Aircraft Drag Polar Estimation Based on a Stochastic Hierarchical Model

Junzi Sun, Jacco M. Hoekstra, Joost Ellerbroek
Control and Simulation, Faculty of Aerospace Engineering
Delft University of Technology, the Netherlands

Abstract—The aerodynamic properties of an aircraft determine a crucial part of the aircraft performance model. Deriving accurate aerodynamic coefficients requires detailed knowledge of the aircraft's design. These designs and parameters are well protected by aircraft manufacturers. They rarely can be used in public research. Very detailed aerodynamic models are often not necessary in air traffic management related research, as they often use a simplified point-mass aircraft performance model. In these studies, a simple quadratic relation often assumed to compute the drag of an aircraft based on the required lift. This so-called drag polar describes an approximation of the drag coefficient based on the total lift coefficient. The two key parameters in the drag polar are the zero-lift drag coefficient and the factor to calculate the lift-induced part of the drag coefficient. Thanks to this simplification of the flight model together with accurate flight data, we are able to estimate these aerodynamic parameters based on flight data. In this paper, we estimate the drag polar based on a novel stochastic total energy model using Bayesian computing and Markov chain Monte Carlo sampling. The method is based on the stochastic hierarchical modeling approach. With sufficiently accurate flight data and some basic knowledge of aircraft and their engines, the drag polar can be estimated. We also analyze the results and compare them to the commonly used Base of Aircraft Data model. The mean absolute difference among 20 common aircraft for zero-lift drag coefficient and lift-induced drag factor are 0.005 and 0.003 respectively. At the end of this paper, the drag polar models in different flight phases for these common commercial aircraft types are shared.

Keywords - aircraft performance, drag polar, aerodynamic coefficient, Bayesian computing, MCMC

I. INTRODUCTION

Since the invention of aircraft, researchers have been studying the aerodynamic properties of airfoils and aircraft. Examples of fundamental studies on aerodynamic drag are given in [1], [2]. From the start, the goal has been to optimize the lift over drag ratio for the cruise flight. Much effort has been dedicated to creating designs that would reduce drag and thus increase the lift efficiency of aircraft. While the zero drag coefficient contains the parasitic drag of the whole aircraft, the wing is mainly responsible for the lift-induced drag. Next, to the chosen airfoil, the aspect ratio of the wing plays an important role. The wing can be seen as a drag to lift converter, of which the already high efficiency can be increased further. This is an on-going effort: some examples of current research are boundary layer suction, morphing wings, plasma control, and blended wing-body aircraft shapes.

Aerodynamic lift and drag forces of an aircraft are complicated and computationally intensive to compute. Lift and drag are considered as functions of the wing area, dynamic airspeed, and air density, and the remaining effects of the flow for both the lift and drag are described with coefficients for both forces. The most complicated part is to model these lift and drag coefficients. These parameters depend on the Mach number, the angle of attack, the boundary layer and ultimately on the design of the aircraft shape. For fixed-wing aircraft, these coefficients are presented as functions of the angle of attack, i.e., the angle between the aircraft body axis and the airspeed vector. In air traffic management (ATM) research, however, simplified point-mass aircraft performance models are mostly used. These point-mass models consider an aircraft as a dimensionless point, where the angle of attack, as well as the side-slip angle, and the effect of the angular rates are not explicitly considered. Hence, the step of calculating these is often skipped and the *drag polar* is used instead.

The relationship between drag coefficient and lift coefficient is the main factor determining the aircraft performance. Knowledge of the drag polar is therefore essential for most ATM research such as trajectory prediction, fuel optimization, parameter estimation.

Many methods exist to explore the aircraft performance during the preliminary design phase, often with a focus on the modeling of the aerodynamics. Hence, one source of open information regarding drag polar models comes from textbooks [3], [4], [5], [6]. However, only older aircraft models are available in the literature. In [7], an empirical model for estimating zero-lift drag coefficients was proposed using existing literature data based on several aircraft models. In general, open data on drag polar model is rare, especially for modern commercial aircraft.

The aircraft manufacturers who design the aircraft do have accurate aerodynamic data. However, these data are rarely publicly available due to commercial competition. The most comprehensive collection of drag polar data is the Base of Aircraft Data (BADA) developed by Eurocontrol [8]. It contains the drag polars for nearly all common aircraft types. BADA is the default "go-to" aircraft performance model for current ATM researchers. However, it imposes a strict license in terms of sharing and redistribution of the model data. The project-based license for newer versions makes it even harder for the same researcher to reuse the model, which will also apply for new users of older versions.

The goal of this paper is to propose an alternative path to estimate the drag polar models for modern fixed-wing commercial aircraft, as well as share the drag polar models that we have obtained. We approach this estimation problem using a

novel stochastic total energy (STE) model. The STE approach treats the parameters of the standard total energy model as random variables. Then, we try to solve the parameter estimation using Bayesian computing, specifically, Markov chain Monte Carlo (MCMC) approximations. Finally, a database of drag polar models for different common commercial aircraft, which were produced using this method, is provided.

The structure of this paper is as follows. In section two, the fundamentals of the point-mass drag polar model are introduced. In section three, we focus on the hierarchical model approach. In section four, experiments are conducted to examine and obtain drag polar data of multiple aircraft types based on this method. Section five and six are dedicated to the discussions and conclusions of this research.

II. THEORY OF AERODYNAMIC MODELING

A. Drag polar in point-mass models

While an aircraft flies, the drag force is produced by the airflow interacting with the aircraft body. The lift force is produced due to the pressure difference between the upper and lower surface of the lifting devices (wings). With the same airspeed and altitude conditions, control of lift is performed by re-configuring the aircraft angle of attack and/or modifying the surface shape of lifting devices. By changing the elevator settings, the pitch angle and the angle of attack can be controlled. On the other hand, a change of the lifting device surface is primarily performed by re-configuring flaps.

In general, the lift and drag forces of an aircraft that is traveling in the free stream can be computed as:

$$L = C_L \; \frac{1}{2} \rho V^2 S$$
$$D = C_D \; \frac{1}{2} \rho V^2 S$$
(1)

where C_L and C_D are lift and drag coefficients, respectively. ρ, V, and S are air density, true airspeed, and the lifting surface area of the aircraft. In practice, C_L and C_D can be modeled as functions of the angle of attack (α), Mach number(M) and flap deflection(δ_f):

$$C_L = f_{cl}(\alpha, M, \delta_f)$$
$$C_D = f_{cd}(\alpha, M, \delta_f)$$
(2)

In aerodynamic models, multi-dimensional table interpolation or higher order polynomials are used. However, in many ATM studies, the six-degree of freedom of aircraft flight dynamic is simplified to the three-degree of freedom point-mass model. Leaving out the aerodynamic angles and pitch, yaw, and roll rates means the aerodynamic models also need to be adapted. In point-mass models, the relation between the aerodynamic coefficients C_D and C_L is simplified to the drag polar. It is commonly represented using a quadratic function in one of the two forms below:

$$C_D = C_{D0} + \frac{C_L^2}{\pi A e}$$
(3)

where A is the aspect ratio of the wing (span divided by the average chord) and where e is the Oswald factor, which lies typically in the range 0.70-0.90. This equation is often written as:

$$C_D = C_{D0} + kC_L^2$$
(4)

with:

$$k = \frac{1}{\pi A e}$$
(5)

These two parameters, C_{D0} and k are the *zero-lift drag coefficient* and *lift-induced drag coefficient factor* respectively. The values of both parameters are considered as constants under a specific aerodynamic configuration of the aircraft. Fig. 1 illustrates an example of drag polar by using a computational fluid dynamics (CFD) simulation for an aircraft with a clean (no flaps or extended landing gear) configuration.
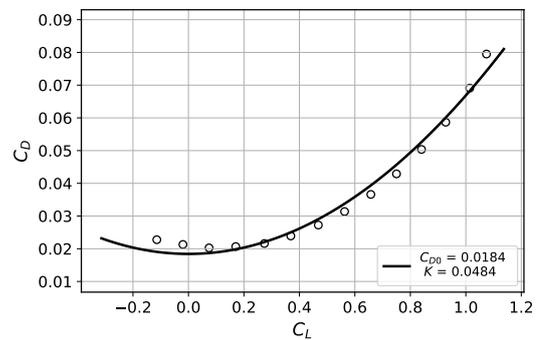
Figure 1. Drag polar (B747-400 simulated, clean configuration, $\alpha < 10°$)

B. Aircraft aerodynamic configurations

Besides the angle of attack that affects the values of the lift and drag coefficients, the change in the shape of the aircraft can alter these values. The most notable change in aircraft is flaps (and slats), speed brakes and landing gear. Each structural setting also has its own corresponding drag polar model.

Flaps are common aircraft surfaces deployed in order to provide an increase in the maximum lift coefficient. They are deployed to be able to fly at lower speeds, typically at low altitudes (for example, during takeoff, initial climb, and approach). Different aircraft types have different configurations of flaps and flap settings. An increase in flap angle leads to an increase in the lift coefficient under the same angle of attack, at the expense of a higher drag. Slats are similar to flaps, but on the leading edge of the wing and they increase the maximum lift coefficient by increasing the stall angle of attack. Slats are automatically extended when selecting a flap setting and are considered as part of this configuration. Different flap designs have been adopted by aircraft manufacturers. In Table I, a list of common flap options on airfoils and their approximated maximum lift coefficients are listed. These values are produced by [9, p.107]. It is worth noting that the $C_{L,max}$ values of an airfoil are larger than the values of the aircraft with the same shaped wing, especially for swept wings [3, p.263].

TABLE I
EXAMPLE FLAP SETTINGS (AIRFOILS)

Flap types	$C_{L,max}$	Illustration
airfoil only	1.5	
leading-edge slat	2.4	
plain flap	2.5	
split flap	2.6	
Fowler single-slotted flap	2.9	
Fowler multi-slotted flap	3.0	
with leading-edge slat	3.3	
with boundary layer suction	3.9	

In this paper, based on the data from [6, p.253], the increase in lift coefficient due to flap deployment is modeled. These values are shown in Table II, where TO and LD represent the take-off and landing configuration respectively. Extended flaps also increase the drag. The increase of drag coefficient due to flaps can be computed using the model proposed by [10]

$$\Delta C_{D,flap} = 0.9 \left(\frac{c_{fp}}{c} \right)^{1.38} \left(\frac{S_{fp}}{S} \right) \sin^2 \delta \qquad (6)$$

where c_{fp}/c and S_f/S are flap to wing chord ratio and surface ratio. δ is the flap deflection angle. When these aircraft characteristics are not available, simplified empirical values from [6, p.253] can be used, which are listed in Table III.

TABLE II
INCREASE OF LIFT COEFFICIENT WITH FLAPS

Trailing	Leading	α^{TO}	α^{LD}	$C_{L,max}^{TO}$	$C_{L,max}^{LD}$
plain flap		20°	60°	1.60	2.00
single-slotted flap		20°	40°	1.70	2.20
Flower single-slotted flap		15°	40°	2.20	2.90
Flower double-slotted flap		20°	50°	1.95	2.70
Flower double-slotted flap	w/ slat	20°	50°	2.60	3.20
Flower triple-slotted flap	w/ slat	20°	40°	2.70	3.50

TABLE III
INCREASE OF DRAG COEFFICIENT WITH FLAPS

$\Delta C_{D,flap}$	Typical flap angle	Flight phase
0.02	$10° \sim 20°$	take-off
0.04	$20° \sim 30°$	take-off
0.08	$30° \sim 40°$	landing
0.12	$40° \sim 50°$	landing

The landing gear adds a significant amount of drag to the aircraft when it is extended. The landing gear is retracted as soon as the aircraft becomes airborne and only extended shortly before landing. There is limited research data that quantifies the drag coefficient of aircraft landing gears. An empirical model proposed by [11] formulates the increased drag coefficient by landing gears as:

$$\Delta C_{D,gear} = \frac{W}{S} K_{uc} \, m_{max}^{-0.215} \qquad (7)$$

where W/S is the wing loading, m_{max} refers to the maximum mass of an airplane, and K_{uc} is a factor depending on the flap deflection. The value of K_{uc} is lower when more deflection is applied. This is because the flow velocity along the bottocm of the wing decreases when flaps are deployed, thus leading to a lower drag on the landing gear. The values of K_{uc} during different flight phases are shown in Table IV.

TABLE IV
VALUE OF K_{uc}

K_{uc}	flap deflection	flight phase
5.8×10^{-5}	none	taxing
4.5×10^{-5}	medium *	take-off
3.16×10^{-5}	full	landing

* interpolated based on existing values

Considering these structural variations, we can model the zero-lift drag coefficient C_{D0}^* as:

$$C_{D0}^* = C_{D0} + \Delta C_{D,flap} + \Delta C_{D,gear} \qquad (8)$$

A different flap setting causes a less significant change to k than to C_{D0}. Hence in this paper, we assume k remains constant for all different flap settings.

III. ESTIMATING DRAG POLAR WITH A STOCHASTIC TOTAL ENERGY MODEL

Although it is simple to explain the drag polar that is described in the previous section, existing performance models that rely on manufacturer data generally come with restricting licenses (for example, BADA 3 and BADA 4). Open data on models included in textbooks and literature are often based on old models. In order to construct an open aerodynamic model, we intentionally refrain from using manufacturer data (unless they are publicly available) and search for an alternative path to model the drag polar.

In this section, we first explore the principle of a novel hierarchical model that describes the total energy model in stochastic fashion, where the model parameters are considered as random variables. Using Bayesian computing, we then try to infer the drag polar based on open flight data from ADS-B and Enhanced Mode-S surveillance communications. The clear benefit of this stochastic total energy (STE) model is that the process can be applied to any aircraft, as long as enough flight data and the basic performance parameters of the aircraft are known. The following part of the section describes the process in detail.

A. Stochastic model

Commonly, the total energy model describes the change of energy by multiplying each force with speeds in the same direction. This results in the following equation:

$$(T_t - D_t) \, V_t = m_t a_t V_t + m_t g V S_t \qquad (9)$$

where T and D are the thrust and drag of the aircraft. m is the aircraft mass. Parameters a, V, and VS are the acceleration, airspeed, and vertical speed respectively. These are three variables that can be derived from aircraft surveillance data. Subscript t indicates the data is a time series. In general, thrust can be modeled as a function of velocity and altitude (h). In this research, we use the model proposed by [12]. Let f represent the set of functions that compute the thrust:

$$T_t = f(V_t, h_t) \tag{10}$$

Combining the previous two equations, we can compute the drag and the drag coefficient based on trajectory data as follows:

$$D_t = f(V_t, h_t) - m_t a_t - m_t g VS_t / V_t$$
$$C_{D,t} = \frac{D_t}{\frac{1}{2} \rho_t V_t^2 S} \tag{11}$$

where ρ and S are the air density and the aircraft wing surface.

From the equilibrium of forces in the direction perpendicular to the airspeed, we find the relation between the lift coefficient and the mass:

$$C_{L,t} = \frac{L_t}{\frac{1}{2} \rho_t V_t^2 S} = \frac{m_t g \cos \gamma_t}{\frac{1}{2} \rho_t V_t^2 S} \tag{12}$$

On the other hand, we can also derive the drag coefficient (denoted as C_D^*) using the drag polar equation:

$$C_{D,t}^* = C_{D0} + k C_{L,t}^2 = C_{D0} + \frac{k(m_t g \cos \gamma_t)^2}{\left(\frac{1}{2} \rho_t V_t^2 S\right)^2} \tag{13}$$

Here γ is the path angle, which can be computed using the time derivative of the altitude and the ground speed. In Equation 5, the coefficient k is defined as a function of aspect ratio (AR) and span efficiency factor (e, close to one). The density of the air can be computed based on the temperature and barometric altitude under ISA conditions up to 36,090 ft using the constants in Table V:

$$\rho_t = \rho_0 \left(\frac{1 - \lambda h_t}{\tau_0} \right)^{-\frac{g}{\lambda R} - 1} \tag{14}$$

TABLE V
LIST OF ISA CONSTANTS

Parameter	Value	Unit	Description
ρ_0	1.225	kg/m^3	air density at sea level
τ_0	288.15	K	temperature at sea level
g	9.80665	m/s^2	sea level gravity acceleration
λ	-0.0065	K/m	troposphere temperature gradient
R	287.05	J/(kg · K)	gas constant at sea level

Assuming a perfect system and perfect observations, the two drag coefficients, C_D and C_D^*, obtained in two different ways should be the same at each time step, with the following relationship:

$$\Delta C_{D,t} = C_{D,t} - C_{D,t}^* = 0 \tag{15}$$

Although noise is inevitably present, Equation 15 should be the condition that we want our estimator to approach. In other words, if we consider the goal as an optimization process, we would try to minimize the $\Delta C_{D,t}^2$, which could have been a possible approach. However, this approach is too ambitious since there are too many unknown system parameters.

Three unknown parameters occur in this system, which are C_{D0}, k, and m_t. In order to have a good estimation of C_{D0} and k, we need to have more knowledge on the aircraft mass m. Unfortunately, it is not available directly from the surveillance data. In our previous paper [13], we were able to estimate the mass and thrust setting, but it was based on drag polar provided by the BADA model. In order to derive the drag polar independently, the new method also needs to address the uncertainty of other parameters.

Similar to in [13], we try to solve the estimation problem from a Bayesian point of view. The difference is that now we can use multiple flights of each aircraft. Even though we don't know the exact mass and thrust setting of each flight, there are some hypotheses that we are confident about:

1) The values for C_{D0} and k are constant and the same for all flights that belong to the same aircraft model, under clean configuration.
2) Based on aerodynamic theory, it is possible to know the value ranges of C_{D0} and k.
3) We have some knowledge on possible distributions for aircraft mass and thrust setting.
4) We are able to obtain accurate surveillance data for a sufficient number of flights, including trajectories, velocity, temperature, and wind conditions.

In the proposed STE model, we consider all parameters as random variables. The observable parameters are defined as follows:

$$
\begin{aligned}
V_t &\sim \mathcal{N}(\tilde{V}_t, \sigma_v^2) \\
a_t &\sim \mathcal{N}(\tilde{a}_t, \sigma_a^2) \\
VS_t &\sim \mathcal{N}(\tilde{VS}_t, \sigma_{vs}^2) \\
\tau_t &\sim \mathcal{N}(\tilde{\tau}_t, \sigma_\tau^2) \\
h_t &\sim \mathcal{N}(\tilde{h}_t, \sigma_h^2) \\
\gamma_t &\sim \mathcal{N}(\tilde{\gamma}_t, \sigma_\gamma^2)
\end{aligned} \tag{16}
$$

where each parameter is assumed to be drawn from a normal distribution. $\tilde{V}_t$, $\tilde{a}_t$, $\tilde{VS}_t$, $\tilde{\tau}_t$, $\tilde{h}_t$, and $\tilde{\gamma}_t$ are the observed values at each time step respectively. σ_v^2, σ_a^2, σ_{vs}^2, σ_τ^2, σ_h^2, and σ_γ^2 are variances for each variable.

The models for the three unknown system parameters can be constructed similarly:

$$
\begin{aligned}
C_{D0} &\sim \mathcal{N}(\mu_{cd0}, \sigma_{cd0}^2) \\
e &\sim \mathcal{U}(e_{min}, e_{max}) \\
m_t &\sim \mathcal{N}(\mu_m, \sigma_m^2)
\end{aligned} \tag{17}
$$

where C_{D0} and m_t are defined with a normal prior. e is defined with a uniform prior, which can be translated to k using Equation 5. Random variable m_t is defined for the mass at time step t. For each time step, a different random variable with the same mean and variance is constructed.

With these variables defined, we convert the deterministic total energy model to a stochastic model. The hierarchical STE model considers each parameter as random variables, and, at the same time, preserves the underlying structure among the observable parameters $(a, V, VS, \tau, h, \gamma)$, forces (T, D) and system parameters (m, C_{D0}, e) as defined in total energy model. In Fig. 2, the dependencies among all those parameters are visualized. This hierarchical model includes eleven random variables, which are indicated in gray.
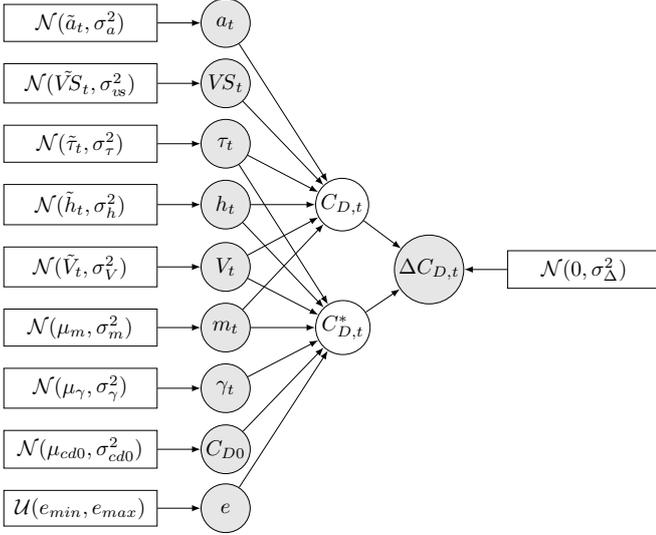
Figure 2. Hierarchical relationships of model parameters

B. Bayesian computation

Instead of estimating the drag polar (C_{D0} and k) directly, we have converted the original estimation to a Bayesian optimization problem. That is, given the prior probability of all model parameters, how can we compute the posterior probabilities based on the constraints in Equation 15? To simplify the expression, let us define $\boldsymbol{\theta}$ (a common way to express the parameters) and $\boldsymbol{y}$ (a common way to express the observations) as follows:

$$\begin{aligned} \boldsymbol{\theta} &= \left\{ C_{D0}, e, [m_t], [V_t], [a_t], [VS_t], [\tau_t], [h_t], [\gamma_t] \right\} \\ \boldsymbol{y} &= \left\{ [\Delta C_{D,t}] \right\} \end{aligned} \tag{18}$$

where parameters in brackets are with multiple dimensions. The goal is to compute the joint posterior probability $p(\boldsymbol{\theta}|\boldsymbol{y})$. It is difficult (and impractical) to compute the analytical solution of the target distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$ in such a high-dimensional space. Instead, we use a Markov Chain Monte Carlo (MCMC) simulation to compute this numerically.

The MCMC method uses sequential sampling, drawing a large number of values of $\boldsymbol{\theta}$ from approximate distributions

and then correcting these draws based on the distribution constructed from previous values drawn. $p(\boldsymbol{\theta}|\boldsymbol{y})$ will be obtained once the Markov chain converges to a unique stationary distribution.

The simplest and most popular form of the MCMC sampling method is the Metropolis algorithm [14]. It can be applied to almost any Bayesian computation problem. The algorithm can be explained in simple steps, which also demonstrates the principles of MCMC sampling. These steps are as follows:

1) Choose a initial set of $\boldsymbol{\theta}$ as $\boldsymbol{\theta}^0$.
2) for iteration $i = 1, 2, \cdots$:
   a) Sample a proposal $\boldsymbol{\theta}^*$ from a *jumping distribution* $J(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{t-1})$, which is a conditional probability dependent on the current set of values. It is usually chosen as a normal distribution.
   b) Calculate the ratio using the Bayes' rule:

$$r = \frac{P(\boldsymbol{\theta}^*|\boldsymbol{y})}{P(\boldsymbol{\theta}^{i-1}|\boldsymbol{y})} = \frac{P(\boldsymbol{\theta}^*)P(\boldsymbol{y}|\boldsymbol{\theta}^*)}{P(\boldsymbol{\theta}^{i-1})P(\boldsymbol{y}|\boldsymbol{\theta}^{i-1})} \tag{19}$$

   c) set $\boldsymbol{\theta}^i$ as $\boldsymbol{\theta}^*$ with probability of $\min(r, 1)$, otherwise keep as $\boldsymbol{\theta}^{i-1}$

When the MCMC sampling process reaches the desired number of iterations, it is stopped. The values from each iteration stored in the Markov Chain are used to construct the posterior distributions of the model parameters.

The metropolis sampler is simple to implement and can be applied to both continuous and discrete probability distributions. The drawback is that it can lead to a very long convergence runtime. This is due to the random-walk behavior. When model parameters are continuous, the Hamiltonian Monte Carlo (HMC) method [15] can be used to suppress the random-walk. It converts the sampling problem into a simulation of Hamiltonian dynamics. In this paper, a new sampler called No-U-Turn Sampler (NUTS) [16] is used. The NUTS sampler is able to converge quickly and requires fewer sampling iterations.

IV. EXPERIMENTS AND RESULTS

The experiment consists of three parts. The first part uses a Boeing 747-400 aircraft as an example and describes the process of MCMC sampling with the STE model. The second part applies the method to common aircraft types. Coefficients C_{D0} and k for all these aircraft types are constructed based on a large number of flights. The results obtained are compared with the existing BADA 3 model. Finally, combining the drag polar under non-clean settings, the final table consisting of drag polar for all flight phases is given in part three.

A. Estimate drag polar using STE model and MCMC sampling

A data-set consisting of around 100 flights per aircraft type is gathered. These are climbing flights from 3,000 ft to 10,000 ft. All data are collected using the Mode-S receiver setup in the Delft University of Technology. The flights observed were taking off from Amsterdam Schiphol airport. In figure 3, the ground tracks and vertical profiles of the Boeing 747-400 flights are illustrated.
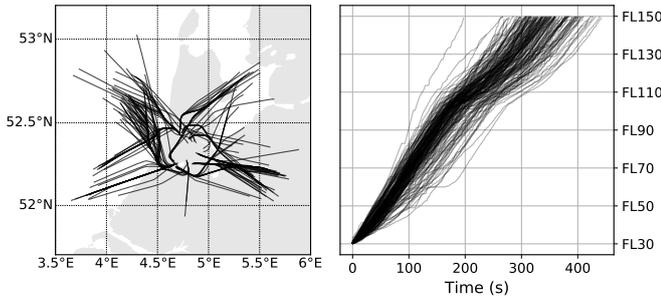
Fig. 3. Boeing 747-400 flights

Besides the ground speed obtained from ADS-B, the meteo-particle model is used to generate real-time wind data [17], [18]. Combining ground speed and accurate wind information, the airspeed of the aircraft is computed. This allows us to describe the performance more accurately.

For each flight, we apply MCMC sampling with the STE model to obtain the posterior probability distribution of the parameters C_{D0} and k. The parameter values of the prior distributions defined in Equation 16 and 17 (also by Fig. 2) are listed in Table VI.

Using the MCMC implementation in the PyMC3 [19] library, we can conveniently perform the sampling of the STE model with the defined prior distributions. To ensure the consistent convergence, we utilize four independent chains, which are sampled in parallel. When the maximum of 2,500 iterations is reached, the last 2,000 iterations are kept for analysis.

First, the results of one example flight are shown. In Fig. 4, the trace of four chains (on the right-hand side) and the estimated posterior density (on the left-hand side) are shown. Each chain is marked with a distinct color. The stable convergence can be observed by comparing the posterior density of all four chains. Combining all chains, we can obtain the mean value C_{D0} and k from the posterior distribution, which is shown in Fig. 5.
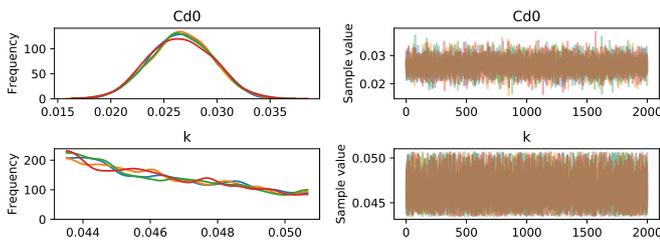
Figure 4. MCMC sampling of a Boeing 747-400 flight
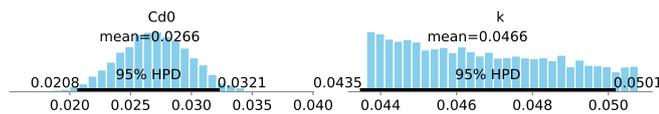
Figure 5. Posterior distribution of C_{D0} and k from MCMC sampling

B. Multiple aircraft types

In this part, we extend the experiment from one flight example to a multiple flight scenario. Based on a one-month data set containing flights for 20 aircraft types, we are able to obtain the drag polar parameters for these common commercial aircraft types. In Fig. 6, the distributions of C_{D0} and k are shown in box plots.
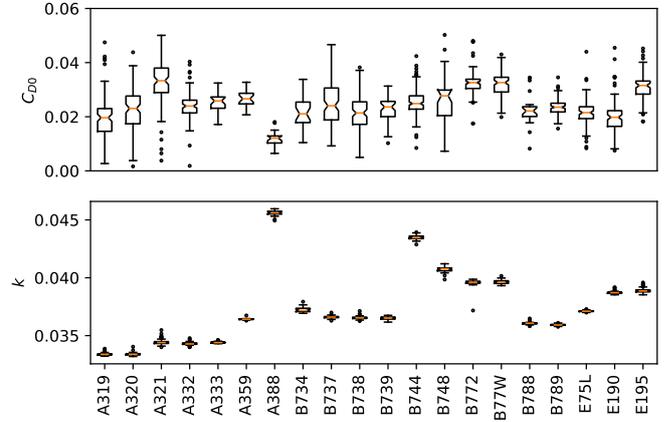
Figure 6. Drag polar of common aircraft models based on multiple flights

We can observe that most of the C_{D0} values are between 0.02 and 0.04. These values are close to values obtained from empirical data that is described in [7]. Lift-induced drag coefficient k has smaller variations than the zero-lift drag coefficient C_{D0} in each aircraft model. This is due to the small range of e as defined in Equation 5. In our model, the primary influential factor for k is the aspect ratio of the aircraft, which is assumed to be constant.

Due to the restriction of the BADA license, exposure of specific BADA coefficients is not permitted. Instead of showing the exact difference between BADA and our model for each aircraft type, the overall statistics for all available aircraft are shown in Fig. 7. The mean absolute difference for all C_{D0} and k are found to be 0.005 and 0.003 respectively.
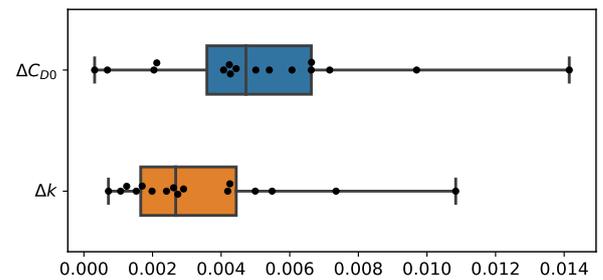
Figure 7. Absolute difference of drag polar between BADA (version 3.12) and model derived in this paper under clean configuration. (Note: A359, B789, E75L, and E195 are not presented in BADA model)

C. Drag polar under non-clean configurations

In Section II-B, we explained the theoretical models that define the additional drag caused by structural changes (flaps

TABLE VI
STE MODEL HIERARCHICAL PARAMETERS

Parameter	σ_v	σ_a	σ_{vs}	σ_τ	σ_h	σ_γ	μ_{cd0}	σ_{cd0}	e_{min}	e_{max}	μ_m	σ_m
Value	5	0.5	2	2	20	2	0.02	0.1	0.85	0.99	289600	107200
Unit	m/s	m/s^2	m/s	K	m	deg	-	-	-	-	kg	kg

and land gears). For the increase of C_{D0} due to flap deployment, a generic flight phase dependent ΔC_{D0} is used. The values for take-off, initial climb, approach, and landing are 0.03, 0.02, 0.03, and 0.08 respectively.

The increase of C_{D0} due to extension of landing gear is computed using Equation 7. The results are shown in Table VII. The increase of C_{D0} due to flaps and landing gears are incorporated with drag polar under clean configurations using Equation 8. The final drag polar models for all 20 aircraft types in different flight phase are also derived. They are listed in Table VIII.

TABLE VII
INCREASE OF ZERO-LIFT DRAG COEFFICIENT DUE TO LANDING GEAR

Aircraft	Aircraft model	taxing	take-off	landing
A319	Airbus A319	0.031	0.024	0.017
A320	Airbus A320	0.032	0.025	0.017
A321	Airbus A321	0.037	0.029	0.020
A332	Airbus A330-200	0.025	0.020	0.014
A333	Airbus A330-300	0.025	0.020	0.014
A359	Airbus A350-900	0.024	0.019	0.013
A388	Airbus A380-800	0.022	0.017	0.012
B734	Boeing 737-400	0.034	0.026	0.018
B737	Boeing 737-700	0.029	0.023	0.016
B738	Boeing 737-800	0.032	0.025	0.017
B739	Boeing 737-900	0.034	0.026	0.018
B744	Boeing 747-400	0.027	0.021	0.015
B748	Boeing 747-8	0.028	0.022	0.015
B772	Boeing 777-200ER	0.026	0.020	0.014
B77W	Boeing 777-300ER	0.029	0.023	0.016
B788	Boeing 787-8	0.024	0.019	0.013
B789	Boeing 787-9	0.026	0.020	0.014
E75L	Embraer E175 (LR)	0.031	0.024	0.017
E190	Embraer E190 (LR)	0.030	0.023	0.016
E195	Embraer E195 (LR)	0.030	0.024	0.017

TABLE VIII
DRAG POLAR COEFFICIENTS

Aircraft	$C_{D0,clean}$	$C_{D0,taxi}$	$C_{D0,to}$	$C_{D0,ic}$	$C_{D0,ap}$	$C_{D0,ld}$	k	e
A319	0.020	0.051	0.074	0.040	0.050	0.116	0.0334	0.9222
A320	0.023	0.055	0.078	0.043	0.053	0.120	0.0334	0.9220
A321	0.033	0.069	0.091	0.053	0.063	0.133	0.0344	0.9244
A332	0.024	0.049	0.074	0.044	0.054	0.118	0.0343	0.9230
A333	0.026	0.051	0.075	0.046	0.056	0.120	0.0344	0.9199
A359	0.027	0.051	0.075	0.047	0.057	0.120	0.0364	0.9210
A388	0.012	0.034	0.059	0.032	0.042	0.104	0.0456	0.9269
B734	0.021	0.060	0.081	0.041	0.051	0.122	0.0372	0.9339
B737	0.024	0.053	0.077	0.044	0.054	0.120	0.0366	0.9213
B738	0.021	0.053	0.076	0.041	0.051	0.119	0.0365	0.9231
B739	0.024	0.057	0.080	0.044	0.054	0.122	0.0365	0.9233
B744	0.025	0.052	0.076	0.045	0.055	0.119	0.0435	0.9271
B748	0.028	0.056	0.079	0.048	0.058	0.123	0.0408	0.9237
B772	0.033	0.059	0.083	0.053	0.063	0.127	0.0396	0.9269
B77W	0.033	0.062	0.085	0.053	0.063	0.129	0.0396	0.9269
B788	0.022	0.046	0.071	0.042	0.052	0.115	0.0361	0.9202
B789	0.024	0.050	0.074	0.044	0.054	0.118	0.0359	0.9241
E75L	0.022	0.053	0.076	0.042	0.052	0.119	0.0371	0.9223
E190	0.020	0.050	0.073	0.040	0.050	0.116	0.0387	0.9219
E195	0.032	0.062	0.085	0.052	0.062	0.128	0.0389	0.9183

V. DISCUSSION

In this section, experimental results are further analyzed. Assumptions that were made to simplify the problem are reflected upon. Finally, we also address the limitations and uncertainties in the proposed methods.

A. The lift-induced drag coefficient

For different flap settings, we assume a constant lift-induced drag coefficient (k) in our model. Theoretically, the wet surface of the aircraft increases when flaps are deployed, which, in turn, changes the aspect ration (AR) slightly. At the same time, the Oswald factor (e) may also differ. However, the magnitude of the change to k due to flap deployment is small compared to C_{D0}. When computing the aircraft drag using the drag polar, the influence of k is also much smaller than C_{D0}. Considering all these factors, the model for k in this paper is simplified.

In Fig. 6, it can be observed that the variances for k are much smaller than C_{D0}. This is an expected result. Based on Equation 5, the only varying parameter for the STE model is the span efficiency e. The change caused by e is much smaller than the aspect ration (AR) for different aircraft types.

B. The STE model, MCMC sampling, and uncertainties

It is important to understand that in the proposed hierarchical model, all parameters are considered as random variables (described by probability density functions) instead of scalar values. Most parameters (except C_{D0} and k) are also expressed as time-varying random variables. The time series variables are constructed as multi-dimensional probability density functions. The solution of this hierarchical stochastic model is only possible thanks to numerical approximation using the Markov chain Monte Carlo (MCMC) techniques.

We discussed the principles of the MCMC in Section III-B. The application of this Bayesian data approach is slowly being brought to the engineering domain. In [20], we can see the application of the Bayesian hierarchical model and the use of the MCMC method for inverse problems in the aerospace domain. The STE model and solutions proposed in this paper relate closely to this line of research.

In the proposed STE aircraft performance model, two major sources of uncertainty exist, which are aircraft mass and thrust. We cannot estimate the mass using the method from [13] since accurate estimation would require knowledge of the aircraft drag polar. Instead, we consider the mass as a bounded normal distribution. This means that all possible masses are considered, while sampled together with all other parameters. When evaluating the sampling results, we can illustrate the mass posterior distribution at different flight time steps, as shown in Fig. 8. Similarly, the uncertainties can also be illustrated in a box plot as shown in Fig. 9.
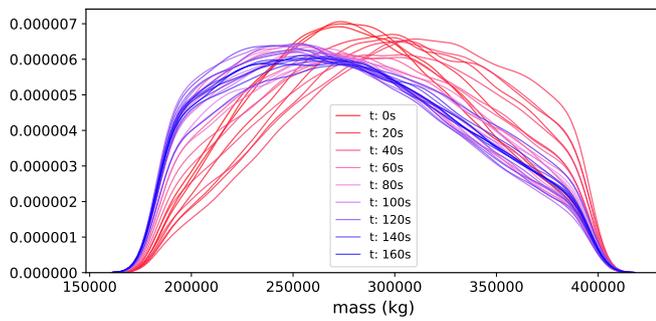
Figure 8. The kernel density approximation of mass at different flight time step
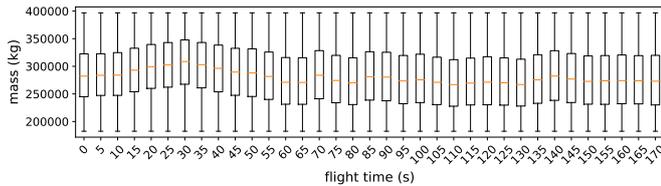
Figure 9. The distributions of mass at different flight time step

The distributions show the probabilities of mass at different flight time step. The goal is to estimate the drag polar coefficients C_{D0} and k, not the mass of an aircraft. The posterior distributions can be thought of as the results that optimize all different parameters at all time steps.

The aircraft thrust model is the factor that influences the results the most while using the STE model, as the thrust is strongly correlated to the drag of the aircraft. An increase in the uncertainty of the thrust could directly lead to a higher uncertainty of C_{D0}. When there is more than one possible engine configuration for a specific aircraft type, we selected the one with the highest thrust. It can be observed that the uncertainty of C_{D0} increases with the number of engine options, as shown in Fig. 6. For example, Airbus A320 and Boeing B737 series have the most number of engine options. On the contrary, aircraft that have only one or two engine options have shown less C_{D0} uncertainty in general, for example, A350, A380, and B787.

In the STE model, we implement the thrust model based on two-shaft turbofan engines [12], which is one of the most accurate open models to the best of our knowledge. The thrust model has an uncertainty of 4.5% on computing thrust when compared to engine performance data. This uncertainty in thrust model undoubtedly passes on to the final uncertainty of the drag polar estimates.

VI. CONCLUSIONS

In this paper, we propose a method to construct point-mass drag polar models, offering a solution to the current lack of an open aerodynamic model for air traffic research. The proposed stochastic total energy (STE) model is able to infer the aircraft drag polar using ADS-B and Mode-S surveillance data. The STE model requires basic information about the aircraft, which is publicly available. Thanks to Bayesian computing (MCMC

sampling), the STE method can generate a drag polar for almost any fixed-wing aircraft (with turbofan engine), as long as sufficient flight data is available.

The methods and results of this paper are not without their own limitations. The level of model fidelity depends greatly on the level of uncertainty in the thrust model. In this paper, we used an open thrust model proposed in the literature. It is a generalized empirical model based on existing engine performance data. The implications of different uncertainties are discussed in the paper.

Finally, as one of the main contributions of this paper, we produced a comprehensive list of drag polar models for the 20 most common commercial aircraft types. To the best of our knowledge, this is the first time such models have been derived and shared publicly with the research community.

REFERENCES

[1] S. F. Hoerner, *Aerodynamic Drag: Practical Data on Aerodynamic Drag Evaluated and Presented*. Otterbein Press, 1951.
[2] S. F. Hoerner, *Fluid-dynamic Drag: practical information on aerodynamic drag and hydrodynamic resistance*. Hoerner Fluid Dynamics, 1958.
[3] J. D. Anderson, *Aircraft performance and design*. McGraw-Hill Science/Engineering/Math, 1999.
[4] G. Ruijgrok, *Elements of airplane performance*, vol. 2. Delft Academic Press, 2009.
[5] E. Obert, *Aerodynamic Design of Transport Aircraft*. IOS Press, 2009.
[6] E. Torenbeek, *Synthesis of subsonic airplane design*. Delft University Press, 1976.
[7] G. W. H. van Es, "Rapid estimation of the zero-lift drag coefficient of transport aircraft," *Journal of aircraft*, vol. 39, no. 4, pp. 597–599, 2002.
[8] A. Nuic, "User manual for the base of aircraft data (bada) revision 3.12," *EUROCONTROL*, 2014.
[9] L. K. Loftin, *Quest for performance: The evolution of modern aircraft*. Scientific and Technical Information Branch, National Aeronautics and Space Administration, 1985.
[10] B. W. McCormick, *Aerodynamics, aeronautics, and flight mechanics*, vol. 2. Wiley New York, 1995.
[11] W. Mair, *Aircraft performance*, vol. 5. Cambridge University Press, 1996.
[12] M. Bartel and T. M. Young, "Simplified thrust and fuel consumption models for modern two-shaft turbofan engines," *Journal of Aircraft*, vol. 45, no. 4, pp. 1450–1456, 2008.
[13] J. Sun, H. A. Blom, J. Ellerbroek, and J. M. Hoekstra, "Aircraft mass and thrust estimation using recursive bayesian method," in *8th International Conference on Research in Air Transportation*, 2018.
[14] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
[15] R. M. Neal, "An improved acceptance procedure for the hybrid monte carlo algorithm," *Journal of Computational Physics*, vol. 111, no. 1, pp. 194–203, 1994.
[16] M. D. Hoffman and A. Gelman, "The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo.," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1593–1623, 2014.
[17] J. Sun, H. V, J. Ellerbroek, and J. M. Hoekstra, "Weather field reconstruction using aircraft surveillance data and a novel meteo-particle model," *PLOS ONE*, vol. 13, pp. 1–33, 10 2018.
[18] J. Sun, H. Vû, J. Ellerbroek, and J. Hoekstra, "Ground-based wind field construction from mode-s and ads-b data with a novel gas particle model," in *Seventh SESAR Innovation Days*, SESAR, 2017.
[19] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, "Probabilistic programming in python using pymc3," *PeerJ Computer Science*, vol. 2, p. e55, 2016.
[20] J. B. Nagel and B. Sudret, "Bayesian multilevel model calibration for inverse problems under uncertainty with perfect data," *Journal of Aerospace Information Systems*, vol. 12, no. 1, pp. 97–113, 2015.