Machine Learning in Aviation Environmental Impact Analysis

Subjects: Engineering, Aerospace | Transportation | Mathematics, Interdisciplinary Applications Contributor: Zhenyu Gao , Dimitri N. Mavris

The rapid growth of global aviation operations has made its negative environmental impact an international concern. Accurate modeling of aircraft fuel burn, emissions, and noise is the prerequisite for informing new operational procedures, technologies, and policies towards a more sustainable future of aviation. Due to the advances in big data technologies and effective algorithms, the transformative data-driven analysis has begun to play a substantial role in aviation environmental impact analysis. The integration of statistical and machine learning methods in the workflow has made such analysis more efficient and accurate.

data-driven methods statistics machine learning aviation environmental impact

1. Introduction

With the rapid growth of global air traffic operations in the past five decades, the aviation industry has grown to become an integral part of the global economy. While the global air transportation operations at scale have greatly facilitated people's travel and business, their negative environmental impact, also identified by some entities as the most significant adverse impact of aviation ^[1], has emerged as a major concern internationally. The three primary aspects of aviation's negative impacts on the environment are: (1) local air quality impacts that can exacerbate health-harming air pollution, (2) climate change impacts that can accelerate global warming, and (3) community noise impacts that can undermine affected population's mental well-being ^[2]. Undoubtedly, the aviation industry must keep the development momentum to meet the needs of a growing economy while simultaneously being more environmentally sustainable. The system must operate harmoniously within the constraints imposed by requisites such as clean air and water, limited noise impacts, and a livable climate.

Aircraft, ground vehicles, Ground Support Equipment (GSE), and other stationary sources at the airport produce emissions as a result of the combustion of fuel. Aircraft engines mainly produce carbon dioxide (CO2), which comprises around 70% of the exhaust, and water vapor (H2O), which comprises around 30% of the exhaust. Less than 1% of the emissions is composed of nitrogen oxides (NOx), carbon monoxide (CO), sulfur oxides (SOx), partially combusted or unburned hydrocarbons (HC), particulate matter (PM), other trace compounds, soot and sulfate aerosols, and increased cloudiness due to contrail formation ^{[2][3]}. These emissions undergo complex interactions among themselves and with the changing background atmosphere ^[4]. Among the aircraft pollutant emissions, around 10% are emitted near the surface of earth (below 3000 ft above ground level) while the remaining 90% are emitted at above 3000 ft, mostly at cruise altitudes within the Upper Troposphere and the Lower Stratosphere (UTLS). While aviation technologies have become more fuel efficient, the overall emissions from

aviation has risen due to the rapidly increasing volume of air travel. Statistics reveal that the annually averaged growth rate in global aviation CO2 emissions was 2.2% per year over the period 1970 to 2012 and 5% per year for 2013 to 2018. In 2018, global aviation CO2 emissions exceeded 1000 million tonnes per year for the first time, which accounts for approximately 2.4% of all anthropogenic emissions of CO2 (including land use change) ^[3]. These observations indicate that aviation emissions remains a challenging issue towards a more sustainable future of aviation.

Aircraft noise pollution refers to the "unwanted sound" produced by aircraft or its components in flight. In general, aircraft noise is produced by three main sources. Engine noise is the main source of aircraft noise. For propeller aircraft and helicopter, engine noise includes both aerodynamically induced noise from the propeller and mechanically induced noise from other moving parts of the engine. For jet aircraft, engine noise is dominated by jet noise from the gas turbine engines, which is responsible for much of the aircraft noise during takeoff and climb. Jet noise is caused by the high speed flow leaving the exhaust of the engine which is highly unstable and turbulent. Aerodynamic noise is the second source of aircraft noise. Aerodynamic noise arises from airflow around the aircraft fuselage and control surfaces and increases with aircraft speed and air density. Supersonic aircraft, such as fighter jets, often creates intense aerodynamic noise called sonic boom due to the formation of shock waves during supersonic flight. Aerodynamic noise can sometimes be mitigated by designing the shape of airframe. The third source of aircraft noise is the aircraft systems. Some examples include noises from the Auxiliary Power Unit (APU) and the cabin pressurization and conditioning systems. Aircraft noise can disrupt sleep, cause community annoyance, adversely affect academic performance of children, and could increase the risk for cardiovascular disease of people living in the vicinity of airports ^[5].

2. Data Reduction

Performing complex analysis and computation on large datasets can be impractical or infeasible. In such cases, data reduction is applied to obtain a reduced representation of the dataset that is much smaller in volume, yet still closely maintains the integrity of the original dataset ^[6]. Applying a reduced dataset in analysis and computation trades accuracy for speed in response to the need of obtaining quick approximate answers to queries on large datasets. The development of data reduction techniques for science and engineering applications has gained increasing interest in the community. The motivation behind the trend is that contemporary operations, scientific observations, experiments, and simulations are generating unwieldy amounts of data which are beyond people's capacity to store, stream, analyze, and archive. In the meantime, these massive datasets almost always contain redundancies and trivialities.

Data reduction strategies mainly include three broad categories: dimensionality reduction, numerosity reduction, and data compression. Dimensionality reduction techniques reduce the number of attributes/features p under consideration. Some dimensionality reduction methods, such as Principal Components Analysis (PCA) and wavelet transform, aim to transform or project the original data onto a lower-dimensional space. Other methods such as attribute subset selection detect and remove non-informative, irrelevant, and redundant attributes from the full feature set. Numerosity reduction, on the other hand, reduces the number of data points n in the original dataset.

Numerosity reduction can be classified as either parametric or nonparametric. For parametric methods, the data is represented by a parametric model which consists of model form and model parameters. After the modeling process, only the model parameters are stored instead of the actual dataset, thus reducing the size of the data. Some examples of parametric data reduction methods include regression models, log-linear models, and graphical models. For nonparametric methods, the data reduction process does not assume a specific parametric model for the data. Therefore, nonparametric methods are overall more flexible yet more challenging. Some typical examples of nonparametric data reduction methods which first transforms the original data into a compressed representation, then reconstructs the data in a later recovery process. Data compressed representation without any information loss. In general, the computational time of a data reduction process should not outweigh the amount of time saved by analyzing the reduced dataset.

In most cases, data reduction is a step before data-driven simulation for estimating aviation environmental impacts, on either the aircraft-level or the fleet-level. The data-driven simulation relies on aircraft flight data to ensure that the simulation result can closely reflect real-world operations. Nevertheless, conducting computation and evaluation for a massive amount of operations and models can be infeasible. Therefore, data reduction is necessary to extract a small amount of representative data and models for efficient yet accurate analysis. Overall, there have been three primary usages of data reduction in aviation environmental impact analysis: representative data, representative models, and representative operations. Representative data refers to a small subset of data points which can closely maintain certain characteristics of the population. For example, in probabilistic analysis and many common scenarios, the small subset should retain the same data distribution as the complete dataset. Reference ^[7] proposes a distributional data reduction method PREM which outperforms random sampling at very small sample sizes. PREM enables efficient simulation-based uncertainty propagation in the uncertainty quantification of aircraft fuel burn and emissions in real-world operations. Representative models involve both numerosity reduction and dimensionality reduction. The need for numerosity reduction is due to the fact that there exists a substantial amount of aircraft types to model, where each aircraft type is a unique combination of airframe and engine. Because building aircraft noise and performance (ANP) model for each aircraft type takes a long and rigorous process, Reference ^[8] selects a small proportion of representative aircraft models that can sufficiently cover the richness and complexity in the population for detailed modeling. References [9][10] select representative aircraft types for efficient fleet-level noise contour and emissions computation. On dimensionality reduction, Reference [11] conducts a feature selection study to find a reduced set of aircraft features which are most influential to different environmental impact metrics. Representative operations refer to the flight procedures, trajectories, or profiles that can be utilized to model aircraft fuel burn, emissions, and noise. References [9][12][13][14] apply clustering on large datasets to group flight trajectories and extract the most representative trajectories. Some works, such as [12][13], take a step forward to convert the representative flight profiles into parameterized forms. Reference [15] also applies probabilistic modeling on the representative mission profiles and account for uncertainty in the process. These representative information from real-world operations have made aviation environmental impact modeling closer to the reality in an efficient manner.

3. Efficient Computation

Since computer simulation/experiment became an indispensable part in contemporary engineering design optimization and systems analysis problems, computational efficiency has been a major concern in such processes due to two main reasons. First, high-fidelity simulation and analysis models are typically computationally expensive and time consuming. One common approach to tackle this challenge is the Surrogate-Based Analysis and Optimization (SBAO) ^[16] approach which relies on surrogate models to provide fast approximations of the sophisticated high-fidelity models. Second, the design optimization of a complex system inevitably involves the exploration of a broad design space. This usually translates to a large number of candidate design points and simulation runs, depending on the actual size of the design space and the number of design parameters. Some statistical and data-driven approaches can further reduce the number of simulation runs to improve the efficiency. Overall, these two facts can lead to excessive computational costs and prohibitive run times in engineering design and analysis process. Two typical examples of the computationally expensive simulations in the design and analysis of aerospace systems are Computational Fluid Dynamics (CFD) and multidisciplinary vehicle modeling.

Aviation environmental impact analysis can also be computationally expensive, because performing high-fidelity fuel burn, emissions, and noise analyses for a air transportation system is a massive task. Take aircraft noise modeling as an example, depending on factors such as the number of aircraft operations, the size of the region, the length of the time interval, and the fidelity level of the models, the current state-of-the-art noise modeling capabilities could require long setup and computational times for a single case study. A previous study ^[17] reported that running the high-fidelity Integrated Noise Model (INM) to perform airport-level noise study for a four-parallel-runway airport in crossflow takes between two days to two weeks to finish. Another example is the Aviation Environmental Design Tool (AEDT), a software system that models aircraft performance in space and time to estimate noise, fuel consumption, emissions, and air quality consequences ^[18]. A study ^[19] reported that on the AEDT, a national-level noise study comprising a moderate number of airports and flights could take several days to complete.

These representative works involving efficient computation can be classified into three generic groups. The first group employs surrogate models or reduced-order models (ROMs) to reduce the computational complexity of the complicated models and therefore reducing the computational time. In some literature, this is also referred to as "meta modeling". The authors of ^[20] construct a response surface model to approximate the computationally expensive Community Multiscale Air Quality (CMAQ) modeling system for fast evaluations of aviation's impacts on air quality. The authors of ^[19] apply ROM on AEDT's noise model to develop a rapid noise prediction capability. The second group builds rapid integrated analysis capabilities for fleet-level aviation environmental impact modeling. Such rapid fleet-level analysis capabilities could consist of elements such as simplified models, generic aircraft and operations (with connections to data reduction), and some pre-computed outcomes. The authors of ^[17] develop the airport noise grid integration method (ANGIM) which uses simplified methods and offline computational results for generic aircraft operations to enable rapid fleet-level noise modeling. The authors of ^[21] propose the GENERICA method which leverages methods such as classification algorithms, designs of experiments, surrogate models, and multi-criteria decision-making to identify better baseline models than the traditional representative-in-class vehicles,

also called "average generic vehicles", for more realistic approximation of fleet-level environmental impact results. The authors of ^[22] develop the Rapid Environmental impact on Airport Community Tradeoff (REACT) environment to conduct rapid tradeoff by modeling different noise mitigation strategies' noise exposure on the airport community. The third group is hybrid data-driven approaches for efficient modeling. The authors of ^[23] use performance and acoustic data from flight and wind tunnel tests to develop an efficient analytical model for helicopter Blade–Vortex Interaction (BVI) noise during maneuvering flight. The authors of ^[24] combine physics-based model and aircraft performance data to build an efficient and accurate "data-enhanced surrogate model" for aircraft fuel consumption. The authors of ^[25] develop Fuel Estimation in Air Transportation (FEAT), a rapid analysis framework, by using a high fidelity flight profile simulator and a reduced order fuel burn model. These efficient models have contributed to the aviation environmental analysis tool-suite that enable rapid assessment and evaluation, which is crucial especially for preliminary analysis.

4. Predictive Modeling

With the ultimate objective of making accurate predictions, predictive modeling is one of the most typical tasks of machine learning. In contrast to the specific machine learning or data mining techniques which uncover patterns in data, predictive modeling encompasses the entire process of developing a mathematical model in a way that we can understand and quantify the model's accuracy on predicting future, yet-to-be-seen data ^[26]. Steps such as data pre-processing, model tuning, performance measurement, and model selection are of critical importance in the predictive modeling process. Therefore, to a certain degree, predictive modeling is highly similar to the machine learning engineering process. Although the foremost objective of predictive modeling is to make accurate predictions, a secondary interest is to interpret the model and understand how it makes prediction. On certain problems, interpretation could be just as important, and this involves a tradeoff between accuracy and interpretation. Overall, a more accurate model is often associated with higher model complexity and lower interpretability.

Predictive models mainly include regression models and classification models, which predict continuous and categorical responses, respectively. Under each category, depending on whether a model is based on linear combinations of the predictors, different models can be further divided into linear models and nonlinear models. Although different models differ by the model form, number of parameters, and overall complexity, no predictive model is universally superior in every problem. Practitioners are encouraged to explore a diverse set of models for any given problem and identify the best predictive model ^[26]. A key foundation to the success of predictive signal exists in a dataset, even a naive model can capture some degree of predictive power. The domain knowledge applied to the modeling process is what distinguishes a great model from good models. In a serious decision making process, neither data-driven predictive model nor expert intuition will do better than a combination of both.

Most of these representative papers were published after year 2018. Although classification models can find applications in many data-driven analysis tasks in aerospace and transportation domains (e.g., flight risk identification), the predictive models for aviation environmental impact analysis are mostly regression models.

Some recent works have started to adopt advanced model architectures on more complex data forms to predict aircraft fuel burn, emissions, and noise. Among some of the earlier works, Reference ^[27] uses Gaussian Process Regression (GPR) and Probabilistic Graphical Model (PGM) to develop a wind forecasting model, which informs improved flight route planning to reduce environmental impact of aviation. Some advanced set-ups in statistics and machine learning are used to more accurately estimate aircraft fuel consumption. In a series works to improve aircraft fuel efficiency, Reference ^[28] first applies ensemble learning to improve the prediction of discretionary fuel and construct uncertainty intervals for the predictions. After that, Reference ^[29] utilizes quantile regression to estimate the Statistical Contingency Fuel (SCF) from a large fuel burn dataset from airline. The rest of the papers adopt different types of deep learning models. To minimize transport aircraft emissions and save fuel, Reference ^[30] applies neural network whose topology is optimized by genetic algorithm on flight data to predict fuel consumption. The authors of [31] use a type of feedforward neural network called covariance bidirectional extreme learning machine (CovB-ELM) to predict aircraft trajectory and the associated fuel consumption. There is also a significant trend which employs Recurrent Neural Network (RNN) to model sequences of data. The authors of ^[32] apply Long Short Term Memory (LSTM) neural network on Flight Data Monitoring (FDM) data records to estimate aircraft on-board parameters such as the fuel flow rate for enhancing the system's efficiency. The authors of [33] apply sequence-to-sequence LSTM on large radar and noise datasets to predict ground level aviation noise and evaluates the model using real-world noise measurements. The authors of ^[34] use a combination of LSTM and extreme gradient boosting (XGBoost) to predict short-term flight emissions within enroute airspace. The last two papers apply the more advanced physics-informed learning approaches which combine data-driven model with physical model to predict specific problems more effectively. The authors of [35] use physics-guided deep learning to model aircraft fuel burn. To outperform both the traditional physics-based models and the common supervised learning approaches, the authors: (1) guide the neural network with fuel flow dynamics equations, and (2) embed physical knowledge as extra losses in the model training process. The authors of ^[36] use physics-quided neural networks to predict propeller tonal noise with less experimental data, which can be difficult to collect. In some other works, even applying methods like ordinal regression and neural network can achieve satisfactory results on certain problems. Yet the more advanced more architectures and considerations have opened the door for a wider variety of problems.

5. Uncertainty Quantification

Uncertainties related to imprecise assumptions, natural variability, and the presence of unknowns is not only an unavoidable part of the real-world, but also a significant factor that could determine the success or failure of a decision or system. At the intersection of mathematics, statistics, and engineering, Uncertainty Quantification (UQ) is an interdisciplinary field that addresses the problems associated with incorporating real-world variability and probabilistic behavior into the design and analysis of complex systems. UQ provides uncertainty information about the Quantities of Interest (QoI) through characterizing, propagating, and managing uncertainties in a computational or real-world system. The high degree of complexity and uncertainty associated with aviation environmental impact analysis have driven practitioners towards the use of UQ in the modeling process.

Different sources of uncertainty can be generally categorized as either epistemic uncertainty or aleatory uncertainty. The former is caused by a lack of knowledge and is possible to be reduced by collecting more information. Aleatory uncertainty, on the other hand, results from the intrinsic randomness of nature. Therefore, it is beyond people's ability to reduce aleatory uncertainty through gathering additional information. Some previous research efforts ^{[37][38]} explore how to deal with both types of uncertainties. Common sources of uncertainty can be classified into four categories:

- **Inputs uncertainty**: The inputs of a model/system may have inherent uncertainty and substantial variation around a deterministic value.
- Model uncertainty: All models are "wrong" because they inevitably include assumptions, approximations, and errors and are therefore not exact representations of reality. Two aspects of uncertainties related to model are model-form uncertainty and uncertainty about parameters within the model.
- **Computational and numerical uncertainty**: Normally numerical errors from running simulations or solving mathematical models, including simplified equations, convergence error, truncation, etc.
- **Physical testing uncertainty**: A result of uncontrolled or unknown inputs, measurement errors, and limitations in the design and implementation of tests.

In aviation environmental impact analysis, UQ is generally used to: (1) understand the uncertainties inherent in a complex model or software systems, (2) predict system responses across uncertain inputs and quantify confidence intervals, (3) understand the key contributors to model output variations, and (4) inform researchers of directions for future model development and enhancement. A summary of representative papers in uncertainty quantification is given in **Table 5**. Several works, such as ^[39], encompass surrogate modeling in their UQ approach as well, because it is common to use simulation-based methods, such as MCS, for uncertainty propagation. This process learns the distributions of the nondeterministic outputs through a large number of experiment runs on the sophisticated analysis code and is impractical without surrogate model. Hence, surrogate model is a vital enabler to efficient UQ and design optimization. Some earlier works perform a complete UQ process on a complex aviation environmental model. Both [40][41] conduct a UO study on AEDT to better understand the uncertainties in AEDT estimations and identify priority aspects for future research and development. Also based on AEDT, Reference [42] conducts a sensitivity analysis for fleet-level environmental impacts to changes in operational uncertain factors for the optimization of flight operations to mitigate aviation environmental impacts. The authors of [43] perform rapid computation and UQ on the global fleet-wide simulation of aviation emissions for rapid and robust policy analysis. On the data-driven approach, Reference ^[44] uses Gaussian Process Regression (GPR) to guantify uncertainty in a data-driven 4D flight trajectory prediction problem and gain insights on uncertainty reduction-an important objective of UQ. Some works focus on novel UQ methodologies and make contributions in methodology development. Inspired by multidisciplinary analysis and optimization, Reference [45] proposes a decompositionbased approach to quantify uncertainty in multi-component systems and applies the method to perform uncertainty analysis and sensitivity analysis for the environmental impacts of new aircraft technologies and operations. When only limited data is available for UQ, Reference ^[46] develops a nonparametric approach to characterize and propagate uncertainty, which is more flexible and does not introduce unwarranted assumptions into the process. Some latest works initiate the trend of performing UQ on the environmental impact of future aircraft configurations. The authors of ^[47] perform UQ on the noise of a Hybrid Wing–Body (HWB) aircraft configuration at the noise certification locations. The authors of ^[48] perform system noise assessment and UQ for a conceptual supersonic aircraft and identifies factors that could significantly affect the concept's Landing and Takeoff noise (LTO) noise. Since robustness is a key consideration in the design and analysis of complex aerospace systems, UQ will continue to play a substantial role in the design and analyze of sustainable aviation system.

6. Pattern Discovery

Pattern discovery (or knowledge discovery) is a term commonly seen in data mining. A pattern generally refers to some useful information in the data that can guide action or decision-making. Some simple patterns include total, ratio, correlation, variation, etc. Examples of the (slightly) more complex patterns include emerging trend, receding signal, alternating behaviour, and spatiotemporal variation. Statistical and machine learning methods provide techniques to discover patterns from data. The following types of non-chaotic patterns can be found in data:

- **Descriptive patterns**: The identification of these patterns usually do not involve advanced algorithms. They are obtained through descriptive statistics or sometimes the direct results of data collection.
- Associative patterns: These patterns are mainly about co-occurring phenomena. A typical statement of associative pattern is: "If A happens, then B is also likely to happen".
- **Periodic patterns**: These patterns repeat themselves with a specific period, which can be found in time series data, sequence data, and spatiotemporal data.
- **Structural patterns**: These patterns are extracted summary information represented in terms of a structure that can be reasoned about. There are different structural forms such as graphs, trees, sets, clusters, etc.
- Abnormal patterns: A substantial divergence from normal behaviour is considered abnormal. These abnormalities could be signals of risk or opportunities for novel discoveries.

In aviation environmental impact analysis, the discovered patterns on aircraft fuel burn, emissions, and noise can provide insights for aviation and environmental analysts to make decisions and plans for mitigating aviation's environmental impact and achieving sustainable air traffic growth. Some patterns can be directly discovered by analyzing existing datasets on aviation environmental impact. Most of the time, however, the dataset for a target study is not available. For example, if a researcher wants to obtain the quantity and distribution of a certain emission type over a continent, it would be impractical to obtain such measurements via sensors and instruments. In such cases, data-driven flight simulation become a key enabler for computation and pattern discovery. Data-driven flight simulation utilizes real-world flight operations data, such as ADS-B, and "flies" the aircraft in an

computational environment. The aircraft furn burn, emissions, and noise are computed using aircraft performance models, such as Base of Aircraft Data (BADA), and noise models. Since the real-world flight data reflects how aircraft operate in time and space, with reliable performance and noise models it is possible to obtain decent approximations of the real-world situations. The estimations resulting from data-driven flight simulations are then ready to be analyzed by statistical and machine learning methods to discover useful patterns.

In one of the pioneering works, Reference ^[49] develops an aviation emission inventory and discovers the disparity of CO2 concentration in different parts of Australia. The authors of [50] use flight track data and fast noise approximation model to observe the variability in noise patterns on evolving airport runway configuration at Boston Logan International Airport (KBOS). The authors of ^[51] use ADS-B data and OpenAP emission models to obtain cruise-level flight emissions for different airlines, geographic regions, altitudes, and timeframe. The authors of [52] employ ADS-B and flight performance model to study aviation emissions at altitude and finds out that NOx and water vapour emissions concentrate around tropospheric altitudes only for long-range flights. The authors of [53] use similar approach to analyze fuel burn and emissions for a network of short-haul commuter flights in Europe. Through analyzing fuel burn and emissions as function of distance, altitude, city pair, the conclusion is that flight range is the most significant discriminator in emissions. The authors of [54][55] extend such simulations to globalscale. Together with a clustering step, Reference [54] studies the transport patterns and climate impacts of aviationemitted NOx and highlights the spatially and temporally heterogeneous nature of the NOx-O3 chemistry in different regions and seasons around the globe. The authors of ⁵⁵ estimate global emissions from aircraft operations between 2017–2020 and quantifies the impact from COVID-19. Deep learning is a powerful tool in finding aircraft emissions patterns in a more complex setting. The authors of [56] apply Convolutional Neural Network (CNN) on satellite images to detect aircraft contrails—a contributor of climate warming effect. The project estimates that contrails cover an average of 0.55% of the contiguous U.S. and discovers detailed patterns of contrail coverage. The above findings are key information for people to understand the status and patterns of regional and global aviation environmental impact. With continued advancements in data quantity, aircraft performance models, and analytical techniques, such data-driven approaches can make even better contributions to sustainable aviation.

7. Verification and Validation

Verification and validation (V&V) are evaluation procedures throughout the development phase to assess whether a system, product, or process meets the requirements and specifications that are initially set in the proposal. V&V are an integral part of the systems engineering processes to ensure the success of a project. Sometimes such procedures need to be executed by a disinterested third party and are referred to as Independent Verification and validation (IV&V). Verification is the procedure of comparing the solution to the requirements. Verification uses examination, demonstration, analysis, and testing to answer the query "are you building it correctly?". A verification procedure takes as input a system/product/process A and the requirements Q, and returns whether A is satisfactory (all behaviors of A meet Q) or unsatisfactory (at least one behavior of A

violates *Q*). Data-driven verification is a novel research area that combines numerical simulation with sensitivity analysis to provide bounds on how much the states of a system/product/process can change in a non-deterministic

setting ^[57]. In contrast, validation is the procedure of checking whether a system/product/process meets the needs of the user and other stakeholders. Validation involves review, demonstration, and testing to answer the query "are you building the correct thing?". Validation is of vital importance because the cost of fixing a user requirement error is very higher—usually much higher than fixing an implementation error.

V&V has become an essential part in the development of a complex model or software, especially for a field that is safety-critical in nature. Because of the complexities in modeling the environmental impact of aircraft and air transportation system, relevant computational tools are complex and multi-module systems as well. The computational models for aviation environmental impact could leverage data processing, flight simulation, aircraft performance models, emissions models, noise models, large aircraft databases, Geographic Information System (GIS), and extensive system databases cover airports, airspace, and fleet information for accurate modeling. Therefore, it is indispensable to conduct V&V in all levels of the system to make sure that the environment functions properly.

Enabled by data-driven simulation and statistical analysis, two common practices are seen in these works: (1) comparing the predictions between different models, and (2) comparing the model predictions with real-world measurements. The results of V&V can guide actions to further enhance the analysis capabilities. Among these works, Reference [58] performs a V&V study on AEDT's emission inventory and air quality modeling capability and investigates causes behind the deviation between AEDT and the legacy tool Emissions and Dispersion Modeling System (EDMS). Still on AEDT, Reference ^[59] provides a structured and repeatable framework for validating AEDT's noise model using detailed airline flight data records, weather data, and noise monitoring data from stations around airport. The authors of [60] compare thousands of the actual single flight noise exposure measurements with predictions from three noise models: AEDT, FLULA2, and sonAIR. To understand a helicopter noise prediction system's limitations, Reference ^[61] compares its Sound Exposure Level (SEL) noise contours with the acoustic flight test data for a range of flight conditions. The authors of [62] conduct a validation for an integrated aircraft environmental simulation software's acoustic and engine exhaust emissions modules using the microphone field measurements at Manchester airport for a range of aircraft types. The authors of [63] compare predictions from the "Dutch aircraft noise model" to measured values from the NOise MOnitoring System (NOMOS) around Amsterdam Airport Schiphol between 2012 and 2018 and observes how the model accuracy has changed overtime. The authors of ^[64] conduct a sensitivity analysis on semi-empirical noise models and compares the predictions to flyover measurements of A320, A330, and B777. The authors of [65] present a validation methodology for the noise impact of delayed deceleration approach, a new procedure, using ground-noise-monitor measurements and radar data for several aircraft types. Most of these studies confirm that the aviation environmental impact models can achieve satisfactory accuracy on their predictions. Some works also identify reasons behind the mismatch and modify the models accordingly to obtain better agreement between modeled and measured values.

8. Infrastructure and Tools

The research outcomes from all the previous themes have direct impacts on the methodology and/or actual findings of aviation environmental impact analysis. This theme devoted to infrastructure and tools is special because the relevant efforts may not lead to immediate breakthroughs on the more efficient and accurate models of aviation environmental impact. Instead, they lay the foundations for data-driven researches and make them happen. Undeniably, the efforts on building infrastructure and tools to collect, integrate, clean, and process data are an indispensable part of the ecosystem and have notably streamlined data-driven analysis in aviation.

Different types of works under this theme can be classified into: (1) hardware system, (2) data repository, and (3) data integration and pre-processing tools. Even these infrastructures and tools are built for aviation researchers and data scientists to perform more complex analyses, such as machine learning tasks, most of them are also capable of preliminary data analysis and data visualization. On the hardware design side, Reference ^[66] presents the system architecture, design, and capabilities of a modern hardware/software infrastructure called the Metroplex Overflight Noise Analysis (MONA). MONA is a system to measure, analyze, and archive the ground noise data from aircraft overflights for a variety of purposes, such as V&V of improved noise prediction methods. It also has a strong data visualization capability. The authors of ^[67] propose DV8—an interactive data visualization framework which provides visualized aviation-oriented insights for capacity planning, flight route prediction, and fuel consumption. Data repository is another crucial part of the infrastructure. Threaded Track [68] integrates radar trajectory data from a variety of surveillance sources to produce an optimal representation of an aircraft's end to end trajectory. Since its inception, Threaded Track has facilitated data-driven analyses for aviation safety and environmental impact. WRAP [69] is an open-source database which includes extracted full-flight aircraft performance parameters from large scale open ADS-B data. Apart from the aircraft performance parameters, WRAP also provides the first set of open parametric performance models for common aircraft types. Flight DNA ^[20] is one of the latest aviation data repositories. It is a common database with anonymized data on aviation components, systems, technologies, and operations. On the data pre-processing tools, there has been a significant trend to convert them into open-source libraries for popular programming languages such as Python and R. traffic ^[71] is a Python toolbox for pre-processing and analyzing aircraft trajectories data so that they are better prepared for statistical modeling and machine learning. pyModeS ^[72] is another open-source library in Python. The focus of pyModeS is to decode the Mode-S Comm-B replies and provide researchers broader access to accurate aircraft state updates that are transmitted via Enhanced Mode-S. openSkies ^[73] is the first R package for processing public air traffic data. It has an interface to resources in the OpenSky Network, standardized data structures, and functionalities to analyze and visualize data. In the future, continued development of infrastructures and tools for aviation data analytics is a key to promoting data-driven transformation for mitigating aviation environmental impact.

References

1. Waitz, I.; Townsend, J.; Cutcher-Gershenfeld, J.; Greitzer, E.; Kerrebrock, J. Aviation and the Environment, A National Vision Statement, Framework for Goals and Recommended Actions;

United States Congress: Washington, DC, USA, 2014.

- 2. FAA (Federal Aviation Administration). Aviation Emissions, Impacts and Mitigation: A Primer; Federal Aviation Administration: Washington, DC, USA, 2015.
- Lee, D.; Fahey, D.; Skowron, A.; Allen, M.; Burkhardt, U.; Chen, Q.; Doherty, S.; Freeman, S.; Forster, P.; Fuglestvedt, J.; et al. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. Atmos. Environ. 2021, 244, 117834.
- Brasseur, G.P.; Gupta, M.; Anderson, B.E.; Balasubramanian, S.; Barrett, S.; Duda, D.; Fleming, G.; Forster, P.M.; Fuglestvedt, J.; Gettelman, A.; et al. Impact of Aviation on Climate: FAA's Aviation Climate Change Research Initiative (ACCRI) Phase II. Bull. Am. Meteorol. Soc. 2016, 97, 561–583.
- 5. Basner, M.; Clark, C.; Hansell, A.; Hileman, J.; Janssen, S.; Shepherd, K.; Sparrow, V. Aviation Noise Impacts: State of the Science. Noise Health 2017, 19, 41–50.
- 6. Han, J.; Pei, J.; Kamber, M. Data Mining: Concepts and Techniques; Elsevier: Amsteradm, The Netherlands, 2011.
- 7. Gao, Z.; Puranik, T.G.; Mavris, D.N. Probabilistic REpresentatives Mining (PREM): A Clustering Method for Distributional Data Reduction. AIAA J. 2022, 60, 2580–2596.
- 8. Gao, Z.; Li, Y.; Puranik, T.G.; Mavris, D.N. Minimax and Multi-Criteria Selection of Representative Model Portfolios for Complex Systems Analysis. AIAA J. 2022, 60, 1505–1521.
- Jensen, L.; Thomas, J.; Brooks, C.; Brenner, M.; Hansman, R.J. Development of Rapid Fleet-Wide Environmental Assessment Capability. In Proceedings of the AIAA Modeling and Simulation Technologies Conference, London, UK, 23–27 September 2017; p. 3339.
- 10. Torija, A.J.; Self, R.H. Aircraft classification for efficient modelling of environmental noise impact of aviation. J. Air Transp. Manag. 2018, 67, 157–168.
- 11. Gao, Z.; Kampezidou, S.I.; Behere, A.; Puranik, T.G.; Rajaram, D.; Mavris, D.N. Multi-level aircraft feature representation and selection for aviation environmental impact analysis. Transp. Res. Part C Emerg. Technol. 2022, 143, 103824.
- Pagoni, I.; Psaraki-Kalouptsidi, V. Calculation of aircraft fuel consumption and CO2 emissions based on path profile estimation by clustering and registration. Transp. Res. Part D Transp. Environ. 2017, 54, 172–190.
- 13. Sun, J.; Ellerbroek, J.; Hoekstra, J. Flight extraction and phase identification for large automatic dependent surveillance–broadcast datasets. J. Aerosp. Inf. Syst. 2017, 14, 566–572.
- 14. Murca, M.C.R.; Hansman, R.J. Identification, characterization, and prediction of traffic flow patterns in multi-airport systems. IEEE Trans. Intell. Transp. Syst. 2018, 20, 1683–1696.

- 15. Kadyk, T.; Schenkendorf, R.; Hawner, S.; Yildiz, B.; Römer, U. Design of fuel cell systems for aviation: Representative mission profiles and sensitivity analyses. Front. Energy Res. 2019, 7, 35.
- 16. Queipo, N.V.; Haftka, R.T.; Shyy, W.; Goel, T.; Vaidyanathan, R.; Kevin Tucker, P. Surrogatebased analysis and optimization. Prog. Aerosp. Sci. 2005, 41, 1–28.
- 17. Bernardo, J.E.; Kirby, M.; Mavris, D. Development of a Rapid Fleet-Level Noise Computation Model. J. Aircr. 2015, 52, 721–733.
- Lee, C.; Thrasher, T.; Hwang, S.; Shumway, M.; Zubrow, A.; Hansen, A.; Koopmann, J.; Solman, G. Aviation Environmental Design Tool (AEDT) User Manual Version 3c; Technical Report; Federal Aviation Administration: Washington, DC, USA, 2020.
- 19. Kim, S.; Lim, D.; Lee, K. Reduced-order modeling applied to the aviation environmental design tool for rapid noise prediction. J. Aerosp. Eng. 2018, 31, 04018056.
- 20. Ashok, A.; Lee, I.H.; Arunachalam, S.; Waitz, I.A.; Yim, S.H.; Barrett, S.R. Development of a response surface model of aviation's air quality impacts in the United States. Atmos. Environ. 2013, 77, 445–452.
- 21. LeVine, M.J.; Bernardo, J.E.; Kirby, M.; Mavris, D.N. Average generic vehicle method for fleetlevel analysis of noise and emission tradeoffs. J. Aircr. 2018, 55, 929–946.
- 22. Monteiro, D.J.; Prem, S.; Kirby, M.; Mavris, D.N. React: A rapid environmental impact on airport community tradeoff environment. In Proceedings of the 2018 AIAA Aerospace Sciences Meeting, Kissimmee, FL, USA, 8 January 2018; p. 263.
- 23. Greenwood, E.; Schmitz, F.H.; Sickenberger, R.D. A semiempirical noise modeling method for helicopter maneuvering flight operations. J. Am. Helicopter Soc. 2015, 60, 1–13.
- 24. Yanto, J.; Liem, R.P. Aircraft fuel burn performance study: A data-enhanced modeling approach. Transp. Res. Part D Transp. Environ. 2018, 65, 574–595.
- Seymour, K.; Held, M.; Georges, G.; Boulouchos, K. Fuel Estimation in Air Transportation: Modeling global fuel consumption for commercial aviation. Transp. Res. Part D Transp. Environ. 2020, 88, 102528.
- 26. Kuhn, M.; Johnson, K. Applied Predictive Modeling; Springer: New York, NY, USA, 2013; Volume 26.
- Kapoor, A.; Horvitz, Z.; Laube, S.; Horvitz, E. Airplanes aloft as a sensor network for wind forecasting. In Proceedings of the IPSN-14 13th International Symposium on Information Processing in Sensor Networks, Berlin, Germany, 8–11 April 2014; IEEE: Piscatway, NJ, USA, 2014; pp. 25–33.
- 28. Kang, L.; Hansen, M. Improving airline fuel efficiency via fuel burn prediction and uncertainty estimation. Transp. Res. Part C Emerg. Technol. 2018, 97, 128–146.

- 29. Kang, L.; Hansen, M. Quantile Regression–Based Estimation of Dynamic Statistical Contingency Fuel. Transp. Sci. 2021, 55, 257–273.
- 30. Baklacioglu, T. Modeling the fuel flow-rate of transport aircraft during flight phases using genetic algorithm-optimized neural networks. Aerosp. Sci. Technol. 2016, 49, 52–62.
- Khan, W.A.; Ma, H.L.; Ouyang, X.; Mo, D.Y. Prediction of aircraft trajectory and the associated fuel consumption using covariance bidirectional extreme learning machines. Transp. Res. Part E Logist. Transp. Rev. 2021, 145, 102189.
- Jarry, G.; Delahaye, D.; Feron, E. Approach and landing aircraft on-board parameters estimation with lstm networks. In Proceedings of the 2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT), Singapore, 3–4 February 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
- Vela, A.E.; Oleyaei-Motlagh, Y. Ground level aviation noise prediction: A sequence to sequence modeling approach using LSTM recurrent neural networks. In Proceedings of the 2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC), San Antonio, TX, USA, 11–16 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–8.
- Wan, J.; Zhang, H.; Lyu, W.; Zhou, J. A Novel Combined Model for Short-Term Emission Prediction of Airspace Flights Based on Machine Learning: A Case Study of China. Sustainability 2022, 14, 4017.
- 35. Uzun, M.; Demirezen, M.U.; Inalhan, G. Physics Guided Deep Learning for Data-Driven Aircraft Fuel Consumption Modeling. Aerospace 2021, 8, 44.
- 36. Wiedemann, A.; Fuller, C.; Pascioni, K. Constructing a physics-guided machine learning neural network to predict tonal noise emitted by a propeller. In Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Seoul, Republic of Korea, 23–26 August 2022; Institute of Noise Control Engineering: Washington, DC, USA, 2022; Volume 264, pp. 151– 162.
- 37. Kiureghian, A.D.; Ditlevsen, O. Aleatory or Epistemic? Does It Matter? Struct. Saf. 2009, 31, 105– 112.
- 38. Guo, J.; Du, X. Sensitivity Analysis with Mixture of Epistemic and Aleatory Uncertainties. AIAA J. 2007, 45, 2337–2349.
- 39. Allaire, D.; Willcox, K. Surrogate modeling for uncertainty assessment with application to aviation environmental system models. AIAA J. 2010, 48, 1791–1803.
- 40. Allaire, D.; Noel, G.; Willcox, K.; Cointin, R. Uncertainty quantification of an aviation environmental toolsuite. Reliab. Eng. Syst. Saf. 2014, 126, 14–24.

- 41. Lim, D.; Li, Y.; LeVine, M.J.; Kirby, M.; Mavris, D.N. Parametric uncertainty quantification of aviation environmental design tool. In Proceedings of the 2018 Multidisciplinary Analysis and Optimization Conference, Atlanta, GA, USA, 25–29 June 2018; p. 3101.
- Behere, A.; Lim, D.; Li, Y.; Jin, Y.C.D.; Gao, Z.; Kirby, M.; Mavris, D.N. Sensitivity Analysis of Airport level Environmental Impacts to Aircraft thrust, weight, and departure procedures. In Proceedings of the AIAA Scitech 2020 Forum, Orlando, FL, USA, 23–27 January 2020; p. 1731.
- 43. Simone, N.W.; Stettler, M.E.; Barrett, S.R. Rapid estimation of global civil aviation emissions with uncertainty quantification. Transp. Res. Part D Transp. Environ. 2013, 25, 33–41.
- 44. Graas, R.; Sun, J.; Hoekstra, J. Quantifying accuracy and uncertainty in data-driven flight trajectory predictions with gaussian process regression. In Proceedings of the 11th SESAR Innovation Days, Online Conference, 7–9 December 2021.
- 45. Amaral, S.; Allaire, D.; Blanco, E.D.L.R.; Willcox, K.E. A decomposition-based uncertainty quantification approach for environmental impacts of aviation technology and operation. AI EDAM 2017, 31, 251–264.
- 46. Gao, Z.; Lim, D.; Schwartz, K.G.; Mavris, D.N. A nonparametric-based approach for the characterization and propagation of epistemic uncertainty due to small datasets. In Proceedings of the AIAA Scitech 2019 Forum, San Diego, CA, USA, 7 January 2019; p. 1490.
- 47. June, J.C.; Thomas, R.H.; Guo, Y. System Noise Prediction Uncertainty Quantification for a Hybrid Wing–Body Transport Concept. AIAA J. 2020, 58, 1157–1170.
- 48. Akatsuka, J.; Ishii, T. System Noise Assessment and Uncertainty Analysis of a Conceptual Supersonic Aircraft. Aerospace 2022, 9, 212.
- 49. Van Pham, V.; Tang, J.; Alam, S.; Lokan, C.; Abbass, H.A. Aviation emission inventory development and analysis. Environ. Model. Softw. 2010, 25, 1738–1753.
- 50. Jansson, M. Development of a Fast Method to Analyze Patterns in Airport Noise. Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2021.
- 51. Sun, J.; Dedoussi, I. Evaluation of aviation emissions and environmental costs in Europe using OpenSky and OpenAP. Eng. Proc. 2021, 13, 5.
- 52. Filippone, A.; Parkes, B.; Bojdo, N.; Kelly, T. Prediction of aircraft engine emissions using ADS-B flight data. Aeronaut. J. 2021, 125, 988–1012.
- 53. Filippone, A.; Parkes, B. Evaluation of commuter airplane emissions: A European case study. Transp. Res. Part D Transp. Environ. 2021, 98, 102979.
- 54. Maruhashi, J.; Grewe, V.; Frömming, C.; Jöckel, P.; Dedoussi, I.C. Transport Patterns of Global Aviation NOx and their Short-term O3 Radiative Forcing–A Machine Learning Approach. Atmos. Chem. Phys. Discuss. 2022, 22, 14253–14282.

- 55. Quadros, F.D.; Snellen, M.; Sun, J.; Dedoussi, I.C. Global Civil Aviation Emissions Estimates for 2017–2020 Using ADS-B Data. J. Aircr. 2022, 35, 1–11.
- 56. Kulik, L. Satellite-Based Detection of Contrails Using Deep Learning. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2019.
- 57. Fan, C. Formal Methods for Safe Autonomy: Data-Driven Verification, Synthesis, and Applications. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Champaign, IL, USA, 2019.
- Li, Y.; Lim, D.; Kirby, M.; Mavris, D.N.; Noel, G. Uncertainty Quantification Analysis of the Aviation Environmental Design Tool in Emission Inventory and Air Quality Modeling. In Proceedings of the 2018 Aviation Technology, Integration, and Operations Conference, Reston, VA, USA, 25–29 June 2018; p. 3050.
- 59. Gabrielian, A.B.; Puranik, T.G.; Bendarkar, M.V.; Kirby, M.; Mavris, D.; Monteiro, D. Noise Model Validation using Real World Operations Data. In Proceedings of the AIAA Aviation 2021 Forum, Online Conference, 2–6 August 2021; p. 2136.
- 60. Meister, J.; Schalcher, S.; Wunderli, J.M.; Jäger, D.; Zellmann, C.; Schäffer, B. Comparison of the aircraft noise calculation programs sonAIR, FLULA2 and AEDT with noise measurements of single flights. Aerospace 2021, 8, 388.
- Botre, M.; Brentner, K.S.; Horn, J.F.; Wachspress, D. Validation of helicopter noise prediction system with flight data. In Proceedings of the Vertical Flight Society 75th Annual Forum & Technology Display, Held, Ukraine, 13–16 May 2019; Volume 13.
- 62. Filippone, A.; Zhang, M.; Bojdo, N. Validation of an integrated simulation model for aircraft noise and engine emissions. Aerosp. Sci. Technol. 2019, 89, 370–381.
- 63. Simons, D.G.; Besnea, I.; Mohammadloo, T.H.; Melkert, J.A.; Snellen, M. Comparative assessment of measured and modelled aircraft noise around Amsterdam Airport Schiphol. Transp. Res. Part D Transp. Environ. 2022, 105, 103216.
- 64. Vieira, A.; von den Hoff, B.; Snellen, M.; Simons, D.G. Comparison of Semi-Empirical Noise Models with Flyover Measurements of Operating Aircraft. J. Aircr. 2022, 1–14.
- 65. Huynh, J.L.; Mahseredjian, A.; John Hansman, R. Delayed Deceleration Approach Noise Impact and Modeling Validation. J. Aircr. 2022, 59, 1–13.
- 66. Jackson, D.C.; Rindfleisch, T.C.; Alonso, J.J. A System for Measurement and Analysis of Aircraft Noise Impacts. Eng. Proc. 2021, 13, 6.
- 67. Omidvar-Tehrani, B.; Nandi, A.; Meyer, N.; Flanagan, D.; Young, S. Dv8: Interactive analysis of aviation data. In Proceedings of the 2017 IEEE 33rd International Conference on Data

Engineering (ICDE), San Diego, CA, USA, 19–22 April 2017; IEEE: Piscatway, NJ, USA, 2017; pp. 1411–1412.

- 68. Eckstein, A.; Kurcz, C.; Silva, M. Threaded Track: Geospatial Data Fusion for Aircraft Flight Trajectories; Technical Report; MITRE Corporation: McLean, VA, USA, 2012.
- 69. Sun, J.; Ellerbroek, J.; Hoekstra, J.M. WRAP: An open-source kinematic aircraft performance model. Transp. Res. Part C Emerg. Technol. 2019, 98, 118–138.
- 70. NREL (National Renewable Energy Laboratory). Flight DNA: An Anonymized Aviation Data Tool and Repository; NREL: Golden, CO, USA, 2022.
- 71. Olive, X. Traffic, a toolbox for processing and analysing air traffic data. J. Open Source Softw. 2019, 4, 1518.
- 72. Sun, J.; Vû, H.; Ellerbroek, J.; Hoekstra, J.M. pymodes: Decoding mode-s surveillance data for open air transportation research. IEEE Trans. Intell. Transp. Syst. 2019, 21, 2777–2786.
- 73. Ayala, R.; Ayala, D.; Vidal, L.S.; Ruiz, D. openSkies-Integration of Aviation Data into the R Ecosystem. R J. 2021, 13, 590–599.

Retrieved from https://encyclopedia.pub/entry/history/show/92532