

Aircraft Fuel Flow Modelling with Ageing Effects: From Parametric Corrections to Neural Networks

Gabriel Jarry*, Ramon Dalmau*, Philippe Very*, Junzi Sun†

*EUROCONTROL
Aviation Sustainability Unit (ASU)
Brétigny-sur-Orge, France

†Faculty of Aerospace Engineering
Delft University of Technology
Delft, the Netherlands

Abstract—Accurate modelling of aircraft fuel-flow is crucial for both operational planning and environmental impact assessment, yet standard parametric models often neglect performance deterioration that occurs as aircraft age. This paper investigates multiple approaches to integrate engine ageing effects into fuel-flow prediction for the Airbus A320-214, using a comprehensive dataset of approximately nineteen thousand Quick Access Recorder flights from nine distinct airframes with varying years in service. We systematically evaluate classical physics-based models, empirical correction coefficients, and data-driven neural network architectures that incorporate age either as an input feature or as an explicit multiplicative bias. Results demonstrate that while baseline models consistently underestimate fuel consumption for older aircraft, the use of age-dependent correction factors and neural models substantially reduces bias and improves prediction accuracy. Nevertheless, limitations arise from the small number of airframes and the lack of detailed maintenance event records, which constrain the representativeness and generalization of age-based corrections. This study emphasizes the importance of accounting for the effects of ageing in parametric and machine learning frameworks to improve the reliability of operational and environmental assessments. The study also highlights the need for more diverse datasets that can capture the complexity of real-world engine deterioration.

Keywords—Aircraft fuel flow, engine deterioration, aircraft ageing, Quick Access Recorder, neural networks, parametric models.

I. INTRODUCTION

Aviation is a significant contributor to global greenhouse gas emissions, accounting for roughly 2–3% of worldwide CO₂ output [1]. This impact is expected to grow as air traffic increases, spurring a strong motivation to improve aircraft fuel efficiency and reduce emissions. Accurate fuel flow prediction is essential not only for airline cost and fuel planning but also for evaluating new operational procedures and technologies aimed at reducing fuel burn. However, achieving high accuracy in fuel flow estimation is challenging due to the complex interplay of aircraft performance, operations, and environmental factors.

The aviation community has developed standardized performance models to estimate fuel consumption. A prominent example is EUROCONTROL’s Base of Aircraft Data (BADA) model [2], which provides generic fuel flow and performance profiles for various aircraft types. Such classical models are parametric, built on physics-based principles (e.g., energy balance and specific fuel consumption curves) and calibrated to manufacturer data. They enable quick fuel burn estimates for trajectory simulations and fleet-level analyses. Nonetheless,

these models have known limitations: they assume “nominal” aircraft performance and do not explicitly account for aircraft-to-aircraft variability or changes over an aircraft’s service life. In practice, factors like engine degradation or maintenance actions can cause an aircraft’s fuel efficiency to deteriorate as it ages. For instance, studies have indicated that an aged engine can require several percent more fuel than a fresh engine for the same flight mission [3]. Standard models like BADA, which lack a representation of ageing effects, may therefore systematically underestimate fuel burn for older airframes. This gap is increasingly important as operators consider longevity and retrofit programs: without incorporating ageing, long-term emission forecasts and fuel planning for older fleets may be optimistic.

Integrating aircraft ageing or deterioration effects into fuel flow models is crucial for improving prediction fidelity. If we can quantify how fuel flow trends change over an aircraft’s lifecycle, we can correct baseline models to better reflect real-world performance. A more accurate fuel burn model would enhance environmental impact assessments by accounting for fleet age mix. Furthermore, the industry push for greener aviation (through measures like optimized flight trajectories, new engine technologies, and alternative fuels) requires robust benchmarks; these benchmarks should consider that a 20-year-old aircraft will not perform like it did when new. In summary, there is a pressing motivation to improve fuel flow prediction by bridging the gap between idealized models and the reality of ageing aircraft.

In this paper, we present a comprehensive investigation into methods for integrating engine deterioration and aircraft ageing effects into fuel flow modeling. Using a rich dataset of nearly nineteen thousand Quick Access Recorder (QAR) flights from nine Airbus A320-214 airframes spanning a wide range of service ages, we systematically evaluate classical parametric models, empirical age-dependent corrections, and modern machine learning architectures. Our approach includes augmenting neural networks with explicit age information either as an input feature or through structured inductive biases that encode expected ageing behaviour.

We demonstrate that standard baseline models, such as EUROCONTROL’s BADA, tend to underestimate fuel consumption as aircraft age, reflecting their assumption of nominal, time-invariant performance. Incorporating ageing effects significantly reduces bias and improves predictive accuracy, al-



though challenges remain due to limited maintenance data and the relatively small number of individual aircraft represented. The contributions of this work are twofold: (i) quantifying the impact of ageing on fuel flow prediction accuracy across diverse modelling paradigms, and (ii) proposing novel hybrid approaches that embed ageing effects within data-driven frameworks to enhance robustness and interpretability.

The remainder of the paper is structured as follows: Section II reviews relevant literature on fuel flow modelling and ageing effects; Section III describes the dataset and its characteristics; Section IV details the methodological approaches; Section V presents results and comparative analysis; Section VI discusses limitations and implications; and Section VII concludes with perspectives for future work.

II. STATE OF THE ART

Before delving into our contribution, it is essential to situate our work within the broader landscape of fuel burn modelling. The estimation of aircraft fuel consumption has been the subject of extensive research over several decades, evolving from purely physics-based formulations to sophisticated data-driven and hybrid approaches. In this section, we review the main strands of the literature, highlighting how classical parametric models, machine-learning methods, and physics-informed approaches have each contributed to the state of the art. We also examine studies on engine deterioration and ageing effects, which provide crucial context for understanding long-term performance trends. This overview lays the foundation for identifying the remaining research gaps that our study aims to address.

A. Classical Parametric Models

Classical fuel burn estimation relies on simplified physics-based formulations that capture the primary energy balance and specific fuel consumption (SFC) characteristics of turbofan engines. The EUROCONTROL Base of Aircraft Data (BADA) model exemplifies this category, offering phase-specific fuel flow equations parametrized by manufacturer-provided coefficients [2], [4]. While BADA underpins many trajectory simulators and global emissions inventories due to its computational efficiency and transparency, it assumes “nominal” aircraft performance, neglecting within-type variations and operational factors beyond flight phase and mass. Extensions such as the FEAT reduced-order model illustrate how BADA outputs can be combined with regression techniques to estimate fleet-level fuel use within 5% of reported values [5]. It provides an empirical formula that can be applied to adjust BADA fuel flow for older aircraft:

$$\text{Seymour_coeff}(\text{age}) = \frac{100}{100 - 1.28 \times \log(\text{age} + 1)}.$$

OpenAP further democratizes parametric modelling by providing an open-source, research-focused implementation of energy-balance equations [6], and recent aerodynamic-theory methods have proposed an alternate aircraft open-source modelling for cruise and other phases [7], [8]. Despite these advances, classical models typically treat aircraft performance as

static over time, leading to systematic biases when additional factors like ageing and deterioration effects alter real-world fuel flow.

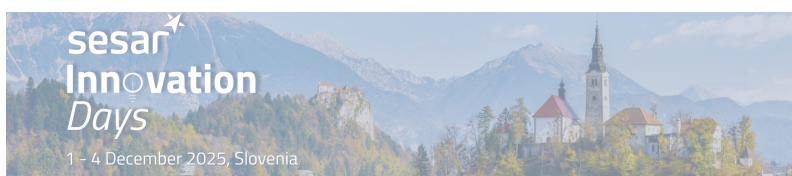
B. Pure Data-Driven Models

Driven by the proliferation of flight-recorded and simulated data, purely statistical methods have been developed to learn fuel burn relationships directly from observations. Gaussian process regression has been applied to operational data to refine emissions inventories, capturing nonlinear dependencies with quantified uncertainty [9]. Tree-based ensemble methods, such as gradient boosting and random forests, further improved per-flight fuel estimates by integrating sensor-derived features [10]. Support vector machines and component-level simulations have been explored to diagnose engine deterioration through fuel flow anomalies [11], [12]. Neural network architectures, including radial basis function networks and feed-forward multilayer perceptrons, demonstrated strong performance across flight phases, outperforming parametric baselines on full-flight sensor datasets [13]–[16]. Time-series models such as long short-term memory (LSTM) networks have been tailored to predict contingency fuel requirements and emissions profiles, showing promise for real-time applications [17], [18]. However, data-driven methods often require extensive, high-quality training data, and their black-box nature can obscure failure modes when encountering out-of-distribution operating regimes or ageing-related drift.

Scaling neural networks to millions of flight-record samples has enabled the capture of complex, non-linear trends in fuel flow predictions. Jarry et al. trained deep architectures on thousands of Automatic Dependent Surveillance-Broadcast (ADS-B) derived flights spanning over 100 aircraft types, incorporating domain generalization techniques, such as stochastic perturbations of performance parameters to reduce error on unseen types [19]. Huang et al. demonstrated that publicly available ADS-B data can act as a surrogate for proprietary QAR data when estimating fuel consumption, by combining classification trees with neural regressors [20]. Despite robust accuracy, deep models typically define each aircraft by static type identifiers, and variability in service life or individual maintenance histories is not accounted for in these studies.

C. Hybrid and Physics-Informed Methods

Hybrid approaches seek to synergize the interpretability of physics-based models with the flexibility of machine learning. One strategy involves data-driven calibration of BADA coefficients against QAR measurements: Uzun et al. showed that per-aircraft thrust and SFC coefficient tuning reduces fuel prediction errors from around 3% down to under 1% for individual flights [21]. Neural networks augmented with physics-based loss functions during training have further improved extrapolation robustness to parameter variations [21], [22]. These hybrid frameworks enhance accuracy and interpretability but still generally assume time-invariant performance parameters.



D. Engine Deterioration and Ageing Effects

Empirical analyses quantify the impact of component wear on fuel efficiency. In 1980, Mehalic and Ziemianski first documented increases in SFC due to compressor fouling and turbine erosion [23]. Decades later, Arrieta et al. projected a 4.5% fuel penalty over CF34-series service life via detailed cycle modelling [3]. Naeem et al. developed deterioration cycle models showing that in-service degradation shifts engine operating points, requiring higher fuel flow to sustain performance [24]–[26]. Advanced anomaly detection using neural networks and fuzzy logic achieves high accuracy in identifying early-stage degradation [27], [28]. Navaratne et al. incorporated degraded-engine performance into trajectory optimization, revealing that using clean-engine optimized paths on aged engines incurs significant fuel and emissions penalties [29]. Another research proposed a logarithmic formula based on aircraft age that computes a deterioration coefficient to multiply the baseline fuel consumption, thereby accounting for engine efficiency loss over time [5]. Environmental factors, such as reduced air density at cruise, exacerbate degradation impacts, though alternative fuels may partially offset these losses due to their higher heating values [30], [31].

E. Research Gap

While classical, statistical, deep learning, and hybrid methods provide a comprehensive toolkit for fuel flow estimation, the explicit incorporation of aircraft ageing remains limited. Most models presume static performance parameters, neglecting within-type and within-aircraft degradation trends that evolve over service life. This omission can introduce systematic over or underestimation of fuel burn, compromising cost planning, emissions forecasting, and operational decision support. Our work addresses this gap by systematically analysing QAR data across aircraft lifecycles and developing integrated modelling strategies that explicitly account for age-related performance drift.

III. DATASET

The dataset comprises approximately 19,000 Quick Access Recorder (QAR) flight records from a single aircraft type, the Airbus A320-214, covering nine distinct airframes over a period of around two years. Although detailed information on engine cleaning cycles is unavailable, the dataset includes the age of each aircraft measured from its first commercial flight. Importantly, the dataset does not span the full service life of each airframe: it contains about two years of flights, during which the aircraft ages at the time of recording range from 3 to 14 years. This age distribution across the fleet is visualized in Figure 1, highlighting the representation of both relatively new and mature aircraft within the dataset. The dataset is split for each aircraft based on the date (half of the overall calendar period) and represents a split of around 2/3 in train and 1/3 in test. This design choice and associated limitations are discussed in the Section VI.

Each flight record contains 4-second-rate measurements of various operational and environmental parameters, enabling

a granular analysis of fuel flow behaviour over time. The dataset’s diversity in aircraft age and operational conditions offers a unique opportunity to explore the impact of ageing on fuel consumption and to calibrate both parametric and data-driven models accordingly. Table I summarizes the set of input features extracted from QAR data and metadata, along with the target variable used for model training.

TABLE I: Main model variables: QAR sensors, metadata, and target.

Variable	Unit	Variable	Unit
Pressure altitude	ft	True airspeed (TAS)	kt
$\Delta\text{TAS} / \Delta t$	kt/s	Vertical speed	ft/min
Ground speed	kt	Mass	kg
Aircraft age	yr	Static air temperature	K
Fuel flow (target)	kg/h		

IV. METHODOLOGY

This study investigates multiple approaches to incorporate aircraft ageing effects into fuel flow modelling, benchmarking them against two established baseline models that do not explicitly account for engine deterioration or age-related performance changes. The models are trained jointly across the entire fleet, with aircraft identity and age-related variables used as inputs. The goal is to evaluate how incorporating ageing information improves prediction accuracy and robustness.

A. Baselines

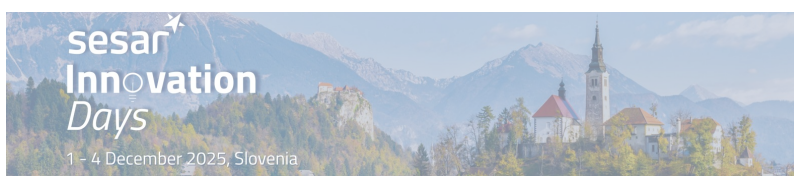
Baseline 1 corresponds to a methodology based on the BADA framework [4] (hereafter referred to as BADA for simplicity) — The EUROCONTROL BADA 4.2 model provides physics-based parametric fuel flow estimations widely used in trajectory simulation and emissions inventories. While highly interpretable and computationally efficient, BADA assumes nominal, time-invariant aircraft performance and does not model ageing or engine degradation effects. To compute the fuel flow we are using the BADA 4.2.1 model (via the *pyBADA* library). Maximum thrust is assumed during the climb phase without any thrust derate and idle thrust is used during descent. True airspeed and vertical speed are smoothed using Savitzky–Golay filtering, and acceleration is then empirically derived from the smoothed true airspeed (TAS).

Baseline 2: Acropole [16] — The model is a feed-forward neural network with multiple dense layers and optional dropout for regularization. All inputs are normalized and outputs are denormalized using fixed statistics to ensure consistency with physical units. L2 regularization is applied to all dense layers, and the final output is constrained to be non-negative via a ReLU activation.

B. Ageing-Inclusive Modeling Approaches

To address the limitations of the baselines, we propose and compare several methods to integrate aircraft age or engine deterioration effects into fuel flow prediction:

BADA Seymour [5] — A parametric correction based on a logarithmic deterioration coefficient proposed by Seymour et



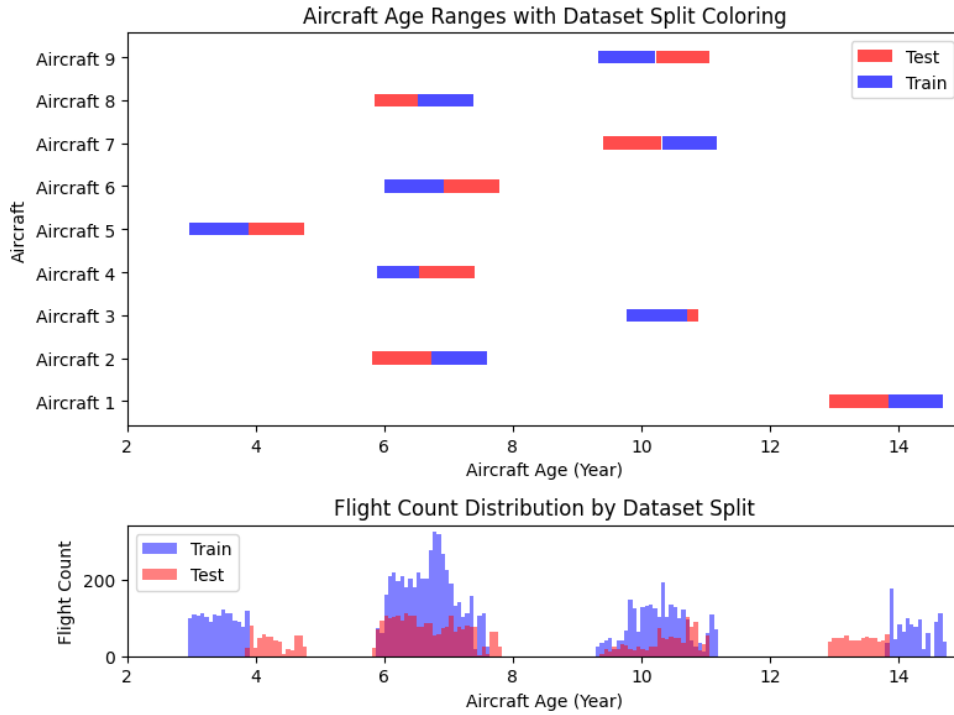


Figure 1: Distribution of aircraft ages in the dataset, illustrating the range of service years covered by the nine A320-214 airframes.

al. that scales baseline fuel consumption according to aircraft age, capturing the empirical fuel penalty due to ageing.

BADA Calibrated — An adaptation of the Seymour formula in which the deterioration coefficient is calibrated using observed QAR data and BADA prediction, allowing for improved fit and better representation of real-world ageing trends.

Acropole Age — A data-driven regression model built upon the Acropole baseline architecture, augmented by explicitly including aircraft age as an additional input feature to capture nonlinear interactions between ageing and other flight parameters.

Acropole Inductive Bias — Compared to the standard neural architecture that treats all input features identically, the proposed model introduces a structured parametrization to explicitly model the effect of aircraft age as an inductive bias. Both models employ input and output normalization, multiple dense layers with L2 regularization, and ReLU-constrained outputs. However, in this approach, the "age" feature is separated from the other normalized inputs before being fed into the neural network at the final layer. The output of the network, denoted as \hat{y}_{base} , is multiplied by an age-dependent coefficient $\text{Coeff}(\text{age}) = 1 + a \cdot \log(\text{age} + 1)$, parametrized as a differentiable function of age. The final prediction is $\hat{y} = \max(0, \hat{y}_{\text{base}} \times \text{Coeff}(\text{age}))$.

This explicit multiplicative correction enforces the effect of ageing directly in the output, embedding domain knowledge about performance deterioration into the model structure, rather than relying solely on generic feature learning.

C. Evaluation Metrics and Testing

Models are trained on disjoint training, validation, and test sets, with final evaluation on the test set. Performance is measured with several metrics, where \mathcal{D} is the dataset of pairs (x, y) and h the model:

- **MAE**: average absolute deviation, $\text{MAE}(h, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{(x,y)} |h(x) - y|$.
- **MAPE**: relative error, $\text{MAPE}(h, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{(x,y)} \frac{|h(x) - y|}{y}$.
- **ME**: signed bias, $\text{ME}(h, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{(x,y)} (h(x) - y)$.
- **MSE**: squared deviation, $\text{MSE}(h, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{(x,y)} (h(x) - y)^2$.
- **Bias ratio**: ME^2 / MSE , fraction of error due to systematic bias.

D. Training

The neural network models were implemented and trained using TensorFlow on a workstation equipped with two NVIDIA RTX 6000 GPUs. The model architecture comprised five hidden layers with 128 units per layer, ReLU activations, and L2 regularization ($\lambda = 1 \times 10^{-5}$). Optimization was performed using AdamW with a warm-up cosine learning rate schedule, where the learning rate increased linearly during the first 10% of the 200 training epochs, followed by a cosine decay towards zero. The batch size was selected to be 4096. Model performance was monitored on a held-out validation set using mean absolute error (MAE) and mean absolute percentage error (MAPE) metrics. Training checkpoints were saved automatically, with model parameters corresponding to the lowest validation loss retained for subsequent evaluation.

V. RESULTS

A. Comparison of Baselines

In this section we analyze the predictive tendencies of the baseline models and the impact of incorporating ageing corrections on fuel flow estimation accuracy.

The BADA model demonstrates a systematic bias that favors newer aircraft: specifically, it increasingly underestimates fuel consumption as aircraft age. This trend, illustrated by the blue line in Figures 2b and 2c, is expected given that the model was originally trained on manufacturer data. As previously shown in [15], the BADA model exhibits an average mean absolute error (MAE) of approximately 100 kg/h (Figure 2b) and an average underestimation bias of 50–60 kg/h. Notably, the contribution of this bias to the mean squared error (MSE) ranges between 5% and 20% (Figure 2a), indicating that the remaining error is primarily due to modeling limitations rather than age-related bias.

Incorporating the Seymour correction coefficient (green lines) reduces the bias contribution to the MSE to less than 5%, and to nearly 0% when using the calibrated coefficient (cyan lines). Correspondingly, the direct bias decreases to -30 kg/h with the Seymour correction and to approximately -5 kg/h with calibration. It should be noted, however, that even with this age correction, the combined MAE remains around 85 kg/h, which is still significantly higher than that observed with the Acropole model.

Analysis of the curves corresponding to the Acropole model (red lines) reveals that the model effectively learns an average fleet age. As a result, it tends to overestimate the fuel consumption of younger aircraft (by approximately $+20$ kg/h) and underestimate that of older aircraft (by approximately -15 kg/h), as shown in Figure 2b. Nevertheless, the bias contribution to the mean squared error (MSE) remains below 10%, with a mean absolute error (MAE) of about 45 kg/h, indicating lower modeling error and enhanced robustness to age-related effects.

Incorporating age as an explicit input (orange lines) or as an inductive bias (gold line) further reduces the age-related bias to nearly zero. However, since the initial bias contribution was already below 5%, the resulting improvement in overall performance is moderate, with the average MAE decreasing only slightly to approximately 44 kg/h. While the overall gain for the fleet is not substantial, it is noteworthy that the modeling of age is particularly beneficial for younger aircraft, for which the MAE decreases from 50 kg/h to 40 kg/h.

Additionally, Figure 3 presents the evolution of the correction coefficients as a function of aircraft age. The coefficient for *Acropole Age* (orange), which corresponds to a non-explicit internal value derived from the neural network, was estimated as follows: for each data point, the predicted fuel flow was first computed by setting the age to zero, and this baseline value was then used to normalize subsequent predictions for ages ranging from 0 to 25 years.

Several observations can be drawn from the correction coefficient curves. First, the Seymour coefficient appears to

underestimate engine deterioration with increasing age. This may be attributed to its derivation from a dataset comprising different aircraft types and a limited number of trajectories. In contrast, the three other calibration methods exhibit consistent trends, showing an increase in fuel consumption of up to 6% for aircraft at 20 years of age. Notably, the *Acropole Age* approach, where age is explicitly provided as an input, yields a less smooth correction curve that adapts closely to observed training values. However, since the dataset contains only nine distinct aircraft, this modelling approach is likely subject to overfitting due to limited representativeness—a limitation discussed further in Section VI.

Moreover, for ages greater than 15 years (for which no observed data are available), the model exhibits a non-monotonic increase in the correction coefficient then decrease, further highlighting that, in situations with limited data representativeness, the introduction of inductive bias can provide more robust model characteristics.

B. Impact on Fuel Consumption Assessment

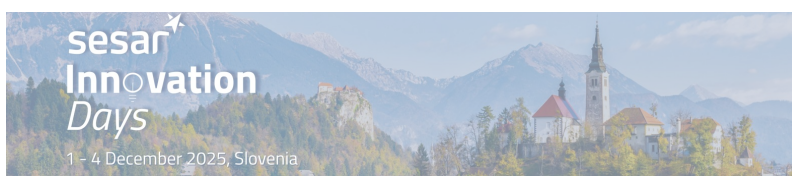
1) *Current fleet impact*: To evaluate the practical implications of the different modelling approaches on environmental impact assessments, we computed the total fuel consumption over the entire test dataset for each model. This cumulative consumption provides insight into how each model's prediction differences translate into estimated fuel use at fleet scale, a key factor in operational planning and emissions accounting.

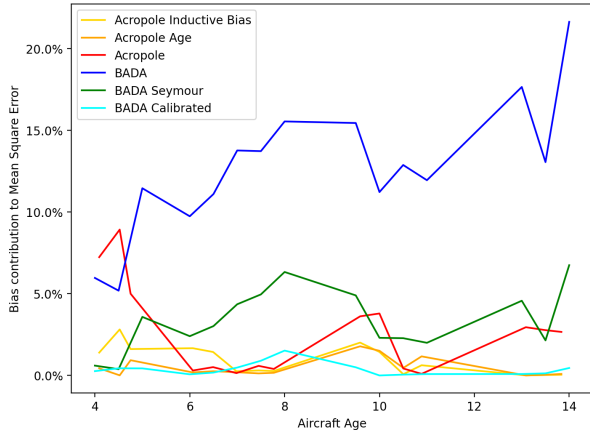
Table II summarizes the total fuel consumption (in tonnes) predicted by each model alongside the ground truth measured consumption. The raw BADA model underestimates total fuel burn by approximately 4.81% (1,115 tonnes) compared to observed values, reflecting its nominal performance assumption without ageing effects. Incorporating the Seymour parametric correction reduces this underestimation by more than half, to around 2.17% (503 tonnes), while the calibrated BADA variant further narrows the gap to a marginal 0.26% (61 tonnes).

The data-driven Acropole models provide even closer estimates, with differences below 0.22% (52 tonnes) across all three age-aware variants. Among these, the Acropole Inductive Bias approach achieves the smallest discrepancy of 0.15% (34 tonnes), indicating the benefit of explicitly embedding ageing effects into the model structure.

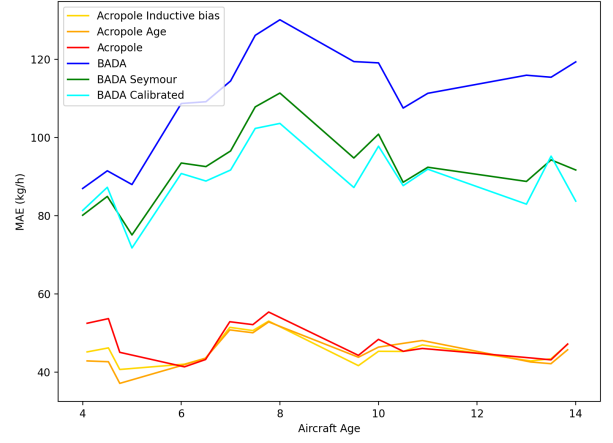
These results highlight that failing to account for engine deterioration and ageing in fuel flow models can lead to significant underestimation of fleet fuel consumption and associated emissions. Conversely, integrating age-dependent corrections or data-driven ageing representations materially improves fuel burn estimates, enhancing the reliability of environmental impact assessments. This has important ramifications for regulatory compliance, airline operational planning, and long-term sustainability strategies.

While the overall absolute differences may appear modest relative to total fuel use, even small percentage errors translate to substantial quantities of fuel and CO₂ emissions when aggregated over the scale of global commercial aviation. Therefore, precise modelling of ageing effects is critical to

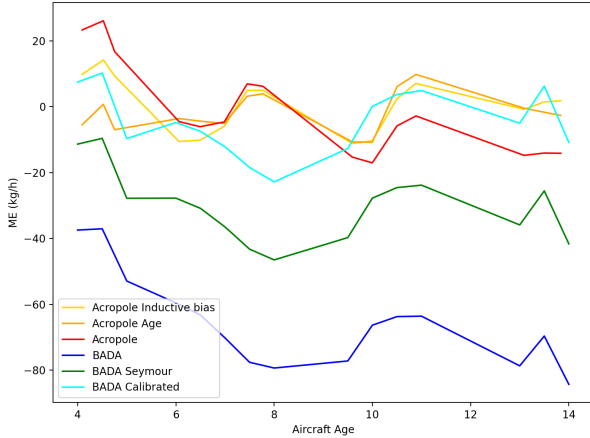




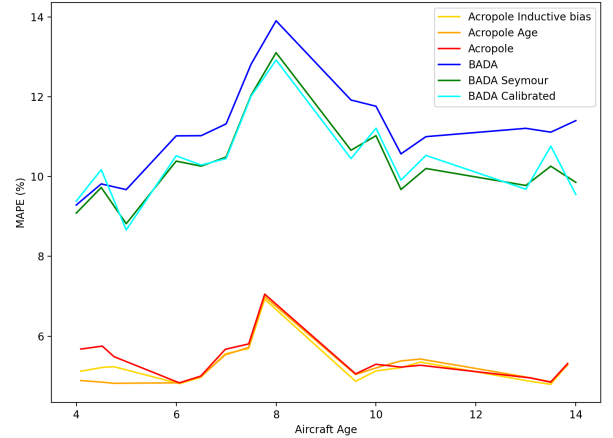
(a) Bias contribution to Mean Squared Error (%) vs. aircraft age



(b) Mean Absolute Error (MAE) in kg/h vs. aircraft age



(c) Mean Error (ME) in kg/h vs. aircraft age



(d) Mean Absolute Percentage Error (MAPE) (%) vs. aircraft age

Figure 2: Error metrics as a function of aircraft age for baselines and different modeling approaches

TABLE II: Computed consumption over test set in Tonnes of fuel

	Ground Thruth	BADA	BADA Seymour	BADA Calibrated	Acropole	Acropole Inductive Bias	Acropole Age
Consumption (tonnes)	23176.4	22061.9	22673.5	23115.2	23124.8	23142.2	23141.2
Difference (tonnes)		-1114.5	-502.9	-61.2	-51.6	-34.3	-35.2
Difference Ratio		-4.81%	-2.17%	-0.26%	-0.22%	-0.15%	-0.15%

underpin robust policy and operational decisions aimed at mitigating aviation’s environmental footprint.

2) *Current fleet in the future*: To evaluate the long-term impact of engine ageing on fleet fuel consumption, we simulate cumulative fuel use over a horizon of 1 to 15 years, assuming a constant fleet size and no new aircraft produced. The simulated fleet corresponds to the nine A320-214 airframes in our dataset, whose ages at the start of the simulation range from 3 to 14 years (average age 8 years).

The baseline reference consumption curve was defined as the average of the correction coefficients obtained from the two best-performing models in our study, namely the *BADA Calibrated* and the *Acropole Inductive Bias* approaches. Since the true functional form of engine degradation with age is not directly observable, we cannot rely on a single “ground

truth” curve. Instead, we approximate the underlying trend by combining the results of these two complementary models. This averaging ensures a more robust baseline, as it reduces the influence of model-specific biases and reflects the central tendency of the most accurate methods available. The resulting function, $1 + 0.0231 \log(1 + \text{age})$, therefore represents our best estimate of the expected long-term increase in fuel consumption due to ageing effects at the fleet level.

Figure 4 presents the projected differences in fuel consumption (in tonnes) for each modelling approach relative to this baseline scenario.

The simulation results show that the raw BADA model substantially underestimates the additional fuel burn compared to the reference deterioration baseline, with the discrepancy increasing from approximately 1,172 tonnes at 1 year to over

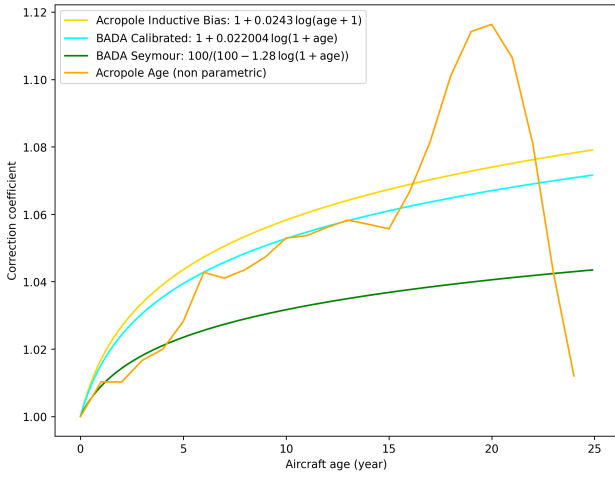


Figure 3: Correction coefficient function of aircraft age for the different modelling approaches

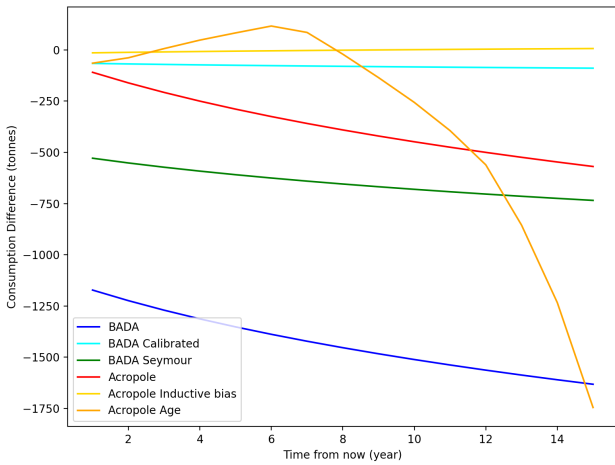


Figure 4: Projected fuel consumption differences (tonnes) over a 1–15 year horizon under the assumption of a constant fleet.

1,600 tonnes at 15 years. These figures correspond to the nine-aircraft fleet considered in this study, i.e. an average deviation of about 130–180 tonnes per aircraft over the 15-year horizon. A similar trend is observed for the BADA Seymour correction, albeit with smaller deviations.

Data-driven models exhibit more varied behaviour: the Acropole model slightly overestimates consumption increases during the initial years, gradually converging towards the BADA Seymour predictions. The Inductive Bias model and BADA Calibrated model demonstrate smaller errors and minor divergence, consistent with their averaging in the chosen baseline correction coefficient.

Finally, the Acropole Age model displays more oscillatory behaviour and, as previously discussed, tends to substantially under-predict fuel consumption once the average fleet age exceeds the maximum age observed during training.

These projections highlight the sensitivity of long-term fuel consumption forecasts to the selection of ageing models and

calibration methodologies. Despite the limited representativeness of the training dataset, the introduction of inductive bias appears to enhance model robustness. Overall, this analysis confirms that incorporating age-related corrections is essential and is best achieved using calibrated coefficients.

VI. DISCUSSION

A key limitation of this study is the restricted number of individual airframes available, with only nine aircraft of the same type. This narrow sampling constrains the representativeness of the age parameter, particularly when attempting to model the complex relationship between fuel consumption and aircraft ageing using fully data-driven neural network approaches. To partially mitigate this, the dataset was split for each aircraft into two temporal subsets based on calendar dates, effectively doubling the temporal coverage and improving the model's ability to capture ageing trends within the limited fleet. Nevertheless, this approach cannot fully compensate for the limited number of distinct airframes.

In real operational fleets, maintenance events such as engine washes, overhauls, or component replacements periodically reset the effective engine degradation, producing non-monotonic patterns of performance loss that are not captured when age is measured solely as time since entry into service. Without access to detailed maintenance or cleaning cycle histories, our models may conflate true age-related deterioration with aircraft-specific idiosyncrasies, leading to over-fitting and poor generalization outside the observed age range. In this context, the 'age' parameter serves as a pragmatic, aggregated proxy for the cumulative effects of degradation over service life, necessitated by the unavailability of proprietary maintenance records.

As demonstrated in the fleet projection analysis (Section V), the choice of ageing model significantly affects long-term fuel consumption forecasts. Under the assumption of a constant fleet, the raw BADA based methodology tends to substantially underestimate additional fuel burn, while calibrated and inductive bias approaches yield more consistent and realistic predictions. The observed oscillations in the Acropole Age model further underscore the risks of extrapolating beyond the training data without robust inductive constraints.

Consequently, while our age-parameterized neural network models can capture ageing trends present in the available data, their ability to generalize and robustly represent ageing effects across a more diverse fleet remains limited if no inductive bias is introduced. Expanding the dataset to include a greater number of aircraft, ideally with detailed records of engine cleaning and maintenance events, would be essential to improve both the accuracy and generalizability of age-based corrections in data-driven fuel flow models. Introducing inductive bias in modelling ageing effects appears to be a promising direction to mitigate data scarcity and improve model robustness, offering a practical solution when perfect maintenance data are unavailable.

VII. CONCLUSIONS

This study has provided a systematic evaluation of methods for incorporating ageing and engine deterioration effects into aircraft fuel-flow modelling. Using a large QAR dataset covering nine A320-214 airframes with a wide range of service ages, we have shown that standard parametric models, such as BADA based methodology, systematically underestimate fuel consumption for older aircraft. The application of empirical correction coefficients from the literature, as well as coefficients calibrated on observed data, substantially reduces bias.

Neural network approaches that explicitly model aircraft age, either as an additional input feature or through an inductive bias parameterization, offer improved performance, especially in reducing mean error and enhancing robustness to within-fleet variability. However, the limited number of distinct airframes constrains the representativeness of age and can result in overfitting or limited generalization. The absence of maintenance history data, such as cleaning or engine overhaul events, further limits the ability to distinguish true age-related effects from aircraft-specific variations.

The simulation of future fleet fuel consumption under a constant fleet size assumption reveals that models without age correction may substantially misestimate long-term fuel burn, potentially biasing operational and environmental analyses. Models incorporating calibrated correction coefficients and inductive biases provide more realistic projections, emphasizing the importance of robust ageing representations.

Future work should focus on expanding datasets to include broader and more diverse fleets, ideally with detailed maintenance records, to strengthen the reliability and applicability of age-based corrections in data-driven fuel-flow models. Nevertheless, this study demonstrates that integrated modelling strategies, particularly those incorporating a physics-informed inductive bias, offer a powerful and practical approach to accurately correct systemic bias.

REFERENCES

- [1] D. S. Lee, D. W. Fahey, A. Skowron, M. R. Allen, U. Burkhardt, Q. Chen, S. J. Doherty, S. Freeman, P. M. Forster, J. Fuglestedt *et al.*, “The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018;” *Atmospheric Environment*, vol. 244, p. 117834, 2021.
- [2] A. Nuic, D. Poles, and V. Mouillet, “Bada: An advanced aircraft performance model for present and future atm systems;” *International journal of adaptive control and signal processing*, vol. 24, no. 10, pp. 850–866, 2010.
- [3] M. D. J. Gurrola Arrieta, R. Botez, and A. Lasne, “An engine deterioration model for predicting fuel consumption impact in a regional aircraft;” *Aerospace*, 2024.
- [4] A. Nuic and V. Mouillet, “User manual for the base of aircraft data (bada) revision 4;” *EUROCONTROL Experimental Centre, EEC Technical/Scientific Report No. 12/11/22-58 v1.3*, 2016.
- [5] 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.
- [6] J. Sun, J. M. Hoekstra, and J. Ellerbroek, “Openap: An open-source aircraft performance model for air transportation studies and simulations;” *Aerospace*, vol. 7, no. 8, p. 104, 2020.
- [7] D. Poll and U. Schumann, “An estimation method for the fuel burn and other performance characteristics of civil transport aircraft in the cruise. part I fundamental quantities and governing relations for a general atmosphere;” *The Aeronautical Journal*, vol. 125, no. 1284, pp. 257–295, 2021.
- [8] —, “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, no. 1284, pp. 296–340, 2021.
- [9] Y. S. Chati and H. Balakrishnan, “A gaussian process regression approach to model aircraft engine fuel flow rate;” in *8th International Conference on Cyber-Physical Systems*. ACM, 2017, pp. 131–140.
- [10] —, “Statistical modeling of aircraft engine fuel flow rate;” in *30th Congress of the International Council of the Aeronautical Science*, 2016.
- [11] Z. Weili and G. Long, “Performance deterioration evaluation analysis of aircraft engine based on simulation;” in *2014 Prognostics and System Health Management Conference*, 2014, pp. 520–524.
- [12] W.-L. Zhao, C. Hou, and Q. Wang, “Diagnosis of aircraft engine performance deterioration based on support vector machines;” in *2014 10th International Conference on Reliability, Maintainability and Safety (ICRMS)*, 2014, pp. 44–48.
- [13] T. Baklacioglu, “Predicting the fuel flow rate of commercial aircraft via multilayer perceptron, radial basis function and anfis artificial neural networks;” *The Aeronautical Journal*, vol. 125, no. 1285, pp. 453–471, 2021.
- [14] G. Jarry, D. Delahaye, and E. Feron, “Approach and landing aircraft on-board parameters estimation with lstm networks;” in *2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT)*. IEEE, 2020, pp. 1–6.
- [15] G. Jarry and D. Delahaye, “Toward novel environmental impact assessment for ansps using machine learning;” in *Climate change and the role of air traffic control Research Workshop*, Vilnius, 2021.
- [16] 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.
- [17] L. Li, S. Yuan, Y. Teng, and J. Shao, “A study on sustainable consumption of fuel—an estimation method of aircraft;” *Energies*, vol. 14, no. 22, p. 7559, 2021.
- [18] S. Metlek, “A new proposal for the prediction of an aircraft engine fuel consumption: a novel cnn-bilstm deep neural network model;” *Aircraft Engineering and Aerospace Technology*, vol. 95, no. 5, pp. 838–848, 2023.
- [19] G. Jarry, R. Dalmau, P. Very, and J. Sun, “On the generalization properties of deep learning for aircraft fuel flow estimation models;” *Transportation Research Part C: Emerging Technologies*, vol. 176, p. 105143, 2025.
- [20] C. Qu, H. Cao, S. Sun, and M. Xu, “Modelling of fuel flow in climb phase through multiple linear regression based on the data collected by quick access recorder;” *Journal Européen des Systèmes Automatisés*, 2019.
- [21] M. Uzun, M. U. Demirezen, and G. Inalhan, “Physics guided deep learning for data-driven aircraft fuel consumption modeling;” *Aerospace*, vol. 8, no. 2, p. 44, 2021.
- [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] C. M. Mehalic and J. A. Ziemianski, “Performance deterioration of commercial high-bypass ratio turbofan engines;” *SAE Technical Paper*, 1980.
- [24] M. Naeem, R. Singh, and D. Probert, “Implications of engine deterioration for creep life;” *Applied Energy*, vol. 60, pp. 183–223, 1998.
- [25] M. Naeem, “Impacts of low-pressure (lp) compressor’s deterioration of a turbofan engine upon fuel-usage of a military aircraft;” *The Aeronautical Journal*, vol. 112, no. 1128, pp. 33–45, 2008.
- [26] —, “Implications of high-pressure turbine’s erosion for a military turbofan’s fuel consumption;” *Journal of Aerospace Engineering*, vol. 25, pp. 108–116, 2012.
- [27] B. Kurt, “Prediction of performance degradation in aircraft engines with fuel flow parameter;” *Neural Computing and Applications*, pp. 1–10, 2023.
- [28] K. Amrutha, Y. K. Bharath, and J. Jayanthi, “Aircraft engine fuel flow parameter prediction and health monitoring system;” in *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, 2019, pp. 39–44.
- [29] R. Navaratne, V. Sethi, and C. Lawson, “Impact of engine degradation on optimum aircraft trajectories: Short range;” in *Volume 1: Aircraft Engine; Fans and Blowers; Marine; Honors and Awards*, 2019, presented at ASME Turbo Expo 2019.



- [30] W. C. Koh and N. M. Mazlan, "Performance analysis of deteriorated engine powered by alternative fuels," *International Review of Mechanical Engineering (IREME)*, 2018.
- [31] W. C. Koh, N. M. Mazlan, P. Rajendran, and M. A. Ismail, "A

computational study to investigate the effect of altitude on deteriorated engine performance," in *IOP Conference Series: Materials Science and Engineering*, vol. 370, 2018.

