ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

PSI

Benchmarking Methods for Fuel Consumption Estimation of Commercial Aircraft

Semester Thesis

Noe Marty
20-929-600
MSc D-ITET

nmarty@student.ethz.ch

Paul Scherrer Institute | Laboratory for Energy Systems Analysis
ETH Zürich | Chair of Energy Systems Analysis

Supervisors:
Michael Weinold, Prof. Dr. Russell McKenna
D-ITET Supervisor: Prof. Dr. Florian Dörfler

January 17, 2026

Acknowledgements

I would like to thank my academic supervisor, Michael Weinold, for guidance and constructive feedback throughout this thesis. I am grateful to Dr. Peter Wild for providing access to data, technical support, and domain expertise relevant to airline operations.

Abstract

Accurate estimation of aircraft fuel burn is essential for aviation climate impact assessment, regulatory reporting, and sustainability benchmarking. Numerous fuel-burn estimation models exist, ranging from statistical and reduced-order approaches to high-resolution closed-form models. However, a systematic, side-by-side comparison of these model classes under consistent assumptions and identical input conditions has not yet been established in the literature.

This thesis presents a systematic benchmarking of widely used aircraft fuel-burn estimation models using operational flight plan (OFP) data from a major European airline as reference. The analysis compares statistical, reduced-order, range-equation, closed-form and machine learning models under a unified set of assumptions, including a no-wind condition and consistent mission definitions, across five representative medium- and long-haul routes.

The results reveal substantial differences in accuracy between models. Increased model complexity does not consistently yield better performance. Several simplified models achieve competitive accuracy with significantly lower input requirements. These findings highlight the trade-off between model fidelity and data availability and provide guidance for selecting appropriate fuel-burn estimation methods in research and applied sustainability contexts.

Contents

Chapter 1

Introduction

Global air transport plays a central role in the modern, globalized economy by enabling the movement of people and goods on an unprecedented scale. However, this growth has come with a substantial and steadily increasing environmental impact. Aviation has experienced sustained multi-decade growth in emissions, with CO_2 output rising at an average rate of approximately 2.2% per year between 1970 and 2012, and accelerating to around 5% per year between 2013 and 2018. In 2018, global aviation CO_2 emissions exceeded 1 Gt, representing roughly 2.4% of total anthropogenic CO_2 emissions [1, p. 4]. Remarkably, half of all aviation CO_2 since 1940 has been emitted in just the past 20 years [1, p. 4]. Reliable estimates of aircraft fuel consumption are therefore essential for understanding the sector's climate impact and for assessing potential decarbonization pathways. Fuel burn directly determines CO_2 emissions and serves as a key indicator for operational efficiency, economic performance, and technological improvements. Although aviation produces significant non-CO_2 climate effects, such as impacts from nitrogen oxides, water vapor emissions, aerosol particles, and aerosol-induced cloud modifications (including contrail cirrus), these effects do not scale linearly with fuel consumption. They are highly dependent on altitude, atmospheric conditions, location and time, making them considerably more complex to quantify in a consistent manner. In this work, we therefore restrict the scope to fuel burn and its associated CO_2 emissions because they are the most consistently measurable and widely reported components of aviation's climate impact.

Aviation fuel consumption can be estimated using a broad range of modeling approaches, each differing in complexity, data requirements, and intended application. Traditional analytical formulations, such as the Breguet range equation, offer simple, closed-form estimates but provide limited accuracy because they cannot capture the distinct fuel requirements of different flight phases. More sophisticated physics-based models, including those built on EUROCONTROL's Base of Aircraft Data (BADA), simulate detailed climb, cruise, and descent profiles using aircraft-specific performance tables. These models form the basis of high-fidelity tools such as the Aviation Environmental Design Tool (AEDT)[2] and the Advanced Emission Model (AEM)[3], which are widely used by regula-

tory bodies and researchers. However, their reliance on proprietary performance data and their substantial computational cost constrain their applicability for large-scale analyses.

To enable global or long-term assessments, several reduced-order and data-driven alternatives have been proposed. These include regression-based distance–fuel relationships, simplified mission models, and surrogate modeling approaches such as the Fuel Estimation in Air Transportation (FEAT) framework proposed by Seymour et al.[4], which leverages reduced-order flight dynamics to approximate fuel burn at a fraction of the computational cost of traditional performance simulations. The primary advantage of these approaches lies in their substantially reduced computational complexity, enabling large-scale analyses that would otherwise be infeasible with detailed trajectory-based models. However, differences in underlying assumptions, such as payload estimation, atmospheric representation, or trajectory abstraction, can lead to non-negligible discrepancies in predicted fuel consumption. As a result, it remains unclear which modeling approach provides the most reliable estimates under different conditions, highlighting the need for a systematic, side-by-side comparison to assess their relative performance, limitations, and domain of applicability.

In this work, we conduct a systematic benchmarking of several established fuel-burn estimation methods against airline operational reference data. All models are evaluated on an identical set of real-world example flights for which detailed operational flight plans (OFPs) are available, ensuring a fair and controlled comparison across approaches. To isolate structural differences between models, all evaluations are performed under a consistent no-wind assumption. Airline operational inputs from the OFPs are used consistently across models, but no model is recalibrated or adapted to better match the airline reference values. Rather than focusing on model-to-model comparisons alone, the analysis primarily assesses how different levels of model abstraction approximate real airline operational fuel-burn figures across short-, medium-, and long-haul missions. By quantifying systematic deviations and relating them to model structure, input requirements, and simplifying assumptions, the study highlights the trade-offs between accuracy, data availability, and computational complexity. This thesis does not propose a new fuel-burn model, but instead provides a transparent, operation-focused evaluation of existing approaches, contributing to a clearer understanding of their practical performance in real airline dispatch contexts.

Chapter 2

Theory

2.1 General Parameters that affect Fuel Burn

2.1.1 Aircraft Type

The aircraft type fundamentally determines baseline fuel efficiency through its aerodynamic design, mass properties, and structural constraints. Key characteristics such as wing aspect ratio, wetted area, and the resulting drag polar directly influence the achievable lift-to-drag ratio. In steady cruise, this aerodynamic efficiency governs fuel burn through the specific air range relationship, which links fuel efficiency to airspeed, lift-to-drag ratio, propulsion efficiency, and aircraft mass [5, Sec. 13.2, Eqs. (13.3)–(13.9)]. As a consequence, different aircraft types operate within distinct efficiency envelopes, constraining the extent to which fuel burn can be reduced through operational measures alone.

2.1.2 Engines

Engine characteristics strongly influence aircraft fuel burn through the thrust specific fuel consumption (TSFC), which is defined as the fuel flow rate divided by the net engine thrust, or equivalently using a weight-based formulation, as discussed in [5, Sec. 8.4, Eqs. (8.30a)–(8.30b)]. TSFC provides a compact measure of how efficiently a jet engine converts fuel into propulsive thrust and reflects both thermodynamic efficiency and propulsive efficiency. Modern high-bypass turbofan engines achieve low TSFC primarily by accelerating a large mass flow of air at relatively low exhaust velocity, thereby improving propulsive efficiency, while core cycle parameters such as overall pressure ratio and turbine entry temperature govern thermal efficiency within material and durability constraints [5, Sec. 8.3–8.4].

Beyond design characteristics, in-service engine condition has a direct impact on fuel consumption. Ageing mechanisms such as compressor fouling, turbine erosion, and seal wear progressively increase TSFC over time, leading to measurable increases in mission fuel burn, particularly during cruise. Depending

on operating conditions and maintenance practices, these effects can result in fuel-burn penalties of several percent over typical on-wing intervals [6].

Furthermore, a given aircraft type may be offered with multiple engine options from different manufacturers, each exhibiting distinct baseline efficiencies, deterioration behaviour, and maintenance profiles. Accounting for both engine selection and in-service degradation is therefore important when estimating fuel consumption, as neglecting these effects may introduce errors on the order of several percent.

2.1.3 Distance Flown

Fuel burn does not scale linearly with flight distance because a significant portion of energy is consumed during distance-independent phases, such as taxi, take-off, and climb. For short-haul flights, these fixed components dominate the total energy budget, resulting in high fuel consumption per kilometer. While efficiency typically improves as the cruise phase becomes more dominant over longer distances, this trend reverses for ultra-long-haul flights. For such flights, the fuel mass penalty becomes significant; the massive fuel load required at departure increases the aircraft's total mass, which in turn induces higher fuel burn during the initial hours of flight. This non-linear relationship is implicit in the Breguet range formulation, which shows that required fuel mass increases exponentially with distance. Consequently, the weight penalty of carrying additional fuel eventually outweighs the aerodynamic advantages of a long cruise, leading to diminishing efficiency at very long flight lengths [5]. Empirical analyses suggest that fuel efficiency peaks at an intermediate stage length, though the exact optimum distance varies significantly by aircraft type and operational assumptions. For a representative long-range aircraft, Egelhofer et al. identified an optimal stage length of approximately 4 300 km based on data from Airbus [7, p. 176].

In practice, realised fuel burn is further influenced by routing inefficiencies. Actual flight trajectories are, on average, approximately 5% longer than the great-circle distance between origin and destination airports, although this deviation varies by region and flight distance [8]. Such deviations increase effective stage length and therefore fuel consumption relative to idealised great-circle assumptions, particularly in congested airspace or regions with constrained routing.

2.1.4 Weight and Payload

Aircraft weight, comprising the airframe, payload, and fuel, serves as a fundamental driver of fuel consumption. During steady, level flight, the lift generated must equal the aircraft weight; consequently, a heavier mass necessitates a higher lift coefficient, which directly increases induced drag. Furthermore, as fuel is depleted throughout the mission, the total mass decreases continuously. This

creates an inherently coupled relationship where fuel burn and weight evolution are mutually dependent on the specific flight profile.

2.1.5 Center of Gravity

The longitudinal position of the center of gravity (CG) affects fuel burn by influencing aircraft trim and stability requirements. Deviations from the optimal CG position require additional control surface deflections, particularly of the horizontal stabilizer, which generate trim drag and increase overall aerodynamic drag [5, Sec. 7.5.3]. During flight, the CG typically shifts as fuel is burned and redistributed across tanks, an effect that is especially pronounced in wide-body aircraft. Maintaining the CG within an optimal range, typically toward the aft end of the certified envelope, can therefore yield measurable fuel savings over long missions [5, Sec. 19.5.1].

2.1.6 Altitude

Cruise fuel efficiency generally improves at higher altitudes due to reduced air density and associated decreases in aerodynamic drag. For a given aircraft weight, Mach number, and atmospheric condition, there exists an optimum altitude that maximises the specific air range (SAR), and thus minimises fuel burn per distance flown. As fuel is consumed and aircraft weight decreases during flight, this optimum altitude typically increases, motivating a gradual climb profile over the cruise segment [5].

In practice, jet transport aircraft approximate this idealised cruise-climb by executing discrete step climbs, as continuous climbs are constrained by air traffic control procedures and airspace structure. On short-haul flights, operational limitations may prevent the aircraft from reaching the altitude corresponding to maximum SAR before descent is initiated. Additional constraints such as weather avoidance, congestion, and operational policies further influence realised cruise altitudes and can lead to deviations from the fuel-optimal profile.

2.1.7 Wind

Atmospheric wind conditions influence fuel burn primarily through their effect on ground speed and, consequently, flight time and ground-referenced efficiency. For a given airspeed and fuel flow, tailwinds increase the specific ground range, reducing flight time and total fuel consumption, while headwinds reduce ground range and increase fuel burn for a given mission distance. Importantly, wind does not directly alter aerodynamic efficiency, but modifies the relationship between air distance and ground distance through its impact on ground speed.

Wind effects further interact with cruise altitude and speed selection. Because the altitude that maximizes still-air specific air range depends on aircraft weight and Mach number, vertical wind gradients can shift the altitude at which maximum ground-referenced efficiency is achieved. In practice, this leads to operational strategies such as step climbs and wind–altitude trade-offs, where aircraft may deliberately cruise above or below the still-air optimal altitude to exploit favourable wind conditions, provided that the resulting fuel savings outweigh the additional climb or drag penalties. Modern flight planning systems therefore explicitly account for three-dimensional wind fields when optimizing cruise profiles to minimize fuel consumption [5].

2.1.8 Cruise Speed

Cruise speed selection represents a trade-off between fuel efficiency and time-related operating costs. The speed that minimizes fuel burn per unit distance corresponds to the condition of maximum specific air range and is referred to as the maximum range cruise (MRC) speed. Operating at this speed minimizes fuel consumption in still-air conditions but results in longer flight times. In practice, airlines typically cruise at higher speeds to reduce time-dependent costs such as crew, maintenance, and aircraft utilization.

This trade-off is formalized through the cost index, which expresses the relative importance of time-related costs versus fuel costs [5, Sec. 18.3.2, Eq. (18.1)]. The resulting economy cruise (ECON) speed minimizes total operating cost rather than fuel burn alone and is therefore generally faster than the MRC speed. A commonly used approximation is the long range cruise (LRC) speed, which is defined as the speed yielding a small reduction (typically around 1%) from the maximum specific air range while providing a shorter flight time. Consequently, commercial flights rarely operate at minimum-fuel conditions, instead selecting cruise speeds that balance fuel efficiency and economic objectives [5, Sec. 18.3].

2.1.9 Operational and Allocation Effects

When fuel burn or emissions are expressed on a per-passenger basis, additional allocation effects become relevant. Per-passenger emissions depend strongly on the seat load factor, which influences both the total fuel consumption, through increased payload mass, and the number of passengers over which emissions are distributed. As a result, higher load factors can simultaneously increase absolute fuel burn while reducing emissions per passenger. Furthermore, the apportionment of fuel burn between passenger transport and belly cargo introduces additional uncertainty, particularly on routes with significant freight demand. These factors make per-passenger emissions inherently more sensitive to operational assumptions than aircraft-level fuel burn estimates.

2.2 Different Fuel Burn Models

The highest fidelity fuel burn estimation model would account for all parameters listed in section 2.1. In practice, however, many of these parameters are only partially known or entirely unavailable, and must therefore be approximated or neglected. Fuel-burn modeling thus inherently involves a trade-off between model complexity, data availability, and achievable accuracy, which strongly depends on the intended application context.

A wide range of fuel-burn estimation approaches has been proposed in the literature. These methods can be broadly grouped into four distinct categories, which differ substantially in their underlying assumptions, data requirements, computational complexity, and predictive fidelity.

Wherever possible, the models evaluated in this study are implemented using *JetFuelBurn*, an open-source Python package for aircraft fuel-burn estimation. The package is primarily designed for environmental impact assessments of air transport and provides a unified framework for atmospheric calculations, mission modeling, and fuel-burn allocation across different cabin classes. It implements a diverse set of fuel-burn models derived from established textbooks, regulatory methodologies, and peer-reviewed scientific literature, thereby enabling a consistent and transparent comparison across modeling approaches [10].

2.2.1 Statistical Models

Statistical fuel-burn models estimate aircraft fuel consumption based on aggregate, empirically derived relationships, such as fuel per distance flown or per seat-kilometre. These models do not simulate detailed flight dynamics or trajectories. Instead, they rely on large-scale datasets that collect fuel uplift and passenger transport statistics reported by airlines. Using these datasets, the models can produce statistically averaged fuel-burn values, making them well suited for high-level emissions inventories or comparisons across large numbers of flights when detailed performance or trajectory data are unavailable.

US DOT

The JetFuelBurn US DOT statistics model uses Form 41 Schedule T–100 data reported by large, certified U.S. air carriers[1]. This dataset includes total fuel consumption, passenger-distance, and weight-distance flown for each aircraft type. From the T–2 summary tables, the model derives average fuel-per-passenger-kilometre and fuel-per-weight-kilometre factors by dividing annual fuel use by the corresponding distance metrics. These coefficients are then combined with a

[1] https://www.transtats.bts.gov/Tables.asp?QO_VQ=EGD&QO

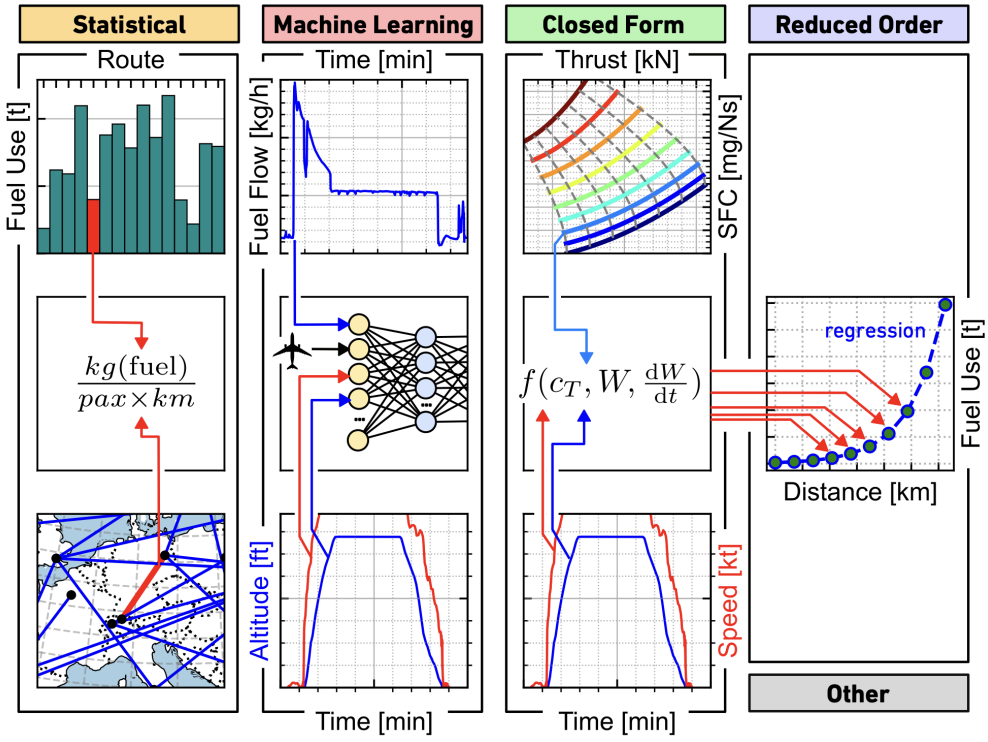

Figure 2.1: Conceptual overview of common fuel-burn modeling approaches used in aviation studies. Statistical models estimate average fuel use at the route or network level, typically normalizing total fuel consumption by passenger-kilometers. Machine learning models infer fuel flow directly from high-resolution flight data, such as time histories of altitude, speed, and aircraft state, requiring large volumes of operational data for training. Closed-form models compute fuel burn from first-principles relations based on the equations of motion and engine performance characteristics, linking thrust demand, aircraft weight, and fuel flow through aerodynamic and thermodynamic models. Reduced-order models approximate higher-fidelity models using regression techniques, enabling fast evaluation of fuel consumption as a function of distance or mission parameters. Hybrid approaches may combine elements from multiple model classes depending on data availability and application scope. Reproduced from Weinold et al. [9].

user-specified flight range (and payload weight, if applicable) to estimate representative fuel burn for that aircraft type in a given reporting year [10].

EU Flight Emissions Label

Created under Article 14 of Regulation (EU) 2023/2405 (ReFuelEU Aviation) [11], the EU Flight Emissions Label (FEL) is a forthcoming voluntary programme that will disclose the CO_2-equivalent emissions of each flight from take-off to landing for commercial flights departing from or arriving at EU airports. Participating operators must consistently apply the label across all eligible flights, following the methodology outlined in Implementing Regulation (EU) 2024/3170 [12]. This methodology primarily relies on verified operational data, such as reported fuel burn, passenger load, freight share, and the life-cycle carbon intensity of uplifted Jet A–1 and sustainable aviation fuel (SAF). Each label corresponds to a unique combination of operator, aircraft type and configuration, route, season, and year, and reports metrics such as CO_2e per passenger and per passenger-kilometre, as described by EASA [13].

Although conceptually related to statistical fuel-burn models, the EU FEL has not yet been launched publicly and is therefore not benchmarked or evaluated in this thesis.

ICAO Carbon Emissions Calculator

The ICAO Carbon Emissions Calculator is a statistical emissions estimation model developed by the International Civil Aviation Organization to provide standardized CO_2 emissions values per passenger for commercial flights. The methodology is based on aggregated airline operational data, including aircraft type, stage length, typical load factors, and cabin configuration, and applies average fuel burn values derived from global traffic statistics rather than flight-specific trajectories or aircraft performance models [14]. Emissions are reported directly as kg(CO_2) per passenger and are differentiated by cabin class using predefined allocation factors that account for seat density and the space occupied by each class.

2.2.2 Reduced-Order Models

Reduced-order models approximate aircraft fuel burn by representing the essential behaviour of high-fidelity performance simulations in simplified mathematical forms. They are typically constructed by running large sets of detailed, physics-based mission simulations, most commonly using the Base of Aircraft Data (BADA) performance framework from EUROCONTROL or the Piano-X aircraft performance model, and fitting surrogate relationships to the resulting

fuel-burn data. By design, reduced-order models retain the dominant physical dependencies of fuel consumption while remaining computationally efficient, making them well suited for large-scale analyses or applications where full mission simulations are impractical. Depending on their formulation, such models may rely on a single predictor, such as flight range, or incorporate multiple variables, including payload, cruise altitude, and the lengths of climb, cruise, and descent segments.

This thesis benchmarks and compares the reduced-order models implemented in the `JetFuelBurn` package as part of the overall evaluation framework.

AIM2015

The AIM2015 reduced-order fuel-burn model is derived from an extensive set of mission simulations generated with Piano-X[2]. Fuel consumption in the climb, cruise, and descent phases is represented through polynomial regressions where coefficients are fitted to these high-fidelity simulations. The resulting model expresses block fuel - defined as the total fuel consumed from engine start at the departure gate to engine shutdown at the arrival gate - as a polynomial function of mission distance and payload, scaled by an aircraft-size-specific inefficiency parameter. AIM2015 provides calibrated parameters for eight representative aircraft size classes, allowing the model to serve as a surrogate for a broad range of commercial aircraft types [15].

EEA Emission Inventory 2009

The EEA Emission Inventory 2009 reduced-order model estimates aircraft fuel burn using discrete fuel–range values provided in the 2009 EMEP/EEA Guidebook [16]. These tables contain fuel-burn simulations derived from Piano-X for 19 representative aircraft types at a set of fixed mission distances. For a given flight range, the model performs a linear interpolation between the nearest tabulated points to obtain total mission fuel consumption, with climb, cruise, and descent aggregated into a single flight-phase category. The approach does not model payload variation, detailed flight-profile effects, or mass changes and is valid only within the distance intervals covered by the tables. Because later EEA guidebook editions rely on EUROCONTROL's Advanced Emission Model (AEM), which is not freely redistributable, the 2009 version is the only one suitable for open reduced-order modeling [10].

[2] https://www.lissys.uk/PianoX.html

myclimate

The *myclimate* fuel-burn model[3] estimates aircraft emissions using a simple quadratic approximation of cruise fuel consumption combined with a fixed landing–take-off (LTO) component. For a given mission distance x, the model computes fuel burn as a polynomial regression

$$f(x) + \mathrm{LTO} = ax^2 + bx + c,$$

where the coefficients a, b, and c are fleet-average parameters published for a limited set of aircraft or size categories. Aircraft payload is assumed to take an average value, and climb and descent fuel consumption is implicitly accounted for in the LTO term, which itself is not made public [17].

It should be noted that the implementation of the *myclimate* reduced-order model in the `JetFuelBurn` package computes fuel burn only [10]. In contrast, the official *myclimate* "CO_2 Flight Calculator" reports total CO_2-equivalent emissions, including both direct and indirect effects, thereby accounting for non-CO_2 climate impacts. As a result, numerical outputs from the two approaches are not directly comparable.

Seymour et al.

The Seymour reduced-order model [4] combines detailed physics-based simulations with a simplified, distance-based representation suitable for large-scale air transportation modeling. High-fidelity fuel-burn data are first generated using EUROCONTROL's Advanced Emission Model (AEM), based on representative trajectories derived from BADA-based performance modeling for specific aircraft–engine combinations. These results are subsequently approximated by aircraft-specific quadratic regression functions of great-circle distance, yielding an explicit reduced-order relationship between mission fuel burn and stage length. The fitted regression coefficients capture both fixed fuel components and non-linear distance-dependent effects associated with fuel load. Fuel consumption for arbitrary flight distances is then obtained by direct evaluation of the regression model, enabling efficient estimation without explicit trajectory simulation.

Yanto et al.

The reduced-order fuel-burn model of Yanto and Liem [18] approximates aircraft mission fuel consumption using flight range and payload as predictors. To construct the model, a large fuel-burn database was generated using medium-fidelity mission simulations: climb and descent segments were evaluated using EUROCONTROL's BADA trajectory model, while cruise fuel burn was computed using

[3] https://co2.myclimate.org/en/flight_calculators/new

a Breguet-based formulation corrected with empirically derived factors. For each aircraft type, simulated missions with varying distances and payloads were used to fit a linear regression of the form

$$W_f = c_R R + c_P \mathrm{PL} + c_C,$$

where W_f denotes mission fuel burn, R the flight range, and PL the payload. The model is designed for large-scale policy or fleet analyses, where the priority is the computationally efficient evaluation of fuel consumption sensitivities to payload and range.

OpenAP

OpenAP (Open Aircraft Performance Model and Toolkit) is an open-source Python library developed at TU Delft for aircraft performance and emission analysis [19, 20]. The framework provides data-driven, polynomial fuel-flow and emission models derived from the ICAO Engine Emissions Databank[4] and the Acropole performance model.[5] Although not a reduced-order model in the classical regression sense, OpenAP employs surrogate relationships that approximate high-fidelity engine and performance behaviour and can therefore be interpreted as a data-driven reduced-order approach.

Fuel consumption is computed using the `FuelFlow` module, which estimates instantaneous fuel-flow rates for all flight phases as a function of aircraft mass, true airspeed, and altitude. Pollutant emissions are obtained via the `Emission` model, which converts fuel-flow rates into species-specific emission rates using engine-specific ICAO data. Given a complete flight trajectory represented as a time series, for example derived from ADS-B data or synthetically generated using the `FlightGenerator` module, OpenAP integrates these instantaneous rates over time to obtain total mission fuel burn and emissions.

In contrast to purely distance-based reduced-order models, OpenAP requires detailed trajectory information. This makes it well suited for applications where realistic flight paths are available and a balance between physical fidelity and computational efficiency is required.

Google Travel Impact Model

Google's Travel Impact Model (TIM) implements a reduced-order fuel-burn and emissions estimation approach based on the EEA Tier 3 methodology (Annex 1.A.3.a, 2023) [21]. The Tier 3 methodology is a flight- and aircraft-type-specific "bottom-up" approach that calculates emissions based flight-specific activity data. It

[4] https://www.easa.europa.eu/en/domains/environment/icao-aircraft-engine-emissions-databank
[5] https://github.com/DGAC/Acropole

segments flights into two distinct phases: Landing-Take-Off (LTO) for activities below 3000 ft and Climb-Cruise-Descent (CCD) for operations above that altitude. In TIM, fuel-burn and emissions factors for these phases are derived from detailed trajectory simulations using EUROCONTROL's Base of Aircraft Data (BADA) in combination with the Advanced Emissions Model (AEM). While these high-fidelity simulations provide the granular data characteristic of Tier 3, they are not exposed directly; instead, their results are provided as aircraft- and distance-specific lookup tables.

TIM estimates fuel consumption by combining tabulated landing–take-off (LTO) and cruise–climb–descent (CCD) fuel-burn values and applying linear interpolation or extrapolation with respect to great-circle distance to obtain fuel estimates for arbitrary stage lengths. Additional processing steps then convert fuel burn to CO_2-equivalent emissions and allocate these emissions on a per-passenger basis using assumptions on passenger load factors and cabin configurations.

At present, TIM is accessible through a free public API, but its use is limited to future scheduled commercial flights. While the model conceptually aligns with statistical and reduced-order fuel-burn approaches, its internal data sources and aggregation procedures remain partially opaque. According to Google, future versions of TIM are expected to integrate results from the EU Flight Emissions Label for flights where such data become available. Overall, TIM provides a scalable and operationally efficient framework for consumer-facing and large-scale emissions reporting, albeit with limited transparency regarding its underlying performance assumptions.

TIM outputs are provided exclusively as CO_2-equivalent emissions per passenger, which precludes a direct comparison of total mission fuel burn. Although the model calculates these values internally, the intermediate fuel-burn data are not publicly accessible. Furthermore, because TIM incorporates proprietary data regarding seat load factors and the allocation between passenger and cargo shares, the total mission fuel burn cannot be reliably back-calculated from the per-passenger emissions.

2.2.3 Machine Learning–Based Models

In recent years, there has been an increasing number of proposals for machine learning-based fuel estimation models. Their development, however, typically depends on access to high-fidelity operational datasets, such as Quick Access Recorder (QAR) or flight data recorder time series, which are difficult to obtain and often subject to strict confidentiality constraints. These approaches learn data-driven mappings from aircraft state variables, commonly including altitude, airspeed or ground speed, vertical speed, and in some cases aircraft mass, to fuel-flow rates or total fuel burn, using recorded fuel-flow measurements as supervision.

Early neural-network-based approaches include the work of Baklacioglu, who models aircraft fuel flow rates across individual flight phases using genetic-algorithm-optimised neural networks trained on high-resolution operational flight data. [22]

Beyond standard regression architectures, authors have proposed a broad range of modeling techniques, including classical statistical learning methods such as Gaussian Process Regression and tree-based models, as well as neural networks, including multilayer perceptrons, long short-term memory (LSTM) sequence models, and hybrid architectures. For example, Zhao et al. [23] propose a radial basis function neural network trained on operational flight parameters to predict total fuel consumption, demonstrating improved generalization compared to conventional feedforward models. Model performance is typically evaluated using phase-dependent error metrics across climb, cruise or level flight, and descent [24]. A key practical focus in more recent work is the transfer of such models to data sources with reduced feature availability, such as ADS-B–based inference, where parameters like aircraft mass and true airspeed may be missing and must either be approximated or treated as latent during training [25]. Physics-guided learning approaches have also been proposed to improve robustness with respect to parameter variations and out-of-distribution operating conditions [26].

While numerous machine learning–based fuel estimation models have been proposed, most rely on proprietary, high-resolution flight recorder data, which limits their applicability for independent benchmarking. This study therefore evaluates only the *Acropole* model, as it is explicitly designed to operate on ADS-B–derived parameters and publicly available aircraft data, without requiring confidential airline inputs.

Acropole

The Acropole model estimates engine fuel flow using supervised regression trained on Quick Access Recorder (QAR) data, with the objective of producing a generic model applicable to ADS-B trajectories [25]. During training, the model used a set of input features that includes kinematic variables (altitude, ground speed, true airspeed, vertical speed, and derivatives of ground speed and true airspeed), aircraft characteristics (maximum operating speed, maximum operating altitude, and engine type), and aircraft mass normalized by empty weight and maximum take-off weight.

The model is trained to predict single-engine fuel flow, expressed in normalized form using the take-off fuel flow from the ICAO Engine Emissions Databank as a reference. It is implemented as a feed-forward neural network with four fully connected layers and ReLU activation functions, followed by a sigmoid output layer. To enable inference on ADS-B data, where aircraft mass and true airspeed are frequently unavailable, the authors apply data augmentation during training by replacing unavailable parameters with predefined default values. In

particular, the mass input is set to an out-of-range value to force the model to learn representations that do not rely on mass information, while true airspeed is approximated by ground speed under a no-wind assumption.

Performance evaluations indicate that including aircraft mass substantially improves prediction accuracy, particularly during the climb phase, while the model remains functional, albeit with reduced accuracy, when mass information is omitted. This design enables the Acropole model to operate under the practical constraints of ADS-B–based fuel estimation while retaining reasonable predictive performance.

2.2.4 Closed-Form Models

Closed-form fuel-burn models estimate aircraft fuel consumption using analytical expressions derived from physical principles of aircraft performance and propulsion. In contrast to statistical or reduced-order models, closed-form approaches rely on explicit governing equations that relate fuel flow to aerodynamic efficiency, thrust requirements, flight conditions, and aircraft mass. These models do not require large simulation datasets or numerical optimisation and are therefore computationally efficient, transparent, and independently verifiable. Their applicability typically assumes that a representative flight trajectory or operating condition is known a priori, allowing fuel consumption to be evaluated deterministically along the mission profile.

Poll–Schumann Model

The Poll–Schumann model is a physics-based, closed-form approach for estimating the fuel consumption and related performance characteristics of commercial transport aircraft across all flight phases. It combines analytical aircraft performance relations with a parametric representation of overall engine efficiency, linking fuel flow explicitly to lift-to-drag ratio, thrust requirement, Mach number, altitude, and aircraft mass through transparent governing equations. Once identified, aircraft- and engine-specific characteristic parameters are applied consistently to cruise, climb, descent, and holding segments, enabling rapid evaluation of complete flight profiles when the trajectory is specified.

The model was developed and validated across a series of three complementary studies. The first part establishes the fundamental quantities and governing relations for aircraft performance in a general atmosphere [27]. The second part focuses on determining aircraft-specific characteristic parameters during cruise [28]. The third part extends the formulation to full-flight profiles, enabling fuel-burn estimation for complete missions with prescribed trajectories [29]. Together, these works demonstrate that the Poll–Schumann model retains sufficient physical fidelity to reproduce observed fuel-flow trends and flight data recorder measure-

ments across a wide range of aircraft types, while remaining open, computationally efficient, and suitable for independent verification.

An open-source implementation of the Poll–Schumann methodology is available through the `pycontrails` Python library, which provides a modular framework for modeling aviation fuel burn and climate impacts [30]. This implementation facilitates practical application of the model in large-scale studies and enables reproducible benchmarking against alternative fuel-burn estimation approaches.

Chapter 3

Results

3.1 Benchmarking of Models

The fuel-burn estimation models considered in this study were benchmarked using five representative commercial passenger flights. The selected aircraft types correspond to those operated by the airline that provided the reference data, ensuring consistency between the modeled scenarios and real-world operations. For each aircraft, a typical route commonly flown in commercial service was selected in order to cover a broad range of stage lengths, from short-haul (994NM) to long-haul (5601NM) missions. An overview of the selected aircraft, routes, engine configurations, and mission characteristics is provided in Table 3.1.

Table 3.1: Overview of evaluated example flights. The great-circle distance (GCD) deviation indicates the percentage increase of the actual mission distance relative to the great-circle distance, where 0% corresponds to a direct great-circle trajectory.

Aircraft type	Engine configuration	Route	Mission distance [NM]	GCD deviation [%]
Airbus A220-300	2× PW1524G-3	GVA–ARN	997	10
Airbus A320neo	2× PW1127G-JM	ATH–ZRH	994	12
Airbus A330-300	2× Trent 772B-60	JFK–ZRH	3521	3
Airbus A340-300	4× CFM56-5C4	ICN–ZRH	5601	18
Boeing 777-300ER	2× GE90-115BL	SFO–ZRH	5210	3

This selection enables the evaluation of model performance across different aircraft classes and mission profiles, while maintaining a realistic and operationally representative context.

3.2 Airline Reference Values

To enable a detailed and consistent comparison of the different fuel-burn estimation models, a suitable ground-truth reference was required. For this purpose, Operational Flight Plans (OFPs) provided by the dispatch department of a major European airline were used. An OFP is a comprehensive document prepared

prior to departure that supports flight crew fuel decision-making by specifying planned taxi, trip, reserve, and block fuel, along with the associated routing, payload assumptions, and aircraft performance parameters.

The provided OFPs contained all relevant technical information for each example flight, including aircraft type, mission distance, planned routing, and payload assumptions representative of typical operations. All OFPs and the corresponding reference fuel values were generated under a no-wind assumption. This simplification does not represent an inherent limitation of the evaluated models or of the comparison itself, but rather a deliberate methodological choice to ensure a controlled and consistent benchmarking framework. The same no-wind assumption was therefore applied across all fuel-burn estimation models considered in this study. The OFP paramateres were then used as inputs to the various fuel-burn models, for example in the specification of take-off weight, payload, and mission distance. This ensured that differences in estimated fuel burn could be attributed primarily to differences in model structure rather than to inconsistent input assumptions or atmospheric effects.

Payload assumptions used in the OFPs and in the model evaluation were aligned with realistic average load factors. Specifically, the average seat load factor reported by the data-providing airline for the year 2024 was used as the basis for passenger-related payload assumptions. This value is closely aligned with the industry-wide average seat load factor of 83.5% reported by IATA for the same year [31]. This comparison provides contextual validation that the adopted payload assumptions are representative of typical commercial operations.

3.3 Methods for Flight Trajectory Data Acquisition and Pre-processing

This section describes the data sources and processing steps used to construct flight trajectories for use as inputs to the fuel-burn estimation models. Two complementary approaches were investigated: empirical trajectories derived from Automatic Dependent Surveillance–Broadcast (ADS-B) data, and planned trajectories reconstructed from Operational Flight Plans (OFPs) provided by the reference airline. While ADS-B trajectories offer insight into real-world aircraft operations, the final evaluation in this thesis relies on OFP-based trajectories to ensure consistency with the reference fuel-burn values used for benchmarking.

3.3.1 ADS-B Trajectories

To derive empirical flight trajectories as potential inputs for the fuel-burn estimation models considered in this study, ADS-B–based trajectory data were initially collected for the five representative routes. The primary data source was

Flightradar24, which provided largely continuous coverage over the full flight duration across the considered routes. To ensure reproducibility, complementary trajectory data were obtained from the OpenSky Network following approval of academic access to its historical database.

A rigorous matching procedure was applied to ensure that all collected trajectories corresponded to the specific aircraft types used in the benchmarking analysis. For each city pair, five individual flight samples were collected. Data quality assessment revealed pronounced geographic disparities in coverage within the OpenSky dataset. In particular, significant gaps, extending over multiple hours, were observed on the ICN–ZRH route over Central Asia and on the SFO–ZRH route over the North Atlantic. These discontinuities are consistent with the sparse ground receiver density in these regions, as documented by the global OpenSky coverage map. While Flightradar24 provided more continuous trajectories in such cases, its data are not openly licensed for research use.

Trajectory data from the OpenSky Network was accessed via the Python `pyopensky` interface. Since historical OpenSky flight queries cannot be filtered directly by aircraft type, an intermediate identification step was required. First, the tail numbers of all aircraft belonging to the relevant fleets of the reference airline were collected. These tail numbers were then mapped to their corresponding ICAO24 transponder addresses using the OpenSky aircraft database (snapshot from August 2025). This mapping enabled indirect filtering by aircraft type through explicit ICAO24 selection.

Data acquisition was performed by iteratively querying the OpenSky Trino database for flights between 1 October 2025 and 30 October 2025, constrained by specific origin–destination airport pairs and matching callsigns. The resulting raw state vectors were subsequently processed using a dedicated data-cleaning pipeline. Records with missing or invalid geospatial and kinematic information—including latitude, longitude, barometric altitude, or ground speed—were removed. Altitude and velocity profiles were then smoothed to mitigate sensor noise and sporadic measurement artefacts. The resulting dataset consists of cleaned, temporally consistent ADS-B trajectories.

Despite this preprocessing, ADS-B trajectories represent realised operational flights that are influenced by air traffic control interventions, weather avoidance, and tactical speed or altitude changes. As a result, they do not necessarily correspond to the planned trajectories underlying the Operational Flight Plans (OFPs) used to generate the reference fuel-burn values. This mismatch ultimately limited the suitability of ADS-B trajectories for the final benchmarking analysis. Nevertheless, the ADS-B–based methodology is documented in detail to support reproducibility and future extensions of this work. [32, 33]

3.3.2 OFP ATC Flight Plan Trajectories

A key limitation of ADS-B–based trajectories is that real-world flown flights do not necessarily follow the exact routing, altitude profile, or timing assumptions contained in the Operational Flight Plans (OFPs) used by the airline to compute the reference fuel burn. Since the primary objective of this study is to compare fuel-burn estimation models against OFP-based ground truth values, a trajectory representation that is fully consistent with the planning assumptions of the OFP was required.

For this reason, the ATC flight plans contained in the provided OFPs were used to reconstruct representative reference trajectories for the final evaluation. The OFPs specify the lateral routing as a sequence of named waypoints, together with planned flight levels (FLs) at selected points along the route. The geographic coordinates of all waypoints were retrieved using SkyVector, a publicly accessible provider of worldwide aeronautical charts. The lateral trajectory was then constructed by connecting successive waypoints using great-circle segments, which is considered an adequate approximation given the level of detail typically employed in airline flight planning.

While the lateral path could be reconstructed directly from the OFP, the vertical profile was only partially defined. Flight levels were specified at discrete waypoints, whereas continuous altitude information during climb and descent phases was not explicitly available. To address this limitation, a vertical profile reconstruction method was implemented based on aircraft-type–specific performance data from the publicly available EUROCONTROL Aircraft Performance Database [34]. Using representative rates of climb and descent as functions of altitude, missing altitude values were reconstructed by forward integration over time. Hard altitude anchors were enforced at the departure and arrival airports using their respective field elevations, as well as at all waypoints with numeric flight level assignments. During climb and descent segments, the reconstructed altitude was additionally constrained not to exceed (or fall below) the next known flight-level anchor along the route.

The reconstruction relies on the planned time-to-waypoint information provided in the OFP. No explicit speed assumptions were introduced. Instead, the known waypoint times were used to define the temporal evolution of the trajectory, and intermediate positions and altitudes were interpolated linearly onto a uniform one-minute time grid. This approach preserves consistency with the planned mission profile while avoiding additional assumptions regarding airspeed schedules or wind conditions.

The resulting trajectories constitute smooth, continuous four-dimensional flight profiles that are fully consistent with the routing, timing, and planning assumptions of the OFP. Although the reconstructed climb and descent segments represent approximations of real aircraft behaviour, the cruise phase dominates

total mission fuel burn for the considered routes. These simplifications are therefore not expected to materially affect the comparative assessment of the fuel-burn models. Consequently, OFP-based trajectories were used exclusively in the final evaluation for all models requiring complete four-dimensional trajectory inputs, including the Poll–Schumann method, the OpenAP fuel model, and the Acropole machine-learning model.

3.4 Input Parameters per Model

This section summarizes the required input parameters for the fuel-burn estimation models considered in this study.

Table 3.2: Required input parameters for the different fuel-burn estimation models.

Model	Aircraft Type	Distance	Payload / TOW	Detailed Trajectory
US DOT Statistics	✓	✓	✓(Payload)	–
AIM2015	✓	✓	✓(Payload)	–
EEA Emission Inventory 2009	✓	✓	–	–
myclimate	✓	✓	–	–
Seymour et al.	✓	✓	–	–
Yanto et al.	✓	✓	✓(Payload)	–
Poll–Schumann	✓	–	✓(TOW)	✓
OpenAP	✓	–	✓(TOW)	✓
Acropole	✓	–	–(TOW optional)	✓
Google Travel Impact Model	✓	✓	–	–
ICAO Carbon Emissions Calculator	–	✓	–	–

Notes.

- **Aircraft Type** refers to an explicit aircraft or aircraft–engine combination.
- **Distance** denotes the use of mission or great-circle distance as a primary input.
- **Payload / TOW** indicates whether payload or take-off weight must be specified explicitly.
- **Detailed Trajectory** refers to time-resolved state information (e.g. ADS-B or synthetic trajectories).

3.5 Aircraft Type Coverage by Model

This section summarizes the aircraft type coverage of the fuel-burn estimation models that have previously been described in the theory section. Table 3.3 indicates whether a given aircraft type is supported by each model for a representative subset of narrow-body and wide-body aircraft.

Table 3.3: Aircraft type coverage by fuel-burn estimation model.

Model	A220-300	A320neo	A330-300	A340-300	B777-300ER
US DOT Statistics (2024)	✓	✓	✓	–	✓
AIM2015	–	–	✓	–	✓
EEA Emission Inventory 2009	–	–	✓*	– (A342)	✓*
Lee et al.	–	–	– (A332)	–	– (B772)
myclimate	–	–	✓*	–	✓*
Seymour et al.	✓	✓	✓	✓	✓
Yanto et al.	–	–	✓	✓	✓
Google Travel Impact Model	✓	✓	✓	✓	✓
Poll–Schumann (pycontrails)	✓	✓	✓	✓	✓
Acropole	–	–	✓	–	–

Notes.

- ✓ indicates that the aircraft type is supported by the model.
- – indicates that the aircraft type is not supported.
- Values in parentheses (e.g. A332, A342, B772) denote the closest available aircraft variant used as a proxy.
- *Exact subtype not specified in the underlying source data; model coverage is provided at aircraft family or type level.
- The table shows only the subset of aircraft types benchmarked in this thesis; most models support a broader set.

3.6 Results of Benchmarking

3.6.1 Total Fuel Burn Benchmarking

The uncertainty band shown around the airline reference values in Figure 3.1 is derived from the contingency fuel component included in the operational flight plans and is defined relative to the planned trip fuel. In airline dispatch practice, contingency fuel is typically specified as a fixed percentage of trip fuel (commonly 2–5%) or, alternatively, as an equivalent amount of flight time, and is added to absorb uncertainties arising from forecast wind errors, air traffic routing deviations, and other sources of operational variability. As such, it represents a realistic dispatch-level tolerance rather than a statistical confidence interval. In operational use, this fuel margin is applied only on top of the planned fuel burn and therefore primarily acts as insurance against higher-than-expected fuel consumption. For the purpose of visual comparison across models, the resulting uncertainty band is displayed symmetrically around the reference value, acknowledging that it reflects planning uncertainty rather than a probabilistic error distribution.

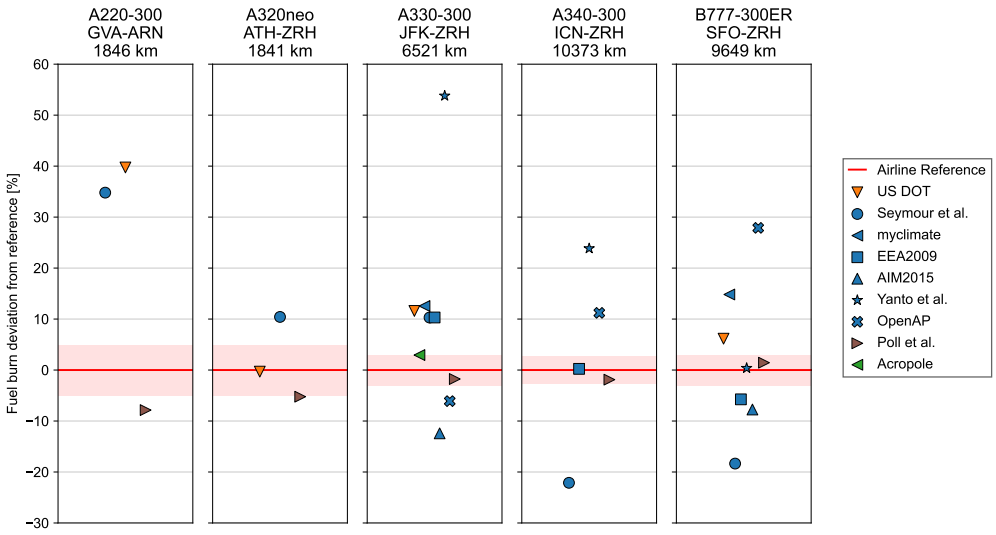

Figure 3.1: Relative deviation of total mission fuel burn from airline operational reference values for five representative commercial flights, evaluated across a range of fuel-burn estimation models. The reference flights correspond to an Airbus A220-300 (GVA–ARN, 1846 km), Airbus A320neo (ATH–ZRH, 1841 km), Airbus A330-300 (JFK–ZRH, 6521 km), Airbus A340-300 (ICN–ZRH, 10 373 km), and Boeing 777-300ER (SFO–ZRH, 9649 km), covering short-, medium-, and long-haul operations. Fuel-burn deviations are reported as percentages relative to airline dispatch flight plan values (red horizontal line). The shaded band around the reference indicates an uncertainty margin approximated from the contingency fuel included in the operational flight plans, reflecting typical dispatch-level allowances for operational variability and planning uncertainty. Model estimates are shown for a selection of statistical, reduced-order, closed-form, and machine-learning-based approaches, including Seymour et al., US DOT, Poll–Schumann, Acropole, myclimate, EEA 2009, AIM2015, Yanto et al., and OpenAP. The figure highlights both systematic biases and inter-model variability across aircraft classes and stage lengths, illustrating how differences in model structure, calibration data, and underlying assumptions influence total fuel-burn predictions.

Table 3.4: Mean absolute error (MAE) and mean absolute percentage error (MAPE) of fuel-burn estimates across all evaluated aircraft.

Model	N	**MAE [kg fuel]**	**MAPE [%]**
Acropole	1	1 189.00	2.94
Poll–Schumann	5	788.80	3.63
EEA 2009	3	3 018.33	5.43
AIM2015	2	5 679.50	10.08
myclimate	2	8 607.50	13.69
US DOT	4	2 877.75	14.45
OpenAP	3	11 313.33	15.07
Seymour et al.	5	7 656.80	19.20
Yanto et al.	3	13 476.33	26.01

Table 3.4 summarizes the aggregate fuel-burn estimation errors for each model, computed across all aircraft for which corresponding airline reference values were available. The number of evaluated cases N varies across models because aircraft type support differs between fuel-burn estimation approaches. As summarised in Table 3.3, not all models are applicable to all aircraft types considered in the benchmarking analysis. Consequently, error metrics for each model are computed only over the subset of aircraft for which the respective model provides compatible estimates. In particular, models with limited aircraft coverage, such as Acropole, are evaluated on a smaller number of representative cases, whereas more general models support a broader range of aircraft types and therefore contribute a larger number of samples to the aggregated error statistics.

Overall, Table 3.4 indicates a clear ranking in aggregate fuel-burn accuracy, which should be interpreted jointly with aircraft coverage (N) and model input requirements. Among the models evaluated on multiple aircraft, Poll–Schumann achieves the lowest MAE and a low MAPE, suggesting that trajectory- and parameter-rich methods can match airline reference values well when sufficient operational detail is available. EEA 2009 also performs comparatively well, but with noticeably higher MAE than Poll–Schumann, consistent with a reduced-order approach that relies on tabulated or simplified aircraft-performance information rather than detailed flight-specific trajectories.

When considering simplicity and data availability, models such as EEA 2009 and US DOT are attractive because they can be applied with substantially fewer and more readily obtainable inputs, but this comes at the cost of reduced accuracy and/or increased variability, as reflected in their higher MAPE values. By contrast, models with higher parameterization or stronger physical detail (e.g., OpenAP in the present setup) do not necessarily improve agreement with airline reference data, highlighting that additional complexity does not guarantee better performance under the assumptions and inputs used in this benchmarking.

The very low error observed for Acropole must be interpreted with care, as it is based on a single evaluated example flight ($N = 1$). While the results indicate strong agreement for the specific aircraft and operational context considered, they do not yet allow conclusions about robustness or generalizability across different aircraft types and mission profiles. Nevertheless, the close alignment observed for the A330-300 suggests that machine-learning-based approaches such as Acropole are promising and merit further validation against a broader set of flights and models.

3.6.2 CO_2 Emissions Benchmarking

In addition to benchmarking total mission fuel burn, the selected models are further evaluated in terms of CO_2 emissions per economy-class passenger. This metric is particularly relevant in a consumer-facing and regulatory context, as per-passenger CO_2 values are commonly reported by airlines, booking platforms, and environmental disclosure tools. To enable a consistent comparison, total mission fuel burn is first converted to CO_2 emissions using the standard conversion factor of 3.16 kg CO_2 per kg of jet fuel, and subsequently allocated at the individual passenger level following the recommended IATA and ICAO cabin-class allocation methodology, in which fuel burn is distributed across cabin classes based on relative seat area. This allocation approach, including aircraft-specific cabin layouts and seating configurations, is implemented consistently using the JetFuelBurn library [10].

A direct comparison of Figures 3.1 and 3.2 shows that, for models based on identical underlying fuel-burn estimates, the relative deviations with respect to the airline reference remain effectively unchanged when transitioning from total mission fuel burn to per-passenger CO_2 emissions. This is expected, as the CO_2 figures are derived deterministically from the same fuel-burn values using a fixed conversion factor and a consistent cabin-class allocation scheme. Consequently, the per-passenger representation does not introduce additional model-dependent variability, but rather preserves the relative ranking and systematic biases observed at the mission level.

The main differences observed in the per-passenger CO_2 comparison arise from the inclusion of consumer-facing emissions tools, such as the ICAO Carbon Emission Calculator and the Travel Impact Model (TTW) integrated into Google Flights, which rely on proprietary assumptions regarding load factors, seating configurations, and operational parameters. These tools therefore introduce additional layers of abstraction beyond fuel-burn estimation alone, leading to deviations that are not directly attributable to fuel-burn modeling accuracy. This highlights the importance of distinguishing between fuel-burn estimation models and end-user emissions calculators when interpreting per-passenger CO_2 values in both scientific and consumer-facing contexts.

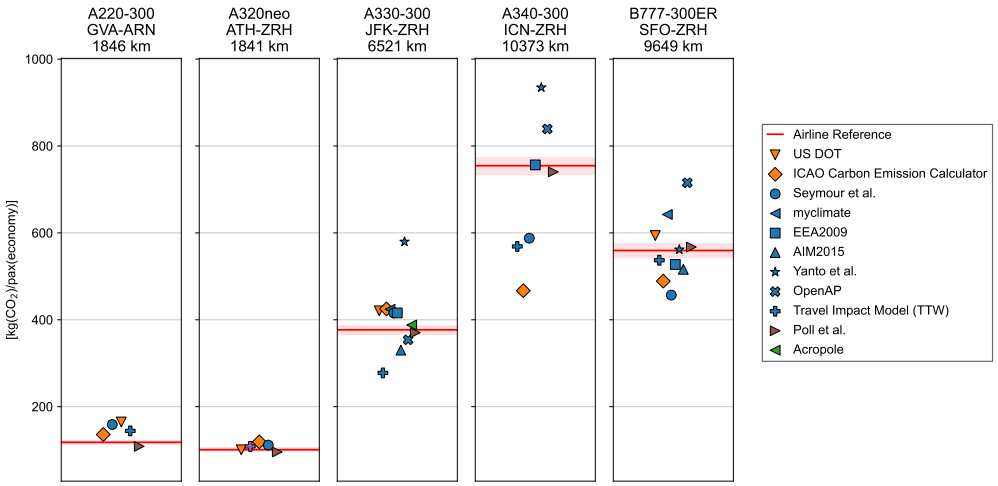

Figure 3.2: Comparison of CO_2 emissions per economy-class passenger for the same five representative commercial flights shown in Figure 3.1, estimated using different fuel-burn and emissions models and benchmarked against airline reference data. The selected routes include an Airbus A220-300 (GVA–ARN, 1846 km), Airbus A320neo (ATH–ZRH, 1841 km), Airbus A330-300 (JFK–ZRH, 6521 km), Airbus A340-300 (ICN–ZRH, 10 373 km), and Boeing 777-300ER (SFO–ZRH, 9649 km), spanning short-, medium-, and long-haul operations. Emissions are reported in kg CO_2 per economy-class passenger and are derived from total mission fuel burn using a standard CO_2 conversion factor of 3.16 kg CO_2 per kg of jet fuel. The resulting emissions are allocated to individual passengers using the recommended IATA and ICAO cabin-class allocation methodology, in which fuel burn is distributed across cabin classes based on relative seat area. The allocation accounts for the actual cabin layouts and seating configurations of the reference airline's fleet. Model estimates are shown for a range of reduced-order, closed-form, statistical, and machine-learning-based approaches, including ICAO Carbon Emissions Calculator, US DOT, EEA 2009, AIM2015, Seymour et al., Yanto et al., OpenAP, Poll–Schumann, Acropole, and the Travel Impact Model integrated into Google Flights. For consistency with the other models evaluated, the Travel Impact Model results shown correspond to *tank-to-wake* (TTW) CO_2 emissions only, representing emissions produced by the combustion of jet fuel during takeoff, flight, and landing. While the Travel Impact Model also reports *well-to-wake* (WTW) emissions by default in consumer-facing interfaces, defined as the sum of well-to-tank (WTT) and TTW emissions, the WTT component—covering emissions from fuel production, processing, handling, and distribution—is excluded here to ensure methodological consistency across models. The figure illustrates both inter-model variability and systematic differences across stage lengths, highlighting the influence of model structure, allocation assumptions, and underlying data sources on per-passenger CO_2 estimates.

3.7 Limitations and Scope

The interpretation of the results should be considered in light of the following limitations and scope choices:

- The analysis is based on operational data from a single airline, which may limit generalizability across different fleets, operating practices, and network structures.
- All reference fuel values and model evaluations were conducted under a no-wind assumption, thereby excluding the influence of atmospheric variability on fuel burn.
- The benchmarking dataset comprises five representative flights, selected to span a range of aircraft types and stage lengths, but not to provide exhaustive statistical coverage.
- Engine age, airframe deterioration, and maintenance state were not explicitly modeled and may contribute to residual deviations between models and reference values.
- Aircraft center-of-gravity effects were not considered, as this information is not explicitly represented in any of the evaluated models.

Despite these limitations, the results provide a consistent and transparent comparison of commonly used fuel-burn estimation approaches under controlled and identical input assumptions. To the author's knowledge, this study represents the first systematic, side-by-side benchmarking of statistical, reduced-order, and high-resolution fuel-burn models against airline-provided reference data. As such, the findings offer practical guidance on the relative strengths, limitations, and appropriate application domains of these models.

Bibliography

[1] D. S. Lee, D. W. Fahey, A. Skowron, M. R. Allen, U. Burkhardt, Q. Chen, S. J. Doherty *et al.*, "The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018," *Atmospheric Environment*, vol. 244, p. 117834, 2021.

[2] Federal Aviation Administration, "Aviation environmental design tool (aedt)," https://aedt.faa.gov/, accessed: December 2025.

[3] EUROCONTROL, "Advanced emission model (aem)," https://www.eurocontrol.int/model/advanced-emission-model, accessed: December 2025.

[4] K. Seymour, M. Held, G. Georges, and K. Boulouchos, "Fuel estimation in air transportation: Modeling global fuel consumption for commercial aviation," *Transportation Research Part D: Transport and Environment*, vol. 88, p. 102528, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S136192092030715X

[5] T. Young, *Performance of the Jet Transport Airplane: Analysis Methods, Flight Operations, and Regulations*, 10 2017.

[6] M. d. J. Gurrola Arrieta, R. M. Botez, and A. Lasne, "An engine deterioration model for predicting fuel consumption impact in a regional aircraft," *Aerospace*, vol. 11, no. 6, p. 426, 2024. [Online]. Available: https://www.mdpi.com/2226-4310/11/6/426

[7] R. Egelhofer, C. Marizy, and C. Bickerstaff, "On how to consider climate change in aircraft design," *Meteorologische Zeitschrift*, vol. 17, no. 2, pp. 173–179, May 2008.

[8] Author(s), "Global inefficiencies in aircraft routing and their impact on fuel burn," *EGUsphere*, 2023, preprint. [Online]. Available: https://egusphere.copernicus.org/preprints/2023/egusphere-2023-724/egusphere-2023-724.pdf

[9] M. Weinold, N. Marty, P. Wild, and R. McKenna, "Estimating fuel consumption of commercial aircraft," Conference poster, 2025, published conference poster. Licensed under Creative Commons Attribution–NoDerivatives 4.0 International (CC BY-ND 4.0). [Online]. Available: https://doi.org/10.3929/ethz-b-000749349

[10] M. P. Weinold and R. McKenna, "Jetfuelburn: A python package for calculating fuel burn of commercial aircraft," *Journal of Open Source Software*, vol. 11, no. 117, p. 9280, 2026.

[11] "Regulation (EU) 2023/2405 of the European Parliament and of the Council of 18 October 2023 on ensuring a level playing field for sustainable air transport (ReFuelEU Aviation)," 2023. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2023/2405/oj/eng

[12] "Commission Implementing Regulation (EU) 2024/3170 of 12 November 2024 laying down detailed rules for the calculation of flight emissions under the Flight Emissions Label," 2024. [Online]. Available: https://eur-lex.europa.eu/eli/reg_impl/2024/3170/oj/eng

[13] European Union Aviation Safety Agency (EASA). (2025) EU Flight Emissions Label (FEL). [Online]. Available: https://www.flightemissions.eu/en

[14] International Civil Aviation Organization, "Methodology for the icao carbon emissions calculator," International Civil Aviation Organization (ICAO), Tech. Rep. Version 13.1, Aug. 2024, accessed: 2025-12-20. [Online]. Available: https://icec.icao.int/Documents/Methodology%20ICAO%20Carbon%20Emissions%20Calculator_v13_Final.pdf

[15] L. M. Dray, P. Krammer, K. Doyme, B. Wang, K. Al Zayat, A. O'Sullivan, and A. W. Schäfer, "Aim2015: Validation and initial results from an open-source aviation systems model," *Transport Policy*, vol. 79, pp. 93–102, 2019.

[16] *EMEP/EEA Air Pollutant Emission Inventory Guidebook 2009: Part B, Section 1 (Energy), Subsection 1.A.3.a Aviation*, https://www.eea.europa.eu/en/analysis/publications/emep-eea-emission-inventory-guidebook-2009, European Environment Agency (EEA), 2009. [Online]. Available: https://www.eea.europa.eu/en/analysis/publications/emep-eea-emission-inventory-guidebook-2009

[17] myclimate. Flight emissions calculator – calculation principles. myclimate. [Online]. Available: https://www.myclimate.org/en/information/about-myclimate/downloads/flight-emission-calculator/

[18] J. Yanto and R. P. Liem, "Efficient fast approximation for aircraft fuel consumption for decision-making and policy analysis," in *Proceedings of the AIAA Modeling and Simulation Technologies Conference*. American Institute of Aeronautics and Astronautics, 2017, aIAA Paper 2017-3338. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2017-3338

[19] J. Sun and J. Hoekstra, "Openap: An open-source aircraft performance model for air transportation studies," *Aerospace*, vol. 7, no. 8, p. 104, 2020.

[20] "Openap: Open aircraft performance model and toolkit." [Online]. Available: https://openap.dev

[21] Google, "About the travel impact model," 2023. [Online]. Available: https://travelimpactmodel.org/about-tim

[22] T. Baklacioglu, "Modeling the fuel flow-rate of transport aircraft during flight phases using genetic algorithm-optimized neural networks," *Aerospace Science and Technology*, vol. 49, pp. 52–62, 2016.

[23] Y. Zhao, Z. Wang, X. Wang *et al.*, "Data-driven fuel consumption prediction model for green aviation using radial basis function neural network," *Scientific Reports*, vol. 15, p. 26275, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-11941-8

[24] M. Baumann, U. Klingauf, and M. Strohmeier, "Modeling aircraft fuel consumption using machine learning methods," *Aerospace*, vol. 7, no. 12, p. 174, 2020.

[25] G. Jarry, D. Delahaye, and C. Hurter, "Towards aircraft generic quick access recorder fuel flow regression models for ads-b data," in *International Conference on Research in Air Transportation*, 2024.

[26] M. Uzun, M. Demirezen, and G. Inalhan, "Physics guided deep learning for data-driven aircraft fuel consumption modeling," *Aerospace*, vol. 8, no. 2, p. 44, 2021.

[27] I. D. A. Poll and U. Schumann, "An estimation method for the fuel burn and other performance characteristics of civil transport aircraft in the cruise. part 1: Fundamental quantities and governing relations for a general atmosphere," *The Aeronautical Journal*, vol. 125, pp. 1–39, 2020.

[28] I. D. A. Poll and U. Schumann, "An estimation method for the fuel burn and other performance characteristics of civil transport aircraft during cruise. part 2: Determining the aircraft's characteristic parameters," *The Aeronautical Journal*, vol. 125, pp. 1–45, 2020.

[29] I. D. A. Poll and U. Schumann, "An estimation method for the fuel burn and other performance characteristics of civil transport aircraft. part 3: Full flight profile when the trajectory is specified," *The Aeronautical Journal*, vol. 129, pp. 1–37, 2025.

[30] M. Shapiro, Z. Engberg, R. Teoh, M. Stettler, T. Dean, and T. Abbott, "pycontrails: Python library for modeling aviation climate impacts," Dec. 2025. [Online]. Available: https://github.com/contrailcirrus/pycontrails

[31] International Air Transport Association, "Industry statistics: Fact sheet," Dec. 2025, public fact sheet. Updated December 2025; next update

June 2026. [Online]. Available: https://www.iata.org/en/iata-repository/pressroom/fact-sheets/industry-statistics/

[32] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. Wilhelm, "Bringing up opensky: A large-scale ads-b sensor network for research," in *Proceedings of the 13th IEEE/ACM International Symposium on Information Processing in Sensor Networks*, 2014.

[33] OpenSky Network, "pyopensky: Python interface for the opensky network." [Online]. Available: https://github.com/open-aviation/pyopensky

[34] EUROCONTROL, "Aircraft Performance Database," https://learningzone.eurocontrol.int/ilp/customs/ATCPFDB/default.aspx, accessed via the EUROCONTROL Learning Zone.