org/10.2514/1.D0156

Post-print of journal article accepted by AIAA Journal of Air Transportation, Aug 2019. doi:
http://arc.aiaa.org/doi/abs/10.2514/1.D0156

Trajectory Prediction Sensitivity Analysis Using Monte Carlo Simulations Based on Inputs' Distributions

Julia Rudnyk*, Joost Ellerbroek†, and Jacco M. Hoekstra‡
Control and Simulation, Faculty of Aerospace Engineering,
Delft University of Technology, Delft, the Netherlands

To facilitate the increasing amount of air traffic, current and future decision support tools for air traffic management require an efficient and accurate trajectory prediction. With uncertainty inherent to almost all inputs of a trajectory predictor, the accurate prediction is not a simple task. In this study, Monte Carlo simulations of a ground-based trajectory predictor are performed to estimate the prediction uncertainty up to 20 minutes look-ahead time and to assess the correlation between inputs and prediction errors. Selected inputs are aircraft bank angle, constant calibrated airspeed and Mach number speed settings, vertical speed, temporary level-offs, air temperature, lapse rate, wind, and air traffic control intent. These inputs are provided in the form of their distribution functions obtained from observed data such as surveillance data, weather forecasts, and air traffic controllers' inputs. Simulations are performed for heavy and medium wake turbulence category aircraft. Results indicate that with a 20 minute look-ahead time, when outliers are not considered, along-track errors can reach up to 18 nautical miles, whereas altitude errors can reach up to around 13,000 feet. Cross-track errors in cruise highly depend on the lateral deviations due to ATC instructions, and, in this study, are within 10 nautical miles. Wind conditions, vertical speed, calibrated airspeed, Mach number speed setting, and temporary level-offs are determined to be the most influential inputs.

I. Introduction

It is a prevalent forecast that air traffic will grow in the coming years [1, 2]. To be able to facilitate this growth, numerous decision support tools are being developed with the purpose of assisting Air Traffic Control (ATC) units in such tasks as flight planning, air traffic flow and capacity management, arrival and departure management, conflict detection and resolution, metering, sequencing, and merging.

*PhD candidate, Control and Simulation, Faculty of Aerospace Engineering, Delft University of Technology, the Netherlands
†Assistant Professor, Control and Simulation, Faculty of Aerospace Engineering, Delft University of Technology, the Netherlands
‡Professor, Control and Simulation, Faculty of Aerospace Engineering, Delft University of Technology, the Netherlands, AIAA Member

All of these tasks require the ability to predict the future positions of aircraft. To facilitate this, a (ground-based) Trajectory Predictor (TP) provides aircraft positions in three dimensions for a number of time instances into the future. Trajectory prediction is a complex process which requires information related to aircraft performance, meteorological conditions, and flight intent, and is performed through the integration of aircraft equations of motion. All information that is required for the prediction process and can vary per flight (e.g. aircraft type, wind conditions, vertical and/or horizontal constraints) can be referred to as input parameters or inputs. Information that does not vary per flight (e.g. equations of motion, gravitational acceleration, Earth radius, coordinate system) is referred to as model and model parameters. Assuming that the model and the model parameters are realistic and correct, the accuracy of the prediction depends on the accuracy of the inputs. It occurs often that the availability of particular input values is (temporarily) limited or absent, and that assumptions have to be made to initiate the prediction process. This study focuses on predictions up to 20 minutes look-ahead time. For this time interval, accurate knowledge of the weather conditions, ATC instructions and the aircraft modes of operation is crucial. These are often not known and either assumed to have standard values or are disregarded. Differences between reality and assumptions result in prediction errors, which in turn diminish the added value of decision support tools (e.g. Medium Term Conflict Detection).

An often proposed solution to this problem would be to ensure that inputs are known with high certainty (e.g. by down-linking some data or computed trajectory from aircraft [3–8]). However, not all inputs can be precisely known in advance (e.g. weather, ATC instructions).

Much research has been conducted exploring the problem of trajectory prediction uncertainty [9–20]. Henderson et al. (2013) measured the trajectory prediction accuracy and quantified the sources of errors by conducting field trials, where aircraft were descending at one of three predefined speed profiles [17]. Some studies [10, 13, 14] performed a sensitivity analysis, where one input was varied at a time, while others were kept constant at their reference values. This is known as a local sensitivity analysis, where usually, inputs are varied from their reference value either by some fixed percentage, the standard deviation, or some absolute value. Although this approach is fast, simple, and provides an estimation of the contribution of each input to the output uncertainty, it cannot be used to study the combined influence of all inputs, and results strongly depend on the chosen reference values.

Another way to study uncertainty is to perform a global sensitivity analysis, where all inputs are varied at the same time from their probability distribution functions. In this way a conclusion can be made about individual and combined contributions of inputs to prediction uncertainty. In their work, Mondoloni et al. (2005) performed simulations where input errors were sampled from distributions [13]. These distributions, however, except one, were not estimated from the observed data. To estimate the sensitivity of predictions to errors in inputs to a TP, Mueller et al. (2004) conducted Monte Carlo simulations where it was assumed that errors are normally distributed [12]. Although the mean values and standard deviations used to define the distributions were based on the measurements, the assumption of normality might be incorrect. Casado (2016) developed a framework to estimate prediction uncertainty using arbitrary Polynomial Chaos

Expansion, where the majority of input distributions were assumed to be normal [20], whereas in related work [21, 22] some inputs were described with uniform and triangular distributions. Torres (2015) performed a global sensitivity analysis where the errors of inputs to the TP were modeled using uniform and normal distributions [19]. Whereas this approach provides reasonable bounds on prediction errors, by taking into account the results from prior research on input variability, the assumed distributions might not reflect the probability distribution of the errors in real operations. Furthermore, to assess the contribution of each input to prediction errors, a local sensitivity analysis was performed. As previously mentioned, this approach does not allow to investigate the impact an input has in combination with other inputs (e.g. effect of one input may be diminished or compensated by the effect of another input).

The objective of this paper is therefore to address this issue by performing Monte Carlo simulations with distributions of the inputs that are obtained from observed data. Additionally, these input distributions are compared to commonly made assumptions (e.g. single settings of an input or a range of feasible values). The selection of TP inputs was based on the conclusions of previous research [10, 11, 13–16, 18, 20] about significant input parameters and includes aircraft bank angle, constant Calibrated airspeed (CAS)/Mach number speed settings, vertical speed, temporary level-offs in the climb and descent phases, air temperature and lapse rate, wind, and ATC intent. To take into account the differences between aircraft in terms of mass, Monte Carlo simulations were conducted for aircraft of heavy and medium Wake Turbulence Category (WTC), since these two categories are most common in civil aviation. Prediction uncertainty is estimated as a function of a look-ahead time up to 20 minutes and the contribution of each input to the prediction errors is assessed based on the computed correlation coefficients.

II. Sensitivity Analysis

A sensitivity analysis can be performed to determine the influence of uncertainties for different TP inputs on the quality of the trajectory prediction. In the field of sensitivity analysis, there is a distinction between *sensitivity* and *uncertainty* analysis [23]. A sensitivity analysis attributes uncertainty in the model output(s) to uncertainty in model inputs. Most often, a sensitivity analysis is performed by measuring sensitivity indices or other measures. An uncertainty analysis is meant to quantify uncertainty in the model output(s) due to uncertainty in model inputs and as a result identify sensitive inputs. Uncertainty analysis methods are used to sample inputs and propagate their uncertainty through the model, which allows to estimate sensitivity. In this paper, both uncertainty and sensitivity analyses are performed and the generalized term of Sensitivity Analysis (SA) will be further used to refer to them. Two types of SA can be distinguished: local and global. A short introduction to the approaches of a local (commonly used) and global (this study) SA is given below.

A. Local Sensitivity Analysis

A local SA uses a one-at-a-time (OAT) sampling approach, where only one input parameter is varied at a time, while the rest is kept at their reference or nominal values [23]. By changing one parameter at a time, a conclusion can be made about what inputs contribute to changes in output and to what extent. The biggest limitation of OAT is that it strongly depends on the reference (nominal) value of each input. Additionally, the assumption of inputs at their nominal values might hide interesting results of a combined impact of several inputs and is an oversimplification of reality. For a local SA, one could take derivatives or change inputs with a certain step (e.g. ±5%) around a nominal value. A derivative-based SA is based on computing partial derivatives of output Y with respect to its inputs $\mathbf{X} = X_i, X_j, ..., X_n$. Although it is a straightforward way of expressing sensitivity, the obtained measure is representative only at the point where it is computed. This implies that if one wants to perform a derivative-based SA of a non-linear system and obtain an average sensitivity measure, derivatives must be computed at a set of different points.

B. Global Sensitivity Analysis

Usually, a global SA uses an all-at-a-time sampling approach where all inputs are varied at the same time from their distribution functions, or the entire feasible space [23]. A simultaneous variation of all inputs allows to investigate the sensitivity of the output to each input and a combined impact due to interactions. A common approach to conduct a global SA is to run Monte Carlo simulations. The Monte Carlo method applies sampling, which is a process of randomly drawing variables from their distribution functions. Usually, Monte Carlo simulations include the following steps: 1) sample inputs from the distribution functions; 2) compose input vectors containing sampled inputs; 3) run a model for all input vectors. Using chosen metrics for a model output evaluation, an uncertainty analysis can be performed by computing output distribution and statistics. Then, scatter plots can be used to inspect the sensitivity of output to each input and a sensitivity measure can be obtained by computing a correlation coefficient.

III. Estimation of Input Distributions

A. Probability Distribution Function

Each studied input to a TP is considered as a random variable and the goal is to describe it using a probability function or a set of such functions. The latter is applicable to parameters that require two or more characteristics to be defined (e.g. temporary level-offs are characterized by the flight level at which they occur and their duration).

Although the majority of studied parameters can be represented as continuous random variables, some parameters (e.g. the flight level of a temporary level-off) should be characterized as discrete random variables. A continuous random variable is described by a probability density function (PDF) f, while a discrete random variable is described by a probability mass function (PMF) p. The probability P that a continuous random variable X takes on a particular value x is 0 and, instead, the probability that X falls within a range from a to b is found $P(a \leq X \leq b) = \int_a^b f(x)dx$.

The PDF f is the derivative of the cumulative distribution function $F(x)$. The probability P that a discrete random variable X takes on a particular value x is $P(X = x) = p(x)$.

B. Maximum Likelihood Estimate

The estimation of the distributions of the studied parameters is a two-step process, which consists of the estimation of the distribution parameters (this subsection) and evaluation of the goodness of fit (next subsection).

Unknown parameters of a distribution can be estimated from observations by means of a maximum likelihood estimation (MLE) [24]. MLE finds such values of parameters, which make the observed data most probable, i.e., most likely. A sample $X = x_1, \ldots, x_n$ of n observations comes from a distribution f defined by a vector of parameters θ, $f(X \mid \theta)$. Assuming these observations are independent and identically distributed, their joint distribution can be expressed as:

$$f(X \mid \theta) = f(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) \tag{1}$$

The likelihood $\mathcal{L}$ is a probability of obtaining observed values given certain values of parameters:

$$\mathcal{L}(\theta \mid X) = \prod_{i=1}^{n} f(x_i \mid \theta) \tag{2}$$

This can be simplified to

$$log\,\mathcal{L}(\theta \mid X) = log \prod_{i=1}^{n} f(x_i \mid \theta) = \sum_{i=1}^{n} log\, f(x_i \mid \theta) \tag{3}$$

Then, the maximum likelihood estimate of θ (θ_{MLE}) is that value of θ that maximizes $\mathcal{L}$:

$$\theta_{MLE} = \arg\max_{\theta} log\,\mathcal{L}(\theta \mid X) = \arg\max_{\theta} \sum_{i=1}^{n} log\, f(x_i \mid \theta) \tag{4}$$

C. Kolmogorov-Smirnov Test

After a data sample is fit to a distribution, i.e., its parameters are estimated using MLE, a Kolmogorov-Smirnov (K-S) test is performed to test the goodness of fit. The K-S test determines if a data sample comes from a population with a certain distribution [25]. The test is based on the maximum distance D (Eq. (5)), i.e., K-S statistic, between a hypothetical cumulative distribution function $F(x)$ and an empirical distribution function $F_N(x)$ obtained from the data sample. If the data sample comes from this hypothetical distribution, D should be small.

$$D = \max\left|F(x) - F_N(x)\right| \tag{5}$$

In this study, the K-S test was applied to PDFs obtained using MLE and the distribution that resulted in the smallest K-S statistic D was chosen as the best fit.

IV. Data Used in This Study

A. Data Source

The surveillance data used in this study, including Mode-S information, was provided in ASTERIX Category 062 format [26] by Maastricht Upper Area Control (MUAC). Figure 1 shows the MUAC airspace geographic boundaries in blue and MUAC primary and secondary surveillance radar coverage is indicated by the red area. The selected data covered a period from the first until the seventh of July 2016 with an update interval of four to five seconds. The weather conditions and air traffic management regulations issued in this period were inspected, but are not considered to have a significant impact on the statistics obtained for the studied input parameters.

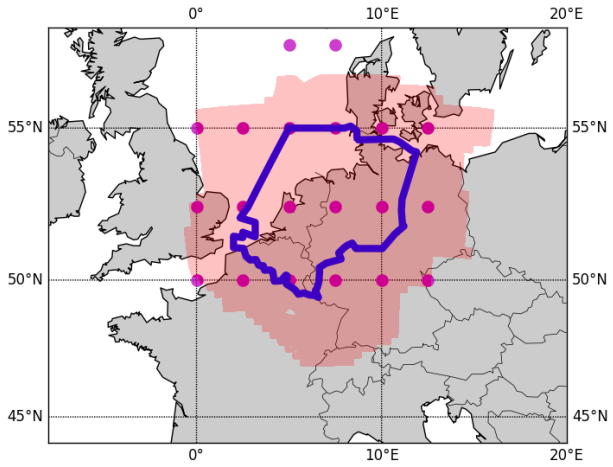


Fig. 1 MUAC airspace boundaries (blue), MUAC radar coverage (red area), weather forecast points (purple).

Weather forecasts were provided by the Royal Netherlands Meteorological Institute and were obtained from MUAC for 2016. These forecasts contain wind direction, wind speed and air temperature for 20 predefined geographic locations at 10 flight levels ranging from flight level (FL) 50 to FL530. These locations are shown in Fig. 1 as purple dots.

Recorded ATC instructions were provided by MUAC. As part of regular operations, ATC instructions are stored in MUAC's Flight Data Processing System. These include, among others, 'direct', heading, speed, and climb/descend instructions.

B. Data Pre-processing

On average, data from approximately 18,000 flights is received by MUAC radar per day, and of those flights, approximately 5,900 flight plans (FPLs) were sent to MUAC. Since it is not obligatory to provide information about

aircraft type, wake turbulence category and flight type in Mode-S, this data is extracted from the MUAC FPL database. Therefore, only those flights for which FPLs were received by MUAC were considered in this study. Furthermore, scheduled and non-scheduled flights were analyzed. Table 1 shows the resulting number and percentage of flights per WTC for 7 days.

Table 1 Number and percentage of flights per WTC for 7 days

WTC	Flights	%
Heavy	7,143	17.6
Medium	33,109	81.4
Light	427	1.0
Total	40,679	100

All light WTC aircraft (a total of 27 different aircraft types) performed non-scheduled flights and amount to only 1% of all observed flights. Based on such a small percentage (i.e., civil aviation mostly comprises medium and heavy aircraft) and a small data set size (fewer than 500 flights) the light WTC aircraft were not included in this study, since the distributions estimated from such data can not be considered as representative. Furthermore, the data used in this study is not geographically and/or vertically limited to the MUAC airspace, all messages within radar coverage were considered. Data records that did not contain a date/time stamp, call sign, position information or altitude were filtered out and discarded. Scattered messages were ordered and organized in flights.

C. Flight Phase Identification

In the analysis of the surveillance data, three flight phases were distinguished: climb, cruise and descent. For each flight, the phases were determined based on the altitude values (down-linked by aircraft) and an indication of the aircraft movement (leveled, climbing or descending) provided by the radar tracker. The top of climb was determined as the first point where an aircraft flew level for more than five consecutive minutes and its altitude either corresponds to the maximum flight level, FL_{max}, for this flight, or it is within the range $[FL_{max} - 6,000$ feet, $FL_{max}]$. This logic allows to include stepped cruise, where aircraft start their cruise at a lower altitude and periodically climb to maintain an optimal altitude. A stepped descent in cruise was detected in the same manner. The top of descent was determined as the first point where an aircraft has started descending continuously. On average, the climb, cruise and descent phases contained 215, 390 and 317 track points, which is about 17, 31 and 25 minutes. Not all flights contained all three phases. This did not present difficulties for this study, since for data analysis, all phases were considered separately from each other.

V. Study Setup

This section outlines the inputs to Monte Carlo simulations, describes the TP implementation used in this study and illustrates the computation of *actual* (ground truth) trajectories and *predictions*. Furthermore, a technique to compute prediction errors is presented and an approach to estimate the required number of simulation runs is demonstrated.

A. Input Data

In this study, the following TP inputs are considered: bank angle, constant CAS and Mach number speed settings, vertical speed, temporary level-offs in the climb and descent phases, air temperature and lapse rate, wind speed and wind direction represented by the wind components, and ATC intent. The distributions of the aircraft-related parameters, such as bank angle, speed settings and vertical speed, were obtained per aircraft type. The ten most frequent aircraft types (Table 2) were selected to represent their wake turbulence category, amounting to 88% and 85% of all flights for medium and heavy WTC. In total, the surveillance data contained 78 and 26 different aircraft types for medium and heavy WTC respectively.

Table 2 Ten most frequent aircraft types per WTC

WTC	ICAO aircraft type designator	% of flights
Medium	B738, A320, A319, A321, E190, B737, CRJ9, E170, B752, B733	88%
Heavy	B77W, A332, B763, B744, B772, A388, A333, B788, B77L, B748	85%

The remainder of this subsection presents the results on the estimated distributions in a graphical form. For sake of clarity, the results for the aircraft-related parameters are illustrated for a single aircraft, the Boeing 737-800 (B738). This model was chosen since it is one of the most common aircraft in civil aviation. The statistics on temporary level-offs are summarized for all medium WTC aircraft. The results for the air temperature, lapse rate and wind components are provided for one season, whereas the results for ATC instructions are summarized for the whole year. All estimated distribution functions (in both graphical and tabular form) as well as their equations can be accessed as a data collection [27]. Furthermore, the average values of CAS and Mach number distributions estimated in this study are comparable to those found in the research on constructing an open-source aircraft performance model utilizing ADS-B data [28]. The average values of the speed distributions were also compared with the BADA [29] values and were found to be within 20 knots for CAS and within 0.02 for Mach number. Within these results, the Embraer E170 and the Boeing 747-800 showed the largest differences.

1. Bank Angle

Considering that the bank angle is not transmitted in Mode-S and it is reasonable to assume that the roll angle value (which is transmitted in Mode-S) is close to that of the bank angle, the distributions were obtained for the roll angle. Figure 2 depicts the roll angle distribution in the climb, cruise, and descent phases of B738. The distributions were

obtained above FL100 to exclude the high frequency of large roll angles occurring when in holding and when flying standard instrument departure and standard arrival routes. The distributions in the climb and descent phases are similar and have several peaks. These peaks occur at around 10 and 25 degrees and could be a result of the manual selection of (preferred) bank angles or flight management system logic. In the cruise phase, smaller roll angles prevail with a peak at around 10 degrees. This can be explained by small course change angles in the cruise phase. When using aircraft performance models (e.g. BADA [29]), the bank angle is often assumed to be a constant value (e.g. 35 degrees).

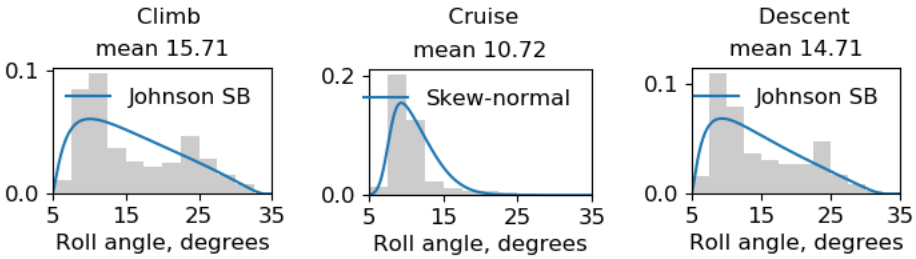


Fig. 2 Roll angle distribution above FL100 in the climb, cruise and descent phase of B738.

2. Speed Profile

Whereas for trajectory prediction purposes a speed profile is usually defined as a 'constant CAS - constant Mach number', CAS is not transmitted in Mode-S, so indicated airspeed (IAS) was used instead. Since CAS and IAS are almost identical only in aircraft clean configuration, the difference between the two is assumed to be insignificant in this study. Figure 3 shows the distributions of IAS and Mach number values in the climb and descent phases and Mach number in the cruise phase for B738. According to the BADA aircraft performance model [29], a B738 climbs at 300 knots CAS / 0.78 Mach, performs cruise at 0.78 Mach, and descends at 0.78 Mach / 290 knots CAS. These values are very close to the average values of the estimated distributions.

In practice, climb, cruise and descent speeds are calculated by the flight management computer based on the given cost index, which is the ratio of the time- and fuel-related costs. The cost index varies among airlines, aircraft models and flights and is not shared with the ground-based TPs. As a measure of the relationship between the IAS and Mach number in the climb and descent phase, the correlation between these two speeds was calculated. Figure 4 depicts an example of moderate (A320) and weak (CRJ9) correlation between the speeds. The red line illustrates the regression line, and the correlation value is given on top of each plot. For the majority of the studied aircraft types, the correlation is stronger in the climb phase, compared to the descent phase. This can be explained by the fact that aircraft mostly perform a fast climb to reach the cruise altitude as soon as possible, usually without speed restrictions from ATC. In the descent phase, however, the top of descent can be determined by ATC and multiple crossing restrictions can be given for sequencing purposes. All this would impact the choice of the descent speed, causing it to be sub-optimal. Although it is evident that the Mach number increases with the IAS, the relationship is not linear or quadratic.

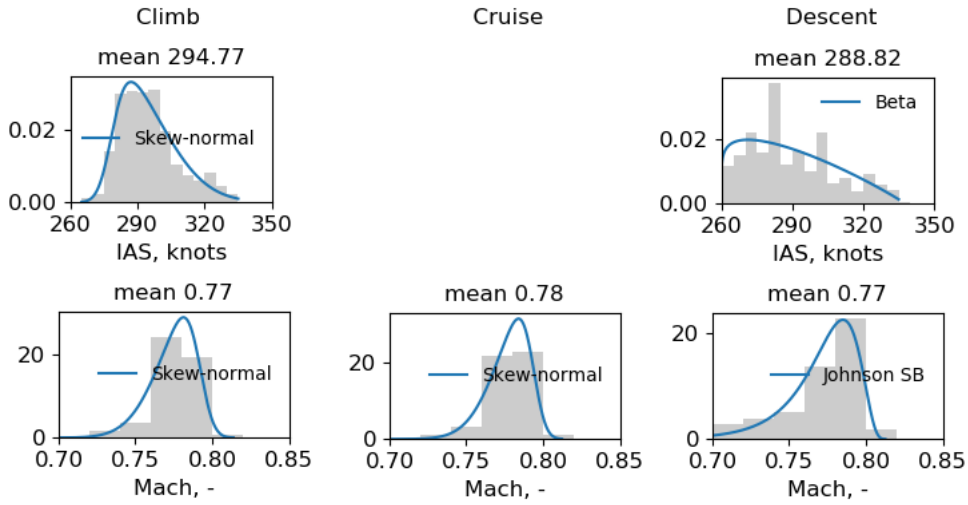


Fig. 3 IAS and Mach number distribution in the climb, cruise and descent phase of B738.

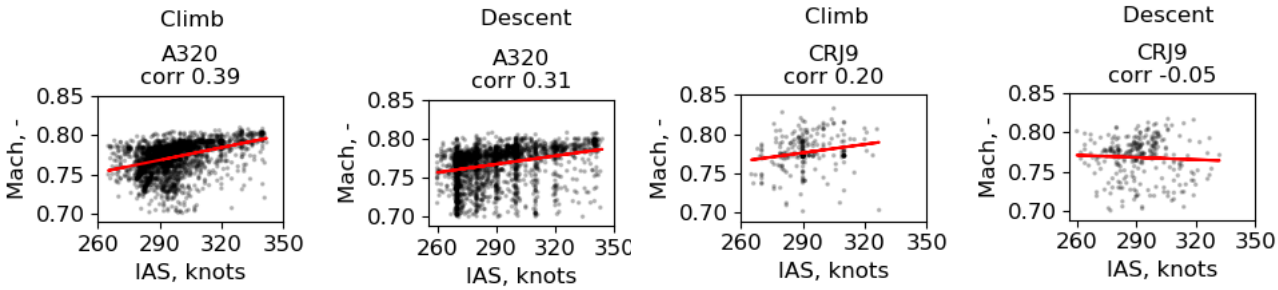


Fig. 4 IAS-Mach number correlation of A320 (left) and CRJ9 (right) in climb and descent phase.

3. Vertical Speed

In the climb phase, the amount of energy available for the climb determines the vertical speed, which decreases with altitude due to the effect of atmospheric conditions on engine performance. In the descent phase, idle thrust is preferred, which together with wind conditions and crossing restrictions determines the vertical speed.

Figure 5 illustrates the relationship between the barometric vertical rate (BVR) (down-linked by aircraft in Mode-S) and the flight level (FL) for the B738. The correlation value is indicated on top of each plot. In the climb phase, there is a strong linear relationship (indicated by the red regression line) between the BVR and altitude, except for altitudes around 3,000 feet and 10,000 feet (FL100). In the observed airspace, noise abatement procedures are in place below 3,000 feet. Above this altitude the noise abatement procedure ends and aircraft can increase power to climb at a higher speed. Above FL100 aircraft are not restricted to 250 knots CAS anymore and can fly at an optimum speed. When accelerating to a higher speed, energy is used on acceleration, which results in less energy being available for climb. In the descent phase, the correlation is weak, however the relationship between the BVR and the altitude can be described with a quadratic equation (as indicated by the red curve in Fig. 5), which mostly corresponds to the average BVR at each

FL. Above FL100, aircraft tend to descend at a fixed vertical speed setting, whereas around and below FL100 aircraft reduce the descent rate for approach.

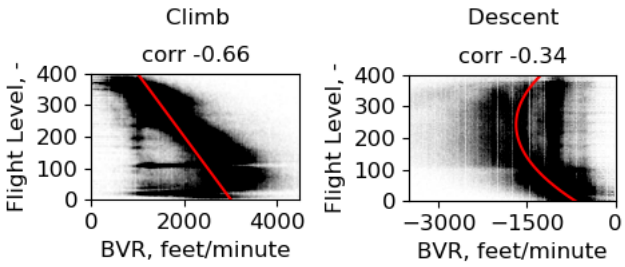


Fig. 5 Relationship between the Barometric Vertical Rate (BVR) and Flight level in the climb and descent phase of B738.

To characterize the vertical speed variation with altitude, the distributions were obtained for two altitude ranges: from 0 to 10,000 feet (FL100) and from 10,000 feet (FL100) to the maximum altitude. Figure 6 presents the results for the climb and descent phases of B738.

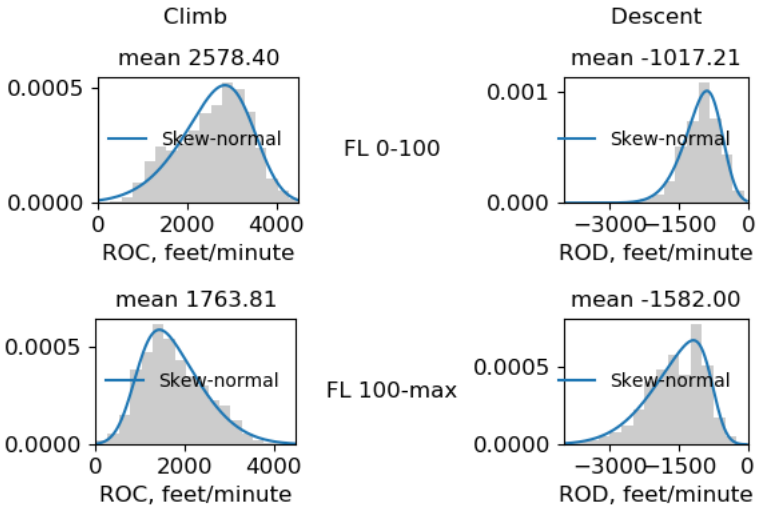


Fig. 6 Vertical speed distribution in the climb and descent phase of B738 for altitude ranges FL0 to FL100 and FL100 to the cruising altitude.

For medium WTC aircraft, the average rate of climb is about 2,500 feet/minute below FL100 and 1,700 feet/minute above FL100, and the average rate of descent is about 1,000 feet/minute below FL100 and 1,600 above FL100.

4. Temporary Level-off

Temporary level-offs have two characteristics: the flight level at which they occur and their duration. The former is a discrete variable, which can be represented by a PMF, and no distribution estimation was performed for it. Figure 7 shows the distributions of the flight levels where temporary level-offs occur and their duration above FL100 in the climb

and descent phases.

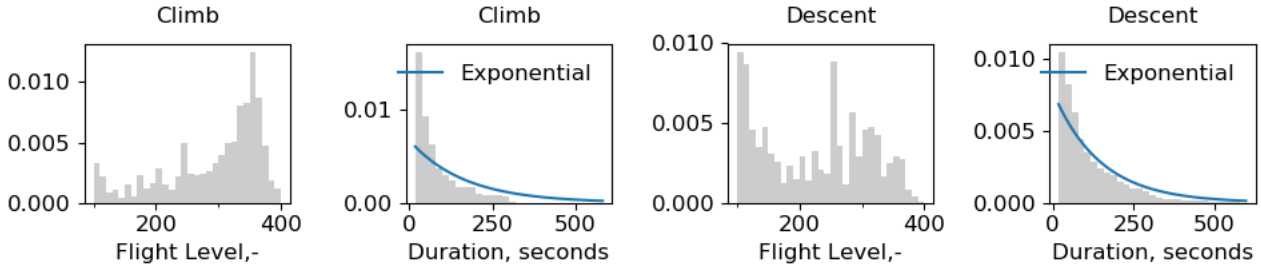


Fig. 7 Distribution of the flight levels and duration of the temporary level-offs above FL100 in the climb and descent phase of medium WTC aircraft.

The obtained distributions are very dependent on the airspace (configuration), implemented procedures and ATC. For instance, in the climb and descent phases, there are peaks at FL240 and FL260 respectively, which are due to aircraft transfer to and from MUAC airspace. Similarly, other temporary level-offs occur as a result of internal procedures to transfer aircraft from a sector to a sector, aircraft waiting for ATC clearances, and ATC conflict resolutions.

5. Air Temperature

To describe the variation of the air temperature, the distributions of two parameters were obtained: the air temperature near the ground and the lapse rate. Figure 8 shows the distributions of the air temperature near the ground (FL50) and the lapse rate below the tropopause in summer.

(a) Air temperature at FL50 in summer **(b) Lapse rate in summer**

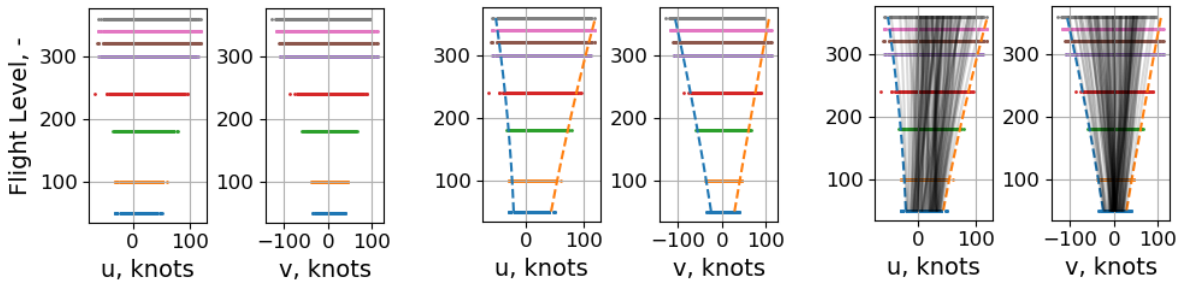
Fig. 8 Air temperature distribution near the ground (FL50) and lapse rate below the tropopause in summer. Red semicircle indicates the standard value (i.e. according to ICAO ISA).

In this figure, red semicircles indicate the air temperature at FL50 and the lapse rate according to ICAO International Standard Atmosphere (ISA) [30]. It can be seen that although the standard (ISA) lapse rate is a good approximation, the ISA air temperature at FL50 deviates substantially from the mean value of the obtained distribution. By taking the difference between the air temperature distribution and ISA temperature, the distribution of the temperature offset from ISA can be obtained.

6. Wind Conditions

When estimating the wind distributions, it is convenient to first split wind into its North-South, v, and West-East, u, components. This allows to avoid analyzing the wind direction as angular data (e.g. 0 and 360 degrees is the same) and to incorporate both wind speed and direction in one variable. Furthermore, to parametrize the distribution of the wind components with altitude, a wind component vertical profile was introduced as a function of a FL.

The estimation of this vertical profile is a three-step process as illustrated in Fig. 9. First, the mean and standard deviation of the wind components were estimated from data for available FLs (Fig. 9a). Next, the lower (i.e., left) and upper (i.e., right) bounds of the vertical profiles were estimated as a mean ±3 standard deviations (Fig. 9b). These margins were subsequently fitted to a function $f(x) = ax^2 + bx + c$, where x is a FL, which allowed to estimate the minimum and maximum of a, b and c coefficients (i.e., for lower and upper bounds).



(a) Scatter **(b) Bounds** **(c) Sampling**

Fig. 9 The process of wind components' vertical profile estimation and sampling.

Then, assuming that a, b, and c coefficients are uniformly distributed*, they were randomly sampled allowing to estimate the distribution of the wind components' vertical profiles (Fig. 9c).

7. ATC Instructions

The distributions of four ATC instruction types, namely 'direct', heading, speed and vertical speed, were estimated. In 2016, on average, 78.6% of the flights in MUAC airspace received 1 or more of these instructions. A 'direct' instruction was given to 76.4% of the flights, a heading instruction to 19.1%, a speed instruction to 6.5%, and a vertical speed instruction to 2.2% of the flights. Note, that these frequencies indicate the average number of flights that received each instruction and are not meant to add up to 100%, since the same flight can potentially receive all four types of instructions.

Two types of 'direct' instructions can be given: direct-to and direct-via. In MUAC airspace, direct-to is given 80% of the time. There are four types of heading instructions: present heading, heading (value), turn left and turn right. 86% of the flights receive a 'maintain present heading' instruction, whereas a specific heading value is given to about 12% of

*The uniform distribution resulted in the best coincidence between the wind speed and direction distributions obtained from the input data (i.e., weather forecasts) and the resultant distributions obtained from the wind components' vertical profile approximation.

the flights. An instruction to turn left or right is given in about 2% of the cases. In MUAC airspace, the direct-to and maintain present heading instructions are commonly used for conflict resolution purposes.

There are four types of speed instructions: assigned, maximum, minimum and present speed. An instruction to fly at the assigned speed is given to almost 50% of flights, whereas an instruction to keep the present speed is almost never given. A maximum and minimum speed to fly at is given in about 26% and 24% of the time. A speed instruction can be given as a Mach number or IAS. Mach number instructions are given nearly 70% of the time. The high percentage of speed instructions given as a Mach number makes sense for MUAC airspace, as the upper airspace encompasses the en-route part of the flights, where aircraft fly at a constant Mach number. The PMF of the Mach number and IAS instructions are presented in Fig. 10 and Fig. 11 respectively.

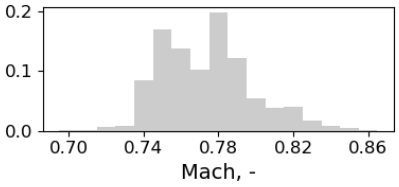


Fig. 10 PMF of Mach number speed instruction.

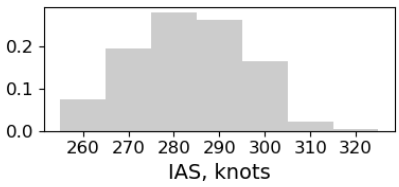


Fig. 11 PMF of IAS instruction.

A vertical speed (VS) instruction is an instruction to climb or to descend at a specific rate to a cleared flight level. An instruction to descend at a given rate is given in slightly over 80% of the cases. Together with the VS value, an instruction can include an indication of the speed limit: minimum or maximum. If such a limit is given, almost always, it is a minimum value. The PMF of the rate of climb and the rate of descent given as the VS instruction is presented in Fig. 12 and Fig. 13, respectively. The most frequent rate of climb is 1,000 feet/minute, whereas the most frequent rate of descent is 1,500 feet/minute.

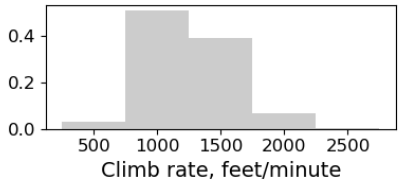


Fig. 12 PMF of climb rate as VS instruction.

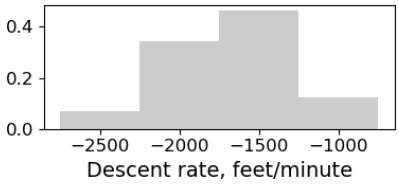


Fig. 13 PMF of descent rate as VS instruction.

B. Input Sampling

All aircraft-related parameters (bank angle, speed profile and vertical speed) are sampled based on the aircraft type, whereas the rest of input parameters are assumed to be independent, meaning that each input is sampled from its distribution function independently from the values of other inputs. When sampling the IAS/Mach number speed

profile, both speed values are sampled independently, since the correlation between them was found not to be strong (see Fig. 4). An independent speed sampling may, however, result in unrealistic speed profiles. For the triple seven aircraft (i.e., B77W, B772, B77L) the down-linked (in Mode-S) data did not contain the roll angle information and, thus, the distributions obtained for B788 were used as a substitute.

Although the input distribution functions were obtained from observed data, when sampling from the estimated distribution functions, values larger/smaller than the maximum/minimum observed values can be sampled. Therefore, constraints were imposed as indicated in Table 3 on some inputs to ensure sensible values.

Table 3 Input parameters' limits

Parameter	Minimum	Maximum	Units
Bank angle	5	35	degrees
Mach number	Minimum speed	Maximum operating Mach number	-
CAS	Minimum speed	Maximum operating speed	knots
Rate of climb	1,000	4,000	feet/minute
Rate of descent	1,000	3,000	feet/minute
Temporary level-off	20	500	seconds

To facilitate the analysis of the Monte Carlo simulations, each input was sampled only once per flight phase as indicated by the + symbol in Table 4. An exception is the vertical speed, which was sampled twice: for the altitude range from 0 to 10,000 feet and from 10,000 feet to the cruise altitude. In this study, three flight phases are distinguished: climb, cruise and descent. Whereas for some inputs a single value might be a reasonable assumption (e.g. a lapse rate or a constant CAS/Mach number speed profile), the majority of the parameters can vary throughout a phase. The air temperature was modeled using a temperature offset from the standard temperature (i.e. temperature according to ICAO ISA). Values of the air temperature offset and the lapse rate were the same in all flight phases.

Table 4 Inputs' sampling per flight phase

Parameter	Climb	Cruise	Descent
Air temperature offset	+	+	+
Lapse rate	+	+	+
Wind speed error	+	+	+
Bank angle	+	+	+
Mach number	+	+	+
CAS	+	-	+
Vertical speed	+	-	+
Temporary level-off	+	-	+
ATC instruction	-	+	-

The wind conditions were modeled as follows. First, a wind forecast grid was constructed as 20 predefined locations

with a cell size of 2.5 by 2.5 degrees (which is the weather grid used at MUAC (see Fig.1)). For each location, u and v wind component vertical profiles were sampled from the estimated distributions. Then, for each FL, u and v wind components were linearly interpolated between the centers of the grid cells. Finally, the wind speed and direction were computed from the wind components. All flights were simulated using the same wind forecast. Although, in reality, u and v wind components are correlated, in this study, their correlation was not taken into account. Figure 14 illustrates the relationship between the wind u and v component in summer and in July. The correlation is stronger when a single month is considered instead of a season. Though it is observable that, for instance, when the u component increases, the v component somewhat decreases, the relationship between them is clearly not linear or quadratic. An independent sampling of the wind components, however, might result in the wind conditions that are unrealistic for the studied geographical area. Nevertheless, for this analysis, we are interested in generalizable findings, for which uncorrelated u and v components are more suitable. To ensure sensible values, the distributions estimated for July were used when sampling wind components.

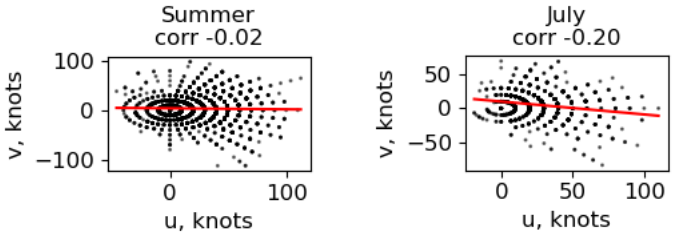


Fig. 14 Relationship between wind u and v components in summer (left) and July (right).

Furthermore, a wind speed error was sampled for each flight (i.e., the error between the predicted wind speed, and the actual wind speed). According to the research conducted for flights above Europe [31], the wind speed error can be considered as normally distributed with a mean of 1.5 knots and standard deviation of 9.2 knots. Although the wind speed error is strongly correlated with the wind speed magnitude, season and wind forecast look-ahead time, these were not considered in this study, and a single wind speed error was sampled from the normal distribution $N(1.5, 9.2)$ for each flight.

Based on the obtained statistics on the ATC instructions, the proportion of flights that received an instruction was as follows: 59.5% received a direct-to instruction; 14.2% had a heading instruction; 4.3% received a speed instruction; 0.9% received a vertical speed instruction. 21.1% of flights did not receive any ATC instruction. Furthermore, the duration of such ATC instructions as heading and speed were sampled from a uniform distribution with a minimum of 30 seconds and maximum of 180 seconds. For each WTC, aircraft types were sampled proportionally to the frequency observed in the surveillance data. Finally, a single route from Paris to Hamburg was modelled for all flights.

C. Trajectory Simulator/Predictor

Since both speed settings and vertical speed are sampled from the distribution functions, a trajectory simulator/predictor was developed employing a kinematic approach (similar to many operational TPs). The equations of motion used for the flight profile computation are given in Eq. (6) to Eq. (9). The earth was assumed to be flat, non-rotating and used as an inertial reference frame. A Cartesian *East, North, Up* coordinate system was used where the origin of the system coincides with the first point of a trajectory.

$$\dot{x} = V_{GS} \sin \psi_g \tag{6}$$

$$\dot{y} = V_{GS} \cos \psi_g \tag{7}$$

$$\dot{h} = VS \tag{8}$$

$$\dot{\psi_g} = \frac{V_{GS}}{r} \tag{9}$$

Here, x and y are the horizontal positions in Eastern and Northern direction respectively, V_{GS} is the ground speed, ψ_g is the track angle, h is the altitude, VS is the vertical speed, ψ_g is the track angle, and r is the turn radius.

The ground speed was obtained as follows:

$$V_{GS} = \sqrt{V_{TAS}^2 + V_w^2 - 2 \cdot V_{TAS} V_w \cos(w - \psi_a)} \tag{10}$$

Where V_{TAS} is the true airspeed, V_w is the wind speed, w is the wind direction, and ψ_a is the aircraft heading.

The turn radius was computed as:

$$r = \frac{V_{GS}^2}{g \cdot \tan \phi} \tag{11}$$

Where g is the gravitational acceleration, $g = 9.81 ms^{-2}$, and ϕ is the bank angle.

The true airspeed of an aircraft was converted from the CAS, V_{CAS}, (Eq. (12)) or the Mach number, M, (Eq. (13)).

$$V_{TAS} = \sqrt{\frac{2\gamma}{\gamma - 1} \frac{p}{\rho} \left[\left(1 + \frac{p_0}{p} \left[\left(1 + \frac{\gamma - 1}{2\gamma} \frac{\rho_0}{p_0} V_{CAS}^2 \right)^{\frac{\gamma}{\gamma - 1}} - 1 \right] \right)^{\frac{\gamma - 1}{\gamma}} - 1 \right]} \tag{12}$$

Here, γ is the adiabatic index of air, $\gamma = 1.4$, p is the air pressure, ρ is the air density, p_0 and ρ_0 are the air pressure and density, respectively, at sea level.

$$V_{TAS} = M\sqrt{\gamma \cdot R \cdot T} \tag{13}$$

Here R is the specific gas constant, $R = 287.05287 m^2 K^{-1} s^{-2}$, and T is the air temperature.

The wind speed and wind direction were computed from the u and v wind components:

$$V_w = \sqrt{u^2 + v^2} \tag{14}$$

$$w = atan2(-u, -v) \tag{15}$$

Here u and v are obtained as a function of a flight level, FL:

$$u = a_u FL^2 + b_u FL + c_u \tag{16}$$

$$v = a_v FL^2 + b_v FL + c_v \tag{17}$$

Where a_u, b_u, c_u and a_v, b_v, c_v are the coefficients for the estimation of the u and v wind components, respectively.

The air temperature offset, ΔT, as well as the lapse rate, λ, were used for the temperature, T, and pressure, p, calculations below the tropopause as follows:

$$T = T_{ISA} + \Delta T + \lambda h \tag{18}$$

$$p = p_0 \left(\frac{T - \Delta T}{T_{ISA}} \right)^{\frac{-g}{\lambda R}} \tag{19}$$

Here T_{ISA} is the ISA air temperature at altitude h.

The only difference between the trajectory simulator and the predictor is the vertical speed estimation. For simulation purposes, the vertical speed was sampled from the estimated distributions, whereas for predictions it was defined as a function of the altitude. Based on the identified relationship between the vertical speed and the flight level, FL, (see Fig.5), the rate of climb, RoC, can be calculated using a linear equation (Eq. (20)), whereas the rate of descent, RoD, can be obtained using a quadratic equation (Eq. (21)). For each aircraft type the coefficients a, b, c were obtained using the least squares polynomial fit. By describing the vertical speed as a function of the altitude, a continuous average vertical speed profile is created.

$$RoC = aFL + b \tag{20}$$

$$RoD = aFL^2 + bFL + c \tag{21}$$

Since temporary level-offs and ATC instructions are *procedural* parameters, their execution was triggered by certain events. Provided that the vertical flight profile includes a temporary level-off in the climb and/or descent phase, an aircraft leveled upon reaching the given FL for a time interval sampled from the temporary level-off duration distribution. ATC instructions were only given in the cruise phase (one instruction per flight) as follows. A 'direct' instruction to a certain waypoint was given at a specific location of a route (i.e., all flights received the same direct-to instruction). A heading instruction (turn left/right 10, 15 or 20 degrees) was issued to an aircraft after the top of climb was reached for a time interval as sampled from its duration distribution. After the heading instruction was completed, an aircraft returned to its intended route. A speed instruction was given as a Mach number for a time interval as sampled from its duration distribution. An instruction to climb or to descend 2,000 feet (because of the hemispherical rule) was issued at around the middle of the cruise phase, which resembles an instruction for conflict resolution in the vertical plane.

D. Actual Trajectories and Predictions

After inputs were sampled from their distribution functions, N actual (i.e., ground truth) trajectories were simulated. These trajectories resemble the possible real flights. Then, for each actual trajectory, predictions were made up to 20 minutes look-ahead time utilizing a time interval moving window. For every minute of the trajectory flight time, the prediction was updated with the correct position and altitude, and trajectory prediction was reinitialized. A one minute update rate mimics the update of surveillance data. Although, surveillance data are usually updated at a rate of 4 to 12 seconds, a one minute update rate was considered to be sufficient for the purposes of this study. For both actual trajectories and predictions, the same trajectory simulator/predictor was used.

To investigate a potential added benefit from estimating input distributions from observed data, two experiments were conducted involving different predictor settings. In the first experiment, all aircraft-related parameters were set using the BADA aircraft performance model [29] and atmospheric parameters according to ISA. These settings resemble those commonly used in operational TPs. In the second experiment, the input parameters were set to the mean values of the distributions obtained from the observed data. These settings would resemble a TP where inputs are set according to the observations (e.g. in a particular ATC center). The same wind model was used for both experiments. Additionally, a third experiment was performed, where the down-linked (in Mode-S) aircraft speed settings were used in predictions. The vertical speed was not set to the down-linked value since this parameter is less stable compared to the speed settings (e.g. due to turbulence). Table 5 summarizes the settings for these three experiments. As often assumed, no temporary

level-offs and ATC instructions were taken into account in the experiments.

The objective of the different experiments is to explore whether the observation-based inputs to a TP (i.e., experiment II) would result in a (significant) reduction of prediction errors compared to the inputs according to BADA and ISA (i.e., experiment I). For experiment III, the goal is to investigate by how much prediction errors can be reduced, when down-linked aircraft data is (partially) used by a TP.

Table 5 Predictor settings in experiments

Parameter / Experiment	I	II	III	Units
Air temperature offset	ISA	Observed average	Observed average	kelvin
Lapse rate	ISA	Observed average	Observed average	kelvin/km
Wind	Forecast	Forecast	Forecast	-
Bank angle	BADA	Observed average	Observed average	degrees
Mach number	BADA	Observed average	Down-linked	-
CAS	BADA	Observed average	Down-linked	knots
Vertical speed	BADA	Observed average	Observed average	feet/minute
Temporary level-off	-	-	-	-
ATC instruction	-	-	-	-

A simple algorithm was used to reinitialize the predictions at an update. The starting position and altitude were set to those of the actual trajectory and the prediction was started. In case the starting position was off-route (e.g. due to a direct-to instruction), the prediction would be performed by rejoining the initial route at the next waypoint downstream.

E. Prediction Errors

For each *prediction*, three errors were computed for 5, 10, 15 and 20 minutes look-ahead time: along-track error (ATE), cross-track error (CTE) and altitude error (AE). The ATE and CTE were calculated as follows:

$$ATE = (x_p - x_t)\sin\psi_{g_p} + (y_p - y_t)\cos\psi_{g_p} \tag{22}$$

$$CTE = (x_p - x_t)\cos\psi_{g_p} - (y_p - y_t)\sin\psi_{g_p} \tag{23}$$

Here x_p and x_t are the predicted and the true horizontal positions in Eastern direction, y_p and y_t are the predicted and the true horizontal positions in Northern direction, and ψ_{g_p} is the predicted track angle.

The AE was computed as the difference between the predicted altitude, h_p, and the true altitude, h_t:

$$AE = h_p - h_t \tag{24}$$

Positive ATE and CTE indicate that the predicted position is ahead and on the right from the actual one, and a

positive AE indicates that the predicted altitude is above the actual one. These errors were subsequently used to assess the prediction uncertainty for a given look-ahead time and to determine the influence of individual inputs. The errors were analyzed for each flight phase separately, meaning, for instance, that an error at 20 minutes look-ahead time in the cruise phase was computed for a prediction initiated also in the cruise phase. This implies that the number of available data points reduces with look-ahead time and is determined by the duration of a respective flight phase.

F. Number of Runs

To determine how many trajectories should be simulated, the number of samples, N, that is required to ensure a small relative error of Monte Carlo sampling, was calculated. When samples are randomly drawn from a distribution function, this approximates the evaluation of an integral [32, 33]. The relative error, e_{appr}, of this approximation can be estimated as:

$$e_{appr} = \frac{1}{\sqrt{N}} \tag{25}$$

In this study, the relative error of 0.01 was assumed. Rearranging Eq. (25) to find N results in $N = 10,000$ samples or simulation runs per each WTC aircraft.

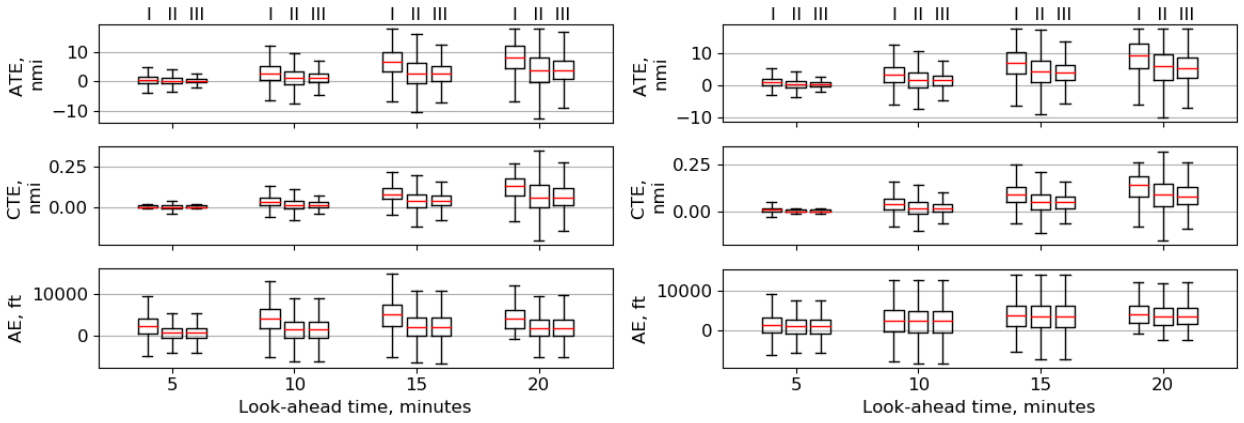
VI. Results

Two types of results were obtained in this study, as a function of a flight phase, for 5, 10, 15 and 20 minutes look-ahead time: the prediction uncertainty (error distribution) and the correlation between inputs and prediction errors.

A. Prediction Uncertainty

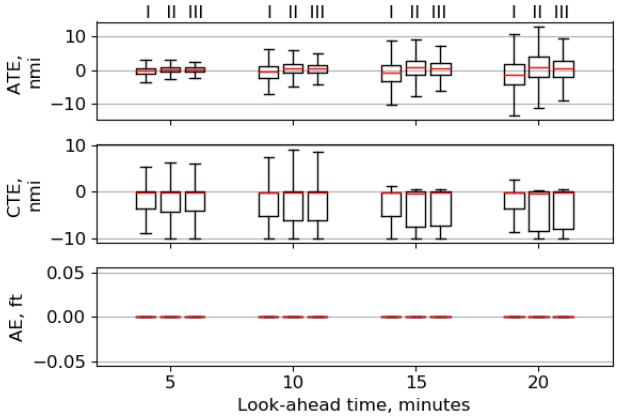
To assess prediction uncertainty, error distributions are presented using box plots. For analysis purposes, the box plots were studied without outliers, which comprised, on average, 1.85% of data. Figure 15 presents the results for heavy and medium aircraft. These box plots give along-track error (ATE), cross-track error (CTE), and altitude error (AE) as a function of flight phase and look-ahead time. For each look-ahead time, a group of three box plots is presented, indicating different predictor experiment settings (I to III from left to right) as explained in Subsection V.D. The ATE and CTE are in nautical miles (nmi) and the AE is in feet (ft).

In general, the magnitudes and the trends of prediction errors are comparable for heavy and medium aircraft. As expected, the ATE and CTE increase with look-ahead time. In the climb phase, the majority of the ATEs is positive, which indicates that the predicted positions are often ahead of the actual ones. This can be partly explained by the high(er) speed settings of the predictor, but is largely due to the effect of the temporary level-offs and vertical speed settings (e.g. due to altitude errors). In experiment III, for instance, the predictor climb CAS and Mach number are set to the down-linked values and are not the error source. When aircraft temporarily level-off their true airspeed does not

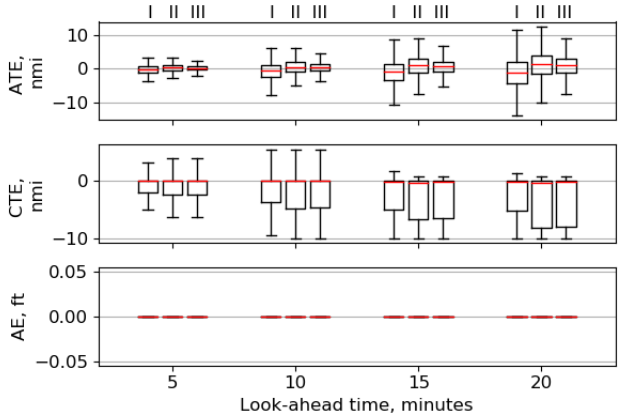


(a) Heavy aircraft. Climb

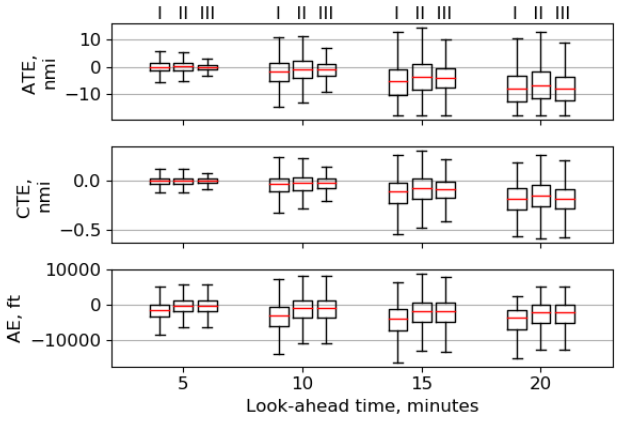
(b) Medium aircraft. Climb



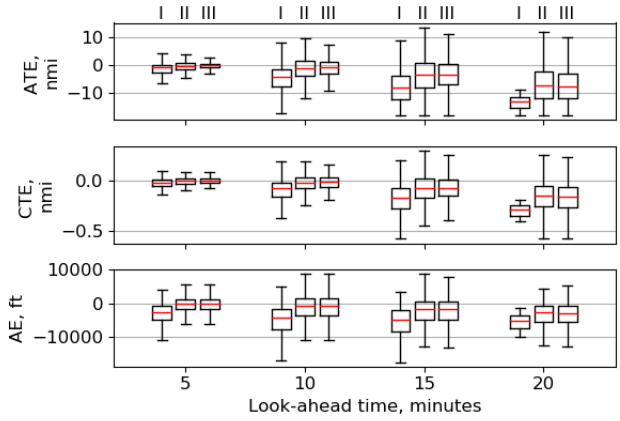
(c) Heavy aircraft. Cruise



(d) Medium aircraft. Cruise



(e) Heavy aircraft. Descent



(f) Medium aircraft. Descent

Fig. 15 Prediction uncertainty with look-ahead time for heavy and medium WTC aircraft in climb, cruise and descent.

increase, in contrast to a predicted continuous climb, where the true airspeed increases with altitude. This is consistent with the observed AEs, which are mostly positive, meaning the predicted altitude is above the actual one. The difference

in the true airspeed results in predictions being ahead of the actual flights. Furthermore, in experiment I, when BADA values and ISA conditions are used, the median of the ATE is shifted further away from 0 (compared to experiments II and III). This indicates that the errors in the atmospheric temperature and average speeds additionally contribute to the ATE.

In the descent phase, the majority of the ATEs is negative (i.e., predicted positions are behind the actual ones), which is mostly the effect of AEs. When a TP has no information on temporary level-offs, it will predict a continuous descent, which would result in predicted altitude below the actual one (i.e., negative AEs). Additionally, the errors in average speeds and atmospheric temperature, as in experiment I, further increase the frequency of large ATEs.

The CTEs are negligible in the climb and descent phases; however, in the cruise phase errors can reach up to 10 nmi. This difference is logical since the deviations from the planned route occur only in the cruise phase. The CTE is predominantly negative, which is due to the direct-to instruction that stretches on the right-hand side from the planned route and results in a maximum lateral deviation of 10 nmi.

In the climb and descent phase, the AE increases with look-ahead time and has the largest spread at 15 minutes. The error is mainly positive (i.e., predicted altitude is above the actual one) in the climb and mainly negative (i.e., predicted altitude is below the actual one) in the descent phase, which can be explained by the temporary level-offs. Assuming an average vertical speed of 2,000 feet per minute, a 5 minute level-off would result in 10,000 feet altitude difference. Additionally, a non-representative average vertical speed used for predictions (i.e., experiment I in the climb phase of heavy aircraft and in the descent phase of both heavy and medium aircraft) shifts the median further from 0 (compared to the results of experiments II and III). In the cruise phase, the effect of ATC instructions to climb or descend is negligible (due to a low frequency) and results in virtually no AEs.

Table 6 summarizes the results on the maximum (absolute) prediction errors at 20 minutes look-ahead time in experiment II. The large ATEs of 18 nmi are infrequent and are induced by the AEs, which largely result from the temporary level-offs. As an example, consider an aircraft in climb which has levelled at FL100 for 5 minutes, whereas a TP predicts a continuous climb. Furthermore, the aircraft is following 290 knots CAS / 0.78 Mach speed profile, whereas the predictor is using 305 knots CAS / 0.8 Mach. Both have an average vertical speed of 1,500 feet per minute. In 5 minutes (i.e., at the end of the temporary level-off), the ATE is about 3.1 nmi and grows to 15.2 nmi in 20 minutes. Assuming an average wind speed error of 10 knots over a 20 minutes period, the ATE can additionally grow by 3.3 nmi, amounting to 18.5 nmi ATE at the end of 20 minutes look-ahead time prediction. Without the temporary level-off, the error at 20 minutes would be 8.8 nmi.

For illustration purposes, Tables 7 and 8 indicate the Root Mean Square Error (RMSE), the mean and the standard deviation (Std) of the errors at 5 and 20 minutes look-ahead time per flight phase for each predictor experiment setup (i.e., I to III). The ATE and CTE are in nautical miles (nmi) and the AE is in feet (ft). The RMSE, mean and standard deviation of the ATE mostly decrease when the average values estimated from the observations are used for predictions

Table 6 Maximum absolute prediction errors at 20 minutes look-ahead time (when outliers are not considered) for heavy and medium WTC aircraft in experiment II

Flight phase	WTC	ATE, nmi	CTE, nmi	AE, ft
Climb	Heavy	18.0	0.4	9,545.4
	Medium	18.0	0.3	11,760.3
Cruise	Heavy	13.1	10.0	0.0
	Medium	12.5	10.0	0.0
Descent	Heavy	18.0	0.6	12,853.2
	Medium	17.9	0.6	12,435.8

(i.e., experiment II), compared to when BADA values and ISA are used (i.e., experiment I). When the down-linked CAS and Mach number are used by a TP in experiment III, the errors decrease even further. For instance, in the climb phase of heavy aircraft at 5 minutes look-ahead time, the RMSE, mean and standard deviation of the ATE are 1.70, 0.68 and 1.56 nmi in experiment I, 1.38, 0.28 and 1.35 nmi in experiment II and 0.88, 0.25 and 0.84 nmi in experiment III.

Table 7 RMSE, mean and standard deviation (std) of prediction errors in experiments I, II and III at 5 and 20 minutes look-ahead time for heavy aircraft. ATE and CTE are in nautical miles and AE is in feet

Heavy WTC	Climb phase																	
	5 minutes look-ahead time									20 minutes look-ahead time								
	Experiment I			Experiment II			Experiment III			Experiment I			Experiment II			Experiment III		
Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
RMSE	1.70	0.01	3342.12	1.38	0.01	1847.69	0.88	0.01	1847.69	9.68	0.14	4956.87	7.06	0.11	3264.61	6.20	0.10	3306.92
Mean	0.68	0.00	2281.08	0.28	0.00	655.04	0.25	0.00	655.08	8.11	0.12	4169.14	4.07	0.06	1929.25	4.07	0.06	1972.83
Std	1.56	0.01	2442.63	1.35	0.01	1727.68	0.84	0.01	1727.67	5.28	0.08	2681.20	5.77	0.09	2633.57	4.68	0.07	2653.99
Heavy WTC	Cruise phase																	
	5 minutes look-ahead time									20 minutes look-ahead time								
	Experiment I			Experiment II			Experiment III			Experiment I			Experiment II			Experiment III		
Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
RMSE	1.16	3.13	0.00	1.13	3.94	0.00	0.91	3.92	0.00	4.43	2.81	0.00	4.35	5.47	0.00	3.21	5.25	0.00
Mean	0.07	-1.63	0.00	0.34	-2.17	0.00	0.30	-2.16	0.00	-1.13	-1.38	0.00	0.94	-3.48	0.00	0.37	-3.27	0.00
Std	1.16	2.67	0.00	1.08	3.29	0.00	0.86	3.28	0.00	4.29	2.44	0.00	4.25	4.22	0.00	3.19	4.10	0.00
Heavy WTC	Descent phase																	
	5 minutes look-ahead time									20 minutes look-ahead time								
	Experiment I			Experiment II			Experiment III			Experiment I			Experiment II			Experiment III		
Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
RMSE	2.07	0.04	2945.20	1.99	0.04	2228.78	1.11	0.03	2229.06	9.94	0.22	5634.37	9.20	0.21	4345.50	9.64	0.22	4252.37
Mean	0.02	0.00	-1626.37	-0.01	0.00	-313.62	-0.03	0.00	-313.90	-7.89	-0.18	-4402.60	-6.60	-0.15	-2780.45	-7.83	-0.18	-2696.65
Std	2.07	0.04	2455.44	1.99	0.04	2206.61	1.11	0.03	2206.85	6.05	0.14	3516.14	6.42	0.14	3339.53	5.63	0.13	3287.96

Experiment I, however, resulted in smaller RMSE, mean and standard deviation of the CTE in the cruise phase, compared to experiments II and III. When the average vertical speed profile estimated from the observations is used for predictions (i.e., experiments II and III), the RMSE, mean and standard deviation of the AE decrease compared to experiment I when BADA is used. This is likely due to large differences between the estimated average vertical speed profile and the average vertical speed profile defined in BADA.

Table 8 RMSE, mean and standard deviation (std) of prediction errors in experiments I, II and III at 5 and 20 minutes look-ahead time for medium aircraft. ATE and CTE are in nautical miles and AE is in feet

Medium WTC	Climb phase																	
	5 minutes look-ahead time									20 minutes look-ahead time								
	Experiment I			Experiment II			Experiment III			Experiment I			Experiment II			Experiment III		
Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
RMSE	1.92	0.02	3067.86	1.50	0.01	2584.43	0.98	0.01	2584.45	10.55	0.15	4912.65	8.15	0.12	4579.14	7.36	0.11	4672.34
Mean	1.15	0.01	1323.47	0.46	0.00	894.48	0.37	0.00	894.69	9.23	0.13	4090.83	5.93	0.09	3616.20	5.68	0.08	3698.16
Std	1.53	0.01	2767.71	1.43	0.01	2424.70	0.91	0.01	2424.65	5.10	0.07	2720.16	5.59	0.08	2809.21	4.68	0.07	2855.59
Medium WTC	Cruise phase																	
	5 minutes look-ahead time									20 minutes look-ahead time								
	Experiment I			Experiment II			Experiment III			Experiment I			Experiment II			Experiment III		
Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
RMSE	1.25	1.30	0.00	1.12	1.68	0.00	0.83	1.71	0.00	4.66	4.29	0.00	4.24	5.58	0.00	3.02	5.50	0.00
Mean	-0.11	-0.50	0.00	0.33	-0.67	0.00	0.26	-0.69	0.00	-1.12	-2.45	0.00	1.25	-3.69	0.00	0.77	-3.61	0.00
Std	1.24	1.20	0.00	1.07	1.54	0.00	0.79	1.57	0.00	4.53	3.52	0.00	4.05	4.18	0.00	2.92	4.15	0.00
Medium WTC	Descent phase																	
	5 minutes look-ahead time									20 minutes look-ahead time								
	Experiment I			Experiment II			Experiment III			Experiment I			Experiment II			Experiment III		
Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
RMSE	2.17	0.05	3821.64	1.61	0.04	1943.04	1.04	0.02	1943.60	13.65	0.31	6212.70	9.18	0.21	4408.78	9.30	0.21	4502.19
Mean	-0.98	-0.02	-2782.82	-0.13	0.00	-287.02	0.02	0.00	-288.02	-13.34	-0.30	-5633.69	-6.83	-0.15	-3088.23	-7.31	-0.16	-3230.84
Std	1.94	0.04	2619.33	1.61	0.03	1921.72	1.04	0.02	1922.14	2.88	0.07	2619.00	6.13	0.14	3146.45	5.74	0.13	3135.50

B. Correlation Between Inputs and Prediction Errors

To determine the relationship between the inputs and prediction errors, Pearson correlation coefficients were computed for each input-error combination (e.g. vertical speed versus AE) as a function of the look-ahead time and flight phase. For this, the average error was estimated per flight phase for each trajectory. Tables 9 and 10 present the results obtained for errors computed in experiment II for heavy and medium WTC aircraft. LVL stands for temporary level-off. Since the average error can be positive or negative, the sign of the correlation (i.e., positive or negative) is somewhat ambiguous. For instance, there is a strong negative correlation between the speed settings (CAS and Mach number) and ATE. This means that when speed increases, the ATE becomes (more) negative (i.e., the predicted position is behind the actual one). Therefore, the sign of the correlation should be treated with caution.

In this study, the correlation coefficients whose absolute value is equal to or greater than 0.23 (marked in bold) were considered as those that show a linear relationship between the input and the error. The value of 0.23 was chosen based on the minimum value of the error variance considered to be of interest in this study, which is 5%. The error variance, i.e., coefficient of determination, can be computed as a squared Pearson correlation coefficient, i.e., $0.23^2 = 0.0529 \approx 0.05$. Therefore, only inputs that contribute to at least 5% of the variance of the prediction errors are of interest.

The ATE is correlated with the wind (error), CAS and Mach number speed settings, vertical speed, and the duration of temporary level-offs. The relationship between the ATE and the speed settings and the ATE and the wind speed error is logical. With increase in the speed settings, the true airspeed increases, resulting in prediction being behind the actual position (i.e., a negative correlation). The wind (speed error) affects the ground speed, which increases with a tail-wind and decreases with a head-wind. However, when the wind is not strictly tail- or head-wind, but its direction is varying, the impact of the wind speed error is less pronounced. In the climb and descent phases, the ATE is correlated with the

Table 9 Input-error correlation for heavy WTC aircraft

Heavy WTC	5 minutes look-ahead time									10 minutes look-ahead time								
	Climb phase			Cruise phase			Descent phase			Climb phase			Cruise phase			Descent phase		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	-0.19	-0.03	-0.01	-0.18	0.01	-0.02	-0.13	-0.03	0.02	-0.16	-0.11	-0.01	-0.18	0.01	-0.02	-0.10	-0.08	0.02
Lapse rate	-0.09	-0.01	-0.01	-0.14	0.00	0.00	-0.06	-0.03	0.00	-0.07	-0.05	-0.01	-0.14	0.02	0.00	-0.05	-0.04	0.00
Wind error	**0.40**	0.01	0.00	**0.54**	0.00	0.01	**0.43**	0.15	-0.01	**0.35**	0.19	0.00	**0.57**	-0.01	0.02	**0.37**	**0.32**	-0.02
Bank angle	0.00	0.00	0.00	0.01	0.03	0.01	0.01	0.01	0.00	-0.01	0.00	0.00	0.01	0.02	0.00	0.01	0.01	0.00
ATC	-0.02	0.02	0.00	0.04	0.08	-0.05	0.00	0.02	0.00	-0.02	-0.01	0.00	0.04	0.08	-0.07	0.00	0.00	0.00
Mach number	-0.01	-0.04	0.03	**-0.50**	-0.05	-0.01	0.00	0.01	0.10	0.01	0.00	0.02	**-0.51**	-0.04	-0.01	0.03	0.02	0.10
CAS	**-0.60**	-0.08	0.00	0.00	0.00	0.00	**-0.78**	**-0.25**	0.00	**-0.53**	**-0.44**	0.00	0.00	0.00	0.00	**-0.71**	**-0.60**	0.02
Vertical speed	**-0.31**	-0.03	**-0.76**	0.00	0.00	0.00	**0.28**	**0.24**	**0.83**	**-0.43**	-0.01	**-0.77**	0.00	0.00	0.00	**0.32**	**0.31**	**0.82**
LVL FL	0.04	0.02	0.15	0.00	0.00	0.00	-0.07	-0.12	-0.22	0.04	0.03	0.14	0.00	0.00	0.00	-0.13	-0.12	-0.20
LVL duration	0.08	0.01	**0.24**	0.00	0.00	0.00	-0.18	**-0.23**	**-0.40**	0.09	0.07	**0.24**	0.00	0.00	0.00	**-0.26**	**-0.26**	**-0.42**
Heavy WTC	15 minutes look-ahead time									20 minutes look-ahead time								
	Climb phase			Cruise phase			Descent phase			Climb phase			Cruise phase			Descent phase		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	-0.15	-0.15	-0.01	-0.19	0.02	-0.01	-0.10	-0.08	0.02	-0.17	-0.15	-0.03	-0.22	0.01	-0.02	-0.07	-0.08	0.05
Lapse rate	-0.06	-0.06	-0.02	-0.14	0.05	0.00	-0.04	-0.03	0.01	-0.07	-0.07	0.00	-0.16	0.09	0.00	-0.06	-0.05	-0.01
Wind error	**0.32**	**0.31**	-0.01	**0.58**	-0.02	0.02	**0.33**	**0.30**	-0.05	**0.29**	**0.28**	0.01	**0.56**	-0.03	0.01	**0.30**	**0.28**	-0.07
Bank angle	-0.02	-0.05	-0.01	0.00	0.00	-0.01	0.01	0.00	0.01	-0.04	-0.04	-0.07	-0.02	0.00	0.02	0.01	0.01	0.01
ATC	-0.01	-0.02	0.01	0.03	0.08	-0.07	0.01	0.01	0.01	-0.02	-0.02	0.01	0.04	0.07	-0.09	0.02	0.02	0.01
Mach number	0.00	-0.01	0.01	**-0.51**	-0.03	-0.01	0.06	0.04	0.12	0.05	-0.01	0.15	**-0.53**	0.01	-0.01	0.04	0.05	0.17
CAS	**-0.46**	**-0.50**	0.04	0.00	0.00	0.00	**-0.63**	**-0.59**	0.08	**-0.49**	**-0.50**	0.06	0.00	0.00	0.00	**-0.60**	**-0.58**	0.10
Vertical speed	**-0.48**	**-0.49**	**-0.63**	0.00	0.00	0.00	**0.49**	**0.25**	**0.76**	**-0.52**	**-0.53**	**-0.50**	0.00	0.00	0.00	**0.44**	0.07	**0.66**
LVL FL	0.01	0.01	0.13	0.00	0.00	0.00	-0.14	-0.13	-0.17	0.00	0.00	0.15	0.00	0.00	0.00	-0.10	-0.09	-0.07
LVL duration	0.08	0.08	**0.25**	0.00	0.00	0.00	**-0.24**	**-0.23**	**-0.39**	0.08	0.08	**0.30**	0.00	0.00	0.00	-0.15	-0.11	**-0.29**

Table 10 Input-error correlation for medium WTC aircraft

Medium WTC	5 minutes look-ahead time									10 minutes look-ahead time								
	Climb phase			Cruise phase			Descent phase			Climb phase			Cruise phase			Descent phase		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	-0.13	-0.01	0.02	-0.15	0.01	0.00	-0.14	-0.04	0.00	-0.11	-0.04	0.01	-0.16	0.01	-0.01	-0.11	-0.09	0.00
Lapse rate	-0.08	-0.01	0.00	-0.13	0.00	0.00	-0.07	0.00	0.00	-0.07	-0.05	0.00	-0.13	0.00	0.00	-0.06	-0.05	0.00
Wind error	**0.42**	0.04	0.00	**0.55**	0.00	0.01	**0.52**	0.13	0.01	**0.37**	0.18	-0.01	**0.57**	0.01	0.01	**0.43**	**0.32**	0.00
Bank angle	0.04	-0.01	0.01	0.03	0.02	-0.01	0.02	0.02	0.00	0.04	0.03	0.01	0.04	0.03	-0.02	0.02	0.02	0.00
ATC	-0.01	0.00	0.01	0.03	0.05	-0.04	0.00	0.02	-0.02	-0.01	0.00	0.01	0.02	0.05	-0.06	-0.01	0.00	-0.02
Mach number	-0.02	-0.02	0.04	**-0.60**	0.00	0.00	-0.06	0.00	-0.02	0.00	0.03	0.04	**-0.63**	0.01	-0.01	-0.05	-0.04	-0.02
CAS	**-0.72**	-0.11	-0.03	0.00	0.00	0.00	**-0.73**	-0.21	-0.02	**-0.64**	**-0.44**	-0.03	0.00	0.00	0.00	**-0.66**	**-0.52**	-0.02
Vertical speed	**-0.33**	-0.13	**-0.81**	0.00	0.00	0.00	**0.30**	**0.28**	**0.84**	**-0.47**	-0.13	**-0.81**	0.00	0.00	0.00	**0.47**	**0.41**	**0.85**
LVL FL	0.03	0.03	0.16	0.00	0.00	0.00	-0.07	-0.12	-0.22	0.03	0.05	0.16	0.00	0.00	0.00	-0.13	-0.12	-0.21
LVL duration	0.07	0.02	**0.25**	0.00	0.00	0.00	-0.17	-0.22	**-0.44**	0.09	0.12	**0.26**	0.00	0.00	0.00	**-0.26**	**-0.27**	**-0.44**
Medium WTC	15 minutes look-ahead time									20 minutes look-ahead time								
	Climb phase			Cruise phase			Descent phase			Climb phase			Cruise phase			Descent phase		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	-0.10	-0.10	0.01	-0.16	0.01	0.00	-0.09	-0.08	0.01	-0.11	-0.10	0.03	-0.16	0.02	0.00	-0.08	-0.08	0.01
Lapse rate	-0.06	-0.05	0.00	-0.13	-0.01	0.00	-0.05	-0.05	0.01	-0.07	-0.07	-0.02	-0.14	0.04	0.01	-0.04	-0.05	0.01
Wind error	**0.35**	**0.31**	-0.01	**0.58**	0.00	0.01	**0.36**	**0.32**	-0.02	**0.28**	**0.27**	-0.02	**0.58**	-0.01	0.01	**0.32**	**0.30**	-0.07
Bank angle	0.03	0.00	0.03	0.04	0.02	-0.03	0.01	0.00	0.00	0.00	-0.02	0.00	0.03	0.01	-0.02	0.03	0.02	0.02
ATC	0.00	0.00	0.01	0.02	0.04	-0.07	0.00	0.00	-0.01	-0.05	-0.06	0.00	0.01	0.04	-0.07	-0.01	0.00	-0.01
Mach number	0.00	-0.01	0.04	**-0.63**	0.01	-0.01	-0.06	-0.06	-0.02	-0.06	-0.04	-0.04	**-0.63**	0.03	-0.01	-0.04	-0.06	-0.01
CAS	**-0.62**	**-0.53**	-0.09	0.00	0.00	0.00	**-0.57**	**-0.53**	0.01	**-0.63**	**-0.55**	-0.11	0.00	0.00	0.00	**-0.52**	**-0.48**	0.07
Vertical speed	**-0.48**	**-0.47**	**-0.67**	0.00	0.00	0.00	**0.52**	**0.29**	**0.83**	**-0.55**	**-0.52**	**-0.55**	0.00	0.00	0.00	**0.42**	0.22	**0.76**
LVL FL	-0.07	-0.06	0.07	0.00	0.00	0.00	-0.17	-0.17	-0.20	-0.07	-0.05	0.06	0.00	0.00	0.00	-0.10	-0.09	-0.07
LVL duration	0.00	0.01	**0.23**	0.00	0.00	0.00	**-0.30**	**-0.29**	**-0.43**	-0.04	-0.04	**0.24**	0.00	0.00	0.00	-0.18	-0.15	**-0.29**

vertical speed. This correlation, however, is induced by the AE. A fast climbing aircraft reaches a higher altitude and a higher true airspeed sooner than a slow climbing aircraft, resulting in the former being ahead of the latter (i.e., a negative

correlation). In the descent phase, a fast descending aircraft reaches a lower altitude and a lower true airspeed faster than a slow descending aircraft, and travels a shorter distance (i.e., a positive correlation). The correlation between the ATE and the temporary level-offs (duration) has been detected only in the descent phase for larger look-ahead times. This relationship is indirect and results from the AEs. As an example, consider an aircraft leveling off while a TP is predicting a continuous descent. The true airspeed of the leveled flight stays constant, and the true airspeed of the predicted flight decreases with altitude. Thus, a shorter distance is travelled in prediction resulting in predicted position being behind the actual one (i.e., a negative correlation).

In the climb and descent phases, the CTE shows a correlation with the wind, speed settings, vertical speed, and the duration of temporary level-offs. These inputs do not directly contribute to the CTE. When there is a cross-wind, aircraft usually correct for the wind to stay on course and this has also been implemented in the predictor used in this study. The speed settings, vertical speed, and temporary level-offs influence the distance travelled along the predefined route (i.e., longitudinal direction), and the resultant ATE causes small CTEs as can be seen in Fig. 15.

The only input (in this study) that directly contributes to the CTE is ATC intent. Instructions such as direct-to and heading deviate aircraft from their route resulting in CTEs, which lead to ATEs. Unfortunately, in this study, the ATC instructions were modeled in a way that does not allow to determine their linear relationship with errors. As an example, Fig. 16 visualizes the relationship between the inputs and the average CTE in the cruise phase for 20 minutes look-ahead time. The data are plotted as a two-dimensional histogram, where the black color represents the highest frequency of data points. The red line shows the correlation between the input and the error, and the correlation coefficient value is indicated at the top of each plot. The ATC instruction *0* stands for no ATC instructions, *D* for the direct-to, *H* for heading instructions, *S* for speed and *VS* for vertical speed instructions. In each plot, the data points are separated into two clouds and its clear that direct-to instructions cause large CTEs.

Thus, due to the discrete nature of ATC instructions modeled in this study, the correlation coefficient cannot be determined for this input. An alternative approach would be to determine the correlation between the CTE and the magnitudes of each instruction separately (e.g. speed values of the speed instruction or degrees of a heading instruction).

The AE is correlated with the vertical speed and temporary level-offs. The correlation between the vertical speed and the AE is negative in the climb phase and positive in the descent phase. When the predicted climb rate is smaller than the actual one, the predicted altitude is below the actual altitude, whereas when the predicted descent rate is smaller than the actual one, the predicted altitude is above the actual altitude. When an aircraft temporary levels off in climb, opposed to a predicted continuous climb, the predicted altitude is above the actual one, and the AE gets bigger for longer level-offs (i.e., a positive correlation). In the descent phase, when an aircraft levels off, whereas a TP predicts a continuous descent, the predicted altitude is below the actual one, and the error increases with the duration of the level-off (i.e., a negative correlation).

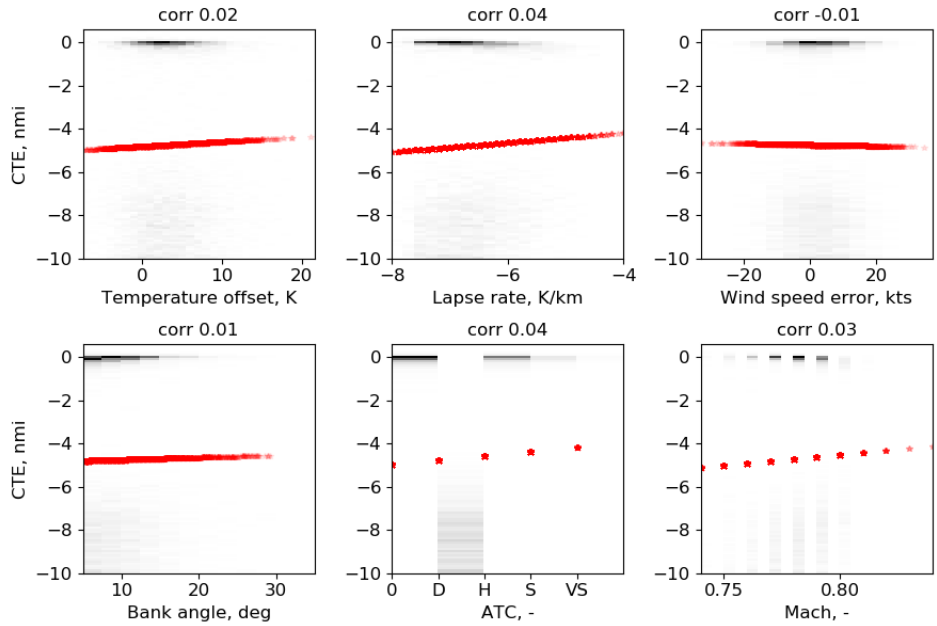


Fig. 16 Cross-track error (CTE) in the cruise phase for 20 minutes look-ahead time.

VII. Discussion

This section discusses the insights gained from the surveillance data analysis and the estimated distributions, and to what extent the obtained results can be generalized. Furthermore, the difficulty of determining the correlation between ATC intent and the prediction errors is addressed. Finally, the relation to results of previous research is summarized.

A. Insights from the Surveillance Data and Input Distributions

A variability of the TP inputs was demonstrated in this study with estimated distributions, which indicates that medium term predictions assuming a constant value setting for these inputs are oversimplifying. For example, the analysis of the surveillance data has shown that about 40% and 37% of flights in the climb and descent phase, respectively, did not follow a 'constant CAS - constant Mach' speed profile. Additionally, on average 20% and 27% of flights in the climb and descent phase, respectively, were not restricted to 250 knots IAS below FL100. On average, 32% of the flights temporarily level off during climb, and 64% temporarily level off in the descent phase. These insights, however, are representative for the operations within the studied area (as indicated in Fig.1). Furthermore, the distribution of vertical speed changes with altitude, which can be explained by the effect of the atmospheric conditions on engine performance, procedures, and practices. Lastly, the results obtained on the frequency of ATC instructions indicate the importance of including this information in the prediction process.

B. Generalization of the Results

The trajectory simulator/predictor used in this study is comparable with a typical TP except for how the vertical speed is obtained. Many operational TPs derive the vertical speed from the equations of a point-mass model as a function of the reference aircraft mass and thrust settings. This study employed an alternative approach (known as a parametric approach), which does not require the knowledge of the aircraft mass and thrust, where the vertical speed is estimated from the surveillance data (similar to ERAM's parametric trajectory predictor which is based on empirical data [34]). Furthermore, the choices of predictor settings are similar to those that are commonly made in both research and operational environments. For medium and heavy aircraft, the input distributions were obtained from a sufficient amount of surveillance data (i.e., 33,109 and 7,143 flights respectively). Generally, meteorological conditions (significantly) vary based on a geographical location (e.g. predominant wind direction). To generalize results, the u and v wind components were modelled as uncorrelated. In practice, the temporary level-offs are highly dependent on the procedures within each center, ATC, the airspace, and letters of agreement between the neighbouring centers. Although this information was not taken into account when sampling temporary level-offs, the obtained results are considered to be representative for the studied operational area, since the modelled route lies within the studied area. However, the magnitudes and frequency of the large prediction errors (due to temporary level-offs) might be different for the airspace that has a different distribution of the temporary level-offs. The results on the identified influential inputs of a TP are deemed to be generic regardless of a geographical location. The strength of the error-input correlation, however, may be affected by the predictor settings.

C. ATC Intent

In this study, when evaluating the impact of ATC on the prediction errors, ATC intent was taken as a discrete variable, which does not allow to obtain a meaningful correlation measure. An alternative approach would be to consider each ATC instruction separately and estimate the correlation between, for instance, the duration of a speed instruction or the heading magnitude of a vectoring instruction and prediction errors. This approach, however, is not applicable to direct-to instructions (i.e., instructions to fly to a certain waypoint of the route), which are commonplace in MUAC airspace. Nonetheless, it is logical to conclude that when vectoring (direct-to and heading) instructions are issued by ATC, these would (often) result in large CTEs, which in turn would contribute to ATEs. ATC speed instructions would, most likely, have a much smaller effect on ATEs compared to the wind.

D. Relation to Results of Previous Research

The results obtained in this study on the importance of TP inputs are similar to the conclusions made in previous research. In their study, Mueller et al. (2004) identified the speed settings and the wind speed as inputs that have the most significant impact on the trajectory prediction of the En Route Descent Advisor system [12]. The top of descent

location, speed settings, wind speed, temporary level-offs and vectoring instructions were shown to have a large impact on the prediction errors in the descent phase in the study performed by Mondoloni et al. (2005) [13]. Torres (2010) has identified that the velocity effects (most likely wind and speed settings) in the cruise phase and vertical modelling effects in the climb and descent phases are the largest sources of prediction errors [16]. In this study, barring the effect of ATC intent, the wind conditions, the CAS and Mach number speed settings, the vertical speed settings (usually as a function of the aircraft mass and thrust settings), and the temporary level-offs were determined to be the most influential inputs. The rest of the studied inputs did not show significantly strong correlation with the prediction errors. However, the global sensitivity analysis has shown that there are complex relationships between different prediction errors and inputs.

Whereas much research has been conducted studying the sensitivity of trajectory prediction to errors in inputs to a TP [12, 13, 19, 20], this study has demonstrated, using observed data, that the inputs are rarely normally or uniformly distributed. Furthermore, ATC instructions, which are recognized as large contributors to lateral deviations [15], and temporary level-offs, both estimated from operational data, were also modelled in this study. Finally, to assess the relationship (i.e., correlation) between the inputs to a TP and prediction errors, simulations were performed varying all inputs at the same time. Although this approach complicates the interpretation of each input individual contribution to resultant uncertainty, it allows to study the combined effect of all inputs while preserving the significance of influential inputs.

VIII. Conclusion

This paper estimated the prediction uncertainty up to 20 minutes look-ahead time and the correlation between trajectory predictor (TP) inputs and prediction errors for heavy and medium weight aircraft. For each category, 10,000 Monte Carlo simulation runs were performed employing a TP based on a kinematic approach where inputs were sampled from their distribution functions obtained from the observed data. Studied inputs included bank angle settings, speed settings (CAS and Mach number), vertical speed settings, temporary level-offs in the climb and descent phases, air temperature, lapse rate, wind and ATC intent. Three prediction errors, namely along-track error (ATE), cross-track error (CTE), and altitude error (AE), were computed for 5, 10, 15 and 20 minutes look-ahead time.

All three errors tend to increase with increasing look-ahead time. When outliers are not considered, at 20 minutes look-ahead time, the ATE can reach up to around 18 nautical miles in the climb and descent phase and around 13 nautical miles in cruise. CTEs are negligible in the climb and descent phase, however, could reach up to 10 nautical miles in the cruise phase. Furthermore, the AE can become as large as 12,000 and 13,000 feet in the climb and descent phase. These large prediction errors are likely to impact the performance of the decision support tools that operate in the time interval up to 20 minutes look-ahead time, such as conflict detection tools. A similar conclusion has been made by Mueller et at. (2004) [12] and Paglione et al. (2009) [15].

Out of all studied input parameters, wind conditions, speed settings, vertical speed and temporary level-offs showed

a correlation with the prediction errors. In most cases, using the average inputs estimated from the operational data, resulted in the reduction of prediction errors, compared to when BADA performance model was used. The ATEs were further reduced when the down-linked speeds from the aircraft (in Mode-S) were included in the predictions.

Although the ATC instructions significantly contribute to the CTE and ATE, their correlation could not be determined under this study setup. A different approach to the modeling of ATC intent is required for a successful correlation estimation. Whereas for an individual trajectory prediction the accuracy of all inputs is important, when studying the prediction accuracy in a generalized form, the effect of many inputs is diminished by the parameters that are more influential. Therefore, the improvement of the less influential inputs is likely to be insufficient to improve the overall prediction accuracy.

Further research could be conducted where the prediction algorithm takes into account the barometric vertical rate down-linked by aircraft in Mode-S (in addition to down-linked speed settings). Additionally, such aircraft intent information as target (cruise) altitude and location of top of climb and top of descent could be also shared with a ground-based TP. All this is expected to decrease AEs and ATEs. Furthermore, the non-linear relationships between the inputs and the errors could be investigated as well.

References

[1] EUROCONTROL, *Eurocontrol seven-year forecast February 2017. Flight Movements and Service Units 2017-2023*, 2017.

[2] Federal Aviation Administration, *FAA Aerospace Forecast. Fiscal years 2017-2037*, 2017.

[3] Torres, S., Klooster, J. K., Ren, L., and Castillo-Effen, M., "Trajectory Synchronization between air and ground trajectory predictors," *IEEE/AIAA 30th Digital Avionics Systems Conference*, Seattle, WA, USA, 2011. doi:10.1109/DASC.2011.6095978.

[4] López Leonés, J., Amo, A., Bronsvoort, J., McDonald, G., Bayraktutar, I., Ángel Pérez Lorenzo, M., and Dierks, H., "Air-Ground Trajectory Synchronization through Exchange of Aircraft Intent Information," *Air Traffic Control Quarterly*, Vol. 20, No. 4, 2012, pp. 311–339. doi:10.2514/atcq.20.4.311.

[5] Bronsvoort, J., McDonald, G., Boucquey, J., Avello, C. G., Vilaplana, M., and Besada, J. A., "Impact of Data-link on Ground-Based Trajectory Prediction Accuracy for Continuous Descent Arrivals," *AIAA Modeling and Simulation Technologies (MST) Conference*, Boston, MA, 2013. doi:10.2514/6.2013-5068.

[6] Schultz, M., Fricke, H., Kunze, T., Gerbothe, T., Grabow, C., De Prins, J., Wimmer, M., and Kappertz, P., "Modelling and Evaluation of Automated Arrival Management Considering Air Traffic Demands," *3rd Eurocontrol SESAR Innovation Days*, 2013.

[7] Bronsvoort, J., McDonald, G., Hochwarth, J., and Gallo, E., "Air to Ground Trajectory Synchronisation through Extended Predicted Profile: A Pilot Study," *14th AIAA Aviation Technology, Integration, and Operations Conference*, Atlanta, GA, 2014. doi:10.2514/6.2014-2290.

[8] Bronsvoort, J., McDonald, G., Paglione, M., Young, C. M., Boucquey, J., Hochwarth, J. K., and Gallo, E., "Real-Time Trajectory Predictor Calibration Through Extended Projected Profile Down-Link," *11th USA/Europe ATM R&D Seminar*, Lisbon, Portugal, 2015.

[9] Mueller, K. T., Sorensen, J. A., and Couluris, G. J., "Strategic Aircraft Trajectory Prediction Uncertainty and Statistical Sector Traffic Load Modeling," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Monterey, California, 2002. doi:10.2514/6.2002-4765.

[10] Gerretsen, A., and Swierstra, S., "Sensitivity of aircraft performance to Variability of input data," Tech. Rep. DOC CoE-TP-02005, EUROCONTROL, 2003.

[11] Swierstra, S., and Green, S. M., "Common trajectory prediction capability for decision support tools," *5th USA/Europe ATM R&D Seminar*, Budapest, Hungary, 2003.

[12] Mueller, K. T., Bortins, R., Schleicher, D., Sweet, D., and Coppenbarger, R., "Effect of Uncertainty on En Route Descent Adviser (EDA) Predictions," *AIAA 4th Aviation Technology, Integration and Operations (ATIO) Forum*, Chicago, Illinois, 2004. doi:10.2514/6.2004-6347.

[13] Mondoloni, S., and Bayraktutar, I., "Impact of factors, conditions and metrics on trajectory prediction accuracy," *6th USA/Europe ATM R&D Seminar*, Baltimore, MD, 2005.

[14] Mondoloni, S., *Aircraft Trajectory Prediction Errors: Including a Summary of Error Sources and Data*, Version 0.2 ed., 2006.

[15] Paglione, M., Bayraktutar, I., Mcdonald, G., and Bronsvoort, J., "Lateral Intent Error's Impact on Aircraft Prediction," *8th USA/Europe ATM R&D Seminar*, Napa, CA, 2009.

[16] Torres, S., "Determination and ranking of trajectory accuracy factors," *IEEE/AIAA 29th Digital Avionics Systems Conference*, Salt Lake City, UT, 2010. doi:10.1109/DASC.2010.5655521.

[17] Henderson, J. M., Vivona, R. A., and Green, S., "Trajectory Prediction Accuracy and Error Sources for Regional Jet Descents," *AIAA Guidance, Navigation, and Control (GNC) Conference*, Boston, MA, 2013. doi:10.2514/6.2013-4535.

[18] Bronsvoort, J., "Contributions to trajectory prediction theory and its application to arrival management for air traffic control," Ph.D. thesis, Universidad Politécnica de Madrid, 2014.

[19] Torres, S., "Trajectory Accuracy Sensitivity to Modeling Factors," *15th AIAA Aviation Technology, Integration, and Operations Conference*, Dallas, TX, 2015. doi:10.2514/6.2015-2599.

[20] Casado Magaña, E. J., "Trajectory prediction uncertainty modelling for Air Traffic Management," Ph.D. thesis, University of Glasgow, 2016.

[21] Casado, E., Goodchild, C., and Vilaplana, M., "Sensitivity of Trajectory Prediction Accuracy to Aircraft Performance Uncertainty," *AIAA Guidance, Navigation, and Control and Co-located Conferences*, Boston, MA, 2013. doi:10.2514/6.2013-5045.

[22] Casado, E., D'Alto, L. P., and Vilaplana, M., "Analysis of the Impact of Intent Uncertainty on the Accuracy of Predicted Trajectories for Arrival Management Automation," *6th International Conference on Research in Air Transportation*, 2014.

[23] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., *Global sensitivity analysis. The Primer*, John Wiley & Sons, Ltd, 2008. doi:10.1002/9780470725184.ch1.

[24] Aldrich, J., "R. A. Fisher and the Making of Maximum Likelihood 1912-1922," *Statistical Science*, Vol. 12, No. 3, 1997, pp. 162–176. doi:10.1214/ss/1030037906.

[25] Massey, F. J. J., "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American Statistical Association*, Vol. 46, No. 253, 1951, pp. 68–78.

[26] EUROCONTROL, *Surveillance Data Exchange. Part 9: Category 062. SDPS Track Messages*, 2010.

[27] Rudnyk, J., "Results of surveillance data, weather forecast and air traffic control operational data analysis," 4TU.Centre for Research Data, `https://doi.org/10.4121/uuid:2f0287d5-5a11-4db0-bb24-b92748761140`, 2019.

[28] Sun, J., Ellerbroek, J., and Hoekstra, J. M., "WRAP: An open-source kinematic aircraft performance model," *Transportation Research Part C: Emerging Technologies*, Vol. 98, 2019, pp. 118–138. doi:10.1016/j.trc.2018.11.009.

[29] Nuic, A., *User Manual for the Base of Aircraft Data (BADA) Revision 3.12*, EEC Technical/Scientific Report No. 14/04/24-44 ed., 2014.

[30] ICAO, *Manual of the ICAO Standard Atmosphere extended to 80 kilometres (262,500 feet)*, 1993.

[31] Robert, E., and De Smedt, D., "Comparison of operational wind forecasts with recorded flight data," *10th USA/Europe ATM R&D Seminar*, Napa, CA, 2013.

[32] Kalos, M. H., and Whitlock, P. A., *Monte Carlo Methods*, 2nd ed., WILEY-VCH Verlag GmbH & Co, KGaA, Weinheim, 2008. pp. 28-36, ISBN: 978-3-527-40760-6.

[33] Bonamente, M., *Statistics and Analysis of Scientific Data*, Springer Science+Business Media, New York, 2013. doi: 10.1007/978-1-4614-7984-0, pp. 177-179.

[34] Torres, S., Dehn, J., McKay, E., Paglione, M. M., and Schnitzer, B. S., "En-Route Automation Modernization (ERAM) trajectory model evolution to support trajectory-based operations (TBO)," *IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*, Prague, Czech Republic, 2015. doi:10.1109/DASC.2015.7311332.