Trajectory Prediction Sensitivity Analysis Using Monte Carlo Simulations

Rudnyk, Julia; Ellerbroek, Joost; Hoekstra, Jacco

DOI
10.2514/6.2018-3669

Publication date
2018

Document Version
Accepted author manuscript

Published in
2018 Aviation Technology, Integration, and Operations Conference

Trajectory Prediction Sensitivity Analysis Using Monte Carlo Simulations

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 efficient and accurate trajectory predictors. With uncertainty inherent to almost all inputs of a trajectory predictor, accurate predictions are not a simple task. In this study, Monte Carlo simulations of a 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 (CAS) 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. Simulations are performed for heavy, medium and light wake turbulence category (WTC) aircraft. Results indicate that at 20 minutes look-ahead time, when outliers are not considered, along-track errors can reach up to 50 nautical miles for heavy and medium WTC and 100 nautical miles for light WTC aircraft. Cross-track errors are mostly within 3 nautical miles for heavy and medium WTC and 8 nautical miles for light WTC. Altitude errors can reach up to 25,000 feet. Wind conditions, vertical speed and CAS/Mach number speed settings are determined to be the most influential inputs.

I. Nomenclature

a, b, c	wind vertical profile function's coefficients	V_{TAS}	true airspeed, kts
AE	altitude error, ft	V_W	wind speed, kts
ATE	along-track error, *nmi*	VS	vertical speed, $ft \cdot min^{-1}$
b_{0i}, b_i	regression coefficients	u	west-east wind direction (East positive), kts
CI_w	confidence interval width	v	north-south wind direction (North positive), kts
CTE	cross-track error, *nmi*	w	wind direction, deg
e_{appr}	relative error of integral approximation	$\mathbf{X}$	set of function inputs (X_i, X_j,..., X_n)
g	gravitational acceleration, $g = 9.81ms^{-2}$	x	horizontal position, Eastern direction, *nmi*
h	altitude, ft	x_p	predicted horizontal position, Eastern direction, *nmi*
h_p	predicted altitude, ft	x_t	true horizontal position, Eastern direction, *nmi*
h_t	true altitude, ft	Y	system output, $Y = f(\mathbf{X})$
M	Mach number	y	horizontal position, Northern direction, *nmi*
N	number of samples	y_p	predicted horizontal position, Northern direction, *nmi*
n	sample size	y_t	true horizontal position, Northern direction, *nmi*
p	air pressure at the considered altitude, Nm^{-2}	$z_{\alpha/2}$	critical value of z (z-score) for a significance level α
p_0	sea level air pressure, $p_0 = 101325Nm^{-2}$	γ	adiabatic index of air, $\gamma = 1.4$
R	specific gas constant, $R = 287.05287m^2K^{-1}s^{-2}$	ΔT	air temperature offset from ISA, K
r	turn radius, *nmi*	μ	population mean
r^2	coefficient of determination	λ	lapse rate, $K \cdot km^{-1}$
r_P	Pearson correlation coefficient	ρ	air density at the considered altitude, $kg \cdot m^{-3}$
r_S	Spearman correlation coefficient	ρ_0	sea level air density, $\rho_0 = 1.225kg \cdot m^{-3}$

*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

S_{Xi}	sensitivity measure of Y to X_i	σ	population standard deviation
T	air temperature at the considered altitude, K	ϕ	bank angle, deg
T_{ISA}	ISA air temperature at the considered altitude, K	ψ_a	aircraft heading, deg
V_{CAS}	calibrated airspeed, kts	ψ_g	track angle (i.e. course), deg
V_{GS}	ground speed, kts	ψ_{g_p}	predicted track angle (i.e. course), deg

II. Introduction

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

All of these tasks require the ability to predict the future positions of aircraft. To facilitate this, a 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 model and model parameters are realistic and correct, the accuracy of 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. For instance, for Medium Term Conflict Detection (MTCD), a prediction is usually made up to 20 minutes look-ahead time. For this, rather large in aviation, time interval the accurate knowledge of the weather conditions, Air Traffic Control (ATC) instructions and aircraft modes of operation is crucial. These are often not known and either assumed to have standard values or disregarded. Differences between reality and assumptions result in prediction errors, which in turn diminish the added value of decision support tools.

An often proposed solution to this problem would be to ensure that inputs are known with high certainty (e.g. by downlinking some data or computed trajectory from aircraft [3–5]). However, not all inputs can be precisely known in advance (e.g. weather, ground delays, ATC instructions).

To study the extent of prediction uncertainty induced by uncertain inputs, sensitivity analyses have been performed, where one input was varied at a time, while the rest were kept constant at their reference values [6–9] (see Table 1 for a list of common inputs studied in the previous sensitivity studies). This is known as local sensitivity analysis, and 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 each input contribution to output uncertainty, it does not allow 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. Casado (2016) developed a framework to estimate prediction uncertainty using arbitrary Polynomial Chaos Expansion and compared its performance with Monte Carlo simulations (both are types of a global sensitivity analysis), where the majority of inputs' distributions was assumed to be normal [10], whereas in related work [11, 12] some inputs were described with uniform and triangular distributions. While the simultaneous variation of inputs is representative when modeling trajectories, the assumed distributions may be not.

The objective of this paper is therefore to address this issue by performing Monte Carlo simulations with distributions of the inputs obtained from observed data. The simulations were conducted per aircraft wake turbulence category (WTC) and the selection of TP inputs was based on the conclusions of previous research [6–10, 13–16] about significant input parameters. Considered inputs included aircraft bank angle, constant calibrated airspeed/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, and ATC intent. Prediction uncertainty is estimated as a function of a look-ahead time up to 20 minutes and inputs' contribution to prediction errors is assessed based on the computed correlation coefficients.

Table 1 Trajectory predictor inputs studied in previous research

Studied inputs	Reference
Wind Gradient	Mondoloni [6, 8, 9], Gerretsen [7]
Wind Field	Mondoloni [6, 8, 9], Gerretsen [7]
Air Temperature	Gerretsen [7], Mondoloni [9]
Weather Forecast Ensembles	Casado [10]
Aircraft Mass	Mondoloni [6, 8, 9], Gerretsen [7], Casado [10]
Speed Settings	Mondoloni [6, 8, 9], Gerretsen [7], Casado [10]
Thrust Settings	Gerretsen [7]
Aircraft Performance	Mondoloni [8], Casado [10]
Aircraft Configuration	Gerretsen [7]
Top of Descent Placement	Mondoloni [6, 8, 9], Casado [10]
Temporary Level-off	Mondoloni [6, 8, 9]
Vector Instruction from ATC	Mondoloni [8]
Aircraft Intent	Casado [10]
Initial Conditions	Casado [10]

III. Sensitivity Analysis

This section serves as a short introduction to a common approach of sensitivity analysis (local sensitivity analysis) and to an approach chosen for this study (global sensitivity analysis).

A sensitivity analysis can be performed to determine which inputs of a TP are more significant or influential than others. In the field of sensitivity analysis, there is a distinction between *sensitivity* and *uncertainty* analysis [17, 18]. A sensitivity analysis allocates 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 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. Usually, uncertainty and sensitivity analyses are performed together, which often results in calling it altogether Sensitivity Analysis. Depending on the purpose of the study and the choice of methods, both sensitivity and uncertainty analysis could be performed or only one of them. 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 summary of common approaches to local and global 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 [17, 18]. 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. For instance, consider a case when aircraft speed settings are varied, while other inputs are kept at their reference values. When the value of drag is calculated, only speed variations contribute to the changes in drag. In reality, however, this is not true. Together with changes in speed, the aircraft mass and atmospheric conditions would also contribute to the magnitude of the drag. The assumption of inputs at their nominal values might hide interesting results of a combined impact of several inputs and is an oversimplification of the reality. For a local SA, one could take partial derivatives or change inputs with a certain step around a nominal value. A derivative-based SA is based on computing a partial derivative of output Y with respect to the i^{th} input X_i (Eq. (1)).

$$S_{Xi} = \frac{\partial Y}{\partial X_i} \tag{1}$$

While Eq. (1) is a straightforward way of expressing the sensitivity of Y to X_i, the obtained measure of sensitivity is representative only at the point where it is computed. The latter implies that if one wants to perform a derivative-based

SA of a non-linear system and obtain an average response of Y to X_i, Eq. (1) must be computed at a set of different points in the space of X_i. A sensitivity measure in Eq. (1) can be normalized to remove the units by multiplying it by X_i/Y.

Another approach of a local SA is to increase and decrease an input by a standard deviation ($\pm SD$) or by a percentage (e.g. ($\pm$5%)) from a nominal value. Then, the output is computed for these local points and compared to the output obtained at a nominal value.

B. Global Sensitivity Analysis

Usually, a global SA uses all-at-a-time sampling approach where all inputs are varied at the same time from their distribution functions or entire feasible space [17, 18]. A simultaneous variation of all inputs allows to investigate the sensitivity of 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. Monte Carlo method applies sampling, which is a process of randomly drawing variables form 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, 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. Sensitivity measure can be obtained by performing a linear regression or by computing a correlation coefficient for input X_i and output Y [17]. For a linear regression analysis, Eq. (2) is assumed to describe a linear relationship between the output and input where coefficients b_{0i} and b_i are obtained using least-square computation.

$$Y = b_{0i} + b_i X_i \tag{2}$$

1. Correlation Coefficient and Coefficient of Determination

There are two types of correlation coefficients that are frequently used to estimate the relationship between two variables: Pearson and Spearman correlation coefficients. The Pearson correlation coefficient, r_P, measures the linear relationship between two continuous variables X and Y:

$$r_P = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{3}$$

Here n is the sample size, $\{X_1, ..., X_n\}$ is the X dataset, $\{Y_1, ..., Y_n\}$ is the Y dataset, $\bar{X}$ and $\bar{Y}$ are the sample mean of X and Y datasets, respectively.

The Spearman correlation coefficient, r_S, measures the monotonic relationship between two variables. In a monotonic relationship, the change of two variables does not have to be at a constant rate. The Spearman correlation coefficient is equal to the Pearson correlation coefficient of the ranked values of the variables. For ranking, the values are ordered (usually in ascending order) and the index of a value in this ordered list is its rank. When all ranks are distinct integers, the Spearman correlation coefficient can be computed as follows:

$$r_S = 1 - \frac{6\sum {d_i}^2}{n(n^2 - 1)} \tag{4}$$

Here d_i is the difference between the ranks of X_i and Y_i.

Both correlation coefficients vary between -1 and +1. 0 implies no correlation, -1 or +1 imply an exact relationship. Positive correlation means that with an increase of X, Y increases, whereas negative correlation means that with an increase of X, Y decreases.

To illustrate the difference between the Pearson and Spearman correlation coefficients, Fig. 1 shows the relation between variables X and Y, where $Y = \exp(X)$, and indicates the computed correlation coefficients. The relationship between X and Y is not perfectly linear (i.e. $r_P = 0.69$), however, it is perfectly monotonic (i.e. $r_S = 1.00$).

Furthermore, a coefficient of determination, $r^2 = {r_P}^2$, can be computed, which measures how much variance in the dependent variable Y is explained by the independent variable X. r^2 ranges from 0 to 1 and when multiplied by 100 indicates the percentage of the variance. 0% means no variance in Y can be explained by X and 100% means all variance in Y can be explained by X.

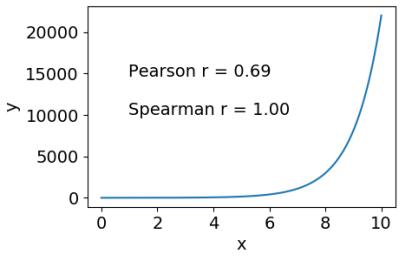


Fig. 1 Pearson and Spearman correlation coefficients for X and Y variables, where $Y = \exp(X)$.

IV. Study Setup

This section outlines the inputs to Monte Carlo simulations and the source of their distribution functions, describes a TP used in this study and illustrates the computation of *true* 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. Lastly, the importance of a suitable random number generator is addressed.

A. Input Data and Sampling

In this study, all input parameters are assumed to be independent, meaning that each input is sampled from its distribution function independently from the values of other inputs. Considered inputs include bank angle, constant calibrated airspeed (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. Here, ATC intent includes direct-to, heading, speed, and climb to instructions. The distribution functions of all parameters, where applicable per WTC, were determined by Rudnyk et al. (2018) [19]. Additionally, ten flight routes within Europe and ten aircraft types for each WTC were sampled assuming a uniform distribution.

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. Moreover, sampled CAS and Mach number could be larger than the maximum operating speed limits (VMO and MMO) of a particular aircraft type or could result in a speed smaller than the minimum speed (V_{min}). Therefore, constraints were imposed (see Table 2) on some inputs to ensure sensible values.

Table 2 Input parameters' limits

Parameter	Units	Minimum	Maximum
Bank angle	degrees	5	35
Mach number	-	V_{min}	MMO
CAS	knots	V_{min}	VMO
Rate of climb	feet per minute	500	4000
Rate of descent	feet per minute	500	3000
Temporary level-off	seconds	20	500

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 3. 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 might vary throughout a phase. The air temperature was modeled using a temperature offset from the standard temperature (i.e. temperature according to ICAO International Standard Atmosphere (ISA)). Values of the air temperature offset and the lapse rate were the same in all phases of a flight. The wind speed and the wind direction were modeled with u and v wind components as a function of altitude. The proportion of flights that received an ATC 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% were instructed to climb 10 flight levels at 1,000 feet per minute. 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.

Table 3 Inputs sampling per flight phase

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

B. 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. The equations of motion used for the flight profile computation are given in Eq. (5) to Eq. (8). 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.

$$\dot{x} = V_{GS} \sin \psi_g \tag{5}$$

$$\dot{y} = V_{GS} \cos \psi_g \tag{6}$$

$$\dot{h} = VS \tag{7}$$

$$\dot{\psi_g} = \frac{V_{GS}}{r} \tag{8}$$

The ground speed V_{GS} and the turn radius r were obtained using Eq. (9) and Eq. (10) respectively. The true airspeed of an aircraft V_{TAS} was converted from the CAS (Eq. (11)) or the Mach number (Eq. (12)).

$$V_{GS} = \sqrt{V_{TAS}^2 + V_w^2 - 2 \cdot V_{TAS} V_w \cos(w - \psi_a)} \tag{9}$$

$$r = \frac{V_{GS}^2}{g \cdot \tan \phi} \tag{10}$$

$$V_{TAS} = \sqrt{\frac{2\gamma}{\gamma - 1} \frac{p}{\rho} \left[\left(1 + \frac{p_0}{p} \left[\left(1 + \frac{\gamma - 1}{2\gamma} \frac{\rho_0}{p_0} V_{CAS}^2 \right)^{\frac{\gamma}{\gamma - 1}} - 1 \right] \right)^{\frac{\gamma - 1}{\gamma}} - 1 \right]} \tag{11}$$

$$V_{TAS} = M \sqrt{\gamma \cdot R \cdot T} \tag{12}$$

The wind speed V_w and the wind direction w were computed from the u and v wind components (Eq. (13) and Eq. (14) respectively), which were determined as a function of a flight level (FL) x, $f(x) = ax^2 + bx + c$.

$$V_w = \sqrt{u^2 + v^2} \tag{13}$$

$$w = atan2(-u, -v) \tag{14}$$

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:

$$T = T_0 + \Delta T + \lambda h \tag{15}$$

$$p = p_0 \left(\frac{T - \Delta T}{T_0} \right)^{\frac{-g}{\lambda R}} \tag{16}$$

Since temporary level-offs and ATC instructions are *procedural* parameters, they are not described with equations, and their execution was triggered by certain events. Provided 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. The 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 following the same route 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 for a time interval as sampled from its duration distribution. After the heading instruction was completed, an aircraft returned to its intended route. Taking into account the MMO, a speed instruction was given as a Mach number for a time interval as sampled from its duration distribution. An instruction to climb 10 FLs at 1,000 feet per minute was issued at around the middle of the cruise phase.

C. True Trajectories and Predictions

After inputs were sampled from their distribution functions, *N true* trajectories were simulated. These *true* trajectories resemble the possible real flights. Then, for each *true* trajectory, *predictions* were made up to 20 minutes look-ahead time utilizing a time interval moving window. For every minute of the *true* 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 *true* trajectories and *predictions*, the same trajectory simulator/predictor was used. For the *predictions*, inputs were set according to the common assumptions and BADA aircraft performance model [20]. Table 4 indicates which setting were chosen for each input. The atmospheric conditions were set to ISA (e.g. ISA temperature and standard lapse rate). The wind conditions were sampled from the same distribution functions that were used for the *true* trajectories (i.e. wind forecast). The bank angle was assumed to be 25 degrees, and the Mach number and CAS speed settings were extracted from BADA according to the aircraft type. The vertical speed was set to the values somewhat lower than those used by the present operational TPs. In this case, a compromise was made between the commonly used values and the median values of the vertical speed distributions estimated from observed data. As often assumed, no temporary level-offs and ATC instructions were taken into account in these predictions.

Table 4 Inputs' settings for predictions

Parameter	Climb	Cruise	Descent
Air temperature offset	0	0	0
Lapse rate	-6.5 K/km	-6.5 K/km	-6.5 K/km
Wind	forecast	forecast	forecast
Bank angle	25 deg	25 deg	25 deg
Mach number	BADA	BADA	BADA
CAS	BADA	-	BADA
Vertical speed	2,500 fpm	-	-1,500 fpm
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 *true* trajectory, and the *prediction* was started. In case the starting position was off-route (e.g. direct-to instruction), the *prediction* would be performed by rejoining the initial route at the next waypoint downstream.

While the expected cruise altitude is not reached, the *prediction* would be done for the climb phase with transitioning to the cruise phase, where applicable. Once the altitude of the *true* trajectory equals the cruise altitude, the *prediction* would be performed for the cruise phase, with transitioning to the descent phase at the precomputed top of descent, where applicable. Once the *true* trajectory is in the descent phase, the *prediction* is performed for the descent phase.

The reinitialization algorithm does not adjust the *prediction* inputs according to the states of a *true* trajectory other than position, altitude and phase. For instance, if, at the moment of an update, the speed of an aircraft is different from the one used in the prediction, the speed used for the prediction won't be updated. Same stands for the vertical speed. Although this results in a somewhat unrealistic prediction process, it is not expected to have a significant influence on the overall trends of the prediction errors' magnitudes.

D. 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 was calculated using Eq. (17) and Eq. (18) respectively. The AE was computed using Eq. (19).

$$ATE = (x_p - x_t)\sin\psi_{g_p} + (y_p - y_t)\cos\psi_{g_p} \tag{17}$$

$$CTE = (x_p - x_t)\cos\psi_{g_p} - (y_p - y_t)\sin\psi_{g_p} \tag{18}$$

$$AE = h_p - h_t \tag{19}$$

These errors were subsequently used to assess the prediction uncertainty with a look-ahead time and to determine the influence of inputs. The errors were analyzed per flight phase, which was determined by the *true* trajectory. For instance, if the errors were computed when the *true* flight was in the climb phase and the *predicted* flight was in the cruise phase, the errors would be attributed to the climb phase. This was done to conform with how inputs are sampled for the *true* trajectories, which is per flight phase.

E. Number of Runs

To determine how many *true* trajectories should be simulated, two numbers were calculated: a number of samples, N, to ensure a small relative error of Monte Carlo sampling and a sample size, n, to ensure statistical significance of the obtained results. When samples are randomly drawn from the distribution function, this approximates the evaluation of an integral [21, 22]. The relative error e_{appr} of this approximation can be estimated as in Eq. (20). In this study, the relative error, e_{appr}, of 0.01 was assumed. Rearranging Eq. (20) to find N results in $N = 10,000$ samples.

$$e_{appr} = \frac{1}{\sqrt{N}} \tag{20}$$

An accuracy requirement can also be imposed on the future results, which requires the computation of a sample size [23]. Assuming the resulting outputs are normally distributed, this can be done by specifying a significance level α and a confidence interval width CI_w:

$$CI_w = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{21}$$

To compute the sample size n, Eq. (21) can be rearranged as:

$$n = \left(\frac{2z_{\alpha/2}\sigma}{CI_w}\right)^2 \tag{22}$$

Equation (22) requires a value of a population standard deviation, σ. Since this value is not known, an assumption could be made or test simulations could be performed with some sample size (50, 100, 500, etc.) to determine the value of a sample standard deviation, which could be assumed equal to the population standard deviation. Instead, it is possible to rewrite Eq. (22) without a population standard deviation by expressing the confidence interval width as a percentage of the population standard deviation. For instance, when CI_w equals 5% of σ, Eq. (22) can be written as:

$$n = \left(\frac{2z_{\alpha/2}\sigma}{0.05\sigma}\right)^2 = \left(\frac{2z_{\alpha/2}}{0.05}\right)^2 \tag{23}$$

The sample size was obtained for $\alpha = 0.05$, which gives a confidence level of 95% and z-score of 1.96. Substituting this in Eq. (23) results in $n = 6,147$ samples. This value means that at least 6,147 samples are necessary to be 95% confident that a sample mean will be within half of the confidence interval width of a population mean (i.e. $\mu \pm CI_w/2 = \mu \pm 0.025\sigma$).

Based on the obtained values for N and n, the largest one (i.e. 10,000) was chosen as a number of Monte Carlo simulation runs for each WTC aircraft.

F. Random Number Generator

The Monte Carlo sampling process makes use of pseudorandom numbers, which resemble truly random numbers. This allows to ensure the repeatability of the simulations' *randomness*. Pseudorandom numbers are generated by a pseudorandom number generator (PRNG) and can be represented as a sequence of uniformly distributed random numbers on the interval [0, 1) [21]. A PRNG can be provided with a seed, which is a number that initializes the sequence of resulting generated numbers. When a PRNG is given the same seed over several simulations, it will produce the same sequence of random numbers in these simulations. A PRNG is characterized by a (sequence) period which determines how many random numbers can be generated before they start repeating. The generator with a longer period is more useful. For instance, if a sequence contains 100 random numbers and sampling is done 10,000 times, the same sequence will be repeated 100 times. This would give the same results as sampling 100 times and does not ensure good sampling. In this study, Python library *numpy* was used to sample inputs from their distribution functions. Numpy uses the Mersenne Twister PRNG, which is a 32-bit random integer generator and has a period of $2^{19937} - 1$ [24].

V. 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 (errors' distribution) and the correlation between the inputs and the prediction errors. Results are presented for heavy, medium and light WTC aircraft.

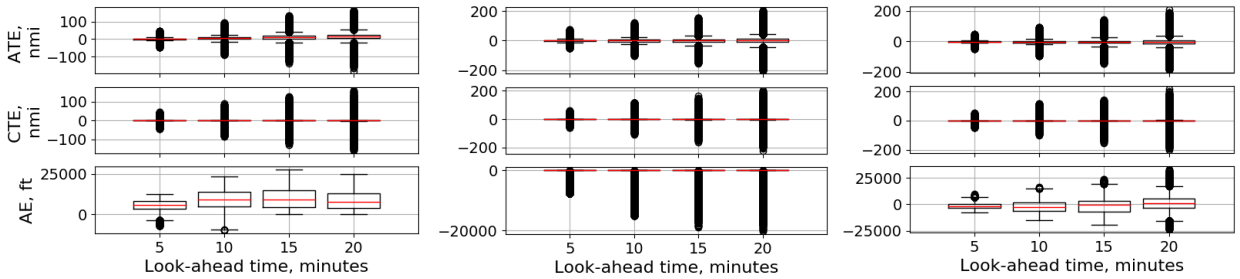
A. Prediction Uncertainty

To assess prediction uncertainty, errors' distributions are plotted as box plots. For analysis purposes, the box plots were studied with and without outliers. The former demonstrates the maximum errors and is provided below mainly for illustrative purposes, whereas the latter allows to investigate the errors' distribution in a greater detail and focuses on the most frequent errors.

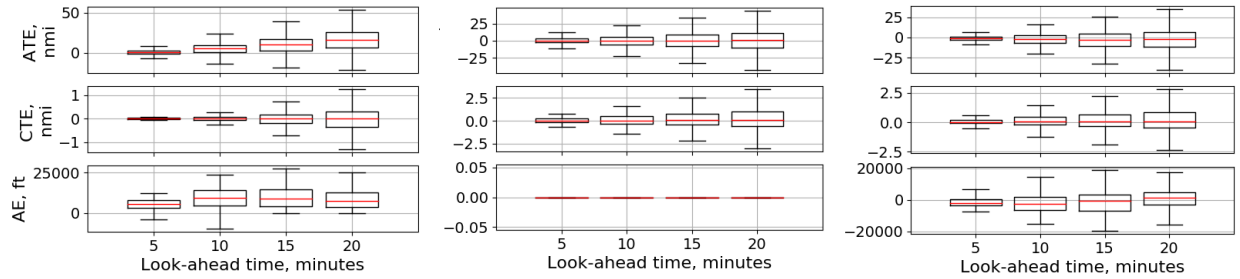
Figures 2, 3 and 4 present the results for heavy, medium and light WTC aircraft, respectively, with (sub-figures a, b, c) and without (sub-figures d, e, f) outliers. As expected, the ATE and CTE grow with a look-ahead time in all phases. Errors for heavy and medium WTC aircraft are similar in magnitudes and demonstrate comparable trends. For 20 minutes look-ahead time, ATEs are mostly less than 50 nmi and CTEs are less than 3 nmi. Larger ATEs and CTEs occur in the cruise phase due to ATC instructions and a prolonged wind impact. Whereas the CTE is predominantly centered at zero, in the climb and descent phases, the ATE median shifts above and below zero, respectively, with an increase of the look-ahead time. This can be a result of errors' computation between the points located on the different segments of a route which bends. Furthermore, consistently different speed settings or wind conditions (e.g. speed settings sampled for the *true* flights are consistently lower than speed settings used for the *predictions*) could also contribute to such errors' trends.

The trends of ATEs for light WTC aircraft are similar to those of heavy and medium WTC aircraft; however, the magnitudes are much larger. At 20 minutes look-ahead time, the errors can reach up to 100 nmi. Most likely, such a difference is the result of frequently different speed settings sampled for the *true* flights compared to the speed settings used for the *predictions*. This indicates that the distributions used for the speed settings of light WTC aircraft are considerably inaccurate. The magnitudes of CTEs in the cruise and descent phases (within 8 nmi) are also larger than those of heavy and medium WTC which is likely to be influenced by the speed settings as well. Additionally, the computation of errors between the points located on the different segments of a route contributes to large errors.

Whereas the ATE and CTE steadily increase with a look-ahead time, the AE has a different trend. In the climb phase, the AE has the largest spread at 10 minutes look-ahead time and in the descent phase at 15 minutes for heavy and

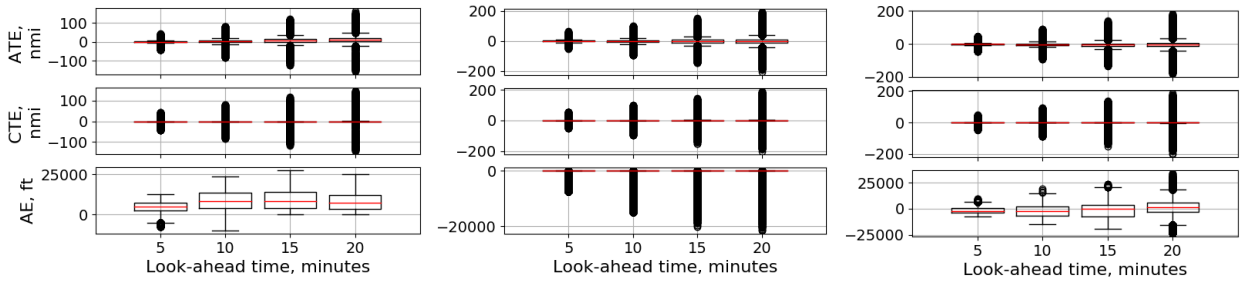


(a) Climb. With outliers **(b) Cruise. With outliers** **(c) Descent. With outliers**

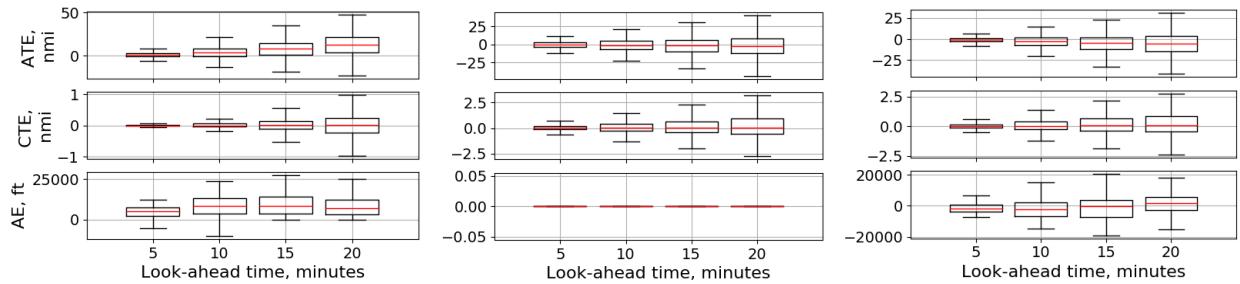


(d) Climb. Without outliers **(e) Cruise. Without outliers** **(f) Descent. Without outliers**

Fig. 2 Prediction uncertainty with look-ahead time for heavy WTC.



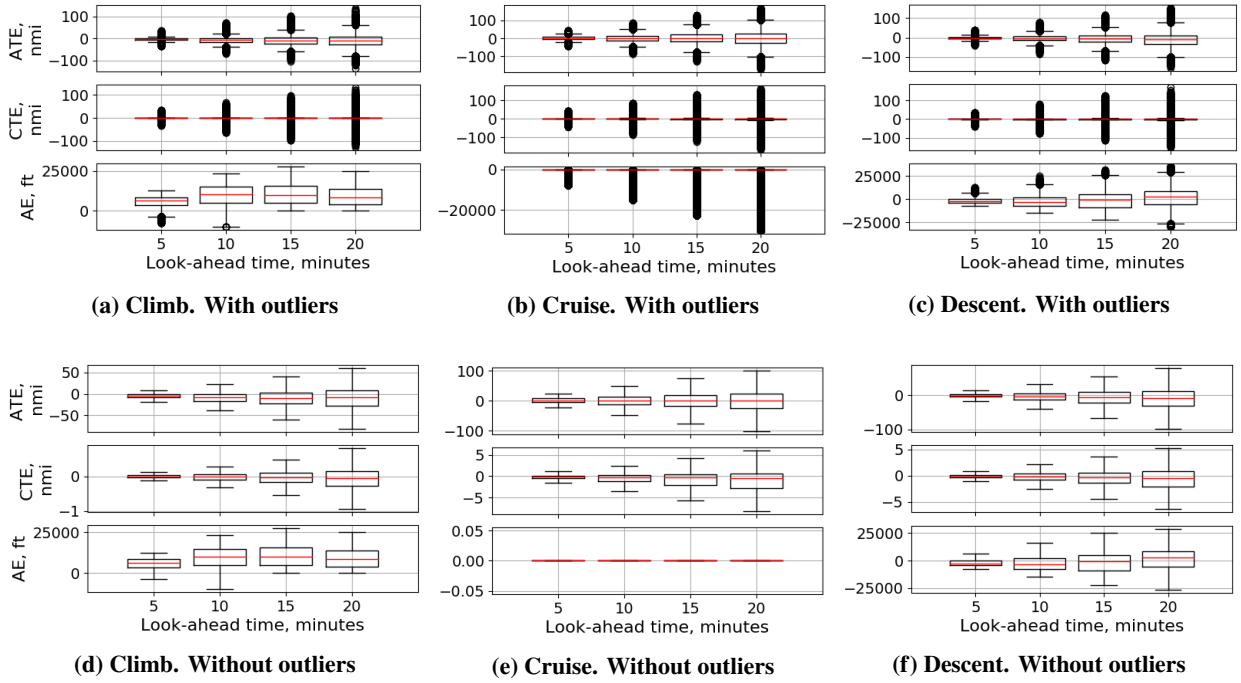
(a) Climb. With outliers **(b) Cruise. With outliers** **(c) Descent. With outliers**



(d) Climb. Without outliers **(e) Cruise. Without outliers** **(f) Descent. Without outliers**

Fig. 3 Prediction uncertainty with look-ahead time for medium WTC.

medium WTC and at 20 minutes for light WTC aircraft. Starting from 15 minutes look-ahead time in the climb phase, the AEs are equal to or larger than 0 which is a result of all *true* trajectories being at the same or lower altitude than predicted (i.e. most likely *predictions* reached the cruising altitude). Additionally, for all look-ahead times in the climb phase, the AE is not centered at zero, which indicates that the choice of the vertical speed settings for the prediction is, in



(a) Climb. With outliers **(b) Cruise. With outliers** **(c) Descent. With outliers**

(d) Climb. Without outliers **(e) Cruise. Without outliers** **(f) Descent. Without outliers**

Fig. 4 Prediction uncertainty with look-ahead time for light WTC.

most cases, inaccurate. In the cruise phase, in general, no AEs occur. The AE outliers in the cruise phase occurred when the *true* flights were in cruise while the *predicted* flights were in the climb phase, when the *true* flights were instructed by ATC to climb or when the *predicted* trajectories started descending while the *true* flights were still in the cruise phase. In the descent phase, AEs are centered at around zero indicating that the vertical speed used for the predictions in the descent phase is often correct. Furthermore, for 20 minutes look-ahead time, the errors can reach up to 25,000 feet.

B. Correlation Between Inputs and Prediction Errors

To determine the relationship between the inputs and prediction errors, correlation coefficients, both Pearson and Spearman, were computed. For this, the next procedure was followed. Firstly, the average error was computed per flight phase for each *true* trajectory. Secondly, the computed errors were mapped with the *true* flights' inputs. Thirdly, the obtained input-error pairs were grouped according to the route. Fourthly, the correlation coefficients were computed for each input-error combination (e.g. vertical speed versus AE) for each route (i.e. route-based correlation coefficients). This was done to ensure that the route-dependent input-error combinations are discovered (e.g. the impact of the predominant wind direction on the predominant route direction). Finally, the overall correlation coefficients were calculated as the average of the absolute values of all route-based correlation coefficients for each input-error combination. The absolute values were considered to avoid the cancellation of positive and negative values of the coefficients. No significant difference was detected between the results for Pearson and Spearman correlation, except for the relationship between the vertical speed and the AE in the descent phase, which is addressed further in this section. Therefore, the correlation coefficients presented in this section are the Pearson correlation coefficients.

As an example, Fig. 5 visualizes the relationship between the inputs and the ATE 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 the heading, *S* for the speed and *VS* for the vertical speed instructions. Due to a discrete nature of ATC instructions, the correlation coefficient cannot be determined for this input under this study setup. A different approach to ATC intent modeling is required for a successful estimation of the correlation coefficients.

Since the correlation coefficients are obtained as the average over all routes, the correlation between the errors and the inputs may be stronger or weaker for different routes. For instance, the correlation between the vertical speed and

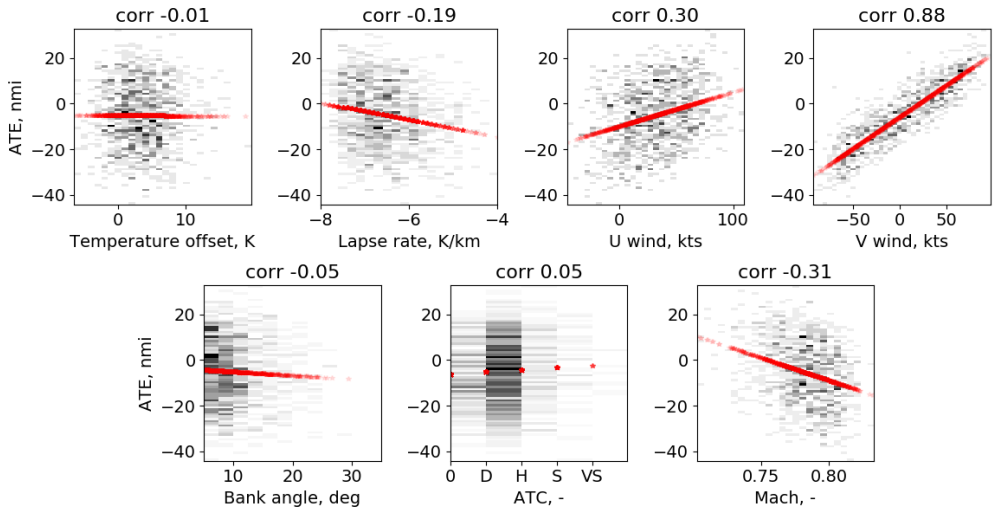


Fig. 5 Correlation between inputs and along-track error (ATE) in the cruise phase for 20 minutes look-ahead time.

the AE is strong for all routes, whereas the correlation between the wind and the ATE depends on the wind and route predominant directions.

There are no strict rules for deciding which magnitude of a correlation coefficient should be considered as a minimum value for concluding on the presence of a linear relationship between two variables. Moreover, the acceptable magnitudes may differ based on the field of study. In this study, five ranges of correlation coefficient's magnitudes were distinguished as indicated in Table 5. The correlation coefficients whose absolute value is equal to or greater than 0.23 (i.e. stronger than very weak) were considered in this study as those that show a minimum 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% (i.e. $r^2 = 0.23^2 = 0.0529 \approx 0.05$). Therefore, only inputs that contribute to at least 5% of the prediction errors' variance are of interest.

Table 5 Ranges of correlation coefficients' strength

Coefficient absolute value	Relationship strength
0.00 - 0.19	very weak
0.20 - 0.39	weak
0.40 - 0.59	moderate
0.60 - 0.79	strong
0.90 - 1.00	very strong

Tables 6, 7 and 8 show the overall correlation coefficients for heavy, medium and light WTC aircraft, respectively. These coefficients are provided for 11 inputs (LVL stands for temporary level-off and VS for vertical speed) and 3 prediction errors as a function of a flight phase and a look-ahead time. Since presented coefficients are average values over all routes, the color coding is used to indicate whether the route-based correlation coefficients are positive, negative or mixed. The red color indicates the negative correlation; the green color is the positive correlation, and the yellow color indicates the mixed correlation (both positive and negative). In Tables 6, 7 and 8 only the color-marked coefficients are of interest (i.e. their absolute value is $\geq$0.23). In the following subsections, the input-error correlations with a look-ahead time are analyzed for each error separately.

1. Along-Track Error

The ATE is correlated with the wind in all phases for all look-ahead times and all three WTCs. The strength of the relationship is between weak and strong for heavy and medium WTC aircraft and between weak and moderate for light

Table 6 Correlation coefficients for heavy WTC aircraft

Heavy WTC	5 minutes look-ahead time									10 minutes look-ahead time								
	Climb			Cruise			Descent			Climb			Cruise			Descent		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	0.09	0.02	0.02	0.05	0.02	0.03	0.08	0.04	0.02	0.08	0.02	0.03	0.04	0.02	0.03	0.06	0.03	0.04
Lapse rate	0.06	0.03	0.02	0.16	0.03	0.03	0.09	0.04	0.02	0.05	0.03	0.03	0.15	0.03	0.04	0.11	0.06	0.03
U wind	0.36	0.06	0.03	0.31	0.05	0.09	0.32	0.11	0.04	0.32	0.05	0.03	0.32	0.06	0.11	0.31	0.12	0.07
V wind	0.62	0.05	0.03	0.77	0.10	0.16	0.74	0.29	0.07	0.55	0.06	0.03	0.77	0.12	0.22	0.74	0.28	0.22
Bank angle	0.03	0.03	0.03	0.06	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.03	0.06	0.03	0.04	0.02	0.03	0.02
ATC	0.00	0.00	0.00	0.04	0.07	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.06	0.03	0.00	0.00	0.00
Mach number	0.05	0.02	0.03	0.22	0.04	0.03	0.09	0.04	0.03	0.03	0.03	0.03	0.21	0.06	0.04	0.08	0.04	0.03
CAS	0.47	0.04	0.03	0.00	0.00	0.00	0.36	0.15	0.03	0.45	0.05	0.03	0.00	0.00	0.00	0.32	0.14	0.09
VS	0.34	0.17	0.97	0.00	0.00	0.00	0.27	0.18	0.94	0.51	0.15	0.96	0.00	0.00	0.00	0.26	0.14	0.73
LVL FL	0.02	0.04	0.07	0.00	0.00	0.00	0.03	0.03	0.11	0.03	0.04	0.05	0.00	0.00	0.00	0.03	0.03	0.12
LVL duration	0.03	0.04	0.13	0.00	0.00	0.00	0.06	0.04	0.21	0.03	0.04	0.11	0.00	0.00	0.00	0.04	0.05	0.17
Heavy WTC	15 minutes look-ahead time									20 minutes look-ahead time								
	Climb			Cruise			Descent			Climb			Cruise			Descent		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	0.09	0.03	0.03	0.05	0.03	0.03	0.06	0.03	0.06	0.09	0.03	0.02	0.05	0.03	0.03	0.06	0.03	0.06
Lapse rate	0.05	0.05	0.02	0.16	0.04	0.05	0.11	0.04	0.06	0.05	0.04	0.04	0.16	0.04	0.05	0.13	0.04	0.08
U wind	0.33	0.04	0.03	0.33	0.07	0.12	0.31	0.10	0.19	0.35	0.07	0.02	0.33	0.07	0.14	0.31	0.09	0.24
V wind	0.57	0.10	0.04	0.78	0.14	0.26	0.76	0.25	0.43	0.61	0.10	0.04	0.78	0.16	0.29	0.77	0.22	0.54
Bank angle	0.02	0.03	0.02	0.06	0.02	0.04	0.02	0.03	0.02	0.03	0.03	0.04	0.05	0.03	0.04	0.02	0.02	0.03
ATC	0.00	0.00	0.00	0.04	0.06	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.06	0.04	0.00	0.00	0.00
Mach number	0.04	0.03	0.03	0.21	0.06	0.06	0.08	0.03	0.04	0.05	0.04	0.05	0.20	0.06	0.07	0.08	0.03	0.03
CAS	0.47	0.08	0.04	0.00	0.00	0.00	0.25	0.09	0.16	0.49	0.11	0.04	0.00	0.00	0.00	0.18	0.06	0.18
VS	0.46	0.19	0.95	0.00	0.00	0.00	0.14	0.09	0.17	0.35	0.24	0.94	0.00	0.00	0.00	0.06	0.10	0.45
LVL FL	0.07	0.05	0.04	0.00	0.00	0.00	0.03	0.03	0.10	0.09	0.06	0.05	0.00	0.00	0.00	0.02	0.04	0.05
LVL duration	0.06	0.05	0.04	0.00	0.00	0.00	0.02	0.05	0.05	0.10	0.05	0.05	0.00	0.00	0.00	0.04	0.05	0.07

WTC aircraft. The latter might be a result of another input influencing or weakening the wind effect. Furthermore, the strength of the correlation does not significantly change with an increase of the look-ahead time. The v wind component has a stronger correlation with the ATE than the u wind component. Since the v wind component represents the North-South wind direction, it would have a larger impact on the ATE if the predominant direction of the flights' routes was also North-South. The correlation coefficients are mixed (i.e. positive and negative), since they highly depend on the wind and route predominant directions.

The correlation between the ATE and the vertical speed is present in the climb and descent phases of heavy and medium WTC aircraft and in the climb phase of light WTC aircraft. In the climb phase, a fast climbing aircraft reaches a higher true airspeed sooner than a slow climbing aircraft, which contributes to the ATE. For heavy and medium WTC aircraft, the correlation is between weak and moderate in the climb phase and weak in the descent phases. The correlation is weak for light WTC. In the descent phase, the correlation reduces with the look-ahead time, most likely due to a stronger influence of the wind conditions. The difference between the results for heavy and medium and light WTC aircraft might be the result of another input influencing or weakening the effect of the vertical speed. Furthermore, the correlation is negative, which signifies that the higher vertical speed results in the decrease of the ATE (i.e. negative ATE).

The results for both medium and heavy WTC aircraft show a correlation between the speed settings and the ATE. In the climb and descent phases, the ATE is correlated with the CAS, whereas in the cruise phase with the Mach number (for medium WTC). The correlation is mostly moderate in the climb phase and its strength slightly increases with the look-ahead time, whereas in the descent phase, the correlation is weak and its strength decreases with the look-ahead time. This implies that in the climb phase, the CAS contributes to the determination of the top of climb, while in the descent phase, the CAS has less impact on the top of descent determination (compared to the wind effect). In the cruise phase of medium WTC, the Mach number has a weak correlation with the ATE for all look-ahead times. Furthermore, speed settings are negatively correlated with the ATE, which means that the increase of the speed results in the negative ATEs (i.e. *predicted* position is behind the *true* position). Surprisingly, the results for light WTC aircraft

Table 7 Correlation coefficients for medium WTC aircraft

Medium WTC	5 minutes look-ahead time									10 minutes look-ahead time								
	Climb			Cruise			Descent			Climb			Cruise			Descent		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	0.09	0.03	0.03	0.05	0.03	0.02	0.07	0.03	0.03	0.09	0.02	0.03	0.05	0.03	0.02	0.06	0.02	0.04
Lapse rate	0.04	0.03	0.03	0.13	0.04	0.01	0.06	0.04	0.02	0.03	0.03	0.03	0.14	0.04	0.02	0.07	0.04	0.03
U wind	0.39	0.04	0.03	0.34	0.05	0.10	0.33	0.14	0.03	0.37	0.03	0.03	0.34	0.05	0.11	0.32	0.13	0.09
V wind	0.66	0.05	0.02	0.79	0.10	0.17	0.75	0.32	0.06	0.60	0.06	0.02	0.79	0.13	0.22	0.75	0.32	0.21
Bank angle	0.03	0.02	0.03	0.03	0.03	0.02	0.04	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.02	0.03	0.01	0.03
ATC	0.00	0.00	0.00	0.03	0.06	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.06	0.04	0.00	0.00	0.00
Mach number	0.03	0.03	0.02	0.28	0.04	0.03	0.09	0.04	0.03	0.03	0.03	0.03	0.27	0.05	0.05	0.09	0.04	0.02
CAS	0.38	0.04	0.03	0.00	0.00	0.00	0.35	0.15	0.03	0.37	0.03	0.03	0.00	0.00	0.00	0.30	0.10	0.07
VS	0.36	0.15	0.96	0.00	0.00	0.00	0.26	0.18	0.94	0.51	0.14	0.94	0.00	0.00	0.00	0.29	0.14	0.73
LVL FL	0.02	0.03	0.12	0.00	0.00	0.00	0.04	0.03	0.13	0.02	0.03	0.09	0.00	0.00	0.00	0.05	0.03	0.14
LVL dur	0.03	0.02	0.18	0.00	0.00	0.00	0.05	0.04	0.21	0.03	0.04	0.16	0.00	0.00	0.00	0.05	0.03	0.18

Medium WTC	15 minutes look-ahead time									20 minutes look-ahead time								
	Climb			Cruise			Descent			Climb			Cruise			Descent		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	0.09	0.03	0.03	0.05	0.02	0.03	0.05	0.03	0.05	0.10	0.05	0.05	0.06	0.02	0.03	0.05	0.03	0.05
Lapse rate	0.04	0.04	0.02	0.13	0.04	0.03	0.09	0.03	0.04	0.05	0.03	0.04	0.13	0.04	0.03	0.11	0.05	0.05
U wind	0.38	0.06	0.04	0.35	0.07	0.13	0.31	0.14	0.18	0.38	0.10	0.03	0.35	0.07	0.14	0.32	0.11	0.23
V wind	0.63	0.09	0.03	0.80	0.15	0.27	0.78	0.29	0.41	0.67	0.17	0.04	0.81	0.17	0.29	0.79	0.27	0.53
Bank angle	0.03	0.04	0.04	0.03	0.03	0.02	0.03	0.03	0.04	0.03	0.04	0.05	0.03	0.03	0.02	0.04	0.02	0.04
ATC	0.00	0.00	0.00	0.03	0.05	0.05	0.03	0.02	0.02	0.00	0.00	0.00	0.03	0.04	0.06	0.00	0.00	0.00
Mach number	0.03	0.02	0.03	0.26	0.05	0.06	0.08	0.03	0.03	0.05	0.03	0.03	0.26	0.05	0.08	0.07	0.03	0.03
CAS	0.40	0.05	0.03	0.00	0.00	0.00	0.24	0.08	0.13	0.42	0.10	0.03	0.00	0.00	0.00	0.17	0.06	0.15
VS	0.43	0.19	0.92	0.00	0.00	0.00	0.19	0.11	0.21	0.33	0.21	0.93	0.00	0.00	0.00	0.07	0.10	0.40
LVL FL	0.11	0.03	0.05	0.00	0.00	0.00	0.05	0.02	0.10	0.12	0.06	0.05	0.00	0.00	0.00	0.03	0.02	0.04
LVL duration	0.10	0.04	0.04	0.00	0.00	0.00	0.03	0.02	0.07	0.14	0.06	0.03	0.00	0.00	0.00	0.01	0.03	0.05

do not show any correlation between the speed settings and the ATE. Most likely, this is the result of using inaccurate speed input distributions (i.e. speeds are too high) for the chosen light WTC aircraft types. In this case, according to the algorithm employed for the *true* flights' simulations, speeds that are larger than the VMO/MMO of a given aircraft type are replaced with the VMO/MMO. This would result in too many flights having maximum speeds and could impact the correlation between the speed and the ATE.

2. Cross-Track Error

The correlation between the vertical speed and the CTE is detected in the climb phase at 20 minutes look-ahead time for heavy and light WTC aircraft. The mixed correlation (both negative and positive) is weak for heavy WTC and moderate for light WTC aircraft. The vertical speed contributes to the CTE indirectly. It induces the ATE, which in turn contributes to the CTE. The difference in the result for all three WTCs might be due to other inputs' influence on the relationship between the vertical speed and the CTE.

The results show a weak correlation between the wind (v wind component) and the CTE in the descent phase of all three WTCs. The correlation is mixed and it slightly decreases with the look-ahead time. Since aircraft, usually, correct for the wind to stay on course, wind does not directly contribute to the CTE, but via the ATE, similar to the effect of the vertical speed.

The only input that directly contributes to the CTE is the ATC intent. Such instructions as direct-to and heading deviate aircraft from the route, which results in the CTE and induces the ATE. 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. 6 shows the CTE in the cruise phase for 20 minutes look-ahead time. The ATC instruction *0* stands for no ATC instructions, *D* for the direct-to, *H* for the heading, *S* for the speed and *VS* for the vertical speed instructions. In each plot, the data points are separated into two clouds and its clear that direct-to instructions cause large CTEs. The relationship between the CTE and the ATC instructions is non-linear and cannot be characterized by a correlation coefficient.

Table 8 Correlation coefficients for light WTC aircraft

Light WTC	5 minutes look-ahead time									10 minutes look-ahead time								
	Climb			Cruise			Descent			Climb			Cruise			Descent		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	0.05	0.03	0.02	0.03	0.02	0.02	0.04	0.04	0.03	0.05	0.02	0.01	0.03	0.02	0.02	0.03	0.02	0.04
Lapse rate	0.03	0.03	0.02	0.06	0.04	0.03	0.04	0.04	0.02	0.03	0.02	0.02	0.06	0.04	0.02	0.04	0.03	0.03
U wind	0.29	0.06	0.03	0.21	0.07	0.07	0.23	0.16	0.03	0.28	0.07	0.03	0.21	0.09	0.08	0.22	0.16	0.09
V wind	0.47	0.09	0.02	0.48	0.14	0.16	0.47	0.32	0.06	0.46	0.12	0.02	0.48	0.18	0.18	0.48	0.32	0.17
Bank angle	0.02	0.03	0.02	0.02	0.03	0.03	0.02	0.03	0.04	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.03	0.04
ATC	0.00	0.00	0.00	0.04	0.08	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.07	0.07	0.00	0.00	0.00
Mach number	0.02	0.02	0.02	0.06	0.04	0.02	0.04	0.04	0.03	0.02	0.02	0.02	0.06	0.03	0.03	0.03	0.04	0.03
CAS	0.07	0.03	0.04	0.00	0.00	0.00	0.06	0.05	0.03	0.07	0.04	0.03	0.00	0.00	0.00	0.05	0.05	0.04
VS	0.14	0.15	0.96	0.00	0.00	0.00	0.09	0.13	0.94	0.25	0.19	0.94	0.00	0.00	0.00	0.12	0.11	0.72
LVL FL	0.04	0.03	0.11	0.00	0.00	0.00	0.02	0.03	0.10	0.04	0.04	0.07	0.00	0.00	0.00	0.02	0.02	0.12
LVL duration	0.04	0.03	0.19	0.00	0.00	0.00	0.04	0.04	0.20	0.06	0.04	0.17	0.00	0.00	0.00	0.04	0.03	0.14
Light WTC	15 minutes look-ahead time									20 minutes look-ahead time								
	Climb			Cruise			Descent			Climb			Cruise			Descent		
Input\Error	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE	ATE	CTE	AE
Temperature	0.05	0.03	0.02	0.03	0.02	0.02	0.03	0.02	0.05	0.05	0.03	0.01	0.03	0.02	0.02	0.03	0.01	0.04
Lapse rate	0.03	0.02	0.02	0.06	0.04	0.03	0.05	0.04	0.04	0.03	0.02	0.03	0.06	0.04	0.03	0.05	0.04	0.03
U wind	0.28	0.07	0.03	0.22	0.10	0.08	0.22	0.15	0.13	0.27	0.08	0.04	0.22	0.10	0.09	0.21	0.13	0.14
V wind	0.46	0.12	0.03	0.49	0.21	0.20	0.48	0.31	0.27	0.46	0.13	0.03	0.49	0.22	0.21	0.49	0.28	0.30
Bank angle	0.03	0.03	0.03	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.03	0.02	0.03	0.02	0.03	0.03	0.02
ATC	0.00	0.00	0.00	0.03	0.08	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.06	0.09	0.00	0.00	0.00
Mach number	0.02	0.03	0.02	0.06	0.03	0.03	0.03	0.05	0.03	0.02	0.03	0.03	0.06	0.04	0.03	0.04	0.04	0.03
CAS	0.07	0.04	0.04	0.00	0.00	0.00	0.04	0.03	0.03	0.06	0.03	0.04	0.00	0.00	0.00	0.04	0.04	0.04
VS	0.27	0.19	0.93	0.00	0.00	0.00	0.10	0.10	0.32	0.20	0.32	0.94	0.00	0.00	0.00	0.07	0.08	0.44
LVL FL	0.05	0.03	0.04	0.00	0.00	0.00	0.02	0.02	0.10	0.04	0.03	0.04	0.00	0.00	0.00	0.02	0.03	0.07
LVL duration	0.04	0.04	0.09	0.00	0.00	0.00	0.03	0.03	0.06	0.03	0.02	0.04	0.00	0.00	0.00	0.03	0.04	0.08

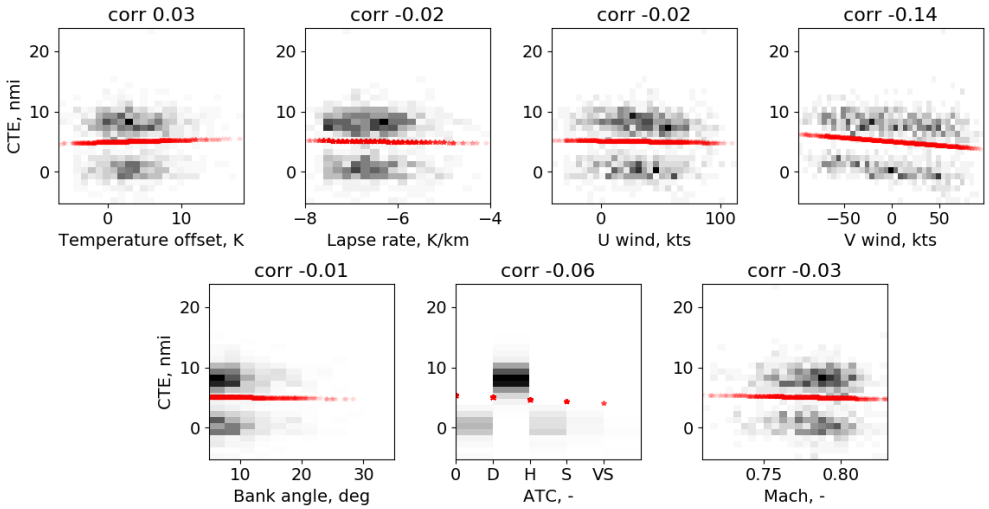
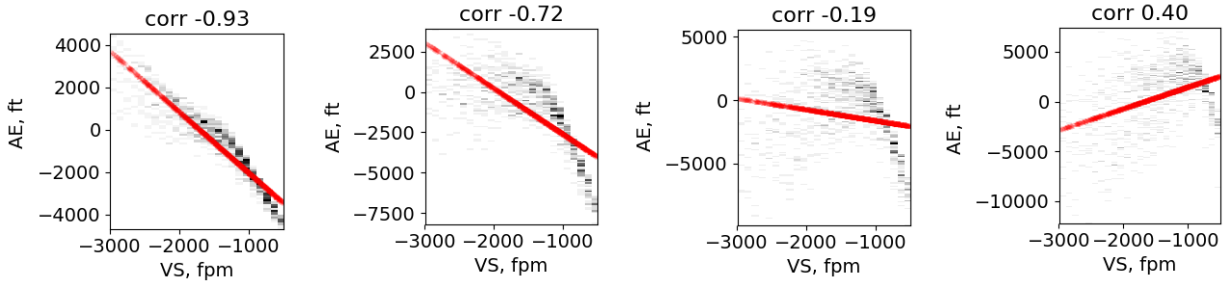


Fig. 6 Cross-track error (CTE) in the cruise phase for 20 minutes look-ahead time.

3. Altitude Error

The AE is very strongly correlated with the vertical speed in the climb phase at all look-ahead times of all three WTCs. The correlation is negative and indicates that with an increase of the vertical speed, the AE becomes more negative (i.e. *predicted* altitude is below the *true* altitude). In the descent phase, the correlation between the vertical speed and the AE decreases with the look-ahead time and the correlation changes from negative towards positive as

illustrated in Fig. 7. With an increase of the look-ahead time, the relationship between the vertical speed and the AE becomes more non-linear and cannot be correctly described with a correlation coefficient. This also resulted in the differences of about 0.1 between the Pearson and Spearman correlation coefficients (Spearman coefficients were lower).



(a) 5 minutes **(b) 10 minutes** **(c) 15 minutes** **(d) 20 minutes**

Fig. 7 Correlation between the vertical speed (VS) and the altitude error (AE) in the descent phase at 5, 10, 15 and 20 minutes look-ahead time.

In Fig. 7 each sub-figure shows two clouds of data points: one is displaying a linear relationship between the vertical speed and the AE and decreases in size with the look-ahead time, whereas the other is a sparse cloud which increases with the look-ahead time. At 5 minutes look-ahead time, a higher rate of descent results in a positive AE (i.e. *prediction* is above the *true* altitude) and, with an increase of the look-ahead time, the errors for higher descent rates become also negative, which is not an intuitive phenomenon. Data analysis showed that such a relationship between the AE and the vertical speed with an increase of the look-ahead time in the descent phase is the result of the different top of descent locations of the *true* and *predicted* trajectories.

Although all look-ahead times include the transition from the cruise to the descent phase (i.e. top of descent), its effect on the AE intensifies with an increase of the look-ahead time. At a high descent rate (e.g. 2,500 to 3,000 feet per minute), the descent phase might take less than 15 minutes (depending on the cruising flight level). As a result, many errors at 15 and 20 minutes look-ahead time were obtained at the beginning of the *true* flight descent phase, while the *predicted* trajectory could have already been in the middle of the descent phase. Furthermore, the top of descent is strongly influenced by the wind, which means that the AE is a result of a combined effect of the vertical speed and the wind conditions. Additionally, the relationship between the AE and the lower descent rates is mostly linear, as expected.

The wind is correlated with the AE in the cruise phase of heavy and medium WTC aircraft and in the descent phase of all three WTCs. No correlation was detected in the climb phase. In the cruise and descent phases, the correlation is present at 15 and 20 minutes look-ahead time and its strength slightly increases with the look-ahead time. This is due to the top of descent location, which is strongly determined by the wind conditions. The error in the top of descent location (i.e. difference between the *prediction* and the *true* top of descent) introduces AE in the cruise and descent phases. The v wind component has a stronger correlation with the AE than the u wind component, similar to the results for the ATE.

C. Error Variance

The variance of the prediction errors can be attributed to inputs by computing the coefficient of determination r^2 from the correlation coefficients and is further presented for each flight phase separately.

1. Climb Phase

On average, 87% of the ATE variance can be explained in the climb phase of heavy and medium WTC aircraft. The wind conditions contribute to about 50%, whereas the vertical speed and the CAS to around 18% and 19% respectively. Only about 32% of the ATE variance can be explained for light WTC aircraft. 29% are attributed to the wind conditions and 3% to the vertical speed. 89% of the AE variance can be attributed to the vertical speed, whereas the CTE variance is unexplained.

2. Cruise Phase

In the cruise phase of heavy WTC aircraft, 70% of the ATE variance can be explained by the wind conditions. For medium WTC aircraft, about 82% of the ATE variance is explained. 75% is attributed to the wind conditions and 7% to the Mach number speed settings. Only 23%, attributed to the wind conditions, of the ATE variance can be explained for light WTC aircraft. Essentially no variance of the CTE and the AE can be explained in the cruise phase.

3. Descent Phase

About 79% of the ATE variance can be explained in the descent phase of heavy and medium WTC aircraft. On average, 67% is attributed to the wind conditions, 5% to the vertical speed and 7% to the CAS. Only around 23% of the ATE variance can be explained for light WTC aircraft, which is attributed to the wind conditions. Around 8% of the CTE variance is explained by the wind conditions for all three WTCs. On average, 51% of the AE variance can be explained in the descent phase, where about 41% is attributed to the vertical speed and about 10% to the wind conditions.

VI. Discussion

This section discusses how general the obtained results are and what are the possible sources of the differences between the results obtained for heavy, medium and light WTC aircraft. Furthermore, the difficulty of determining the correlation between the ATC intent and the prediction errors is addressed as well as how the employed method for errors' computation might affect the results. Finally, the similarities between the results of the previous studies and this study are outlined.

A. Results Generalization

The results obtained in this study are considered to be generic in terms of the estimated ranges of the prediction errors with look-ahead time and the identified influential inputs. The trajectory simulator/predictor used in this study is comparable with a typical TP. The main difference between this study trajectory simulator/predictor and a typical operational TP is how the vertical speed is obtained. In this study, it is sampled from a distribution function, whereas the majority of operational TPs derive the vertical speed from the equations of a point-mass model.

The obtained results for the AE are influenced by the choice of the vertical speed settings used for predictions. When a different value of the vertical speed is used for predictions, it reflects on the AE distribution (i.e. for a box plot the location of the median is shifted) and on the strength of its correlation with the AE in the descent phase. The latter is not very meaningful since the relationship between the vertical speed and the AE in the descent phase becomes more non-linear with an increase of the look-ahead time. Furthermore, the change of the vertical speed settings also affects the strength of the correlation between the vertical speed and the ATE and the correlation between the AE and other inputs (e.g. temporary level-offs).

Additionally, there is a relationship between the wind and flight route predominant directions. When the majority of the routes is mostly perpendicular to the predominant wind direction, the correlation between the wind and the ATE would be weaker compared to when the majority of the routes is aligned with the predominant wind direction. In this study, routes were chosen to have different directions with respect to the predominant wind direction. Therefore, it is not expected that a larger number of routes would significantly influence the strength of the correlation between the wind conditions and ATEs.

Based on the above-mentioned aspects of the correlation between the prediction errors and such inputs as vertical speed and wind conditions, inputs could be additionally classified as *weak* and *strong*. For instance, the vertical speed is a *strong* input since it is correlated with the AE regardless of the flight route. The wind conditions, on the other hand, is a *weak* input since the correlation between the wind and the ATE is conditional on the wind and route predominant directions.

B. Wake Turbulence Category

The results for heavy and medium WTC aircraft are similar in both trends and magnitudes. The results for light WTC aircraft, however, are significantly different from the heavy and medium WTC in both, trends and magnitudes. Most likely, these differences are the result of the inaccurate speed input distributions for light WTC aircraft. In many cases, the sampled speeds were too large for a given aircraft type and had to be replaced with the maximum operating

speed. This could have influenced the resultant ATEs and most likely diminished the correlation between the speed settings and the ATE.

C. ATC Intent

Although, according to the correlation coefficients in this study, the ATC instructions were not correlated with any prediction errors, this does not reflect the reality. In this study, the ATC instructions were modeled in a way that did not allow to determine meaningful values of the correlation coefficients. Nonetheless, it is logical to conclude that the majority of the large CTEs in the cruise phase is the result of the ATC intent. Furthermore, ATC instructions also have an effect on the ATE (i.e. through the CTE) and on the AE (i.e. through the ATE).

D. Error Computation

The employed method for horizontal error (ATE and CTE) calculation can sometimes result in very large incorrect magnitudes of errors when two compared points are located on the different segments of a route with turns. The method takes into account the direction of the route segment where the predicted position is located, but it does not take into account the actual along-track distance between the predicted and true positions. Therefore, the errors are computed with respect to the segment where the predicted position is located.

E. Relation to Results of Previous Research

The results obtained in this study on the importance of the TPs inputs correspond to the conclusions made in previous research. Barring the effect of the ATC intent, the vertical speed settings (usually as a function of the aircraft mass and thrust settings), the wind conditions and the CAS and Mach number speed settings were determined to be the most influential inputs. The rest of the studied inputs did not show much correlation with the prediction errors.

VII. Conclusion

This paper estimated the prediction uncertainty up to 20 minutes look ahead-time and the correlation between the TP inputs and the prediction errors for heavy, medium and light WTC aircraft. For each WTC, 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 real data. Studied inputs included the bank angle settings, the speed settings (CAS and Mach number), the vertical speed settings, the temporary level-offs in the climb and descent phases, the air temperature, the lapse rate, the wind and the ATC intent. Three prediction errors, namely ATE, CTE and AE, were computed for 5, 10, 15 and 20 minutes look-ahead time.

All three errors tend to increase with the look-ahead time. When outliers are not considered, at 20 minutes look-ahead time, the ATE can reach up to 50 nmi for heavy and medium WTC aircraft and up to 100 nmi for light WTC aircraft. The CTE is mostly less than 3 nmi for heavy and medium WTC aircraft and 8 nmi for light WTC aircraft. Furthermore, the AE may be as large as 25,000 feet. The difference between the results for light and heavy and medium WTC is, most likely, the result of the inaccurate speed input distribution for light WTC aircraft. Based on these values of the prediction errors, a conclusion can be made that it is not worthwhile to employ the decision support tools that operate in the time interval up to 20 minutes look-ahead time (e.g. MTCD) until the prediction accuracy is significantly improved.

Out of all studied input parameters, only three inputs, namely the wind, the vertical speed and the speed settings, showed a correlation with the prediction errors. On average, 65% and 25% of the ATE variance can be attributed to the wind for heavy and medium and light WTC aircraft, respectively. The wind conditions explain about 8% of the CTE variance and 10% of the AE variance in the descent phase. The vertical speed is mainly correlated with the AE in the climb and descent phases. It explains 89% of the AE variance in the climb phase and about 41% in the descent phase. The vertical speed explains 18% and 5% of the ATE variance in the climb and descent phase, respectively, of heavy and medium WTC aircraft. It explains 3% of the ATE variance in the climb phase of light WTC aircraft. The CAS explains around 19% and 7% of the ATE variance in the climb and descent phase, respectively, of heavy and medium WTC aircraft. The Mach number speed settings is attributed to 7% of the ATE variance in the cruise phase of the medium WTC aircraft.

Although the ATC instructions significantly contribute to the CTE and ATE in the cruise phase, 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.

It has been identified that the choice of the vertical speed settings used for predictions influences the prediction errors' distribution and the strength of the correlation between some inputs and errors. Future research could be conducted where simulations are performed with different input settings used for predictions. The TP model could be improved to include the vertical speed variation with altitude. Additionally, an adaptive prediction algorithm, where all observed aircraft states are taken into account for prediction reinitialization, is expected to result in a decrease of some prediction errors. Furthermore, the non-linear relationships between the inputs and the errors could be investigated as well as the combined effect of several inputs on the prediction errors.

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] 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.

[4] 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.

[5] 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.

[6] Mondoloni, S., Paglione, M., and Green, S. M., "Trajectory modelling accuracy for air traffic management decision support tools," *International Congress of Aeronautical Sciences*, 2002.

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

[8] 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.

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

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

[11] 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.

[12] 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.

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

[14] 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.

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

[16] 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.

[17] 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.

[18] Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T., "Sensitivity analysis of environmental models: A systematic review with practical workflow," *Environmental Modelling and Software*, Vol. 79, 2016, pp. 214–232. doi:10.1016/j.envsoft.2016.02.008.

[19] Rudnyk, J., Ellerbroek, J., and Hoekstra, J. M., "Derivation of Trajectory Predictor Input Distributions from Observed Data," *AIAA Aviation Technology, Integration, and Operations Conference*, Atlanta, GA, 2018. Accepted.

[20] 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.

[21] Kalos, M. H., and Whitlock, P. A., *Monte Carlo Methods*, 2nd ed., WILEY-VCH Verlag GmbH & Co, KGaA, Weinheim, 2008.

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

[23] Dekking, F., Kraaikamp, C., Lopuhaa, H., and Meester, L., *KANSTAT. Probability and Statistics for the 21st Century*, Delft University of Technology, 2004.

[24] Matsumoto, M., and Nishimura, T., "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator," *ACM Transactions on Modeling and Computer Simulation*, Vol. 8, No. 1, 1998, pp. 3–30. doi: 10.1145/272991.272995.