# Modeling Aircraft Performance Parameters with Open ADS-B Data

Junzi Sun, Joost Ellerbroek, Jacco Hoekstra

Control and Simulation, Faculty of Aerospace Engineering

Delft University of Technology

Delft, The Netherlands

### Abstract

Open access to flight data from ADS-B (Automatic Dependent Surveillance Broadcast) has provided researchers more insights for air traffic management than aircraft tracking alone. With large quantities of trajectory data collected from a wide range of different aircraft types, it is possible to extract accurate aircraft performance parameters. In this paper, a set of more than thirty parameters from seven distinct flight phases are extracted for common commercial aircraft types. It uses various data mining methods, as well as a maximum likelihood estimation approach to generate parametric models for these performance parameters. All parametric models combined can be used to describe a complete flight that includes takeoff, initial climb, climb, cruise, descent, final approach, and landing. Both analytical results and summaries are shown. When available, optimal parameters from these models are also compared with the Base of Aircraft Data and Eurocontrol aircraft performance database. This research not only presents a comprehensive set of methods for extracting different aircraft performance parameters but also provides a first part of open-source parametric performance models that is ready to be used by the ATM community.

## Nomenclature

| | | |
|---|---|---|
| **to** | Aircraft takeoff phase | |
| **ic** | Aircraft initial climb phase | |
| **cl** | Aircraft climb phase | |
| **cr** | Aircraft cruise phase | |
| **de** | Aircraft descent phase | |
| **fa** | Aircraft final approach phase | |
| **ld** | Aircraft landing phase | |
| | | |
| $a_*$ | Acceleration or deceleration | $m/s^2$ |
| $d_*$ | Distance (takeoff or landing) | $km$ |
| $V_{h,*}$ | Vertical speed | $m/s$ |
| $V_{cas,*}$ | Calibrated air speed | $m/s$ |
| $R_*$ | Flight range | $km$ |
| $M_*$ | Mach number | - |
| $H_*$ | Aircraft altitude | $km$ |

## 1    Introduction

Aircraft performance models (APM) for ATM researches are commonly built based on manufacturer performance reports and/or recorded flight data. However, it is difficult for each researcher to

obtain flexible access to these models and related data, as they are either costly or have certain license limitations. Even the most commonly used APM, Base of Aircraft Data family 3 (BADA3), has its license limitations [1]. BADA family 4 comes with even more strict license agreements, which make it very difficult to perform studies that can be easily reproduced and compared within the research community.

Thanks to the ICAO mandates on aircraft ADS-B equipment installation by 2020, an increasing number of commercial aircraft are being equipped with this capability. At the same time, many ground receivers are being installed around the world. Open flight data has become more and more accessible for every ATM researcher. Certain flight data such as position, altitude, velocity, and vertical speed of common aircraft types are broadcast and received constantly around the world by these ground receiver networks. These large amounts of scatted flight data provide us a new way of modeling aircraft performance.

By using carefully designed data mining algorithms and statistical analyses, operational performance and limitation parameters of aircraft can be obtained from ADS-B flight data. In this paper, we are focusing on extracting complete operational and limitation models solely based on this open data approach.

However, the biggest challenge is how to make use of these large quantities of data. As an aircraft performs differently in each flight phase, flight data need to be segmented accordingly. Machine learning and data mining methods, as proposed by Sun et al., have provided some efficient ways to process these loosely connected data points into organized segments based upon trajectory and respective flight phases [2]. As a continuation of this research, output data generated is further divided into detailed segments according to the model definitions that are proposed in this paper.

In real flights, none of the two trajectories are identical. In every aircraft type, different levels of performance deviations are always present, and these values are often differ, even within the same phase of a flight. The main causes include aircraft mass, thrust settings, wind, and flight procedures. Errors in data can also contribute to these deviations. However, there is an underlying statistical model that can be used to describe these performance parameters. To identify the best mode, maximum likelihood estimators can be employed.

The remainder of this paper is structured as follows. Section two introduces the model parameters, definitions, and maximum likelihood estimations for three different distribution functions. Section three explains the data and pre-processing methods. Section four is focused on methods and algorithms for constructing parameters of each flight phase. Detailed results on a single aircraft type and summaries of 17 different aircraft types are shown in section five. Finally, discussions and conclusions are addressed in sections six and seven.

## 2    Model definitions

### 2.1    Parameters

Similar to the flight phase definitions from ICAO Common Taxonomy Team [3], this research aims to derive performance parameters from the following flight phases: takeoff, initial climb, climb, cruise, descent, final approach, and landing. The significance of the performance indicators differs from one flight phase to another. For example, vertical velocity is only of interest during climb and descent, while during takeoff and landing, the runway distance, acceleration, and boundary velocities are of interest. Specifically, during the climb (or descent) phase, a constant CAS/Mach profile is extracted. Throughout the paper, the following parameters, shown in Table 1, are defined and obtained accordingly.

### 2.2    Estimation method

The maximum likelihood estimation (MLE) [4] is used to estimate the best unbiased value of a parameter (or a vector of parameters) based on observations. First, $N$ observations $X$ : $\{x_1, x_2, \ldots, x_n\}$ are drawn from an unknown probability distribution function $p(\cdot)$ and characterized by a parameter vector $\phi$, denoted $\{p(\cdot|\phi), \phi \in \Phi\}$. Then, assuming the sample observations are independent, the joint probability density, also known as the likelihood function of $\phi$, denoted $\mathcal{L}(\phi; \cdot)$, is expressed as follows:

Table 1: Performance model parameters

| Flight phase | Performance parameters |
|---|---|
| Takeoff | $V_{lof}, d_{tof}, \bar{a}_{tof}$ |
| Initial Climb | $V_{cas,ic}, V_{h,ic}$ |
| Climb | $R_{top,cl}, V_{cas,cl}, M_{cl}, H_{cas,cl}, H_{mach,cl}, V_{h,pre\text{-}cas,cl}, V_{h,cas,cl}, V_{h,mach,cl}$ |
| Cruise | $R_{cr}, R_{max,cr}, H_{cr}, H_{max,cr}, M_{cr}, M_{max,cr}$ |
| Descent | $R_{top,de}, M_{de}, V_{cas,de}, H_{mach,cl}, H_{cas,cl}, V_{h,mach,de}, V_{h,cas,de}, V_{h,post\text{-}cas,de}$ |
| Final Approach | $V_{cas,fa}, V_{h,fa}$ |
| Landing | $V_{app}, d_{brk}, \bar{a}_{brk}$ |

$$\mathcal{L}(\phi\,;\,x_1,\ldots,x_n) = p(x_1,x_2,\ldots,x_N|\phi) = \prod_{i=1}^{N} p(x_i|\phi) \tag{1}$$

The maximum likelihood estimate of $\phi$, denoted $\hat{\phi}_{\text{MLE}}$, is the $\phi$ vector that maximizes the likelihood. To simplify the calculation, the log-likelihood function is used:

$$
\begin{aligned}
\hat{\phi}_{\text{MLE}} &= \underset{\phi \in \Phi}{\arg\max} \prod_{i=1}^{N} p(x_i|\phi) \\
&= \underset{\phi \in \Phi}{\arg\max} \left[ log(\prod_{i=1}^{N} p(x_i|\phi)) \right] \\
&= \underset{\phi \in \Phi}{\arg\max} \sum_{i=1}^{N} log[p(x_i|\phi)]
\end{aligned}
\tag{2}
$$

## 2.3 Probability distribution functions

Three continuous probability distribution functions (PDF) are assumed for each of the performance parameters, which are Normal, Gamma, and Beta distributions. Then MLE is applied to obtain the best estimates $\hat{\phi}_{\text{MLE}}$ for each PDF from sample data.

### 2.3.1 Normal distribution

For a Normal distribution, the density function is expressed:

$$p(x|\phi) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

where the unknown PDF parameters $\phi : (\mu, \sigma^2)$ represent the mean and variance. The MLE estimator $\hat{\phi}_{\text{MLE}}$ can easily be found for a Normal distribution, with the mean and variance as the observation data:

$$\hat{\mu} = \quad \bar{X} \tag{4}$$

$$\hat{\sigma}^2 = \quad \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{X})^2 \tag{5}$$

### 2.3.2 Gamma distribution

For a standardized Gamma distribution with $y \in (0, +\infty)$, the density function is expressed:

$$p(y|\phi) = \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y} \tag{6}$$

where the $\alpha$ ($\alpha > 0$) represents the shape of the distribution.

In a general form where $x \in (-\infty, +\infty)$, the location ($\mu$) and scale ($k$) parameters are introduced:

$$p(y|\phi) = \frac{1}{k \cdot \Gamma(\alpha)} y^{\alpha-1} e^{-y}$$
$$y = \frac{x - \mu}{k} \tag{7}$$

The MLE estimator $\hat{\phi}_{\text{MLE}} : (\hat{\alpha}, \hat{\mu}, \hat{k})$ does not have a close form solution, but it can be solved numerically [5].

### 2.3.3 Beta distribution

For a standardized Beta distribution with $0 < x < 1$, the density function is:

$$p(x|\phi) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \tag{8}$$

where $B(\alpha, \beta)$ is defined as:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \tag{9}$$

In a general form where $x \in (-\infty, +\infty)$, the location ($\mu$) and scale ($k$) parameters are also introduced:

$$p(y|\phi) = \frac{1}{k \cdot B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}$$
$$y = \frac{x - \mu}{k} \tag{10}$$

Similar to Gamma distribution, it also does not have a close-form solution for the optimal $\hat{\phi}_{\text{MLE}} : (\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{k})$, which needs to be solved numerically [6].

### 2.3.4 Model fitness evaluation

After all possible optimal probability distribution models are computed, the best model of all three need to be identified. The Kolmogorov-Smirnov (KS) Test [7] is applied to obtain such a model. This test compares the cumulative distribution function (CDF) of a statistical model, denoted $F(x)$, against the empirical distribution function (EDF) from the data, denoted $F_n(x)$, and then generates the KS statistic $D$:

$$D = \max_{x} |F_n(x) - F(x)| \tag{11}$$

which represents the largest distance between $F(x)$ and $F_n(x)$. Of all three KS statistics $D_{\mathcal{N}}$, $D_{\Gamma}$, and $D_B$, the probability distribution that yields the minimum $D$ will be selected. It is worth noting that the KS-test shall be performed on a test dataset that is not been used for for constructing the models. In this paper, the experiment data is split in equal two parts (for fitting and testing) to construct and validate the best parametric model.

An example data set of one parameter (takeoff acceleration) is shown in Fig. 1, where the CDF of all three models obtained from MLE are displayed against the EDF computed from the data. The best model (beta) is obtained by comparing $D$. The final selected PDF are illustrated in the third plot of Fig. 9.
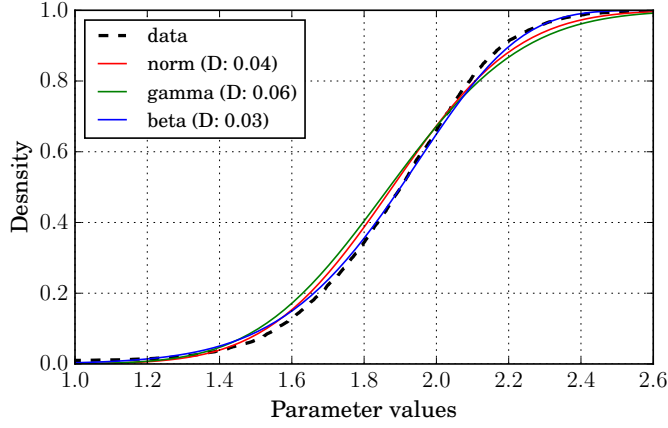
4

Figure 1: Comparison of three CDF with data EDF

## 2.4 Parameter expression

For each parameter in Table 1, the following expression is constructed:

$$\{\hat{\psi}, \psi_{min}, \psi_{max}, *pdf\} \tag{12}$$

where the first three values are the optimal, minimum, and maximum values of such a parameter. The *min* and *max* values are obtained from a confidence intervals: 0.8 for most of the velocity parameters, 0.98 for range parameters, and 0.9 for other parameters. The probability density functions are provided as *pdf* for the cases where parameter values need to be drawn from a probability distribution. Depending on the KS test, optimal *pdf* is expressed as follows:

$$*pdf = \begin{cases} ['norm', \mu, \sigma] & for\ x \sim \mathcal{N} \\ ['gamma', \alpha, \mu, k] & for\ x \sim \Gamma \\ ['beta', \alpha, \beta, \mu, k] & for\ x \sim B \end{cases} \tag{13}$$

For deterministic calculations, $\hat{\psi}$ is recommended. However, any value from the interval $[\psi_{min}, \psi_{max}]$ is also applicable. For simulations that study the uncertainties, multiple values of a parameter can be drawn from the distribution, which would resemble the real world air traffic situation.

It is important to note that certain errors may be present in flight data, such as inaccurate estimated wind speed or errors contained in GPS positions. Trajectory data need to be well processed in the first place.

# 3 Data preparation

## 3.1 Data source

For this research, the input data are primarily based on ADS-B messages that are broadcast by airplanes through Mode-S transponders. However, even with good installation and placement, a single ground based receiver only has a maximum reception range of around $250nm$ ($\sim 500km$). Considering Mode-S line-of-sight availability, it is not possible to capture large quantities of completed flight data with a single ground station. It is especially challenging when dealing with long range aircraft. Thanks to networks of ADS-B ground receivers, such as FlightRadar24 [8], it is possible to gain access to a much larger scale of flight data collected from other ground stations from contributors around the world.

It is also worth noting that in addition to ADS-B data, using the same receiver setup, aircraft positions, and velocities can be obtained from multiple ground stations using Mode-S multilateration. The technology is also deployed by some ground receivers within the FlightRadar24 network.

It is useful for those aircraft that are not yet equipped with ADS-B out equipment. However, the availability is limited to areas where ATC is presented and not useable when aircraft are close to the ground.

## 3.2 Trajectory data process

Data collected from ADS-B are usually scattered. Previous machine learning methods in [2] have made it possible to extract flight trajectory easily. However, this method can only segment flights into four different phases: on-ground, climb, cruise, and descent.

For the research needs of this paper, further processing of those segments is employed. Firstly, an altitude threshold is applied on the climb and descent trajectory to extract initial climb and final approach. Secondly, a cruise buffer is introduced to distinguish the level-flight segments in climb or descent phases from cruise. Lastly, an evaluation process is used to examine the data of all flight segments, ensuring a certain level of completeness and continuity in the given time series data.

The ground velocity information from ADS-B measurements are also integrated with meteorological data using interpolation models described in [2] to produce the true airspeed (TAS) of aircraft at any given location. TAS is then converted to calibrated airspeed (CAS) and Mach number under standard atmospheric conditions.

# 4 Model construction methods

After these flight trajectories are sorted and segmented by proper flight phases, they are ready to be used for constructing desired model parameters. For each aircraft type, at least five thousand trajectories are used to give a good level of confidence in the model. This section discusses in detail the methods for extracting these parameters from the trajectory data.

## 4.1 Takeoff

During the takeoff phase, the parameters of most interest are takeoff distance ($d_{tof}$), average acceleration ($\bar{a}_{tof}$), and lift-off speed ($V_{lof}$). To overcome large noise in velocity measurements during the takeoff, a spline filter is applied to obtain a smoothed velocity sample set.

After the smoothed time series data is computed, distance parameters can be derived from aircraft surface position at the starting and ending positions of the takeoff, using the spherical law of cosines:

$$
\begin{aligned}
d_{tof} &= d_{1,N} \\
&= R \cdot \arccos[\sin\gamma_1 \cdot \sin\gamma_N \\
&\quad + \cos\gamma_1 \cdot \cos\gamma_N \cdot \cos(\lambda_N - \lambda_1)]
\end{aligned}
\tag{14}
$$

where $\gamma$ and $\lambda$ represent the latitude and longitude in radians.

Average acceleration $\bar{a}_{tof}$ can be obtained accurately from velocity measurements:

$$
\bar{a} = \underset{\bar{a}}{\arg\min} \sum_{i=1}^{N} \left( \bar{a} \cdot (t_i - t_0) - v_i \right)^2
\tag{15}
$$

Compared to the previous two parameters at takeoff, lift-off speed is more complicated to estimate, due to the low data update rate. There is usually a gap of several seconds between the last on-ground data and the first in-air data.

To estimate the exact moment of lift-off, as shown in Fig. 2, a bi-directional extrapolation is used. Firstly, the vertical rate $V_h$ and time stamp at the first in-air data are used to estimate the time of lift-off $t_{lof}$. Combining this result with previously calculated $\bar{a}_{tof}$, the lift-off speed can be obtained as follows:
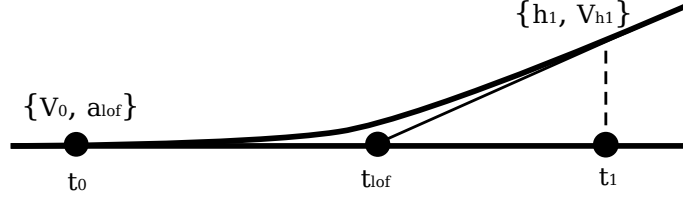
Figure 2: Estimate the lift-off moment

$$t_{lof} = t_1 - \frac{h_1}{V_{h_1}} \tag{16}$$

$$V_{lof} = V_0 + \bar{a}_{tof} \cdot (t_{lof} - t_0) \tag{17}$$

## 4.2 Initial climb

The initial climb phase is defined as the segment from $35\ ft$ until the aircraft reaches around $1500\ ft$. There are several major configuration changes (landing gear and flaps) during this short period of time that can affect the performance. However, the initial climb segment lasts for a short time in respect to the entire flight path. Thus, it suffers from relatively low data samples comparable to the takeoff phase.

The parameters to be studied are aircraft calibrated airspeed ($V_{cas,ic}$) and vertical rate ($V_{h,ic}$). Both parameters can be computed directly from ADS-B data.

## 4.3 Climb

The climb segment starts when the aircraft reaches clean configuration and lasts until the moment when it reaches cruise altitude. As a common practice, aircraft first accelerate to a target CAS and then fly according to this constant CAS. As the altitude of an aircraft increases, the speed of sound decreases, which results in an increasing Mach number. When a certain Mach number is reached, an aircraft will fly according to this constant Mach number until its cruising altitude. During the Mach climb segment, a decreasing CAS will be observed due to the decreasing air density. The climb CAS profile can be generalized in Fig. 3.
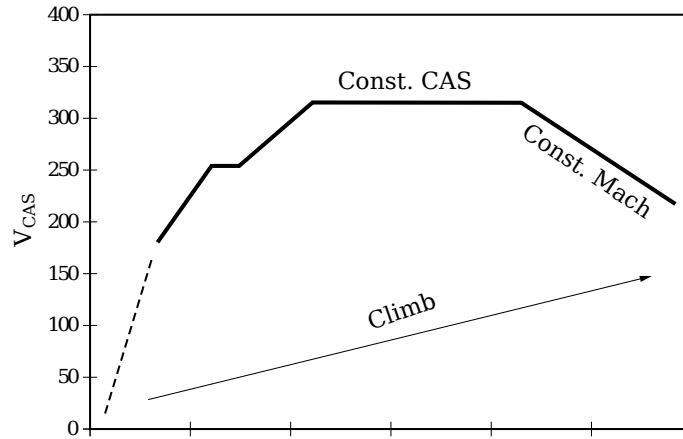


Figure 3: Standard climb profile

The challenge is to identify the crossover of this constant CAS/Mach climb. Knowing the general profile of CAS and Mach number during the climb, it is possible to design a two-step and two-piecewise estimator for extracting this feature from CAS and Mach number profiles.

The estimator consist of two parts: an increasing (quadratic or linear) segment and a constant segment. The quadratic increasing part is designed so that it resembles the general observation

7

of velocity increase with a decreasing acceleration. The constant segment approximates constant CAS or Mach values. The estimator is expressed as follows:

$$f_{mach}(t) \quad = \quad \begin{cases} k_1 \cdot (t - t_{mach})^2 + y_{mach} & t \leq t_{mach} \\ y_{mach} & t \geq t_{mach} \end{cases} \tag{18}$$

$$f_{cas}(t) \quad = \quad \begin{cases} -k_2 \cdot (t - t_{cas})^2 + y_{cas} & t \leq t_{cas} \\ y_{cas} & t_{cas} \leq t \leq t_{mach} \end{cases} \tag{19}$$

To illustrate this knowledge retrieving process, this two step piecewise least squared fitting can be described in the diagram of Fig. 4.
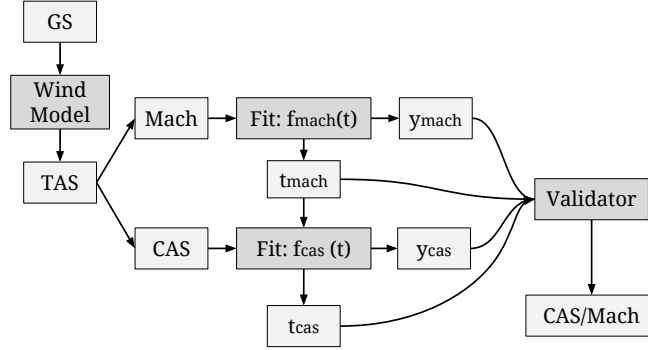


Figure 4: Constant CAS/Mach climb identification process

The entire climbing phase in terms of Mach number can be divided in two parts (increasing and constant). The first estimator applies on the entire climb data and uses the least-square fitting to extract the best CAS/Mach transition time $t_{mach}$ and constant Mach climb number $y_{mach}$.

The second estimator uses $t_{mach}$ as cutoff time, fitting the remainder of the time series where Mach number is increasing. This similarly structured estimator is used to obtain the constant CAS speed $y_{cas}$ and its starting time $t_{cas}$.

The final validation step takes into account $k_1$, $k_2$, and all four output parameters: $t_{cas}$, $y_{cas}$, $t_{mach}$, and $y_{mach}$ to ensure the following boundary conditions:

$$t_0 < t_{cas} < t_{mach} < t_N$$
$$CAS_{min} < y_{cas} < CAS_{max}$$
$$Mach_{min} < y_{mach} < Mach_{max}$$
$$k_1 > 0$$
$$k_2 > 0 \tag{20}$$

To illustrate this entire process, this method is applied on one flight and shown in Fig. 5. The green and red colors represent the Mach and CAS profile during the climb, both with a constant part that has been identified by the model.

After $t_{cas}$ and $t_{mach}$ are determined, transition altitudes $H_{cas,cl}$, $H_{mach,cl}$, at which the constant CAS and constant Mach climbing commence can be obtained. Vertical rates $V_{h,pre\text{-}cas,cl}$, $V_{h,cas,cl}$, $V_{h,mach,cl}$ that correspond to these three segments can also be computed. In addition, range from takeoff to the top of climb $R_{top,cl}$ is calculated.

## 4.4 Cruise

Aircraft generally cruise at selected optimal altitudes. Parameters that are of interest for performance are extracted from the trajectories, such as the standard cruise altitude $H_{cr}$ and Mach
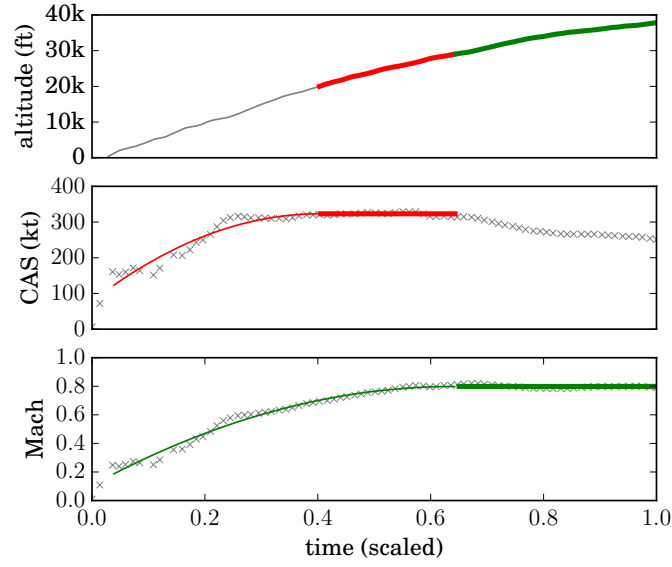
Figure 5: Constant CAS/Mach climb identification example

number $M_{cr}$. The limitations of aircraft such as cruise ceiling $H_{max,cr}$ and maximum cruise Mach number $M_{max,cr}$ can also be found. In addition, typical cruise range $R_{cr}$ is extracted as an operational reference parameter.

Since aircraft rarely flight directly between original and destination airports, the flight range cannot be calculated as the great circle distance between origin and destination. On the other hand, due to the measurement noise inherited from on-board GPS receivers, position reports in ADS-B can contain errors. When integrating positions along the trajectory, the accumulated error grows even larger. Hence, a filter is first implemented, and then the entire trajectory is re-sampled for integrating the cruise range. Such a process can be illustrated in Fig. 6, where the filtered trajectory is shown in red as compared to real position reports in blue and the direct path in dashed black line.
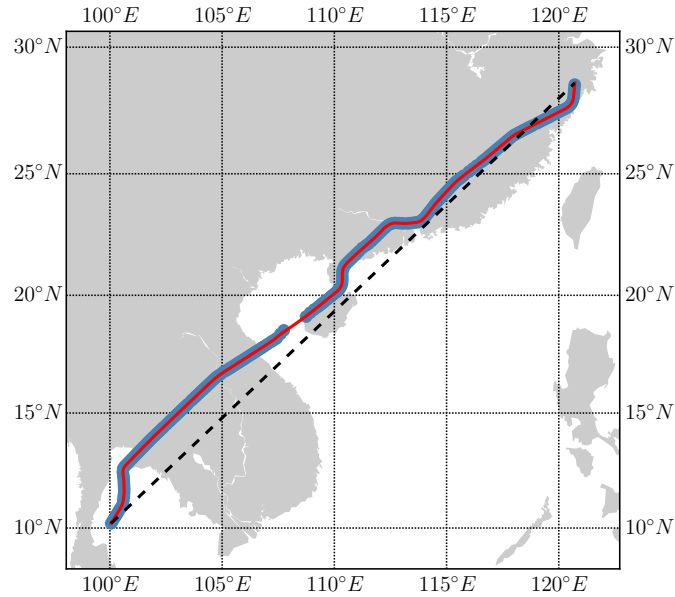


Figure 6: Cruise range spline filtering

## 4.5 Descent

The descent phase of the aircraft is comparable to the climb phase. From the top of descent, the aircraft undergoes a constant Mach and constant CAS descent segment before reaching the approach altitude, as illustrated in Fig. 7.
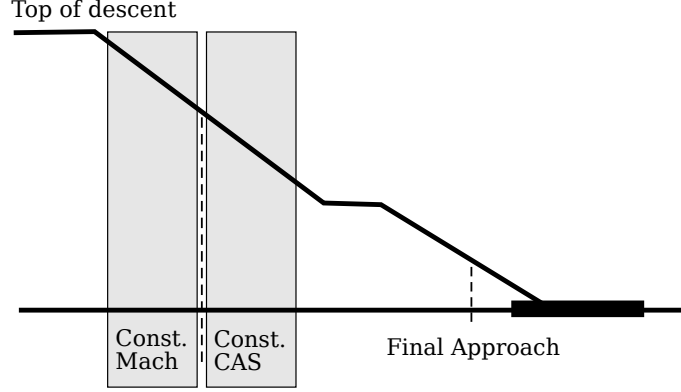


Figure 7: Standard descent profile

The essential parameters to be modeled are: range from top of descent to destination $R_{tod,de}$, constant Mach $v_{mach,de}$, constant CAS $v_{cas,de}$, crossover altitudes $H_{mach,de}$ and $H_{cas,de}$, and vertical speed $V_{h,mach,de}$, $V_{h,cas,de}$, and $V_{h,post\text{-}cas,de}$.

Similar to climb, the constant Mach/CAS descent performance can be modeled by piecewise estimators. It is possible to use the same process as described in Fig. 4. The Mach profile is described using two linear pieces. The CAS profile consists of a linear and quadratic piece due to the high non-linearity in speed at the final approach segment. The model can be described as follows:

$$f_{mach}(t) = \begin{cases} y_{mach} & t \leq t_{mach} \\ -k_1 \cdot (t - t_{mach}) + y_{mach} & t \geq t_{mach} \end{cases} \tag{21}$$

$$f_{cas}(t) = \begin{cases} -k_2 \cdot (t - t_{cas})^2 + y_{cas} & t \leq t_{cas} \\ y_{cas} & t_{cas} \leq t \leq t_{mach} \end{cases} \tag{22}$$

Similarly, the result of such model as applied on one flight is shown in Fig. 8.

During descent, it is common for aircraft to follow a certain Continuous Descent Approach (CDA), eliminating level-flight segments so as to reduce fuel consumption and engine noise. Such an approach affects the result of descent performances. This factor is discussed in section 6.3 in this paper.

## 4.6 Final Approach

Due to different control procedures at each airport, it is not easy to generalize the entire approach segment solely based on flight data. However, the final approach segment from around 1000 ft until landing can be modeled.

The segment of final approach represents the end of a descent, where aircraft operate at a constant airspeed and rate of descent. The approaching speed $v_{cas,fa}$ and Rate of descent $v_{h,fa}$ are modeled.

## 4.7 Landing

The landing model is comparable with the take-off model. Parameters such as: approaching velocity $V_{app}$, braking distance $d_{brk}$, and deceleration $a_{brk}$ of the aircraft are modeled similar to the takeoff phase. Approaching speed $V_{app}$ can be observed from the last in-air velocity. Similar
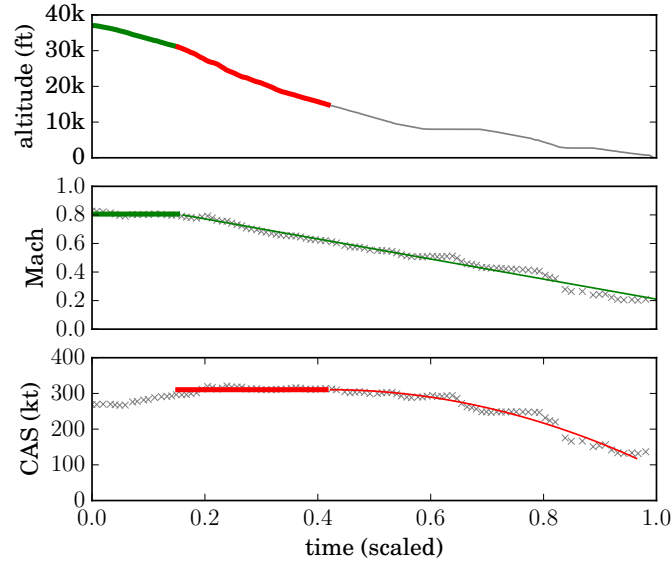
Figure 8: Constant CAS/Mach descent identification example

to takeoff, breaking distance can be calculated using Equation 14, and average deceleration can be calculated from Equation 15.

# 5    Results

A set of 17 common aircraft types are studied in this research. Sufficient data are collected for all aircraft. For each aircraft type, around 5000 sets of data per flight phase are considered. In order to better illustrate the modeling of each individual parameter, detailed results from a single aircraft type (Airbus A320) are described according to the previous model specification. Results for other aircraft types are shown in a combined table.

## 5.1    Airbus A320

For each parameter, all three probability density functions (Normal, Gamma, and Beta) are applied using MLE. The best model is chosen and represented by different colors in all figures, shown in red, green, and blue respectively. $\hat{\psi}$, $\psi_{min}$, and $\psi_{max}$ are also marked in each of the plots accordingly.

### 5.1.1    Take-off

Three performance parameters $d_{tof}$, $V_{lof}$, and $\bar{a}_{tof}$ are shown in Fig. 9, where the optimal values of these parameters are $1.68km$, $84m/s$ ($\sim 163kt$), and $1.88m/s^2$.
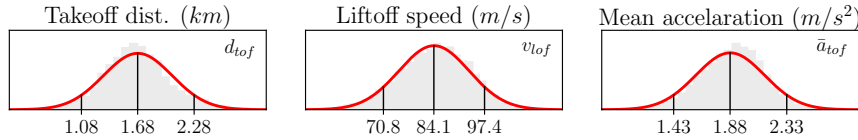


Figure 9: Take-off parameters

### 5.1.2    Initial climb

Two performance parameters $V_{cas,ic}$ and $V_{h,ic}$ of the initial climb, up to the altitude of 1500ft, are shown in the first two plots of Fig. 10, where the most likely values are $81m/s$ ($\sim 157kt$) and
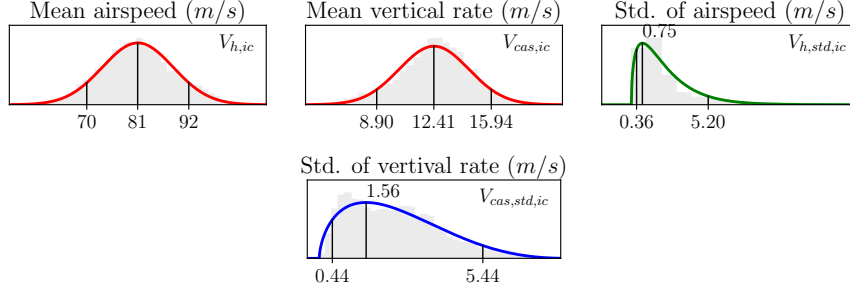
11

$12.4 m/s \ (\sim 2440 fpm)$



Figure 10: Initial climb parameters

The evidence for quasi-constant airspeed and vertical rate assumption can be seen from the standard deviations per trajectory, shown in the last two plots. It should also be noted that, in general, the vertical rate has larger variances. This is due to the fact that data sources for vertical rate commonly contain certain level of uncertainties as suggested in [9].

### 5.1.3 Climb

Within the climb phase, the objective is to model the constant CAS/Mach climbing profile, as well as the vertical rates in each of the segments of the profile. All parameters are shown in Fig. 11. Before the aircraft accelerates to a constant calibrated airspeed of 153 $m/s$ (295 $kt$), it has a mean climbing rate of 10.32 $m/s$ (2000 $fpm$), once reaching the transition altitude of approximately 4.8 $km$ (16$k$ $ft$). The aircraft then climbs at 8.5 $m/s$ (1670 $fpm$) until reaching the constant Mach number 0.77 at an altitude of 7.7 $km$ (25$k$ $ft$). After that, the aircraft climbs at 5.44 $m/s$ (1060 $fpm$) until reaching the cruise altitude. The flight range of the climb phase is also shown in the last plot, which is typically around 246 $km$.
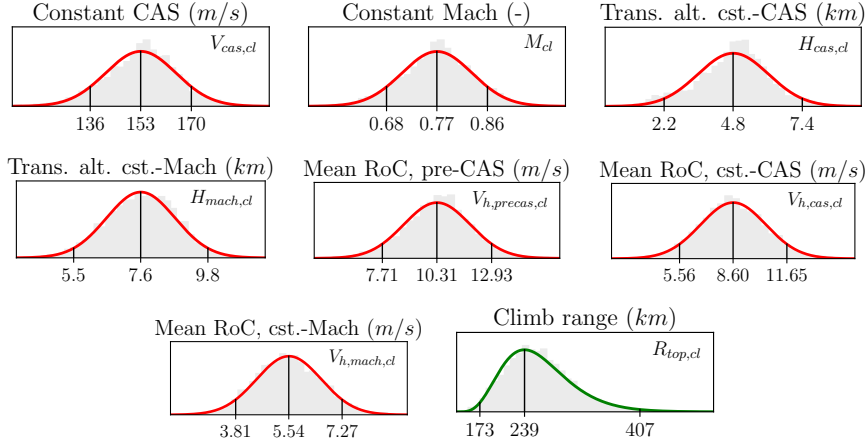


Figure 11: Climb parameters

### 5.1.4 Cruise

During the cruise phase, cruise speed $V_{cas,cr}$, Mach number $M_{cr}$, altitude $H_{cr}$, and cruise range $R_{cr}$ are shown in Fig. 12 respectively. Maximum cruise Mach number $M_{max,cr}$ and maximum cruise altitude $H_{max,cr}$ can be obtained as the maximum value of $M_{cr}$ and $H_{cr}$.

Unlike other performance parameters, the cruise range $R_{cr}$ is spread very widely, where 98% of flights globally range from 473 $km$ to 3819 $km$. This can reflect the wide ranging usage of the A320. However, most common flights cruise with range around 1000 $km$ operationally. The maximum
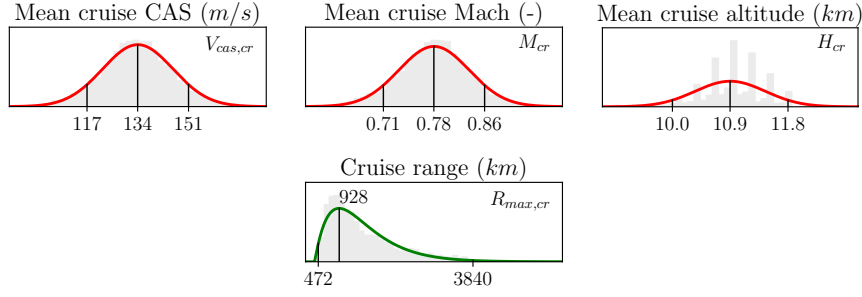
Figure 12: Cruise parameters

range presented in the dataset is around 4700 $km$, which is closer to its claimed maximum flight range (6500km) if taking into account climb, descent range, and fuel reserves.

### 5.1.5 Descent

Similar to climb, the descent phase can also be modeled as a constant Mach descent segment and a constant CAS descent segment. The parameters are shown in Fig. 13. The aircraft starts initial descent at constant Mach number 0.77 and vertical rate $-6.3$ $m/s$ ($-1240$ $fpm$) until reaching an altitude of $9.4km$ ($31k$ $ft$). It then starts a constant CAS descent of 150 $m/s$ (290 $kt$) and vertical rate of $-9.7$ $m/s$ ($-1900$ $fpm$) until reaching the altitude $5.3km$ ($17k$ $ft$). Then, the aircraft descends with a vertical rate at $-5.9m/s$ ($-1160$ $fpm$) until final approach. The last plot shows the range from top-of-descent to destination, typically around $230km$.
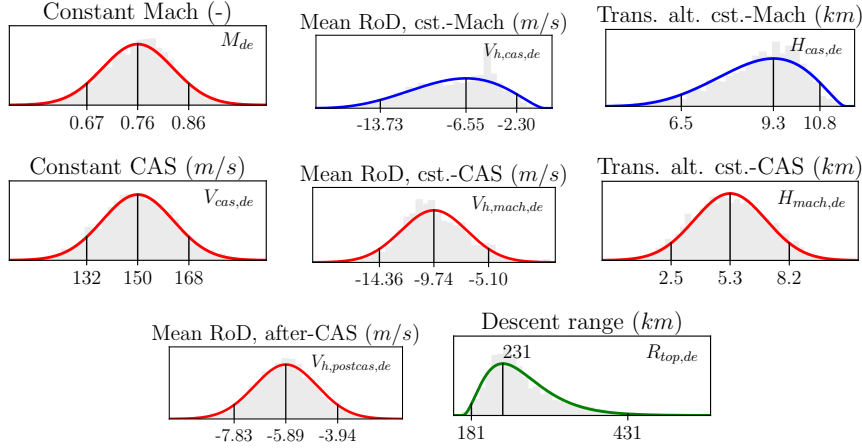


Figure 13: Descent parameters

It is worth taking into consideration that the results obtained from the dataset contain both CDA and non-CDA descent approaches.

### 5.1.6 Final Approach

At final approach, airspeed $V_{cas,fa}$ and vertical rate $V_{h,fa}$ are shown in Fig. 14, which are 69 $m/s$ (134 $kt$) and $-3.54$ $m/s$ ($-700$ $fpm$) respectively. The last two plots show the variance of these two parameters within each trajectory. Similar to the initial climb, the quasi-constant airspeed and vertical rate assumption is still valid.

Figure 14: Final approach parameters

### 5.1.7 Landing

At the final flight landing phase, approach speed $V_{app}$, breaking distance $d_{brk}$, and mean deceleration $\bar{a}_{brk}$ are shown in Fig. 15, with values of 68 $m/s$ (132 $kt$), 1.2 $km$, and $-1.07 m/s^2$. Breaking distanc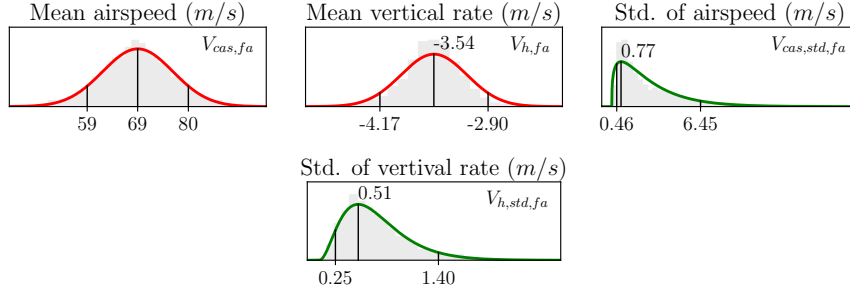e shows a large variance, ranging from around 600 meters to 3km. Different factors such as aircraft weight, wind direction, and runway conditions would all cause such large deviations.
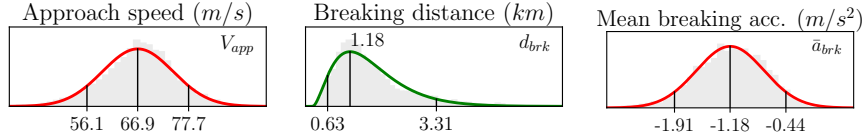


Figure 15: Landing parameters

## 5.2 Multiple aircraft types

The same methods and analysis are applied on a much larger dataset that contains multiple aircraft types. Due to the complexity of displaying all distribution and models from each parameters per aircraft type, the optimal parameter values of 17 aircraft types are shown in Table 2.

# 6 Validations and discussions

## 6.1 Validation of performance parameters

The performance parameters described in this paper are all derived from large numbers of flight data, which resemble the best estimates of how aircraft truly perform. There are a few existing performance models that can be used as references to examine the validity of the estimated parameters.

Within the BADA family 3 [10], similar performance parameters are contained in the OPF (Operational Performance File) and APF (Airline Procedure File). The difference between BADA and our model on ten performance parameters (from 14 aircraft types) are shown in Fig. 16.

It can be observed that most parameters from different aircraft types are very close to what is presented in BADA, with the exception that takeoff distance ($V_{tof}$) is around 20% lower than for BADA. This is likely due to the maximum takeoff weight assumed in the BADA model. If the maximum values of the parameter $V_{lof|max}$ from this paper are used, the result becomes comparable with BADA.

In a similar fashion, the Eurocontrol Aircraft Performance Database [11] can also be used as a source of comparison. It provides more performance parameters that are used primarily for training purposes on a wide range of commercial and military aircraft. A total of 17 parameters from 14 aircraft that exist in both models are compared. The result is shown in Fig. 17:

The parameters that are displayed with large bias are vertical rates during the descent phase. In the Eurocontrol database, it is common that the second descent segment has much higher vertical

Table 2: Performance model for various commercial aircraft

| Phase | Param | Unit | A319 | A320 | A321 | A332 | A333 | A343 | A388 | B737 | B738 | B739 | B744 | B752 | B763 | B77W | B788 | B789 | E190 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TO | $v_{lof}$ | $m/s$ | 80.8 | 85.3 | 89.6 | 91.2 | 91.5 | 86.3 | 88.1 | 81.3 | 85.2 | 90.7 | 91.9 | 88.8 | 92.7 | 98.2 | 89.8 | 95.6 | 87.0 |
| | $d_{tof}$ | $km$ | 1.62 | 1.68 | 1.84 | 2.01 | 2.05 | 2.27 | 2.51 | 1.59 | 1.75 | 2.04 | 2.31 | 1.72 | 1.82 | 2.24 | 2.16 | 2.52 | 1.81 |
| | $\bar{a}_{tof}$ | $m/s^2$ | 1.83 | 1.95 | 1.95 | 1.75 | 1.73 | 1.37 | 1.31 | 1.71 | 1.77 | 1.81 | 1.64 | 1.85 | 1.89 | 1.87 | 1.56 | 1.60 | 1.79 |
| | $V_{h,ic}$ | $m/s$ | 77 | 81 | 85 | 85 | 88 | 82 | 87 | 78 | 85 | 90 | 91 | 86 | 88 | 97 | 87 | 91 | 77 |
| | $V_{cas,ic}$ | $m/s$ | 12.12 | 12.78 | 13.20 | 12.35 | 12.19 | 6.74 | 5.84 | 11.62 | 11.90 | 12.15 | 9.52 | 12.93 | 14.09 | 12.95 | 10.40 | 10.50 | 11.20 |
| CL | $R_{top,cl}$ | $km$ | 214 | 246 | 244 | 282 | 291 | 290 | 297 | 209 | 222 | 243 | 232 | 227 | 208 | 221 | 278 | 274 | 226 |
| | $V_{cas,cl}$ | $m/s$ | 150 | 155 | 156 | 156 | 158 | 158 | 167 | 150 | 153 | 157 | 171 | 160 | 161 | 170 | 162 | 168 | 143 |
| | $M_{cl}$ | - | 0.77 | 0.77 | 0.77 | 0.80 | 0.79 | 0.78 | 0.83 | 0.77 | 0.78 | 0.78 | 0.84 | 0.80 | 0.80 | 0.84 | 0.84 | 0.84 | 0.75 |
| | $H_{cas,cl}$ | $km$ | 4.8 | 4.8 | 5.0 | 5.1 | 4.9 | 5.0 | 5.1 | 5.0 | 5.4 | 5.2 | 5.8 | 5.3 | 5.8 | 5.8 | 5.5 | 5.5 | 4.4 |
| | $H_{mach,cl}$ | $km$ | 8.1 | 7.8 | 7.4 | 8.2 | 7.8 | 7.4 | 7.7 | 8.2 | 8.3 | 7.7 | 7.6 | 7.8 | 7.8 | 7.9 | 8.5 | 8.0 | 8.5 |
| | $V_{h,precas,cl}$ | $m/s$ | 11.27 | 10.32 | 9.50 | 9.71 | 8.94 | 7.15 | 8.17 | 11.90 | 11.06 | 10.42 | 8.87 | 10.44 | 10.47 | 9.45 | 9.69 | 9.32 | 10.64 |
| | $V_{h,cas,cl}$ | $m/s$ | 10.14 | 8.60 | 7.89 | 8.02 | 7.65 | 6.93 | 7.19 | 11.44 | 9.96 | 8.76 | 8.72 | 8.91 | 9.86 | 8.80 | 8.49 | 8.19 | 9.27 |
| | $V_{h,mach,cl}$ | $m/s$ | 6.54 | 5.44 | 5.35 | 5.04 | 4.93 | 4.40 | 5.45 | 7.21 | 6.44 | 5.74 | 6.65 | 6.38 | 7.18 | 6.23 | 6.10 | 5.99 | 5.13 |
| CR | $R_{max,cr}$ | $km$ | 3613 | 3819 | 4065 | 10622 | 8773 | 13263 | 13928 | 4485 | 4417 | 4404 | 12083 | 6528 | 10531 | 14839 | 11338 | 13685 | 2380 |
| | $V_{cas,cr}$ | $m/s$ | 129 | 134 | 141 | 133 | 135 | 140 | 139 | 122 | 130 | 139 | 149 | 134 | 140 | 152 | 135 | 140 | 131 |
| | $V_{cas,max,cr}$ | $m/s$ | 135 | 144 | 151 | 148 | 150 | 160 | 160 | 130 | 142 | 150 | 169 | 147 | 158 | 173 | 153 | 159 | 139 |
| | $M_{cr}$ | - | 0.77 | 0.78 | 0.78 | 0.82 | 0.82 | 0.82 | 0.85 | 0.77 | 0.78 | 0.79 | 0.85 | 0.80 | 0.81 | 0.85 | 0.85 | 0.85 | 0.78 |
| | $M_{max,cr}$ | - | 0.88 | 0.86 | 0.88 | 0.89 | 0.88 | 0.89 | 0.91 | 0.87 | 0.86 | 0.88 | 0.92 | 0.88 | 0.90 | 0.91 | 0.92 | 0.92 | 0.88 |
| | $H_{cr}$ | $km$ | 11.5 | 10.9 | 10.4 | 12.0 | 11.5 | 11.0 | 11.6 | 11.7 | 11.2 | 10.5 | 10.8 | 11.2 | 10.8 | 10.5 | 12.0 | 11.6 | 11.0 |
| | $H_{max,cr}$ | $km$ | 11.4 | 11.1 | 10.6 | 12.0 | 11.8 | 11.6 | 12.1 | 11.9 | 11.4 | 10.7 | 11.3 | 11.4 | 11.2 | 11.0 | 12.3 | 12.0 | 11.2 |
| DE | $R_{top,de}$ | $km$ | 239 | 234 | 239 | 285 | 294 | 282 | 321 | 254 | 249 | 247 | 264 | 255 | 245 | 262 | 298 | 304 | 242 |
| | $M_{de}$ | - | 0.76 | 0.77 | 0.77 | 0.80 | 0.81 | 0.80 | 0.82 | 0.77 | 0.76 | 0.76 | 0.81 | 0.77 | 0.79 | 0.82 | 0.82 | 0.83 | 0.75 |
| | $V_{cas,de}$ | $m/s$ | 148 | 150 | 151 | 153 | 152 | 154 | 154 | 147 | 147 | 149 | 152 | 151 | 153 | 157 | 154 | 156 | 147 |
| | $H_{cas,de}$ | $km$ | 9.5 | 9.4 | 8.7 | 10.1 | 9.5 | 9.3 | 10.1 | 9.8 | 9.7 | 9.4 | 9.8 | 9.0 | 9.1 | 9.1 | 10.4 | 10.3 | 8.9 |
| | $H_{mach,de}$ | $km$ | 5.1 | 5.3 | 5.9 | 5.5 | 5.8 | 5.5 | 5.8 | 5.4 | 5.5 | 6.1 | 6.1 | 6.0 | 6.3 | 6.1 | 6.5 | 6.4 | 5.0 |
| | $V_{h,cas,de}$ | $m/s$ | -5.81 | -6.30 | -5.46 | -7.51 | -6.00 | -6.32 | -7.25 | -6.96 | -7.39 | -6.16 | -5.89 | -6.16 | -6.63 | -6.57 | -8.12 | -7.77 | -6.90 |
| | $V_{h,mach,de}$ | $m/s$ | -9.59 | -9.73 | -9.14 | -9.12 | -8.65 | -9.17 | -8.21 | -9.03 | -9.71 | -8.81 | -8.97 | -9.31 | -9.66 | -9.12 | -9.34 | -8.92 | -9.19 |
| | $V_{h,postcas,de}$ | $m/s$ | -5.79 | -5.88 | -5.89 | -5.43 | -5.53 | -5.38 | -5.22 | -5.71 | -6.03 | -5.73 | -6.07 | -5.89 | -6.00 | -6.04 | -5.88 | -5.87 | -5.77 |
| FA | $V_{cas,fa}$ | $m/s$ | 64 | 69 | 72 | 70 | 72 | 72 | 71 | 68 | 75 | 76 | 78 | 68 | 75 | 77 | 74 | 76 | 68 |
| | $V_{h,fa}$ | $m/s$ | -3.42 | -3.54 | -3.68 | -3.61 | -3.63 | -3.64 | -3.65 | -3.57 | -3.82 | -3.92 | -3.89 | -3.42 | -3.73 | -3.98 | -3.89 | -4.08 | -3.57 |
| LD | $V_{app}$ | $m/s$ | 62.2 | 68.1 | 69.9 | 68.9 | 72.1 | 70.8 | 68.1 | 66.1 | 73.5 | 74.7 | 77.3 | 65.8 | 73.0 | 75.5 | 70.6 | 74.4 | 64.7 |
| | $d_{brk}$ | $km$ | 1.66 | 1.20 | 1.76 | 1.66 | 1.63 | 1.71 | 2.25 | 2.22 | 1.38 | 1.64 | 1.92 | 1.26 | 1.60 | 1.62 | 2.11 | 2.51 | 1.90 |
| | $\bar{a}_{brk}$ | $m/s^2$ | -0.77 | -1.07 | -1.04 | -1.08 | -1.11 | -1.03 | -0.94 | -0.76 | -1.32 | -1.23 | -1.14 | -0.98 | -1.05 | -1.26 | -1.03 | -1.02 | -0.94 |
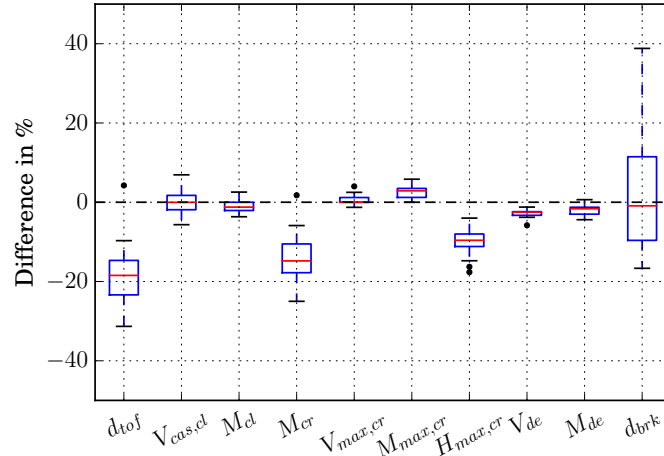


Figure 16: Comparison with BADA

rate. However, this trend is not significant in the flight data as observed. It can also be the case that the transition altitudes obtained from data in this paper are different from the fixed transition altitude from the Eurocontrol database.

Furthermore, compared to both BADA and the Eurocontrol database, the models obtained in this paper also include the minimum value, maximum value, and a parametric model for each parameter. This accuracy comes as an additional advantage for large-scale ATM simulators such as BlueSky simulator [12].
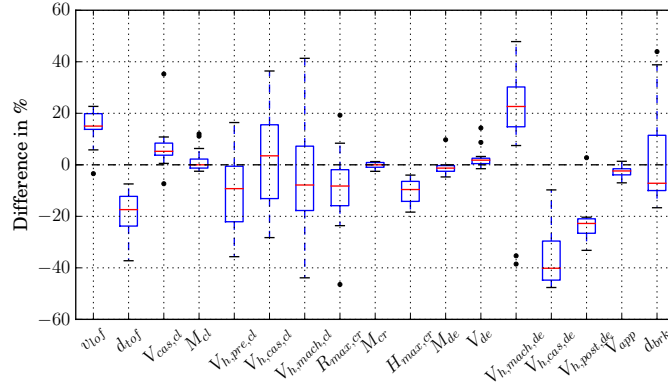
Figure 17: Comparison with Eurocontrol Aircraft Performance Database

## 6.2 Point of lift-off and touch-down

During the takeoff and landing phases, due to a low data update rate, an estimation method in Fig. 2 is used to locate the most likely lift-off or touch-down moment, from which related parameters are interpolated.

Statistically, there are relatively large gaps between the last on-ground data and first in-air data. The time gaps $\Delta t$ are measured from a large number (around 7000 flights) of takeoffs and landings, as shown in Fig. 18.
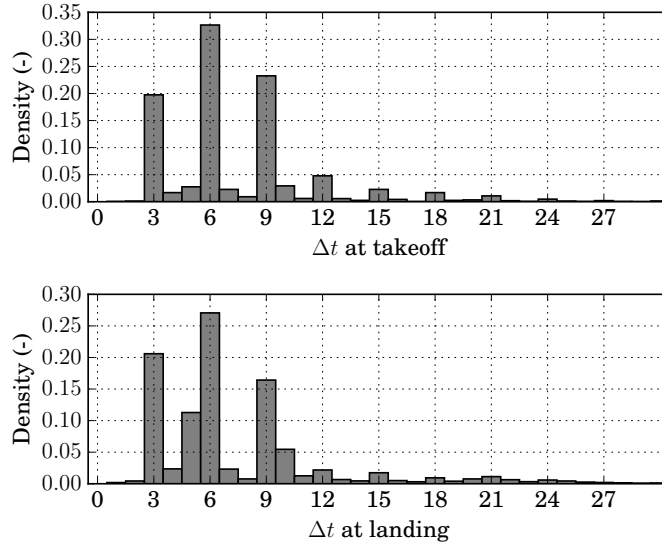


Figure 18: Time gap of data at takeoff and landing

Most commonly, the $\Delta t$ is about three to nine seconds during which the aircraft would have flown, climbed, or descended for a fairly large distance. Thus, when dealing with such data, using the above method improves the accuracy of parameter estimations during the phases where fewer samples are available.

It is also interesting to notice that there is a three-second time gap in both figures of take-off and landing. This is likely due to the update rate of aggregated ADS-B data from FlightRadar24 that is used in this paper. When using raw ADS-B data, this gap can be reduced to around one second when receivers are in good visibility of an aircraft.

16

## 6.3 CDA

To maximize the fuel efficiency and when it is allowed by air traffic controllers, aircraft often perform continuous descent approaches (CDA) in order to save fuel and reduce noise. However, as modeling the performances, CDA has an impact on some parameters during the entire descent phase.

To study the influence of this factor, a similar number of CDA and non-CDA flights (around 3000 flights each) are used. All parameters related with descent are computed and compared in parallel. As shown in Fig. 19, eight parameters in CDA and non-CDA flights are shown. Transitional altitudes and velocities of constant Mach/CAS descent do not change between these two approaches, but other parameters show large differences. In the first sub-figure, it is clear that CDA decreases the range of top-of-descent drastically. From sub-figure six to eight, it is also apparent that the descent rates are all increased under CDA. All of these differences are indeed aligned with the definition of CDA.
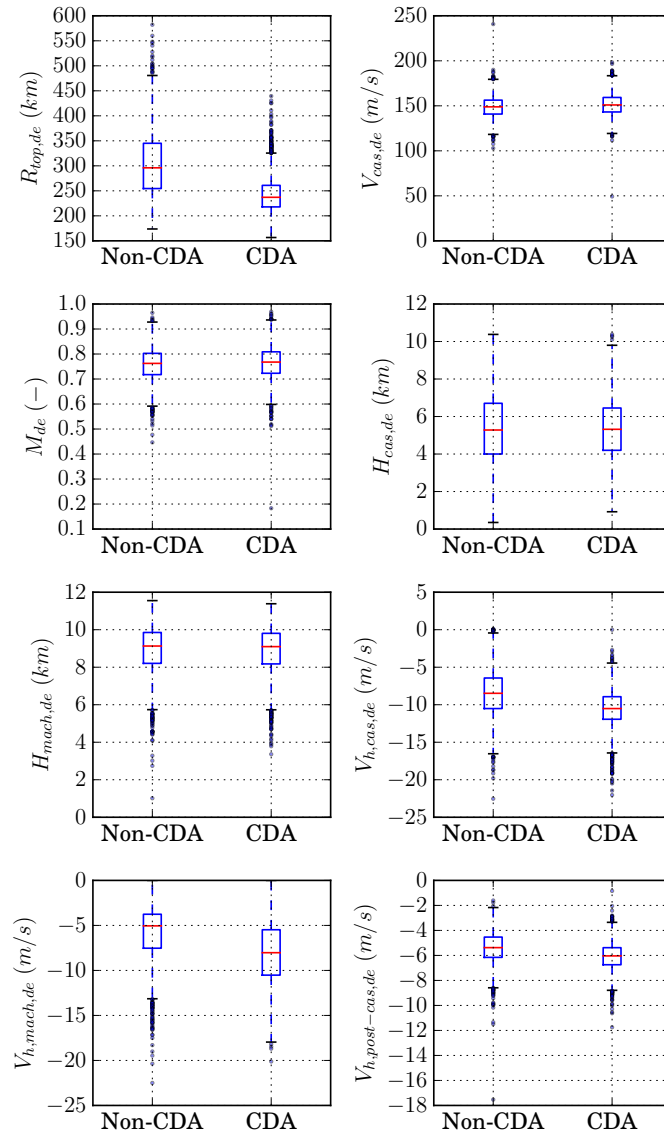


Figure 19: Comparing descent parameters during CDA and Non-CDA

17

## 6.4 Data and models

Models from this paper describe the operational performances and limitations of aircraft. Compared to BADA, elements such as thrust, fuel, and drag polars need to be developed with aid from other open data in future research projects. Some of these parameters may also not be completely independent, and indeed, weak correlations are likely to exist among certain parameters, which could be caused by factors such as the aircraft weight or flight strategy. Further research will also be conducted to study this effect.

ADS-B data are not distributed equally per aircraft type. The same amount of data from most common aircraft types such as Airbus-320 and Boeing-737 can be collected within a week, while for other less common aircraft types such as Airbus-380 and Boeing-777, this process may take several months. While completing this research, a total of 1.6 million flights have been captured during a period of five months for around 30 commercial aircraft types. Models are being continuously improved with new data. All models from this paper are published under the GNU v3 open source license.

# 7 Conclusions

Based on an improved version of previously developed flight phase identification methods, this paper utilizes large amounts of ADS-B data to construct accurate aircraft parametric performance models. From these flight data, maximum likelihood estimations on different distribution functions are used to calculate the best models to describe each parameter. In addition to detailed model construction methods along the flight trajectory, a total of 32 parameters per aircraft are presented, covering most common aircraft types. The results for 17 aircraft types are summarized in this paper. In addition, these detailed models are published under open-source license. Comparison studies are also performed against existing models such as BADA and the Eurocontrol Aircraft Performance Database. Finally, this paper provides a starting point for comprehensive open source aircraft performance models for the ATM research community.

# References

[1] Nuic, A., "Base of Aircraft Data (BADA) Product Management Document," Tech. rep., Tech. Rep. EEC Technical/Scientific Report, 2009.

[2] Sun, J., Ellerbroek, J., and Hoekstra, J., "Flight Phase Identification and Machine Learning for Mining ADS-B and Meteorological Data," *Journal of Aerospace Informatiion Systems*, 2016 - submitted.

[3] Commercial Aviation Safety Team / Common Taxonomy Team, "Phase of Flight: Definitions and Usage Notes," Vol. 1.3, No. April, 2013, pp. 1–11.

[4] Aldrich, J. et al., "RA Fisher and the making of maximum likelihood 1912-1922," *Statistical Science*, Vol. 12, No. 3, 1997, pp. 162–176.

[5] Choi, S. and Wette, R., "Maximum likelihood estimation of the parameters of the gamma distribution and their bias," *Technometrics*, Vol. 11, No. 4, 1969, pp. 683–690.

[6] Gnanadesikan, R., Pinkham, R., and Hughes, L. P., "Maximum likelihood estimation of the parameters of the beta distribution from smallest order statistics," *Technometrics*, Vol. 9, No. 4, 1967, pp. 607–620.

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

[8] FlightRadar24 AB, "FlightRadar24, Terms and Conditions, Item 9," `http://www.flightradar24.com/terms-and-conditions`, January, 2017.

[9] Hempe, D. W., "Airworthiness Approval of Automatic Dependent Surveillance-Broadcast (ADSB) Out Systems," *Advisory Circular 20-165*.

[10] Nuic, A., "User manual for the Base of Aircraft Data (BADA) revision 3.12," *Atmosphere*, Vol. 2014, 2014.

[11] EuroControl, "EuroControl Aircraft Performance Database," `https://contentzone.eurocontrol.int/aircraftperformance`, January, 2017.

[12] Hoekstra, J. and Ellerbroek, J., "BlueSky ATC Simulator Project: an open Data and Open Source Approach," *Proceedings of the 7th International Conference on Research in Air Transportation*, 2016.

## Authors' Biographies

**Junzi Sun** is a Ph.D. candidate at the Delft University of Technology. He received his M.Sc from the Polytechnic University of Catalonia. His doctoral research topic is aircraft performance modeling using open flight data, with focuses on machine learning, data mining, modeling, and statistical inferences.

**Joost Ellerbroek** received the M.Sc. and Ph.D. degrees in aerospace engineering from the Delft University of Technology, The Netherlands, in 2007 and 2013, respectively. His research interests are in the domain of air-traffic management, including analysis of airspace complexity, and the design and analysis of conflict detection and resolution algorithms.

**Jacco M. Hoekstra** is a full professor at the faculty of Aerospace Engineering of the Delft University of Technology. After obtaining his M.Sc. and pilots license in 1990, he started working at the National Aerospace Laboratory (NLR). He then obtained his Ph.D. on Air Traffic Management and headed the NLR Air Transport Operations division, before joining the TU Delft faculty. He served two terms as a dean and now heads the CNS/ATM chair.