Engine Thrust Model Determination from Large Operational Flight Data Base

Christoph Deiler — Scientist, DLR – German Aerospace Center, Institute of Flight Systems, 38108, Braunschweig, Germany. christoph.deiler@dlr.de

ABSTRACT

Different engine thrust models are developed from operational flight data with limited apriori knowledge as part of a novel process for aircraft flight performance model determination. The given big data problem is solved by application of fundamental engineering knowledge and a specific data evaluation strategy. The resulting smart data approach is fundamentally different from existing artificial intelligence methods to solve such big data problems. A linear, a local-linear and a complex nonlinear thrust model are determined on the example of a given large database of operational flights with Airbus A 320neo aircraft. Even with limited information about the actual engine thrust from the available data, the resulting models allow to (well) predict the engine thrust characteristics within the required flight envelope. Finally, the characteristics of the different thrust model implementations and results are discussed.

Keywords: Engine Thrust Model; Flight Data Analysis; System Identification; Aircraft Flight Performance

Nomenclature

A	=	regressor matrix
C_D	=	drag coefficient, 1
d_γ	=	weighted Euclidean norm
g	=	gravitational acceleration, m/s^2
H	=	altitude, m
i	=	data index
J	=	Jacobian matrix
$\mathbb{L}$	=	cost function
Ma	=	Mach number 1
m_{AC}	=	aircraft mass, kg
n	=	number of measurements
N_1	=	engine fan speed, %
N_{BP}	=	number of breakpoints
$n_{\mathrm{x,a}}$	=	load factor in aerodynamic x-direction, 1
p	=	number of parameters
$\overline{q}$	=	dynamic pressure, Pa
r	=	residual
R^2	=	coefficient of determination, 1
S	=	wing surface area, m^2

T_{net}	=	net engine thrust force, N
T_{req}	=	required thrust force per engine, N
$T_{tot,x}$	=	total thrust in body-fixed x-direction, N
$T_{tot,x,a}$	=	total thrust in aerodynamic x-direction, N
y	=	simulated model output
z	=	measured output
$\bar{z}$	=	average of measured output
α	=	angle of attack, rad
β	=	angle of sideslip, rad
ε_{xy}, σ	=	engine installation angles (toe out, inclination), rad
γ	=	Tikhonov weight
Γ_2^*	=	Tikhonov regularization matrix
θ	=	engine model parameter
$\widehat{\theta}$	=	engine model parameter estimate

1 Introduction

Aircraft operations are mainly driven by the aircraft's flight performance. Therefore, new developments always target better aerodynamic performance and propulsion system efficiency with less energy consumption in order to reduce the operational costs and environmental impact of each individual flight. Hence, the optimization of an aircraft's flight performance is key to sustainable future aviation. Europe's Flight Path 2050 [1] describes the vision of technologies and procedures available in the year 2050 that will allow a reduction of 75% of the $CO2$ emissions and 90% of the NOx emissions per passenger kilometer compared to the year 2000. In addition, the perceived noise emission of an aircraft in flight must be reduced by 65%.

Within a short-term perspective, the reduction of emissions and pollution can only be achieved by new technologies enabling today's modern aircraft to be even more environmentally friendly and allow a greener aviation. For the best possible reduction of emissions with current fleets an optimization of aircraft operations is required using smart flight control strategies. One of these strategies is DLR's Low Noise Augmentation System (LNAS) [2, 3], which is able to optimize the aircraft's descend and approach by advising the pilots with optimized autopilot commands and configuration points. It has proven during flight tests that its application can save up to 25% of fuel and emissions for certain approach flight phases. In addition, measurements on ground revealed that a significant noise reduction could be obtained when using LNAS: with an overall reduction on ground along the approach path certain areas show a reduction up to 5dB(A). But the system relies on a high-quality simulation model for aircraft flight performance evaluation which is used within the internal flight path analysis and optimization process.

There are several ways to obtain such simulation models which base on different sources of information about the distinct aircraft type. For example, the simulation model formulation could be based on published engine data, e.g., from handbooks or aircraft manuals [4]. Unfortunately, aircraft manuals do normally not contain the engine thrust values but only engine state information, which makes it impossible to obtain engine thrust models without additional information. Furthermore, one could use high-fidelity aircraft and engine design models to obtain the necessary informations for generating aircraft flight performance models. Such an approach is mainly used by aircraft manufacturers because they have full insight in the aircraft design process. If available, ground test data and wind tunnel results can also be used to determine a propulsion system model, e.g., [5], but the corresponding informations are also mostly limited to engine manufacturers. Another way to obtain the required informations and determine flight performance models is the conduction of special flight test programs with an aircraft of

the specific type and use flight data recordings to extract the information about the flight performance. But even with extensive and expensive flight test programs, additional apriori information on e.g., engine thrust might be necessary to develop reliable flight performance models.

A completely different fourth way is to only use flight data gathered during operational flights, which can be easily recorded daily. But without dedicated flight test procedures, this will pose a big data problem which has to be solved. Nevertheless, the desired information is inside this database, and if the database is large enough, the information can be extracted by application of proper methodologies and algorithms. Big data problems and artificial intelligence methods to solve these are omnipresent today. But to solve a certain big data problem with minimal effort, the simple application of artificial intelligence algorithms is no smart solution. A smarter way to solve a big data problem in engineering is to apply as much fundamental knowledge about the underlying system as possible. This way big data is transferred to smart data and the initially posed problem can be solved with much simpler, faster and presumably more deterministic methods. For operational flight data this means that engineering knowledge about aircraft flight mechanics is applied and the data analysis process is designed accordingly. Doing so, well established model formulations can be used which further allow a direct interpretation of the resulting model, e.g., evaluation of lift-to-drag ratio. In this way, the big data analysis of operational data can directly help to reduce aircraft emissions and make aviation sustainable.

The challenge for such work is the correlation between aerodynamics and engine thrust which is present in the observed aircraft data during daily operations. The major task is to extract the corresponding information from flight data. Hence, the main assumption for this work is that big data allows decorrelating drag and thrust if the aerodynamics do vary due to flap/slat configuration changes and engine thrust is not mainly dependent on these changes. It means that similar engine thrust conditions do exist for different aircraft configurations and consequently drag which then allow determining each part individually. Moreover, using this assumption it is not only possible to estimate aircraft drag and thrust but also get unbiased values for both in terms of nearly correct zero-lift drag and engine idle thrust.

This paper presents a part of the work on solving the described big data problem in a smart and intelligent engineering way in order to reveal the flight performance, aerodynamics and engine thrust from operational flight data of three different Airbus A 320neo (New Engine Option) aircraft. It is based on previous DLR research on the determination of flight performance variations within operational flight data [6, 7] and has a different focus than already established methods for flight performance monitoring, e.g., [8–10]. A distinct process is developed in order to reduce the data size and necessary computational effort. First, a flight data preprocessing allows to select only the information required for the further flight performance analysis and already make all computations only required once in the process before starting any further processing (loop). Second, the data is processed and analyzed based on distinct engineering knowledge which does not require unstructured artificial intelligence algorithms or highly complex evaluation methods. The process does contain state-of-the-art data processing and system identification techniques and is mainly based on aerospace engineering knowledge. Again, this has the advantage that the resulting models can be directly evaluated by analyzing the corresponding parameters (e.g., aerodynamics) and model outputs (e.g., engine thrust).

Fig. 1 gives an illustration of the data evaluation process used for this work. This process contains five individual blocks necessary to process and analyze the data and to determine the desired models:

a) Flight Data Processing: processing the flight data by e.g., performing unit conversions, transfer all relevant measurements to CG, determination of AC configuration, anti-ice system status and flight condition, calculation of the atmospheric parameters using the international standard atmosphere.

b) Aerodynamics Evaluation: calculation of aerodynamics with given model and from flight data depending on available thrust information with simplified or complete equations.

3

c) Engine Thrust Model Identification: extraction of relevant data for thrust model identification, calculation of required thrust for the thrust model parameter estimation, engine thrust prediction in relation to the given flight data measurements with the identified model.

d) Aerodynamics Calculation: calculation of aerodynamic coefficients based on flight data and predicted engine thrust as reference for the aerodynamic model update.

e) Aerodynamics Model Identification: new estimation of aerodynamic model parameters for all different aircraft configurations.
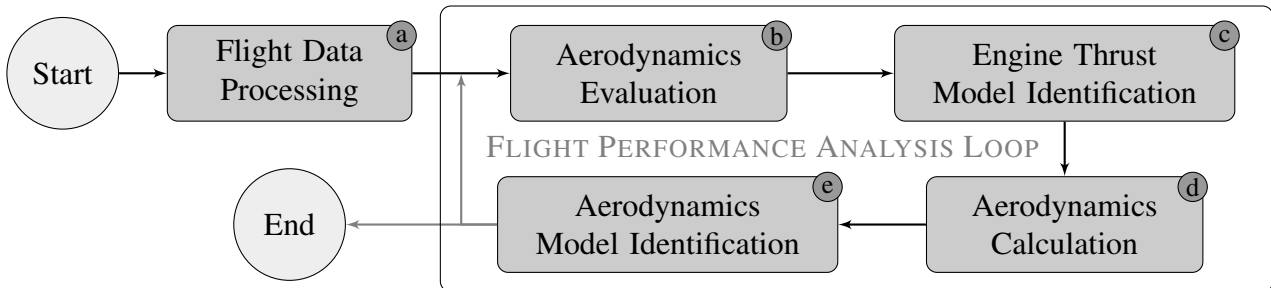
Fig. 1   Basic scheme of the evaluation process used for the flight performance determination (aerodynamic model and thrust model identification).

The herein presented work is focused on the "Engine Thrust Model Identification" in block (c) using given information about the aircraft aerodynamics to obtain the required thrust information. The paper is structured as follows, giving information about:

1) the data base of operational flight data (section 2);
2) a general description of engine thrust modeling for dynamic aircraft simulation (section 3);
3) definition and prediction of a global linear thrust model (section 4);
4) formulation of local linear thrust models (section 5);
5) estimation of a complex nonlinear multi-dimensional thrust model table (section 6);
6) comparison of different engine thrust model approaches (section 7);

A summary and conclusions on the overall results are given in section 8.

2 Flight Data Base

This work is based on flight data recorded during operational flights on the pilots' electronic flight bags (EFBs). The DLR-developed recorder was installed on the EFBs and used with different Airbus A 320neo aircraft of a major European airline in preparation of the SESAR Very Large Demonstration (VLD2) ALBATROSS project[1]. The recorder is able to record messages sent via ARINC 429, ARINC 834 or ARINC 717 using different interfaces depending on the EFBs it is installed on.

The flight data base is limited to reasonable airspeeds and altitudes sufficiently high above ground to guarantee the aircraft flying outside of the ground effect. Therefore, the following initial limits are used for the evaluation presented herein: height above ground above 50 ft and barometric altitude above 500 ft, whichever comes first, as well as true airspeed above 130 kts. These limits do no restrict the evaluation results at all, but allow to circumvent any special conditions of flight near ground. A brief overview of the flight data base as a result of the data pre-processing is given in Tab. 1. The number of data points does only include the data which was used for the evaluation and is therefore less than the total size of all data. Fig. 2 visualizes the distribution of available flight data on an altitude-Mach plane split into the different aircraft flap/slat configurations.

Table 1   Brief overview of flight data base used for evaluation.

number of aircraft		3
number of flights		844
number of data points		55 479 606
data envelope	min	max
baro altitude	500 ft	20 964 ft
Mach	0.1866	0.7311
weight	48.57 t	73.44 t

flap/slat configuration	data points
FLAP 0	43 056 406
FLAP 1	1 914 563
FLAP 1+F	1 547 312
FLAP 2	4 514 991
FLAP 3	1 069 383
FLAP FULL	3 376 951
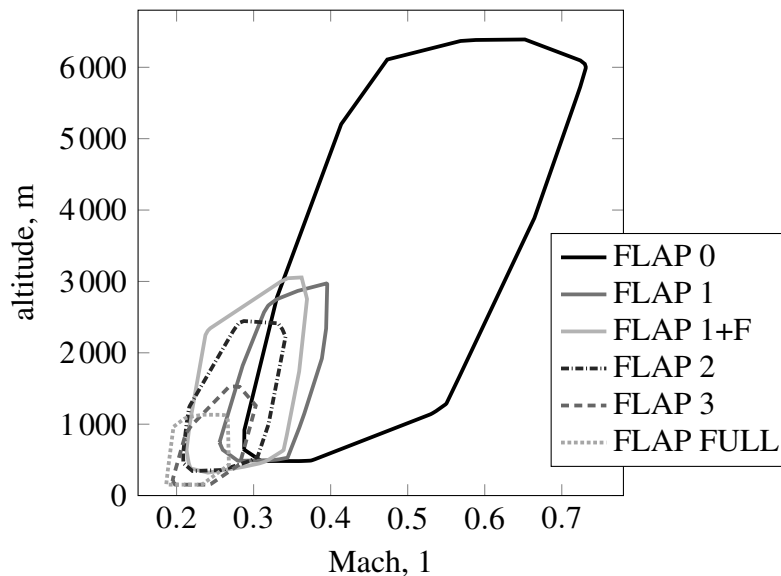
Fig. 2   Flight data envelope: altitude-Mach diagram for all flap/slat configurations.

The flight data sets used for the flight performance model determination are reduced to the relevant information. Initially, the data was recorded directly from aircraft buses/data streams and had to be processed in order to contain individual data channels with a sampling rate of 50 Hz. For each flight, the data sets finally contain:

- time stamp;
- body-fixed accelerations: longitudinal, lateral, vertical;
- angular rates: pitch, yaw, roll;
- inflow/atmosphere: corrected angle of attack, estimated angle of sideslip, true airspeed, Mach number, static temperature;
- altitude: barometric, barometric corrected, GPS, radio height;
- engine fan speed (left/right);
- aircraft gross weight, longitudinal CG position;
- configuration: landing gear extension (main left/right, nose), flap position, slat position;
- control inputs: spoiler deflection (left/right, spoilers 2 to 5), horizontal stabilizer deflection;
- anti-ice system status: wing, engine.

There is one major restriction with the available data sets: the records do not contain any information about the elevator deflection. It does only slightly affect the aircraft flight performance by a minor change of the overall aircraft lift, but is necessary to correctly predict the flight performance. But, with elevators only deflected during a short time during flight with the A 320, it can be considered as an additional

uncertainty within the overall process and is mainly relevant for the aerodynamics prediction outside the work presented herein.

Not all data of the recorded flight data sets can be used for the evaluation and model determination. There are some requirements which are essential for the process shown in Fig. 1. First of all, the aircraft's flap/slat configuration must be fixed in order to successfully determine the corresponding aerodynamics. There is no need to determine the aerodynamics during configuration changes. Moreover, this would introduce further uncertainties into the model determination and estimation process and would not be beneficial to achieve the goals of this work. Therefore, all data marked as containing moving slat, flap or gear is set to invalid for the further data evaluation. In addition, data with low airspeed and altitude is also excluded from the evaluation process as already described above. Also, the data is checked for segments containing unreliable measurements (e.g., not-a-number or completely wrong values[2]), which lead to an exclusion of these segments or the whole flight depending on the severity of this unreliability. Finally, this leads to the selection of the about 55.5 million data points given in Tab. 1.

3 Engine Thrust Model for Dynamic Aircraft Simulation

Engine thrust prediction is essential for any flight performance evaluation. Therefore, the engine thrust model development is one essential apart of the process described herein (see block (c) in Fig. 1), because there was no apriori information about the Pratt & Whitney "PW1100G" engine[3] thrust available for this work. The given flight data base is searched for specific symmetric engine conditions and the results are split into individual data sets for each anti-ice system configuration[4]. An overview of the data selection for the engine thrust model development is given in Tab. 2.

Table 2   Distribution of data in flight data base: anti-ice system status.

anti-ice status	data points	
	all	sym. conditions
off	55 367 617	53 143 789
engine on / wing off	111 989	111 685
engine & wing on	0	0

3.1 Required Thrust Using Aircraft Aerodynamics

In the first step, the required thrust is calculated by using the initial guess for the aerodynamics. Later on in the process, the calculation is repeated with the updated aerodynamics. The sum of forces in longitudinal direction of the aerodynamic frame during flight is composed of the current engine required thrust, the drag force as well as the AC mass and the longitudinal acceleration. Solving this equation for the required total thrust yields

$$T_{\text{tot,x,a}} = m_{\text{AC}} \cdot g \cdot n_{\text{x,a}} + C_D \cdot \bar{q} \cdot S. \tag{1}$$

---

[2]this could e.g., result from a wrong configuration of the recorder during flight

[3]used with the Airbus A 320neo in the given flight data base

[4]the A 320 ice protection system is based on thermal anti-ice requiring engine bleed air which directly influences the engine thrust despite similar engine conditions: for example, similar fan speed, airspeed, altitude or air temperature could lead to different thrust conditions with ice protection systems activated.

The required total thrust is further transferred to the body-fixed coordinate system (neglecting thrust force components in vertical and lateral direction):

$$T_{\text{tot,x}} = \frac{T_{\text{tot,x,a}}}{\cos(\alpha) \cdot \cos(\beta)}. \tag{2}$$

For the thrust model definition symmetric engine conditions are considered in order to allow splitting total thrust on both engines equally. Using engine inclination and toe out angle, the approximated required thrust per engine resulting from acceleration measurements and drag force predicted by the assumed aerodynamic model finally results in

$$T_{\text{req}} = 0.5 \cdot \frac{T_{\text{tot,x}}}{\cos(\sigma) \cdot \cos(\varepsilon_{xy})}. \tag{3}$$

Note that the data also includes different states of the anti-ice system. Hence, the data is selected according to the different anti-ice cases (off, engine anti-ice on, wing and engine anti-ice on) and individual thrust models are determined.

3.2 Thrust Model Regressors

The thrust of a jet engine is depended on several engine states and external parameters. Hence, the model formulation must cover the main influences to reliably predict the engine thrust within the required flight envelope and for all relevant thrust settings. The general influence of different parameters on thrust is shown in Fig. 3. It is clearly visible that the jet engine thrust has a complex nonlinear behavior with the variation of individual parameters. In consequence, the main regressors for the engine thrust model used in this work are defined as follows:

- engine fan speed N_1,
- Mach number Ma,
- barometric altitude H,
- (• temperature offset ΔISA).

The latter is used as model regressor but in a subsequent engine thrust correction, especially to reduce the necessary dimensions of the thrust table (see section 6) and is subject to future work. In general, the engine thrust model is defined by

$$T = T_{\text{net}} \cdot f_{\text{corr}}(\Delta\text{ISA}, \dots) = f(N_1, Ma, H) \cdot f_{\text{corr}}(\Delta\text{ISA}, \dots) \approx T_{\text{req}} \tag{4}$$

with the net thrust being a (arbitrary) function of the above named three main influence parameters. In a first step, the temperature correction function is neglected which leads to following approximation:

$$T_{\text{req}} \approx T_{\text{net}} = f(N_1, Ma, H). \tag{5}$$

4 Linear Thrust Model

The simplest model for net thrust is a linear model with the given three regressors and four parameters:

$$T_{\text{net}} = \theta_0 + \theta_1 \cdot N_1 + \theta_2 \cdot Ma + \theta_3 \cdot H \tag{6}$$

Note that although it is well known (see Fig. 3) that engine thrust is strongly and nonlinearly dependent on the several chosen regressors, from an engineering point of view, a linear model is a good choice to start with. If necessary the model must be enhanced step by step to further include or approximate
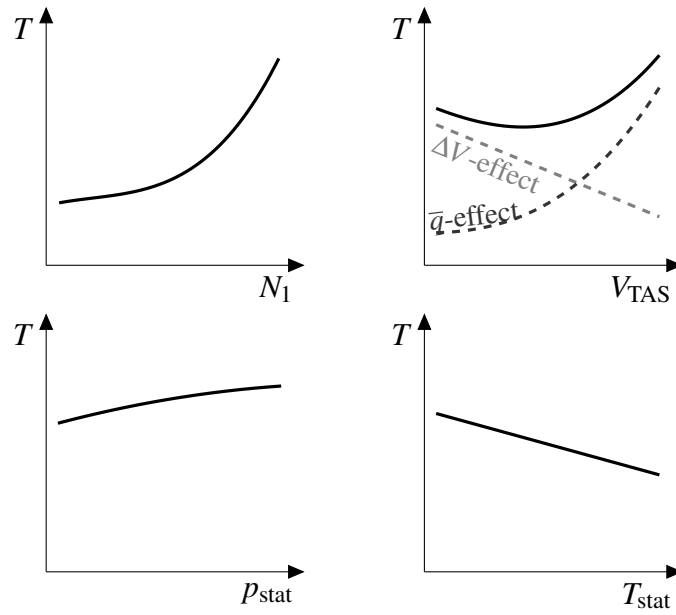
Fig. 3 General influence of different engine/aircraft states and atmospheric parameters on jet engine thrust, according to Refs. [11, 12].

nonlinear thrust characteristics. Hence, the necessary effort to obtain the four parameters is also limited, because the maximum likelihood problem can be solved easily:

$$\widehat{\theta} = \underset{\theta}{\text{argmin}} \left(\theta \to \sum_{i=1}^{n} (z_i - y_i(\theta))^2 \right), \tag{7}$$

with z_i being the required thrust and $y_i(\theta)$ the modeled net thrust T_{net}. Using for example MATLAB®, the linear problem can be formulated and solved directly:

$$A\theta = z \quad \Rightarrow \quad \theta = A \setminus z, \quad \text{with} \quad A_i = [1, N_{1,i}, Ma_i, H_i] \quad \text{and} \quad z_i = T_{\text{req},i}. \tag{8}$$

The standard error vector of the resulting parameters results from the unscaled covariance matrix C and the scaling factor $\widehat{\sigma}^2$

$$\sigma(\theta) = \sqrt{\widehat{\sigma}^2 \cdot \text{diag}(C)}, \quad \text{with} \quad \widehat{\sigma}^2 = \frac{\|z - A\theta\|}{n - p} \quad \text{and} \quad C = \left(A^T A \right)^{-1}. \tag{9}$$

The results for the model parameters and the limits of the linear thrust model are given in Tab. 3. These limits result from the available data within the thrust envelope. The very large standard error for the offset parameter θ_0 and the Mach derivative θ_2 directly indicate the low identifiability of these parameters using the available data set and the low ability of the linear model formulation to represent the given nonlinear engine thrust behavior. An example of the linear thrust model evaluation is given in Fig. 4. The left side plot (Fig. 4a) shows the predicted thrust variation with fan speed and Mach number at fixed altitude, whereas the right side plot (Fig. 4b) illustrates the thrust versus fan speed and altitude at fixed Mach number.

With the linear model the whole nonlinear thrust envelope is covered. But the model delivers wrong predictions in certain areas of the envelope because the nonlinear thrust behavior cannot be predicted sufficiently by the linear model. Furthermore, the linear model will give completely wrong results at the borders of the envelope, where large negative thrust values are predicted (e.g., see Fig. 4 at low N_1 and high Mach or altitude). Hence, the model quality and validity are strictly limited and extrapolation outside the limiting borders will off course give even worse predictions.
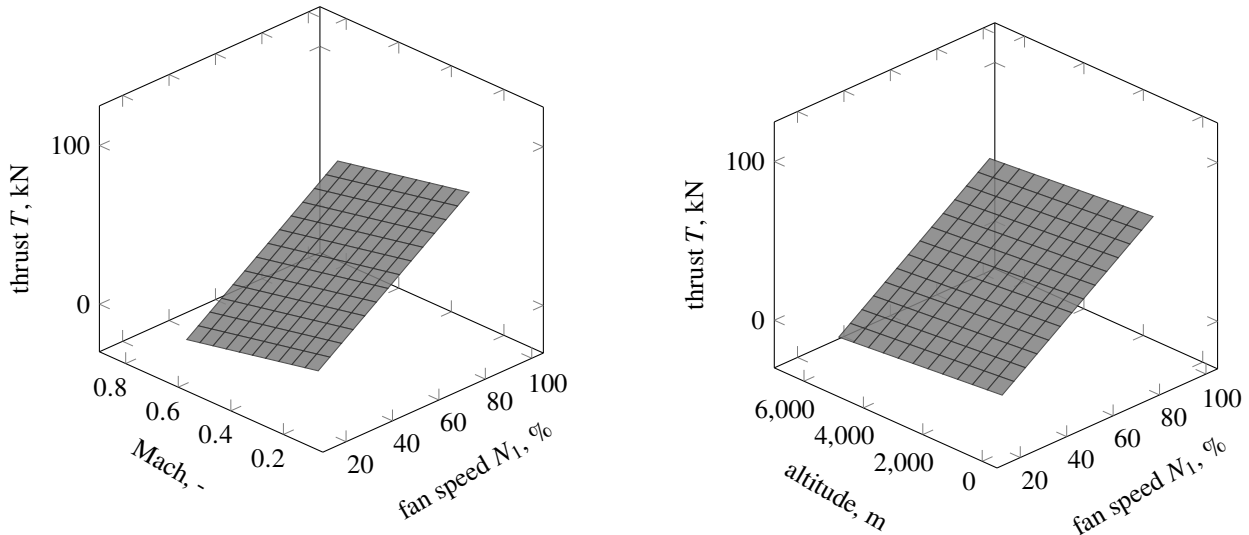
8

Table 3   Linear thrust model parameters and limits

parameter	value	unit	σ in %
θ_0	-14115.48307	-	452.568
θ_1	1076.26550	1/%	3.280
θ_2	-35154.81252	-	1671.046
θ_3	-1.76190	1/m	0.120

dimension	min	max
N_1	24.0 %	90.0 %
Ma	0.19	0.73
H	150.0 m	6390.0 m

(a) variation of engine fan speed and Mach number at $H = 3000\,\text{m}$

(b) variation of engine fan speed and altitude at $Ma = 0.4$

Fig. 4   Linear thrust model evaluations at example conditions; anti-ice off.

5 Local Linear Thrust Models

A set of local linear models allows to better approximate nonlinear functions than one single linear model. Therefore, it is reasonable to use these in the given evaluations. Hence, the complete data is segmented and the regression problem formulated in equation (8) is solved for each data set.

The boundaries of the local linear models used for this evaluation – given in Tab. 4 – result in 48 potential models within the envelope. Note that these boundaries could significantly exceed the given data – e.g., the maximum altitude represented in the data set – but this does not affect the model estimation results. Furthermore, the boundaries are extended for the regression by the values given in the last row of Tab. 4 to potentially reduce the gaps between the individual models: if the local model is forced to also cover the thrust slightly outside the boundaries, the discontinuous behavior at the boundaries is assumed to be reduced to a minimum, accepting that the models might lose some ability to predict the required thrust within the boundaries.

A measure for the quality of the local linear model is the coefficient of determination, which is defined by the ratio between the sum of squares of the residuals and the total sum of squares

$$R^2 = 1 - \frac{\sum_i (z_i - y_i)^2}{\sum_i (z_i - \bar{z})^2}. \tag{10}$$

Table 4 Boundaries of local linear models for breakpoint definition.

fan speed	Mach number	altitude
20.0 % - 50.0 %	0.15 - 0.30	0.0 m - 2000.0 m
50.0 % - 75.0 %	0.30 - 0.45	2000.0 m - 4000.0 m
75.0 % - 100.0 %	0.45 - 0.60	4000.0 m - 6000.0 m
	0.60 - 0.75	6000.0 m - 8000.0 m
Δ 5 %	0.025	250 m

A value of 1 indicates a perfect match of model output and measurement whereas a negative value reveals a very poor measurement prediction. Using this, a set of valid models can be extracted from the local linear models within the thrust envelope. Tab. 5 gives an overview of the obtained models with a coefficient of determination above 0.6. The different numbers of data points used for regression indicate the unbalanced distribution of data within the whole thrust envelope, which further indicates the problem of identifying all potential 48 local models with the same quality. Although some models, e.g., number 13 or 33, show high values of R^2, it is very questionable if these values are meaningful because less than 4000 data points were used for the regression and R^2 itself does not contain any information about the specific data distribution within the local boundaries. Therefore, only results for models with high coefficients of determination and a large data set used for regression should be considered to be reliable.
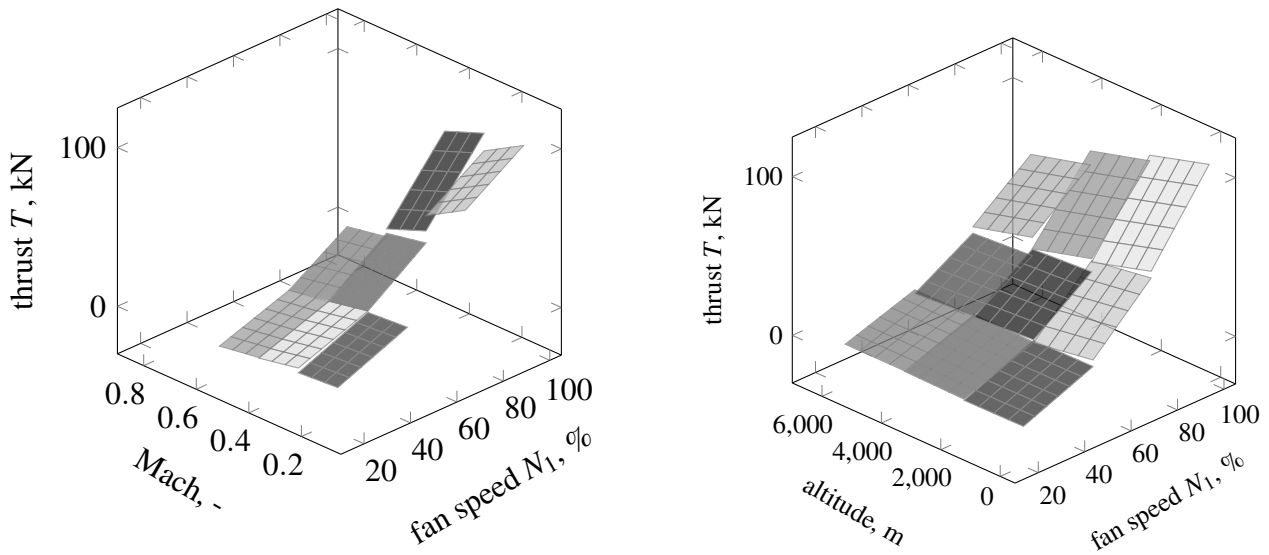
An example of the local linear thrust model evaluation is given in Fig. 5. Similar to Fig. 4, the model evaluations are made and visualized for an example altitude (Fig. 5a) and Mach number (Fig. 5b). The local linear models do better approximate the nonlinear engine thrust behavior with the given regressors than a global linear model. Furthermore, a change of the local model boundary definition could allow an even better approximation of the nonlinear engine thrust behavior. Nevertheless, the number of models is limited to the given data in the way that models can only be determined if enough data is available within a certain part of the envelope and therefore within the given boundaries. With focus on a high model quality, a distinct number of measurements with sufficient (local) variation of regressors must be available to estimate the individual model parameters.

Furthermore, to allow a continuous evaluation of the engine thrust models beyond their local boundaries and to switch from one model to another, model stitching would be necessary. Therefore, the models must be, for example, steady across the boundaries on the one hand or a suitable interpolation between the models over the boundaries must be made on the other. But as shown in Fig. 5 the local linear models are not steady across the model boundaries, which means that a different model parameter estimation technique than a simple, unrestricted regression as given in equation (8) would be necessary to include a local steadiness requirement. But this will reduce the quality of the local linear model fit with the given data because the additional requirement limits its flexibility. Moreover, without a suitable local steadiness at the model boundaries a stitching is to achieve a smooth transition between the multi-dimensional models is highly complex, very difficult or gets even impossible. For example: Fig. 5a shows for low Mach numbers and high fan speeds significantly different gradients of the resulting local linear models. Hence, no "simple" interpolation technique can be applied to stitch the models. Different approaches with e.g., additional linear intermediate models to fill the gaps between the local linear models would act similarly like a finer grid, but there is no guarantee of success. If the models provide local steadiness, e.g., model blending [5, 13] can be used as stitching technique.

Fig. 6 shows two examples for local linear models: one with a coefficient of determination near 1 representing a very good model fit, and another with a negative value of R^2 indicating a poor thrust pre-

Table 5   Valid local linear models after estimation with $R^2 > 0.6$.

#	fan speed range	Mach range	altitude range	R^2	data points
1	20.0 % - 50.0 %	0.15 - 0.30	0.0 m - 2000.0 m	0.94403	6 685 817
2	50.0 % - 75.0 %	0.15 - 0.30	0.0 m - 2000.0 m	0.70906	514 930
3	75.0 % - 100.0 %	0.15 - 0.30	0.0 m - 2000.0 m	0.77658	1 143 763
4	20.0 % - 50.0 %	0.15 - 0.30	2000.0 m - 4000.0 m	0.86447	4 319
5	75.0 % - 100.0 %	0.15 - 0.30	2000.0 m - 4000.0 m	0.74421	4 982
6	20.0 % - 50.0 %	0.30 - 0.45	0.0 m - 2000.0 m	0.89705	6 708 667
7	50.0 % - 75.0 %	0.30 - 0.45	0.0 m - 2000.0 m	0.93024	794 581
8	75.0 % - 100.0 %	0.30 - 0.45	0.0 m - 2000.0 m	0.79404	2 572 856
9	20.0 % - 50.0 %	0.30 - 0.45	2000.0 m - 4000.0 m	0.91042	4 812 148
10	50.0 % - 75.0 %	0.30 - 0.45	2000.0 m - 4000.0 m	0.92853	700 452
11	75.0 % - 100.0 %	0.30 - 0.45	2000.0 m - 4000.0 m	0.76992	2 591 370
12	20.0 % - 50.0 %	0.30 - 0.45	4000.0 m - 6000.0 m	0.92923	46 158
13	50.0 % - 75.0 %	0.30 - 0.45	4000.0 m - 6000.0 m	0.89251	3 200
14	75.0 % - 100.0 %	0.30 - 0.45	4000.0 m - 6000.0 m	0.75458	30 259
15	20.0 % - 50.0 %	0.45 - 0.60	0.0 m - 2000.0 m	0.90324	273 320
16	50.0 % - 75.0 %	0.45 - 0.60	0.0 m - 2000.0 m	0.93493	35 540
17	75.0 % - 100.0 %	0.45 - 0.60	0.0 m - 2000.0 m	0.88264	7 702
18	20.0 % - 50.0 %	0.45 - 0.60	2000.0 m - 4000.0 m	0.91852	6 194 572
19	50.0 % - 75.0 %	0.45 - 0.60	2000.0 m - 4000.0 m	0.93780	1 822 825
20	75.0 % - 100.0 %	0.45 - 0.60	2000.0 m - 4000.0 m	0.77543	2 949 163
21	20.0 % - 50.0 %	0.45 - 0.60	4000.0 m - 6000.0 m	0.91386	6 321 329
22	50.0 % - 75.0 %	0.45 - 0.60	4000.0 m - 6000.0 m	0.94956	1 327 218
23	75.0 % - 100.0 %	0.45 - 0.60	4000.0 m - 6000.0 m	0.88654	3 490 848
24	20.0 % - 50.0 %	0.45 - 0.60	6000.0 m - 8000.0 m	0.90735	14 731
25	50.0 % - 75.0 %	0.45 - 0.60	6000.0 m - 8000.0 m	0.72041	4 991
26	20.0 % - 50.0 %	0.60 - 0.75	2000.0 m - 4000.0 m	0.84552	103 211
27	50.0 % - 75.0 %	0.60 - 0.75	2000.0 m - 4000.0 m	0.97192	71 962
28	75.0 % - 100.0 %	0.60 - 0.75	2000.0 m - 4000.0 m	0.93524	89 165
29	20.0 % - 50.0 %	0.60 - 0.75	4000.0 m - 6000.0 m	0.86634	944 784
30	50.0 % - 75.0 %	0.60 - 0.75	4000.0 m - 6000.0 m	0.95945	606 807
31	75.0 % - 100.0 %	0.60 - 0.75	4000.0 m - 6000.0 m	0.86989	2 210 720
32	20.0 % - 50.0 %	0.60 - 0.75	6000.0 m - 8000.0 m	0.93232	10 014
33	50.0 % - 75.0 %	0.60 - 0.75	6000.0 m - 8000.0 m	0.79030	2 670

(a) variation of engine fan speed and Mach number at $H = 3000\,\mathrm{m}$

(b) variation of engine fan speed and altitude at $Ma = 0.4$

Fig. 5   Local linear thrust model evaluations at example conditions; anti-ice off.

diction. The local linear model visualized in Fig. 6a well predicts the required thrust which is represented by the model surface being very close to the data. Furthermore, the required thrust shows a nearly linear behavior with variation of the regressors which means that the local linear thrust model representation is a good choice in this case. In contrast, Fig. 6b shows a very poor fit of required thrust with the local linear model. Therefore, the model cannot be used to reliably predict thrust for further analysis, because e.g., the model underestimates the required thrust for the given example conditions. It is clearly visible that the (example) data for model estimation is locally concentrated and widely spread, which makes it difficult to well estimate the local linear model parameters in this case.

All in all, for this work it was found that the estimation of a complete thrust table would be more beneficial than a complex local linear model determination for the whole thrust envelope and subsequent extensive model stitching.

6 Nonlinear Engine Thrust Table

For this work, the representation of the engine thrust by a multidimensional table is a better choice than using (local) linear models. By this, the nonlinear behavior of engine thrust with variation of engine states and external influences can be modeled. The choice of breakpoints, the spacing of the grid and the interpolation method between the breakpoints defines the quality of the resulting thrust prediction. A coarse grid and the application of a linear interpolation method will give less accurate results as a very fine grid and/or a higher order interpolation method. Furthermore, a non-uniform grid will also be able to cover strong local nonlinearities, but will pose some additional challenges to the table definition and the entire estimation process. For this work, it was assumed that a moderately fine grid allows covering the nonlinear thrust behavior with the used regressors and a linear interpolation method is sufficient to predict the thrust between the breakpoints of this fine grid. Also, it was kept in mind that the table estimation should be performed on a state-of-the-art desktop computer with large but limited memory: because of the number of free parameter during the estimation directly increases with the number of breakpoints, the table size is simply limited by the available memory. Nevertheless, the grid definition in Tab. 6 was initially found sufficient and finally delivered together with (simple) linear multi-dimensional interpolation a good representation of the required thrust behavior. Note that the breakpoints are chosen
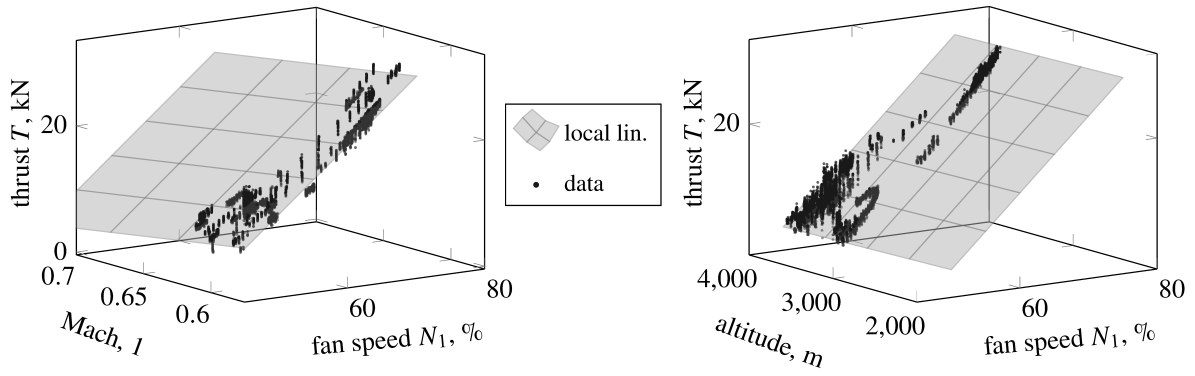
(a) local linear model thrust match: required thrust (data points from measurements) and model prediction (surface) with variation of different regressors; example conditions at $H = 4000\,\text{m}$ (left side) and $Ma = 0.625$ (right side); good model fit, $R^2 = 0.97192$, **Model 27**
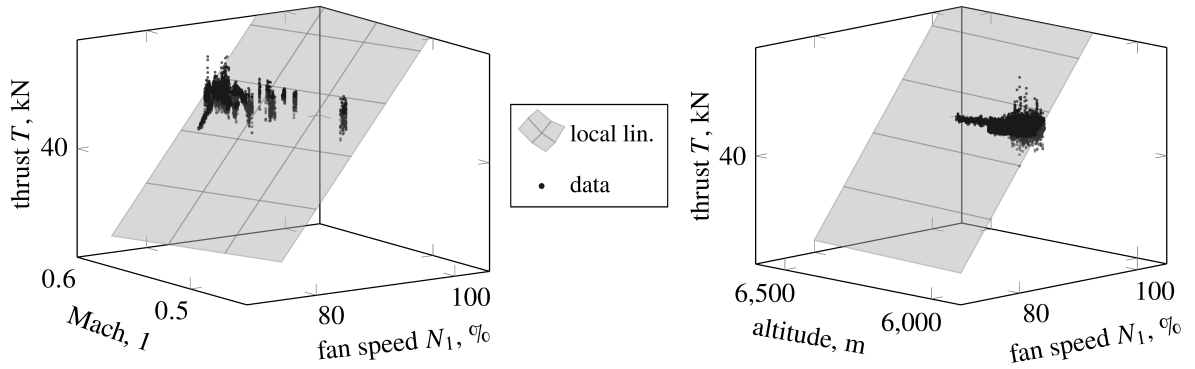
(b) local linear model thrust match: required thrust (data points from measurements) and model prediction (surface) with variation of different regressors; example conditions at $H = 6000\,\text{m}$ (left side) and $Ma = 0.575$ (right side); bad model fit, $R^2 = -0.8662$

Fig. 6 Examples of thrust predictions using local linear model compared to required thrust data: very good and bad model fit.

to cover the operational limits of the aircraft but also consider the limitations of the given data sets, i.e. the altitude limit at 6500 m.

To allow the estimation of the thrust table entries with e.g., the well-known Gauss-Newton method, the table is decomposed into the parameter vector θ. Hence, the number of parameters p respectively the length of the parameter vector is given by the product of lengths in each dimension

$$p = N_{\text{BP}_{N_1}} \cdot N_{\text{BP}_{Ma}} \cdot N_{\text{BP}_H}$$

The linear model (section 4) is used to obtain an initial guess for the table entries respectively θ. Using the breakpoints defined in Tab. 6 the linear model is evaluated for the whole resulting grid and the table is initialized.

Table 6 Nonlinear thrust model table breakpoint definition.

	lower limit	grid size	upper limit	N_{BP}
fan speed	15 %	5 %	100 %	18
Mach number	0.1	0.05	0.85	16
(barometric) altitude	0 m	500 m	6500 m	14

One major requirement for the engine thrust table is that it should have a smooth shape and no significant local changes of curvature. The basic influence in different regressors on engine thrust in Fig. 3 shows that there is no different, discontinuous behavior of thrust with fan speed, velocity or Mach number and altitude or pressure expectable. Furthermore, the fundamental engineering knowledge about jet engines and the way they produce thrust directly imposes the need to obtain "smooth" thrust curves: a jet engine within its normal operation limits will not produce an abrupt change of thrust due to a change of flight condition because the thermodynamic cycle within the engine is running continuously. As long as no major disturbance of this process occurs, like engine stall or any malfunction, there must be no significant change of thrust curvature with the used regressors. Note that this is the type of engineering knowledge which is implemented in the presented approach and differs from unstructured artificial intelligence techniques to solve the given big data problem. Without such assumptions, the resulting thrust model will not be able to reasonably and reliably predict the engine thrust. The optimization problem is now formulated as:

$$\widehat{\theta} = \underset{\theta}{\mathrm{argmin}} \left(\theta \to \sum_{i=1}^{n} (z_i - y_i(\theta))^2 + \|\Gamma_2^* \theta\|^2 \right). \tag{11}$$

In this case, the Tikhonov regularization matrix Γ_2^* [14, 15] is used to penalize the second-order derivatives of the multi-dimensional thrust table in each direction in order to smooth the table shape. It is defined to consider all break point combinations of the table except for the corners of the resulting break point cube. Its definition further considers the fact, that the faces of the cube are represented with the second-order derivative in only two dimensions and the edges in only one dimension. Fig. 7 shows all break points and illustrates the named cube, faces and edges, which must be treated individually to compose the matrix. The matrix size is defined by the number of parameters p (columns) and the number of break-points to be considered for calculation of the curvature (rows) given by

$$\begin{aligned}
Num\ rows(\Gamma_2^*) = \ & 3 \cdot \left(N_{\mathrm{BP}_{N_1}} - 2\right) \cdot \left(N_{\mathrm{BP}_{Ma}} - 2\right) \cdot \left(N_{\mathrm{BP}_H} - 2\right) \cdot \left(N_{\mathrm{BP}_{Ma}} - 2\right) \\
& + 4 \cdot \left(N_{\mathrm{BP}_{N_1}} - 2\right) \cdot \left(N_{\mathrm{BP}_{Ma}} - 2\right) + 4 \cdot \left(N_{\mathrm{BP}_{Ma}} - 2\right) \cdot \left(N_{\mathrm{BP}_H} - 2\right) \\
& + 4 \cdot \left(N_{\mathrm{BP}_{N_1}} - 2\right) \cdot \left(N_{\mathrm{BP}_H} - 2\right) \\
& + 4 \cdot \left(N_{\mathrm{BP}_{N_1}} + N_{\mathrm{BP}_{Ma}} + N_{\mathrm{BP}_H} - 6\right).
\end{aligned}$$

The matrix is based on the coefficients $[-1, +2, -1]$ of the well-known Mexican Hat wavelet filter or the Laplacian-of-Gaussian filter [16] of order two. The filter is scaled with the weighted Euclidean norm d_γ between two break-point combinations BP (including the Tikhonov weights γ in each direction):

$$d_\gamma(\mathrm{BP}_i, \mathrm{BP}_j) = \sqrt{\gamma_{N_1} \cdot \left(\mathrm{BP}_{N_1,i} - \mathrm{BP}_{N_1,j}\right)^2 + \gamma_{Ma} \cdot \left(\mathrm{BP}_{Ma,i} - \mathrm{BP}_{Ma,j}\right)^2 + \gamma_H \cdot \left(\mathrm{BP}_{H,i} - \mathrm{BP}_{H,j}\right)^2}. \tag{12}$$

Note that this is independent of the breakpoints' location, i.e. within the inner cube, on the faces or the edges.

The Tikhonov weights γ are used to define the strength of the penalization against the least squares within the optimization. In the simplest case of a one-dimensional model, the Tikhonov matrix Γ_2^* is given by the (scaled) filter coefficients on the matrix main and secondary diagonals, as given in [17] for a one-dimensional wind profile. But in case of this multi-dimensional thrust table the definition of the matrix entries is more complex. The scaled filter coefficients must be related to the distinct entry in the parameter vector representing the value for the BP_j as neighbor of BP_i. Therefore, it is necessary to apriori connect all breakpoints and calculate their distances to each other as required in order to set-up Γ_2^*.
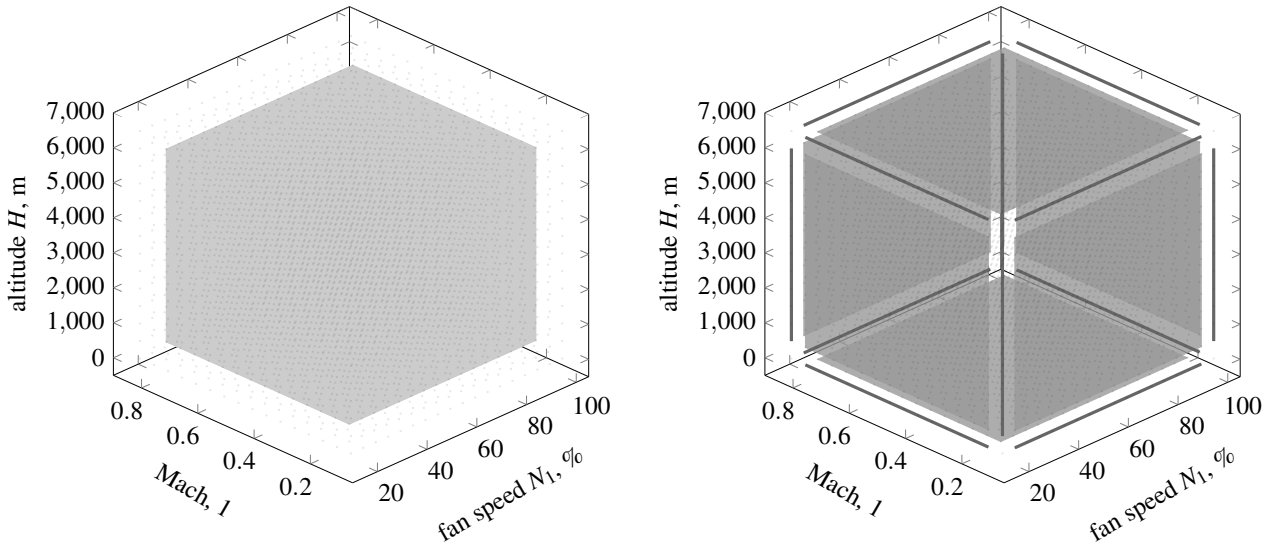
Fig. 7 Illustration of thrust table model breakpoints (gray dots) to be considered differently for Tikhonov matrix regularization during parameter estimation. Left side: inner cube to be considered with three-dimensional distances between the breakpoints for weighting. Right side: outside faces and edges which result in two- and one-dimensional distances for weighting.

The optimization problem in equation (11) can be solved by application of the iterative Gauss-Newton method. The resulting least squares problem can be written as

$$\widehat{\theta} = \underset{\theta}{\text{argmin}} \left(\|r(\theta)\|^2 + \|\Gamma_2^* \theta\|^2 \right), \tag{13}$$

where the vector of residuals is composed of the n residuals resulting from the n measurements (in this case the required thrust values) and the model prediction using the given parameter vector of the current iteration k:

$$r_i\left(\theta^k\right) = z_i - y_i\left(\theta^k\right), \quad \forall i \in [1,n]. \tag{14}$$

The Jacobian matrix entries are defined by the gradient of residuals with parameter variation

$$J_{i,j}\left(\theta^k\right) = \frac{\partial}{\partial \theta_j} r_i\left(\theta^k\right), \quad \forall i \in [1,n], \forall j \in [1,p]. \tag{15}$$

The regularization of the estimation problem requires to extend the residuals vector by additional rows

$$r^*\left(\theta^k\right) = \begin{bmatrix} r\left(\theta^k\right) \\ \Gamma_2^* \, \theta^k \end{bmatrix}, \tag{16}$$

so that r^* does now contain the penalization of too large parameter variations within the vector respectively thrust table. Similarly, the Jacobian matrix is extended by the Tikhonov matrix

$$J^*\left(\theta^k\right) = \begin{bmatrix} J\left(\theta^k\right) \\ \Gamma_2^* \end{bmatrix} \tag{17}$$

and the Gauss-Newton step to update the parameter vector (in each iteration) becomes

$$\theta^{k+1} = \theta^k + \left(\left[J^*\left(\theta^k\right)\right]^T J^*\left(\theta^k\right) \right)^{-1} \left[J^*\left(\theta^k\right)\right]^T r^*\left(\theta^k\right). \tag{18}$$

The iterative algorithm is terminated for this problem if the relative change of the cost function between two steps lies within a range of -10^{-4} and 0. The cost function is defined as the sum of residual squares

$$\mathbb{L}^k(\theta^k) = \sum_i r_i^2(\theta^k). \tag{19}$$

The thrust table estimation described above requires large matrices which have to be inverted. To significantly reduce the size of these matrices, the necessary memory and the corresponding computational effort during the Gauss-Newton step, the data is clustered before the estimation algorithm is applied. This further allows to use a state-of-the-art desktop computer for the whole evaluation process in Fig. 1. During the clustering, similar data is combined to one individual point with larger weight during the table estimation which reduces the number of rows of J^* significantly. For comparable regressor points the required thrust values are averaged and the number of averaged points defines the weight of the resulting residual value in r. Using the clustering grid as defined in Tab. 7 the number of data points can be reduced by a factor of 235 or up to 50.7 million rows of J^* (in case of anti-ice off). If the clustering grid is fine enough, it can be further shown that the table estimation results are (nearly) identical to the case in which the data is not clustered.

Table 7 Definition of clustering grid to reduce the number of data used during the engine table estimation.

	lower limit	grid size	upper limit
fan speed	0 %	0.25 %	105 %
Mach number	0	0.01	0.85
(barometric) altitude	0 m	50 m	7000 m

As an example, Fig. 8 shows some results of the engine thrust table evaluation similar to Figs. 5 and 4. The results reflect the (steady) nonlinear engine thrust behavior which could not be predicted by the linear and local linear models.
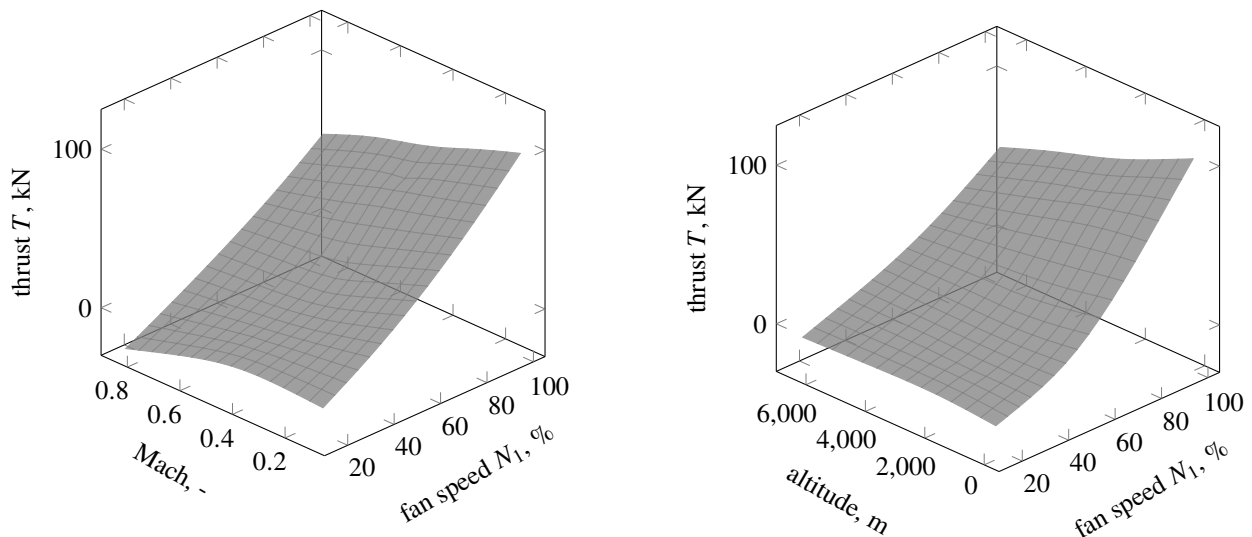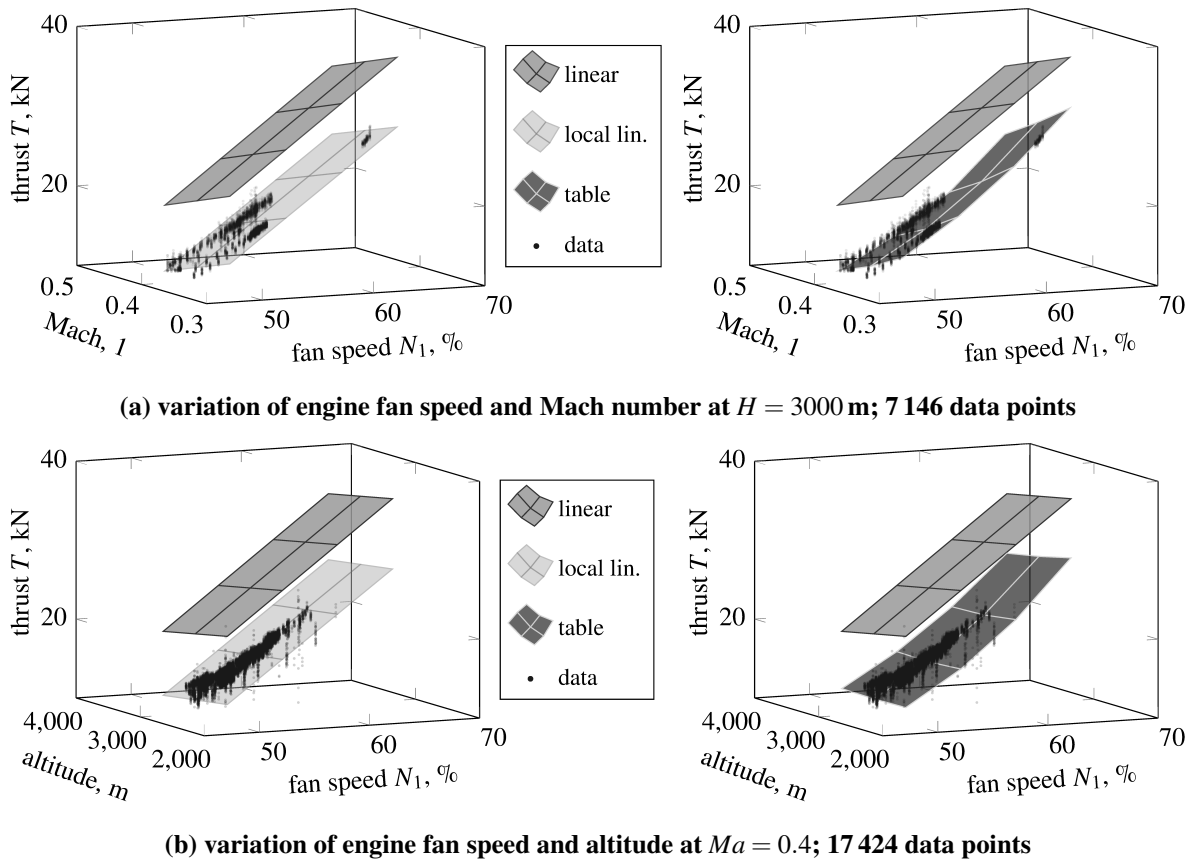

(a) variation of engine fan speed and Mach number at $H = 3000\,\mathrm{m}$

(b) variation of engine fan speed and altitude at $Ma = 0.4$

Fig. 8 Nonlinear thrust table evaluations at example conditions; anti-ice off.

(a) variation of engine fan speed and Mach number at $H = 3000\,\text{m}$; 7146 data points

(b) variation of engine fan speed and altitude at $Ma = 0.4$; 17424 data points

Fig. 9 **Example of different thrust model evaluations for selected space in model envelope and comparison with given required thrust values; anti-ice off.**

7 Comparison of Different Thrust Models

A selection of data around the example conditions ($H = 3000\,\text{m}$ and $Ma = 0.4$) was done to provide a comparison of the different models with the required thrust data. For the fixed Mach number condition a selection criterion of $Ma = 0.4 \pm 0.001$ was applied resulting in 17424 data points. The fixed altitude data was selected by $H = 3000 \pm 25\,\text{m}$ resulting in 7146 points. Fig. 9 shows a comparison of the different models with the required thrust data around the example conditions. In each plot, the linear model results are given together with the required thrust data and the local linear models (left plot) or the nonlinear table (right plot). It is clearly visible that the linear model does not allow to predict the required thrust data as the resulting surface lies far above the data because it fits the whole data across the envelope. Again, this points out that a global linear model for thrust gives a good initial guess for the engine thrust behavior but is not suitable for the given purpose. The other two modeling approaches do better represent the given required thrust, as expected. Fig. 9a shows the results of thrust variation with fan speed and Mach number at fixed altitude and Fig. 9b the variation with fan speed and altitude at fixed Mach number. In both cases, the table model is the (clearly visible) better choice to cover the nonlinear thrust behavior as it follows the curvature of the required thrust. Note that a higher number of local linear models or a change to a higher-order interpolation polynomial of the local models could be used to better approximate the nonlinear thrust behavior, but with the remaining problem of model stitching. The nonlinear table model is still the better choice in this case.

8 Conclusion

The paper presents the estimation of engine thrust models from operational flight data with limited apriori knowledge about the aircraft's flight performance characteristics. It is part of a novel evaluation

process for flight performance model determination which solves the posed big data problem with a smart data approach. The direct comparison of linear, local-linear and nonlinear modeling approaches revealed the necessity of the nonlinear table model to represent the required engine thrust. The specific data reduction and matrix regularization allow to reduce the overall computational effort and formulate a well-posed optimization problem. Therefore, the nonlinear thrust table estimation becomes reliable, efficient and relatively fast on a state-of-the-art desktop computer and does not require any specific hardware. Hence, this work is an important step for the overall flight performance model determination process and provides a significant contribution to solutions of big data problems in aviation.

As future work, the engine thrust prediction will be extended by a temperature correction also estimated from the given flight data to further account for engine thrust variations during flight. Afterwards, the engine model prediction will be implemented in the outlined development process for aircraft performance model determination.

Acknowledgments

References

[1] European Commission. Flightpath 2050 Europe's Vision for Aviation. Katalog KI-31-11-098-EN-C, European Union, Luxemburg, 2011. DOI: 10.2777/50266.

[2] Christoph Georg Kühne, Marco Scholz, and Fethi Abdelmoula. LNAS - a pilot assistance system for energy-optimal approaches using existing aircraft-infrastructure. In *Advanced Aircraft Efficiency in a Global Air Transport System Conference (AEGATS) 2018*, Toulouse, France, October 23rd - 25th 2018. Association Aéronautique et Astronautique de France (3AF).

[3] Fethi Abdelmoula and Marco Scholz. LNAS - a pilot assistance system for low-noise approaches with minimal fuel consumption. Belo Horizonte, Brazil, September 2018. 31st Congress of the International Council of the Aeronautical Sciences (ICAS).

[4] Georges Ghazi and Ruxandra Mihaela Botez. Identification and validation of an engine performance database model for the flight management system. *Journal of Aerospace Information Systems*, 16(8):307–326, August 2019. DOI: 10.2514/1.I010663.

[5] Benjamin M. Simmons. System identification for propellers at high incidence angles, aiaa 2021-1190. Virtual Event, January 11th-15th and 19th -21st 2021. AIAA SciTech Forum, American Institute of Aeronautics and Astronautics, Inc. (AIAA). DOI: 10.2514/6.2021-1190.

[6] Christoph Deiler. Evaluation of aircraft performance variation during daily flight operations. Friedrichshafen, Germany, September 2018. Deutscher Luft- und Raumfahrtkongress, Deutsche Gesellschaft fur Luft- und Raumfahrt (DGLR). DOI: 10.25967/480025.

[7] Christoph Deiler and Nicolas Fezans. Performance-based ice detection methodology. *Journal of Aircraft*, 57(2):209–223, March 2020. DOI: 10.2514/1.C034828.

[8] Karolina Krajěk, Dario Nikolić, and Anita Domitrović. Aircraft performance monitoring from flight data. *Technical Gazette*, 22(5):1337–1344, October 2015. DOI: 10.17559/TV-20131220145918.

[9] V. A. Deo, F. Silvestre, and M. Morales. Flight performance monitoring with optimal filtering applications. *The Aeronautical Journal*, 124(1272):170–188, 2020. DOI: 10.1017/aer.2019.142.

18

[10] Georges Ghazi, Benoit Gerardin, Magali Gelhaye, and Ruxandra Mihaela Botez. New adaptive algorithm development for monitoring aircraft performance and improving flight management system predictions. *Journal of Aerospace Information Systems*, 17(2):97–112, 2020. DOI: 10.2514/1.I010748.

[11] Irwin E. Traeger. *Aircraft Gas Turbine Engine Technology*. Glencoe/McGraw-Hill, Westerville, Ohio, USA, 3rd edition, 1996.

[12] Leland M. Nicolai and Grant E. Carichner. *Fundamentals of Aircraft and Airship Design*, volume 1 - Aircraft Design. American Institute of Aeronautics and Astronautics, Inc. (AIAA), Reston, Virginia, USA, 2010. DOI: 10.2514/4.867538.

[13] Eugene A. Morelli and David G. Ward. Automated simulation updates based on flight data, aiaa 2007-6714. Hilton Head, South Carolina, USA, August 20th-23rd 2007. AIAA Atmospheric Flight Mechanics Conference and Exhibit, American Institute of Aeronautics and Astronautics, Inc. (AIAA). DOI: 10.2514/6.2007-6714.

[14] Andrey Nikolayevich Tikhonov, F. John, and V.J. Arsenin. *Solutions of Ill-posed Problems*. Halsted Press Book. V.H. Winston & Sons, 1977.

[15] Andrey Nikolayevich Tikhonov, A. V. Goncharsky, V.V. Stepanov, and Anatoly.G. Yagola. *Numerical Methods for the Solution of Ill-Posed Problems*. Mathematics and Its Applications. Springer, Dordrecht, 1995. DOI: 10.1007/978-94-015-8480-7.

[16] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 207(1167):187–217, February 1980.

[17] Nicolas Fezans, Hans-Dieter Joos, and Christoph Deiler. Gust load alleviation for a long-range aircraft with and without anticipation. *CEAS Aeronautical Journal*, 10(4):1033–1057, January 2019. DOI: 10.1007/s13272-019-00362-9.