Forschungsbericht 2025-15

Enhancing Techno-Economic Assessments in Aeronautic Product Development with Systematic Uncertainty Management

Ahmad Ali Pohya

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Instandhaltung und Modifikation
Hamburg-Finkenwerder

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Enhancing Techno-Economic Assessments in Aeronautic Product Development with Systematic Uncertainty Management

---

Verbesserung von techno-ökonomischen Bewertungen im Bereich der Produktentwicklung der Luftfahrt mittels systematischem Unsicherheitenmanagement

Von der Fakultät für Maschinenwesen der Rheinisch-Westfälischen Technischen Hochschule Aachen zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften genehmigte Dissertation

vorgelegt von

Ahmad Ali Pohya

Berichter: Univ.-Prof. Dr.-Ing. Eike Stumpf
Prof. Dimitri Mavris, Ph.D.

Tag der mündlichen Prüfung: 26. Februar 2025

Forschungsbericht 2025-15

Enhancing Techno-Economic Assessments in Aeronautic Product Development with Systematic Uncertainty Management

Ahmad Ali Pohya

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Instandhaltung und Modifikation
Hamburg-Finkenwerder

230 Seiten
91 Bilder
25 Tabellen
302 Literaturstellen

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Herausgeber:

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Wissenschaftliche Information
Linder Höhe
D-51147 Köln

ISSN 1434-8454
ISRN DLR-FB-2025-15
Erscheinungsjahr 2025
DOI: 10.57676/ybp5-jp22

Unsicherheitenmanagement, Unsicherheitenquantifizierung, Lebenszyklusmodellierung, Hybridlaminarhaltung

Ahmad Ali POHYA
DLR, Institut für Instandhaltung und Modifikation, Hamburg-Finkenwerder

Verbesserung von techno-ökonomischen Bewertungen im Bereich der Produktentwicklung der Luftfahrt mittels systematischem Unsicherheitenmanagement
Rheinisch-Westfälische Technische Hochschule Aachen

Diese Dissertation untersuchte die Steigerung der Transparenz und Reproduzierbarkeit von techno-ökonomischen Bewertungen (TEAs) in der Entwicklung von Luftfahrtprodukten unter Einbeziehung von Parameterunsicherheiten. Das Hauptziel war es, identifizierte Barrieren bei der Einführung einer systematischen Unsicherheitsmanagementmethodik zu überwinden. Diese umfassten Methoden zur Separation von relevanten und vernachlässigbaren Unsicherheiten, zur Anwendung der Dempster-Shafer-Theorie der Evidenz (DSTE) unter Datenknappheit sowie zur kombinierten Berücksichtigung epistemischer (wissensbasierter) und aleatorischer (variabilitätsbasierter) Unsicherheiten. Durch die Verknüpfung dieser Barrieren mit systematischen und vergleichenden Analysen bieten die Erkenntnisse dieser Dissertation einen robusten Rahmen für das effektive Management von Unsicherheiten in TEAs, fördern das Feld der Entwicklung von innovativen Luftfahrtprodukten und verbessern Entscheidungsprozesse unter Unsicherheit.

Zur Veranschaulichung der entwickelten Unsicherheitsmanagementmethodik wurde eine wiederkehrende Fallstudie zur lebenszyklusbasierten TEA der Hybridlaminarhaltung (HLFC) herangezogen, für die auf Informationen aus zwei Europäischen Projekten zurückgegriffen werden konnte. Diese Fallstudie diente als realistisches und interdisziplinäres Beispiel, um die Quantifizierung von Eingangs- und Ausgangsunsicherheiten sowie die Anwendung weiterer UQ Methoden dieser Arbeit zu demonstrieren.

Ein wesentlicher Beitrag dieser Dissertation war die Untersuchung der Stärken und Schwächen verschiedener Techniken der Globalen Sensitivitätsanalyse (GSA), welche die individuelle Kritikalität von Parameterunsicherheiten quantifizieren. Im Gegensatz zu konventionellen Ansätzen, die GSA-Methoden häufig ohne klare Kriterien auswählen, bewertete diese Forschung systematisch deren Fähigkeiten, Interpretierbarkeit und Recheneffizienz. Die dabei identifizierten und teils signifikanten Unterschiede unterstreichen die Notwendigkeit einer informierten und auf den spezifischen Kontext der Bewertung abgestimmten Auswahl von GSA Techniken.

Zusätzlich wurde das Python-Paket dste entwickelt, um den Bedarf an benutzerfreundlichen Programmierumgebungen zur Behandlung von DSTE-basierten UQs zu adressieren. Zugehörige Analysen demonstrierten die Fähigkeiten des Pakets und diskutierten den Einsatz von DSTE mit Hilfe von systematischen Experteninterviews sowie theorie-spezifischen UQ Metriken. Darüber hinaus wurden die damit verbundenen Interpretationsschwierigkeiten, insbesondere im Hinblick auf die Adressaten der TEA, sowie die Herausforderungen in Bezug auf die Recheneffizienz untersucht.

Die Forschung untersuchte auch Methoden zur Kombination epistemischer und aleatorischer Unsicherheiten und schlug einen neuartigen Ansatz vor, der DSTE-basierte und probabilistische UQ Ansätze mittels verschachtelten Monte-Carlo-Simulationen kombiniert. Dieser Ansatz verbessert die Interpretierbarkeit und rechnerische Effizienz im Vergleich zu einem rein evidenztheoretischen Ansatz und bietet eine nuancierte Darstellung von Unsicherheiten. Entscheidungsträger profitieren von klareren Einblicken durch verständliche Visualisierung und einfache Interpretation, während Nutzer maßgeschneiderte Empfehlungen aufgrund der deutlichen Trennung von epistemischen und aleatorischen Effekten ableiten können. Darüber hinaus bietet dieser Ansatz Wiederholbarkeit, sodass UQ während des gesamten Produktentwicklungsprozesses konsequent angewendet und wiederholt werden kann, wenn neue Informationen zur Verfügung stehen.

Uncertainty management, uncertainty quantification, lifecycle simulation, hybrid laminar flow control

(Published in English)

Ahmad Ali POHYA
German Aerospace Center (DLR), Institute of Maintenance, Repair and Overhaul, Hamburg-Finkenwerder

Enhancing Techno-Economic Assessments in Aeronautic Product Development with Systematic Uncertainty Management

RWTH Aachen University

This thesis investigated the enhancement of transparency and reproducibility in technoeconomic assessments (TEAs) for aeronautical product developments when input parameter uncertainties are present. The primary objective was to overcome identified barriers in the adoption of a systematic uncertainty management methodology. These included methods for the separation of relevant and negligible uncertainties, the application of Dempster-Shafer Theory of Evidence (DSTE) under data scarcity, as well as the combination of epistemic (knowledge-based) and aleatory (variability-based) uncertainties. By linking these barriers with systematic and comparative analyses, the findings of this dissertation provide a robust framework for effective uncertainty management in TEAs, promote the field of innovative aeronautic product development, and improve decision-making processes under uncertainty.

To illustrate the developed uncertainty management methodology, a recurring case study on the lifecycle-based TEA of Hybrid Laminar Flow Control (HLFC) was utilized, drawing on information from two European projects. This case study served as a realistic and interdisciplinary example to demonstrate the quantification of input and output uncertainties, as well as other UQ methods addressed in this thesis.

A significant contribution of this dissertation was the investigation of the strengths and weaknesses of various Global Sensitivity Analysis (GSA) techniques, which quantify the individual criticality of parameter uncertainties. Unlike conventional approaches that often select GSA methods without clear criteria, this research systematically assessed their capabilities, interpretability, and computational efficiency. The identified and partially significant differences underscore the necessity for an informed and context-specific selection of GSA techniques.

Additionally, the Python package dste was developed to address the need for user-friendly programming toolboxes for handling DSTE-based UQ. Related analyses demonstrated the capabilities of the package and discussed the application of DSTE through systematic expert interviews and theory-specific UQ metrics. Furthermore, the associated interpretation difficulties, particularly concerning the recipients of the TEA, and the challenges related to computational efficiency were examined.

The research also explored methods for combining epistemic and aleatory uncertainties and proposed a novel approach that integrates DSTE-based and probabilistic UQ approaches using nested Monte Carlo simulations. This approach enhances interpretability and computational efficiency compared to a purely evidence-theoretic approach and provides a nuanced representation of uncertainties. Decision-makers benefit from clearer insights through understandable visualization and straightforward interpretation, while users can derive tailored recommendations due to the clear separation of epistemic and aleatory effects. Additionally, this approach offers repeatability, allowing UQ to be consistently applied and repeated throughout the product development process as new information becomes available.

Abstract

This thesis investigated the enhancement of transparency and reproducibility in techno-economic assessments (TEAs) for aeronautical product developments when input parameter uncertainties are present. The primary objective was to overcome identified barriers in the adoption of a systematic uncertainty management methodology. These included methods for the separation of relevant and negligible uncertainties, the application of Dempster-Shafer Theory of Evidence (DSTE) under data scarcity, as well as the combination of epistemic (knowledge-based) and aleatory (variability-based) uncertainties. By linking these barriers with systematic and comparative analyses, the findings of this dissertation provide a robust framework for effective uncertainty management in TEAs, promote the field of innovative aeronautic product development, and improve decision-making processes under uncertainty.

To illustrate the developed uncertainty management methodology, a recurring case study on the lifecycle-based TEA of Hybrid Laminar Flow Control (HLFC) was utilized, drawing on information from two European projects. This case study served as a realistic and interdisciplinary example to demonstrate the quantification of input and output uncertainties, as well as other UQ methods addressed in this thesis.

A significant contribution of this dissertation was the investigation of the strengths and weaknesses of various Global Sensitivity Analysis (GSA) techniques, which quantify the individual criticality of parameter uncertainties. Unlike conventional approaches that often select GSA methods without clear criteria, this research systematically assessed their capabilities, interpretability, and computational efficiency. The identified and partially significant differences underscore the necessity for an informed and context-specific selection of GSA techniques.

Additionally, the Python package `dste` was developed to address the need for user-friendly programming toolboxes for handling DSTE-based UQ. Related analyses demonstrated the capabilities of the package and discussed the application of DSTE through systematic expert interviews and theory-specific UQ metrics. Furthermore, the associated interpretation difficulties, particularly concerning the recipients of the TEA, and the challenges related to computational efficiency were examined.

The research also explored methods for combining epistemic and aleatory uncertainties and proposed a novel approach that integrates DSTE-based and probabilistic UQ approaches using nested Monte Carlo simulations. This approach enhances interpretability and computational efficiency compared to a purely evidence-theoretic approach and provides a nuanced representation of uncertainties. Decision-makers benefit from clearer insights through understandable visualization and straightforward interpretation, while users can derive tailored recommendations due to the clear separation of epistemic and aleatory effects. Additionally, this approach offers repeatability, allowing UQ to be consistently applied and repeated throughout the product development process as new information becomes available.

Kurzfassung

Diese Dissertation untersuchte die Steigerung der Transparenz und Reproduzierbarkeit von techno-ökonomischen Bewertungen (TEAs) in der Entwicklung von Luftfahrtprodukten unter Einbeziehung von Parameterunsicherheiten. Das Hauptziel war es, identifizierte Barrieren bei der Einführung einer systematischen Unsicherheitsmanagementmethodik zu überwinden. Diese umfassten Methoden zur Separation von relevanten und vernachlässigbaren Unsicherheiten, zur Anwendung der Dempster-Shafer-Theorie der Evidenz (DSTE) unter Datenknappheit sowie zur kombinierten Berücksichtigung epistemischer (wissensbasierter) und aleatorischer (variabilitätsbasierter) Unsicherheiten. Durch die Verknüpfung dieser Barrieren mit systematischen und vergleichenden Analysen bieten die Erkenntnisse dieser Dissertation einen robusten Rahmen für das effektive Management von Unsicherheiten in TEAs, fördern das Feld der Entwicklung von innovativen Luftfahrtprodukten und verbessern Entscheidungsprozesse unter Unsicherheit.

Zur Veranschaulichung der entwickelten Unsicherheitsmanagementmethodik wurde eine wiederkehrende Fallstudie zur lebenszyklusbasierten TEA der Hybridlaminarhaltung (HLFC) herangezogen, für die auf Informationen aus zwei Europäischen Projekten zurückgegriffen werden konnte. Diese Fallstudie diente als realistisches und interdisziplinäres Beispiel, um die Quantifizierung von Eingangs- und Ausgangsunsicherheiten sowie die Anwendung weiterer UQ-Methoden dieser Arbeit zu demonstrieren.

Ein wesentlicher Beitrag dieser Dissertation war die Untersuchung der Stärken und Schwächen verschiedener Techniken der Globalen Sensitivitätsanalyse (GSA), welche die individuelle Kritikalität von Parameterunsicherheiten quantifizieren. Im Gegensatz zu konventionellen Ansätzen, die GSA-Methoden häufig ohne klare Kriterien auswählen, bewertete diese Forschung systematisch deren Fähigkeiten, Interpretierbarkeit und Recheneffizienz. Die dabei identifizierten und teils signifikanten Unterschiede unterstreichen die Notwendigkeit einer informierten und auf den spezifischen Kontext der Bewertung abgestimmten Auswahl von GSA-Techniken.

Zusätzlich wurde das Python-Paket `dste` entwickelt, um den Bedarf an benutzerfreundlichen Programmierumgebungen zur Behandlung von DSTE-basierten UQs zu adressieren. Zugehörige Analysen demonstrierten die Fähigkeiten des Pakets und diskutierten den Einsatz von DSTE mit Hilfe von systematischen Experteninterviews sowie theorie-spezifischen UQ-Metriken. Darüber hinaus wurden die damit verbundenen Interpretationsschwierigkeiten, insbesondere im Hinblick auf die Adressaten der TEA, sowie die Herausforderungen in Bezug auf die Recheneffizienz untersucht.

Die Forschung untersuchte auch Methoden zur Kombination epistemischer und aleatorischer Unsicherheiten und schlug einen neuartigen Ansatz vor, der DSTE-basierte und probabilistische UQ Ansätze mittels verschachtelten Monte-Carlo-Simulationen kombiniert. Dieser Ansatz verbessert die Interpretierbarkeit und rechnerische Effizienz im Vergleich zu einem rein evidenztheoretischen Ansatz und bietet eine nuancierte Darstellung von Unsicherheiten. Entscheidungsträger profitieren von klareren Einblicken durch verständliche Visualisierung und einfache Interpretation, während Nutzer maßgeschneiderte Empfehlungen aufgrund der deutlichen Trennung von epistemischen und aleatorischen Effekten ableiten können. Darüber hinaus bietet dieser Ansatz Wiederholbarkeit, sodass UQ während des gesamten Produktentwicklungsprozesses konsequent angewendet und wiederholt werden kann, wenn neue Informationen zur Verfügung stehen.

Foreword

The research documented in this dissertation is complemented by a number of scientific papers published in scientific journals and presented at national and international conferences. These publications played a crucial role in shaping and developing the chapters of this dissertation. The following overview aims to provide transparency about which aspects of this thesis were influenced by these publications or which chapters served as the foundation for them:

Publication	**Related Chapters**
A. A. Pohya et al. "Comparison of Direct Operating Cost and Life Cycle Cost-benefit Methods in Aircraft Technology Assessment." In: *2018 AIAA Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics, Jan. 2018. DOI: 10.2514/6.2018-0282	Chapter 2 • Overview of TEA practices in aeronautics. • Scope and applicability of DOC methods.
A. A. Pohya et al. "A Modular Framework for the Life Cycle Based Evaluation of Aircraft Technologies, Maintenance Strategies, and Operational Decision Making Using Discrete Event Simulation." In: *Aerospace* 8.7 (July 2021), p. 187. DOI: 10.3390/aerospace8070187	Chapter 3 • Development of the modular lifecycle simulation framework LYFE.
A. A. Pohya et al. "Cloud Encounter Impact on Operational and Economical Effectiveness of Hybrid-laminar-flow-control Aircraft." In: *Journal of Aircraft* 56.4 (July 2019), pp. 1513–1523. DOI: 10.2514/1.c035205	Chapters 3 and 4 • Quantification of loss of laminar efficacy due to cloud encounter at high altitudes. • Lifecycle-based TEA of HLFC under realistic boundary conditions.
A. A. Pohya et al. "An Impact Assessment of Degrading Elements on the Overall Benefit of Aircraft with Hybrid Laminar Flow Control." In: *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics, Jan. 2019. DOI: 10.2514/6.2019-1589	Chapters 3 and 4 • Lifecycle-based TEA of HLFC with simultaneous consideration of benefit-degrading aspects. • Uncertainty quantification using local sensitivity analysis.
A. A. Pohya et al. "Introducing Variance-based Global Sensitivity Analysis for Uncertainty Enabled Operational and Economic Aircraft Technology Assessment." In: *Aerospace Science and Technology* 122 (Feb. 2022), p. 107441. ISSN: 1270-9638. DOI: 10.1016/j.ast.2022.107441	Chapter 4 • Comparison of various GSA techniques. • Application on HLFC use case.

J. Ramm et al. "Uncertainty quantification in hydrogen tank exchange: Estimating maintenance costs for new aircraft concepts." In: *International Journal of Hydrogen Energy* 68 (May 2024), pp. 159–169. ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2024.04.157

Chapter 5

- Expert elicitation and UQ with evidence theory.
- Combination of SCSA method for GSA with evidence theory.

Contents

List of Figures

List of Tables

List of Codes

List of Abbreviations

1 Introduction

"It is trivial to note that the future is uncertain. It is, however, far from trivial to analyze that uncertainty."

Mäler et al., 2005 [154, p. 574]

For profit-driven businesses, the assessment of a product's performance is a crucial step in the development process. Performed periodically at the end of each design phase, Integrated Product Assessments (IPAs)[1] are used to reduce the number of potential candidates to be developed in more detail at the next stage [277, p. 20]. This is also referred to as "*survival of the fittest*" [108, 169] and is illustrated in Fig. 1.1. The earlier stages are typically characterized by a lack of information about the final product design and the predicted market demand. Simultaneously, the decisions made in these earlier stages define the majority of the program's cost, as Fig. 1.2 shows. While the amount and quality of information in the later phases of the design process is much higher, the leverage of improving economic performance is much lower, and, depending on the sector, a significant amount of money has already been spent [77, p. 13f].

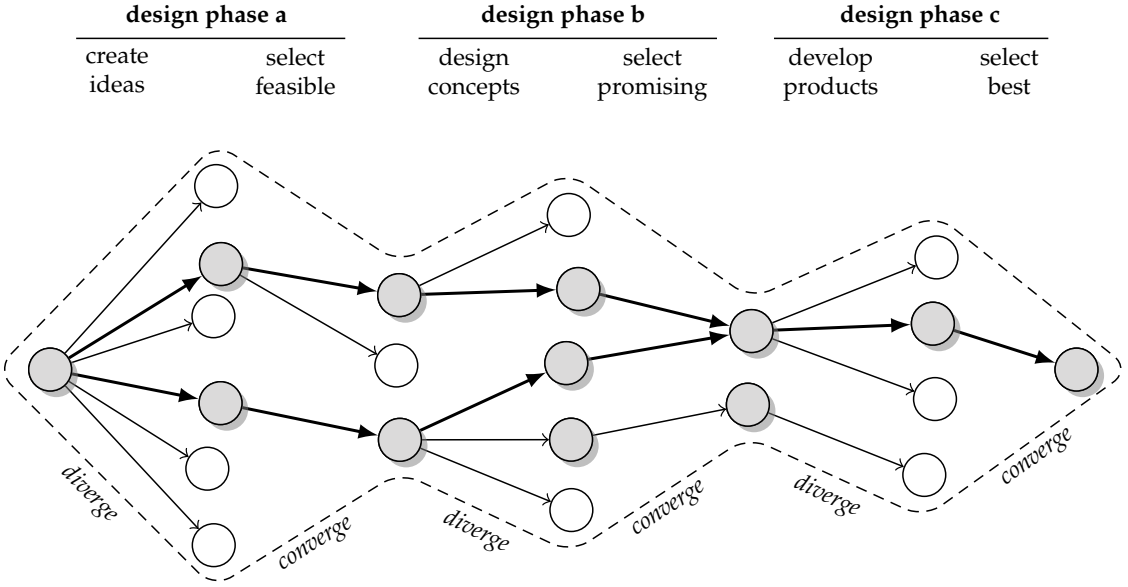


Figure 1.1 Alternating divergence and convergence during product development.

The aeronautics sector, for instance, is facing an ever-increasing economic pressure while operating in a highly complex and competitive environment [46, 65]. Strategic decisions for or against a product typically involve immense investment cost [187, p. 33], often offer low profit margins for operators IATA [113, p. 3], and can impact businesses for decades to come. Therefore, economic assessments of aircraft concepts, operational strategies, or technologies, hereafter referred to as products, have to be thorough, objective, and trustworthy in order to be accepted [153, 301]. As these assessments predict the product's performance in the

[1] The term *integrative* in IPA describes the approach where tools and inputs of various disciplines and domains are combined to estimate the product's performance. It is an essential part of the engineering field of integrated product development.

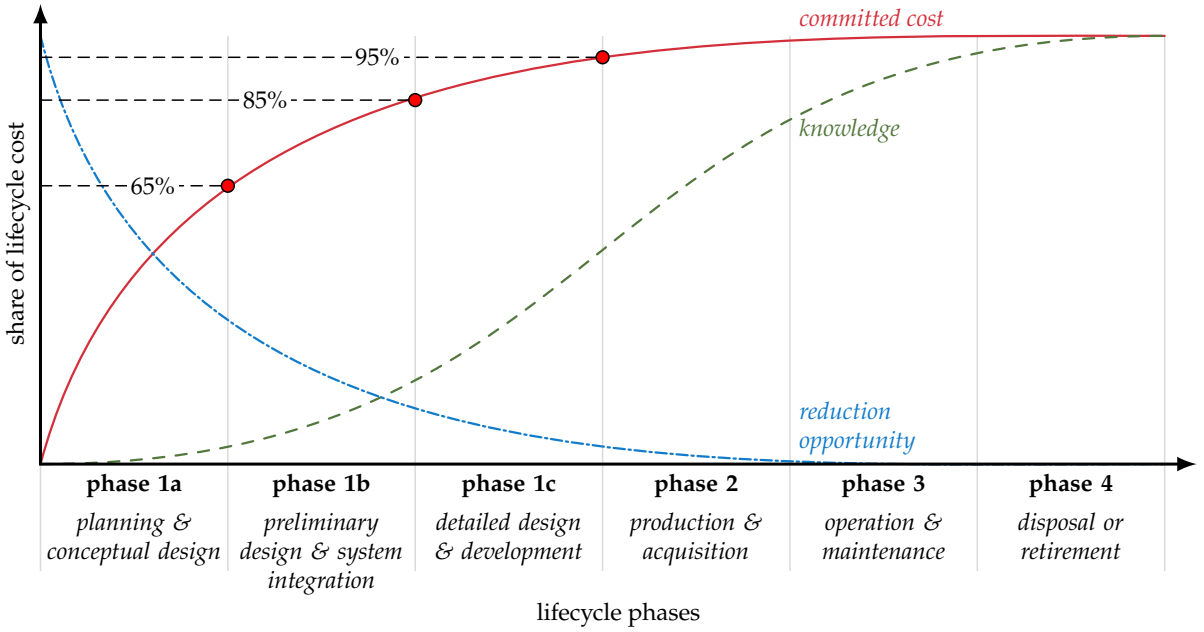


Figure 1.2 Committed lifecycle cost, reduction opportunity, and knowledge over time [77, p.13].

future, the results are prone to a variety of uncertainties, especially when considering the long lifetimes of aircraft. However, there is little to no guidance regarding a systematic inclusion of uncertainties [243, p. 1]. This can lead to a lack of adequate consideration for uncertainty, which in turn negatively impacts transparency [86, p. 435], undermines trust in analyses [235, p. 484], and significantly contributes to the recent reproducibility crisis in mathematical and computational modeling [213, p. 7-8]. Furthermore, it impedes businesses and investors from understanding the true potential of a product [260, pp. 113-115], which is one of the key reasons why two out of three products developed globally never make it from a research project to commercialization[2] [2, p. 2]. Therefore, a proper inclusion of uncertainties can help to bridge the so-called "*valley of death*"[3] [300, p. 90], illustrated in Fig. 1.3, and spur the innovation that the aerospace sector is in dire need of in order to meet both economic and environmental challenges [75, p. 8-9].

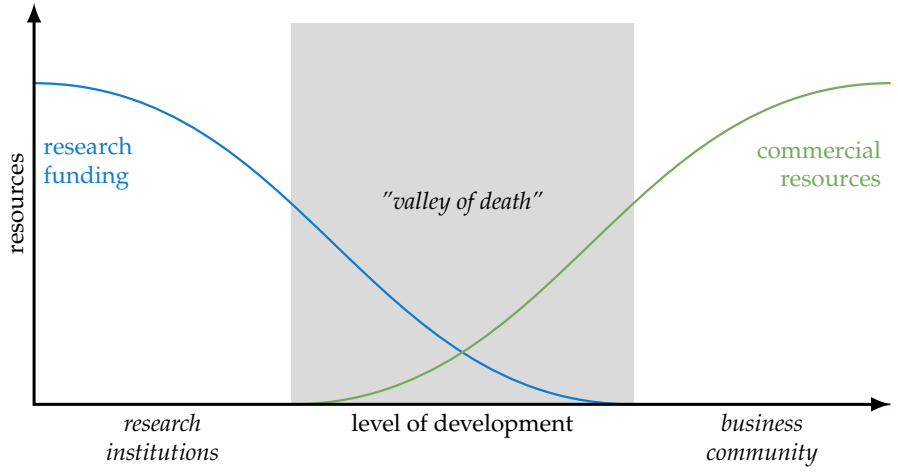


Figure 1.3 The valley of death bridging the gap between research and commercial application [24, p. 371]

[2] This ratio can vary significantly and may exceed 90 %, depending on the phase in the development cycle at which a product idea or concept begins to be recognized as such [257, p. 17].

[3] A phrase coined by Marczewski [156] describing the gap between the stage of research technologies and market entry.

While a variety of sources and types of uncertainties exist[4] and potentially affect the IPA results, the work in this thesis is devoted to uncertainties in the input variables and model parameters rather than uncertainties regarding the model itself or its technical implementation. Due to the recurring nature of IPA in the product development process, it is argued that the input uncertainties are more likely to change during the product development process, e.g., when new information arise, whereas the underlying model and its structure are more likely to remain as is for fair comparison purposes. To facilitate the management of these input uncertainties in aeronautic IPA, a systematic approach is proposed, in which said uncertainties are identified, classified, quantified, combined, propagated, and interpreted in a repeatable manner, as illustrated in Fig. 1.4. This methodology is intended to be used in the intermediate stages of the Integrated Product Development Process (IPDP), where concept or preliminary designs are available and uncertainties about the product and its intended environment are known and quantifiable to some degree.

The methodology begins with defining a model for the assessment, the scope of which should be delineated in consultation with the customer or recipient. This is followed by the identification of uncertainties, where assumptions are carefully reviewed in order to select, classify, and organize uncertain parameters for reference and overview purposes. In this step, the source, nature, and extent of these are distinguished, which is commonly referred to as uncertainty classification. The third step involves the quantification of input uncertainties. This is achieved by utilizing literature, databases, and - where needed - expert elicitation to model the selected input uncertainties, taking into account their classification and the availability of data. Subsequently, the propagation of uncertainty is carried out. A setup for repeated model execution is defined, either with the original model or a surrogate. The fifth step involves the quantification of output uncertainty. It includes an analysis of the statistical moments and a calculation of input parameter criticality to guide future uncertainty reduction efforts. This leads to the sixth step, where sources are revised and uncertainty is reduced wherever possible and suitable. Steps three to six are repeated until the desired accuracy is met or all reduction options are exhausted, marked by the "*minor iteration*" loop. The final step of the methodology is the expression of uncertainty and addresses the recipients of

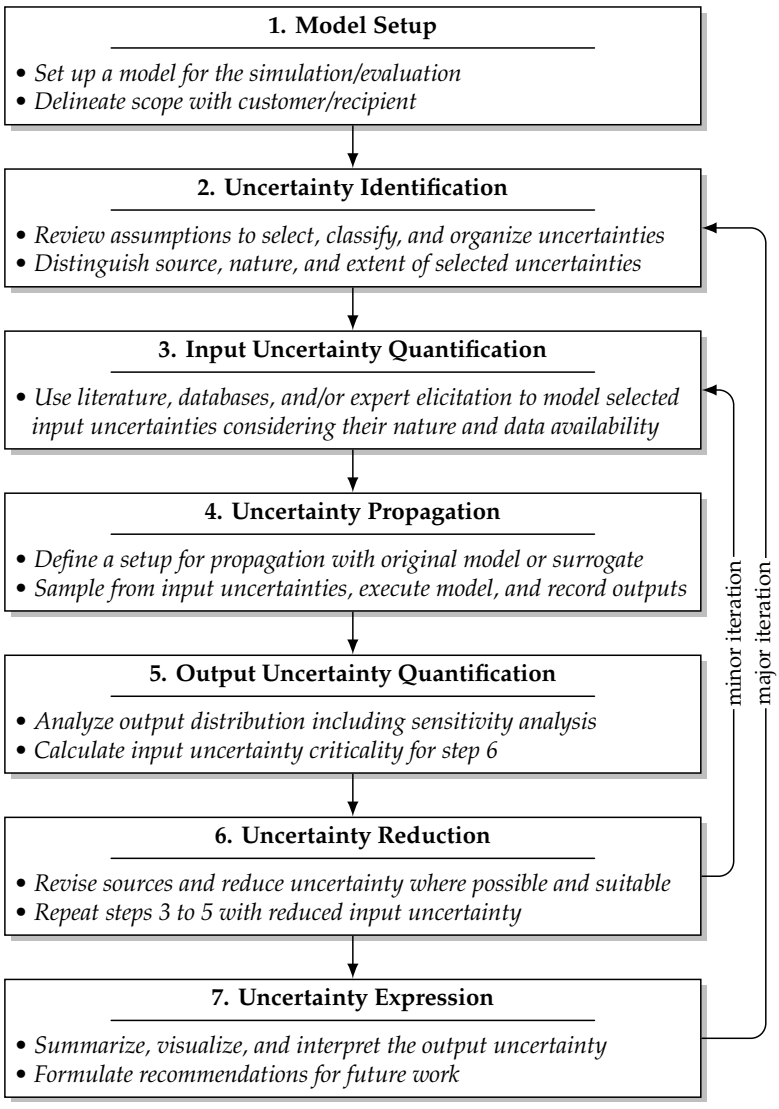


Figure 1.4 Proposed uncertainty management methodology.

[4]See Section 2.2.1 for further details on definitions and distinctions of uncertainty.

the study, who are likely not experts in Uncertainty Quantification (UQ). Therefore, the output uncertainty is summarized and visualized, and interpretation aids are given. Recommendations for future work are then formulated based on these findings. In cases of major product changes, steps 2-7 can be repeated, possibly resulting in a new set of selected uncertainties to quantify, propagate and interpret.

This approach assists practitioners in determining where to invest more research and decision-makers in understanding and correctly interpreting the study results, facilitating the transition of promising technologies from research to successful commercial application. It is posited here that the additional effort of uncertainty inclusion is outweighed by its merit, as it enhances the quality of research [243, p. 3] and addresses the shortcomings of the conventional approach, where decision-makers often rely on their gut feelings [85, p. 55], influenced by past experiences and emotion [248, p. 35], which is neither objective nor reproducible [255, p. 1393].

The research performed here seeks to pave the way for the uncertainty management methodology's successful application by identifying, reviewing, organizing, and comparing different mathematical theories and engineering techniques for UQ, ultimately culminating in informed recommendations throughout this thesis. In order to guide the research and develop a research design, some background information on IPA in aerospace and UQ is given next, followed by a description of the state of the art in both fields in Section 1.2.

1.1 Background

This Section provides some background on the two topics of fundamental interest for this thesis, illustrated in Fig. 1.5: (a) integrative product assessments in aeronautics; and (b) uncertainty quantification, both probabilistic and non-probabilistic. The intention is to create a basic understanding of relevant concepts, definitions, techniques, and key publications to facilitate the identification of the research gap that is intended to be closed by this study, which is detailed further in Section 1.2. A more detailed literature review can be found in Chapter 2.

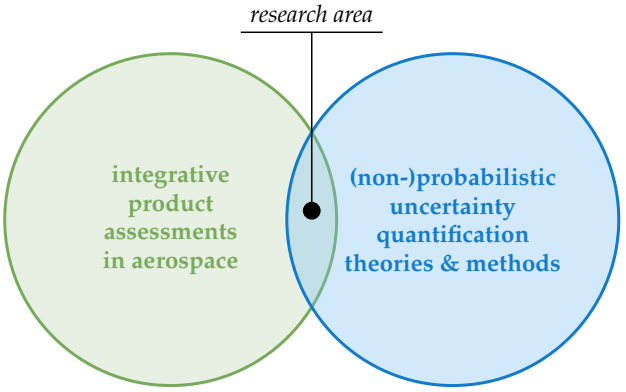


Figure 1.5 The research area as the union of IPA and UQ

1.1.1 Economic Assessments in Aerospace

In product assessments, new aircraft, technologies, processes, or new products in general are typically compared to a well-defined reference. The outcome is either an absolute or relative difference, indicating the superiority of one product over another. The reference may represent an investment in a conventional or competing product, or describe the *status quo alternative*, e.g., in a business sense, not investing in one of the products of interest but in the business itself instead.

Cost Estimation Relationship Methods

Historically, IPAs in aerospace have been performed using Direct Operating Cost (DOC) methods. These methods are based on sets of Cost Estimation Relationships (CERs), which correlate certain aircraft characteristics (e.g., the operating empty weight) or operational parameters (e.g., the average flown distance) with the airline's DOC. These CERs are often based on regression analyses of airline finance databases, which are usually not publicly available, rendering this approach rather non-transparent. Due to their simple and fast evaluation, DOC methods have been widely used and integrated into tools for aircraft design (e.g., in Hansen et al.

[96]), technology evaluation (e.g., in Kyprianidis et al. [143]), and operations assessment (e.g., in Xu et al. [292]). Often enough, this approach is used outside of its scope of applicability, such as in Martinez-Val et al. [158] or Cuerno-Rejado et al. [51], where aircraft designs that are highly unconventional are evaluated[5], even though the underlying regression of the DOC methods arguably does not include the substantial data basis for said applications, which is an inevitable symptom of the lack of method transparency. With the overarching topic of quantifying uncertainties in IPA in mind, it should be noted that most of the DOC methods do not provide any information on the accuracy, error, degree of explained variance (i.e., R^2), or any other indicative parameter for UQ. This lack of consideration for the unknown is not unique to the aerospace sector, but rather a common problem throughout most engineering science endeavors [30, p. 669].

(Whole) Lifecycle based Methods

With the increase in available computational power and a more integrative approach to product development, a shift towards a more holistic and lifecycle-driven approach has been observable [1, 77]. In the aviation sector, these typically take one of two forms[6]:

- **Lifecycle-Based Methods** which focus on the long operational phase of aircraft, where the impacts of aging, wear and tear, and other degradation mechanisms on the overall economic performance are investigated. Thereby, the evaluation perspective is usually fixed to *one* stakeholder, i.e., the operator (e.g., Justin et al. [128] and Pohya et al. [197]).
- **Whole Lifecycle-Based Methods** which aim to include all lifecycle phases with a "*cradle to grave*" approach. These expand the operator's perspective by manufacturing cost and end-of-life cost (e.g., Curran et al. [56] and Marx et al. [159]).

Like DOC methods, lifecycle-based tools have been used to assess various aircraft design alternatives (e.g., Johnson [122]), technologies (e.g., Wicke et al. [286]), or operations (e.g., Langhans et al. [146]). However, there seems to be no industry- or research-wide accepted tool or methodology, with most contributing institutions using some form of custom method. As for the whole lifecycle-based approach, some more established (albeit outdated) methods include Roskam [221] or Aircraft Lifecycle Cost Analysis (ALCCA) [80, 160], which both use CERs to calculate Whole Lifecycle Cost (WLCC) and are useful for early-stage estimations.

In other words, the (whole) lifecycle-based methods differ from the DOC methods in their scope and less in the nature of how the method itself calculates the costs. With respect to UQ, it can be concluded that conventional (whole) lifecycle considering methods, due to their prevalent CER-based nature, are often just as non-transparent as their DOC counterparts.

1.1.2 Uncertainty Quantification

Dealing with what is not known is not a trivial task. Even defining it is difficult, as there is no one kind of uncertainty. Earlier publications refer to it as "*ignorance*" [60, p. 206], some call it "*that what is not precisely known*" [30, p. 670], and more elaborate definitions are, for instance, "*any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system*" [281, p. 5]. Considering this, it is no surprise that there is neither a commonly shared terminology nor an agreement on a generic typology of uncertainties. Instead, some researchers use different terms for the same kind of uncertainty, while others use the same term to refer to different kinds [281, p. 5-6]. However, a frequently used classification distinguishes *aleatory*[7] and *epistemic*[8] uncertainty. This attribute is often referred to as the *nature* of uncertainty.

[5] A blended wing body in Martinez-Val et al. [158] and a joined-wing aircraft in Cuerno-Rejado et al. [51].

[6] The nomenclature of "*lifecycle-based*" and "*whole lifecycle-based*" approaches is inspired by other sectors such as in civil engineering [1, 278] as well as in general cost engineering [15].

[7] Stemming from the Latin term *alea* which translates to *dice* or *chance*.

[8] Stemming from the Greek term *epistémē* which translates to *knowledge* or *understanding*.

As it plays a major role in the uncertainty management methodology, a brief description follows, with a more elaborate introduction provided in Section 2.2.1.

The term "*epistemic*" is used to categorize uncertainties that exist due to a *lack of knowledge*, i.e., epistemic uncertainty may be reduced by further research and empirical efforts[9]. If, for instance, the weight of a product can only be estimated with a range between 100 and 200 kg, this uncertainty is epistemic in nature. With more information, this range can be reduced and finally replaced by a deterministic value, e.g., 155 kg, once it has been developed, manufactured, and physically put on a scale.

The term "*aleatory*" is used to categorize uncertainties that exist due to their inherent variability and hence cannot be reduced by incremental effort. For example, the frequency of how often aircraft surfaces are cleaned by precipitation can be obtained by statistics to a certain degree, but it is impossible to reduce this forecast to a single, invariably true value, e.g., every three days.

Each engineering endeavor, whether it is an experiment, simulation, or design, deals with an individual set of uncertain parameters. These may originate from various sources, such as an imprecise sensor, an uncontrollable environment, or model inputs and assumptions. It is up to the analyst to decide which uncertainties are pertinent to the engineering endeavor. Once identified, the next question the analyst needs to answer is how to model these uncertainties. Thereby, one can distinguish between the mathematical theory and the engineering application.

Mathematical Theories for Uncertainty Quantification

There have been numerous developments in mathematical uncertainty representation. It shall be mentioned right away that there is no method that fits all cases [256, p. 3] and different uncertainties are best described by different theories [243, p. 3]. The most straightforward distinction foresees a separation between conventional Probability Theory (ProT) and non-probabilistic theories.

Probability Theory is one of the best-known theories among engineers. The general term "*probability*" describes the area of study predicting the relative likelihood of various outcomes. With respect to UQ, the field of ProT provides several metrics, for instance, the Probability Density Function (PDF) (out of which the mean, standard deviation, skew, and curtosis can be obtained), or the closely linked Cumulative Distribution Function (CDF) (out of which the Interquartile Range (IQR) can be used as UQ). These metrics are primarily useful when the uncertainties at hand have a numerical character and the number of samples is sufficiently high. Additionally, many authors describe ProT as being suitable to cover only aleatory uncertainties, such as Agarwal et al. [3].

Non-Probabilistic Theories include evidence theory, possibility theory, and fuzzy set theory. The former two are typically used as an alternative to classical ProT, especially for epistemic uncertainties. While ProT uses only one function to characterize uncertainty, i.e., *the probability function*, it is broken down into two functions in evidence theory: *belief* and *plausibility*. These provide bounds on the probability of an event and are especially useful when information from independent sources have to be combined. Possibility theory was developed to deal with imprecision and vagueness in information. Similarly to evidence theory, it distinguishes uncertainty with two functions: *Possibility* and *necessity*. Fuzzy set theory takes a different approach by extending the Boolean logic of classical set theory by a membership function, enabling membership values in the closed interval of $[0, 1]$, i.e., values between *true* and *false*. Fuzzy theory is particularly useful in control systems or when dealing with the lack of confidence in sensor data.

[9] Note that Walker et al. [281, p. 13] correctly conclude that further efforts on reducing epistemic uncertainties may reveal new uncertainties which formerly have been "*deep uncertainties*".

Prevalent Techniques for Uncertainty Quantification

The process of systematically dealing with uncertainties has different names, e.g., "*uncertainty quantification*", "*uncertainty analysis*", "*uncertainty assessment*" or "*uncertainty management*". In addition to the inconsistent terminology, there seems to be no universal methodology available to define each process step, starting from acknowledging their existence over modeling them appropriately up to representing them in an understandable manner. Few authors have made an effort in this direction, such as Booker et al. [30], but applied studies following a comprehensive methodology are rarely seen. Instead, engineers and researchers typically use one of the following techniques to deal with uncertainty [90]:

Treat uncertainties as assumptions: A large portion of studies acknowledge the presence of uncertain parameters but treat them as assumptions. That is, a value or scenario is assumed to be true and fixed for the analysis. These values are often educated guesses but can include intentional or unintentional bias. The set of assumptions as a whole may be either conservative, progressive, or a mixture of both. In the latter case, it can be difficult for the recipient to correctly interpret the likelihood of the results.

Sensitivity Analysis: A typical first step when aiming for uncertainty-enabled studies is to perform sensitivity analyses, which can be classified into Local Sensitivity Analysis (LSA) and Global Sensitivity Analysis (GSA). LSA provides *design insights* by calculating the change of an output variable for a change of an input variable. Thereby, a lower, intermediate, and upper value is defined for each uncertainty. The model is then executed repeatedly for each value. GSA is rather performed to gain *overall insights* by analyzing the output behavior (e.g., variance or other statistical moments) with respect to the entire range of input values, usually represented by PDFs. A possible outcome of a GSA would be that certain input uncertainties barely matter and may be neglected for complexity reduction. In the context of UQ, GSA is often used to apportion the output uncertainty to each input constituent.

Monte Carlo Simulations: MCSs refer to computational algorithms that use iterative random sampling techniques to produce a large number of numerical outcomes. In the domain of UQ, MCSs are employed to capture the inherent variability of a system resulting from uncertain input parameters. By conducting multiple iterations of the model, each time using a different set of randomly selected inputs, a distribution of possible outcomes is generated. The distribution is subsequently examined in terms of its mean and standard deviation, which serve as indicators of the level of uncertainty in the output. Nevertheless, the Monte Carlo method has several limitations. One significant obstacle pertains to the computational intensity that arises when dealing with high-dimensional problems or models characterized by intricate dependencies. Such scenarios often necessitate a substantial number of calculations to obtain accurate estimates. Moreover, the dependability of the approach is intrinsically linked to the Law of Large Numbers (LLN), a principle that asserts that as the number of experiments increases indefinitely, the outcomes tend to converge towards the anticipated value. Dealing with epistemic uncertainties is an additional difficulty when using MCSs. Epistemic uncertainties frequently exhibit a lack of clearly defined probability density functions, primarily attributable to the limited availability of data. In such cases, researchers often use a uniform distribution between a lower and an upper bound. As MCS outcomes heavily depend on the input distributions, this approach may prove inadequate, highlighting the need for alternative approaches to quantify uncertainty in such scenarios.

Surrogate Models: As an effect of MCS execution times quickly reaching and exceeding acceptable computational budgets, recent research has focused on mitigating this issue by substituting the original model with a more computationally efficient surrogate. Techniques such as Polynomial Chaos Expansion (PCE) and Stochastic Collocation (SC) are commonly employed for surrogate model generation [59, 71]. PCEs approximates the model response using a series of orthogonal polynomials, specifically chosen to align with the probability

distribution of the input random variables. Conversely, SC interpolates the original output at carefully selected collocation points, chosen to optimize model accuracy by aligning with key statistical properties of the input space, such as using the roots of orthogonal polynomials for Gaussian quadrature. However, while surrogate models offer computational efficiency, they also introduce new uncertainties. These include model approximation errors, which arise from the difference between the surrogate and the original model, and uncertainties related to the choice of the surrogate model, such as the selection of basis functions in PCE or the collocation points in SC [213, p. 9].

Further techniques for and related to UQ include, but are not limited to:

- **Interval Analysis**, in which uncertain inputs are assigned a lower and upper bound but the distribution of the inputs within these bounds is unknown. The model's response to these input uncertainties is then evaluated by assuming uniform distributions, sampling from these and finding the output values of interest, e.g., minima and maxima. Alternatively, optimization techniques such as bound constrained Newton methods [11, 259] can be used to obtain the model's optima response given the bounded input intervals.
- **Scenario Analysis**, which involves creating and analyzing detailed models of different possible future states to assess the effects of varying uncertainties and decisions. This technique allows engineers to explore how changes in input variables might impact outcomes, using a set of defined scenarios, each representing a plausible future environment. This method is particularly useful in strategic planning and complex decision-making where uncertainties are high.
- **Non-Probabilistic Techniques**, which include Fuzzy Inference System (FIS) from fuzzy set theory or make use of the concepts of belief and plausibility (from evidence theory) or possibility and necessity (from possibility theory).

Now that some background information on the available theories and engineering techniques for UQ has been given, the state of the art of uncertainty-enabled IPA in aerospace is discussed next.

1.2 State of the Art and Key Barriers

This Section briefly reviews relevant literature from Techno-Economic Assessments (TEAs) and related aeronautics fields, focusing on the incorporation of uncertainties. First, interdisciplinary tools are examined, followed by increasingly discipline-specific studies where the use of mathematical theories and employed UQ techniques are discussed.

Within the product-agnostic TEA environment, the first framework to highlight is the Integrated Cost and Revenue Estimation (i-CARE) tool, developed by Justin et al. [126–128]. By combining different modules with a commercial mission simulation tool, i-CARE computes all aircraft related costs and revenues, integrating the effects of flights, maintenance, and degradation. This approach values new aircraft concepts and engine maintenance strategies from an airline's perspective, addressing valuation uncertainties with probabilistic methods such as MCS for Real Option Analysis (ROA). ROA extends traditional methods from discounted cash-flow analysis by incorporating options theory to handle uncertainty and decision-making with a focus on flexibility [84, p. 431], contrasting with input parameter UQ, which primarily assesses the impact of parameter variability on economic Key Performance Indicators (KPIs)[10]. ROA evaluates managerial flexibility and strategic options, such as delaying, expanding, or abandoning projects under different scenarios [127].

[10] Supplemental information about ROA can be found in A.1.

A tool with a similar objective is the economic framework from Thokala et al. [266, 267], which uses a Discrete Event Simulation (DES) to model the complex interactions in aeronautics and calculate the lifecycle cost of different aircraft focusing on the operational phase. Uncertainties within acquisition module are mentioned to be dealt with MCS, but their overall framework is deterministic, resulting in point estimates of the economic KPIs.

ALCCA, which resulted from the combined effort of NASA and GeorgiaTech's Aerospace Systems Design Lab (ASDL), not only models the operator's cash flow, but also the manufacturer's and calculates overall economic metrics such as the Return on Invest (ROI) [163]. In contrast to the aforementioned techno-assessment tools, ALCCA has frequently been used in combination with aircraft design tools such as Flight Optimization System (FLOPS) [174]. This combination allowed for a vast number of technology selection investigations, in which multiple potential technologies are evaluated with respect to their combined impact the overall aircraft design and its predicted efficacy[11].

This TPS process has often been embedded in the Technology Identification, Evaluation and Selection (TIES) methodology [165], which deals with uncertainties using sensitivity analyses and MCS. Following the selection phase managed by TIES, the Technology Impact Forecasting (TIF) methodology takes over to assess the future impacts of the chosen technologies [166]. Due to its forecasting nature, it deals with significant uncertainties about future conditions and technological performance [250]. Typical techniques to model these uncertainties rely on probability theory and comprise scenario analysis and sophisticated sensitivity analyses including non-linear interaction effects [134, 164, 165]. It has since been applied in various contexts [35, 109, 132, 161, 162], commonly focusing on identifying which subset of technologies from a larger pool should be incorporated into overall aircraft design. For a comprehensive overview of TIF and TIES, refer to Soban et al. [250].

Another research stream worth highlighting is the work from Akram and Mavris [7, 8, 10, 11], which stands out for its use of evidence theory in addressing the complexities of UQ in technology selection. They effectively integrate this theory into three critical tools: the Technology Impact Matrix (TIM), Technology Compatibility Matrix (TCM), and Technology Synergy Matrix (TSM). These matrices serve as analytical frameworks that enable a comprehensive evaluation of combining potential technologies. The TIM assesses the effects of different technologies on organizational goals, considering both direct and indirect impacts [133]. The TCM examines how well new technologies integrate with existing systems, ensuring that the selected technologies are compatible with one another [224]. Meanwhile, the TSM identifies and evaluates the synergistic effects of combining multiple technologies, highlighting potential enhanced benefits or unintended negative interactions [9, 299]. Their methodological innovation lies in how these matrices, underpinned by evidence theory, allow decision-makers to navigate the uncertainties of technological advancements with greater confidence, thereby optimizing strategic technology adoption and integration.

Focusing more on aircraft design and performance evaluation and less on the overall TEA is the work by Pfeiffer et al. [190–192] and Krosche et al. [138]. Pfeiffer et al. uses the Remote Control Environment (RCE) from DLR to connect the various disciplines involved in aircraft design. The authors developed an interface where non-intrusive MCS are automatically performed on parameters which have previously been defined as uncertain by the responsible experts. However, these simulations only allow for normal distributions and do not discern the nature of uncertainty nor their effect on the result. Krosche et al., in contrast, does discern the uncertainty nature in their aircraft design study and focused on those that are epistemic. These were treated with probability theory and propagated using a Monte-Carlo setup, resulting in a robust low-noise cruise-efficient short takeoff and landing transport aircraft.

[11] Supplemental information about Technology Portfolio Selection (TPS) can be found in Appendix A.2.

Similar approaches to uncertainty inclusion can be observed in the aircraft noise assessment studies by Bertsch et al. [27] and June et al. [125]. Bertsch et al. investigated the influence of uncertain aircraft design parameters, which are likely to be epistemic, and varying environmental boundary conditions, which are likely to be aleatory, while assuming normal distributions and using MCS for propagation. However, no explicit discernment of the uncertainty nature was provided. Consequently, the different impact epistemic and aleatory uncertainties have on the overall result was not discussed. June et al. focused on the installed aircraft systems (e.g., the main and nose landing gear or the Kruger flap) for the noise assessment. In this study, uncertainties were categorized based on their effect (e.g., whether they are a source for noise or a represent a possibility to reduce noise) for overview purposes. Similarly to Bertsch et al., they quantify uncertainty with probability theory and use MCS, but do not differentiate the uncertainty nature nor its nuanced impact on the assessment outcome.

While the previous integrative and/or multidisciplinary studies almost exclusively use probability to deal with uncertainties, there are a number of intradisciplinary aerospace studies in the that use non-probabilistic approaches. Structural designs under uncertainty through Finite Element Analysis (FEA), for instance, is provided by Chowdhury et al. [45] and Rezaei et al. [216]. Both make use of fuzzy logic as an alternative to probability theory and utilize MCS to design an aircraft wing and T-Tail, respectively. Another example is given by Chen et al. [41], who quantify model uncertainties by applying FIS to predict aircraft taxi times. Further uses of fuzzy set theory can be found in Hawer et al. [100] and Reinhart et al. [214], applied to factory evaluations. The authors combine artificial neural networks with FIS to quantify linguistic input uncertainty and subsequently translate these into PDFs.

Evidence theory has also been used in some intradisciplinary studies to deal with situations where data is limited and experts area available. Worden [287], for example, discussed its use in an evidence-based damage classification study for aircraft structures. Similarly, Bae et al. [22, 23] investigated the structural response of an intermediate complex wing using evidence theory. Agarwal et al. [3] investigated the applicability of evidence theory for an aircraft sizing problem. This type of evidence-based approach to optimization under uncertainty, however, is rarely seen in literature.

Sensitivity analysis is a standard method used to distinguish critical uncertainties from negligible ones, which is essential for targeting uncertainty reduction efforts effectively. However, the methodologies employed in sensitivity analysis differ widely, as documented in the review by Roelofs et al. [220]. For example, the range of methods within the LSA subset includes one-at-a-time parameter variations, as seen in the UQ study of a constrained aircraft design by Xie et al. [289], and extends to more complex approaches that account for interaction effects, such as those reported in the aircraft subsystem design studies by [37, 38]. GSA are also applied in various contexts, such as in the variance-based uncertainty decomposition approach used in the aircraft conceptual design study by Opgenoord et al. [184]. Other notable methods include the entropy-based approach to GSA proposed by Curran et al. [55] in a high pass filter control design, and the application of the Sobol' method in a compressor simulation model by Bilal [28]. Despite the development of new and diverse methods, as highlighted by [212, 213], the choice of methodology often remains unexplained, which may impede their broader adoption in GSA.

Table 1.1 summarizes the key UQ characteristics of the previously mentioned studies. The 12 column-wide attributes cover the theory and technique chosen to deal with uncertainties, and what the nature of the considered uncertainties are. The individual literature entries are clustered into those of the TEA domain, TPS, aircraft design studies and related fields such as aircraft noise assessment, and studies from a more monodisciplinary domain such as aircraft structural design.

Table 1.1 Key UQ characteristics of key publications from various aeronautic fields.

		theory				**technique**					**nature**		
Domain / Description	**References**	deterministic	probability	evidence	fuzzy set	none	MCS	LSA	GSA	non-probabilistic	epistemic*	aleatory*	unspecified
techno-economic assessments													
Thokala et al.	[266, 267]	●				●							
i-CARE	[126–128]		●				●	●					●
technology portfolio selection													
ALCCA & FLOPS (TIES & TIF)	[35, 109, 132, 161, 162]		●				●	●					●
Akram and Mavris (TIM, TCM, TSM)	[7–11]			●			●			●	●		
aircraft design and related fields													
Pfeiffer et al.	[190–192]		●				●						●
Krosche et al.	[138]		●				●				●		
Bertsch et al. and June et al.	[27, 125]		●				●						●
Agarwal et al.	[3]			●						●	●		
Opgenoord et al.	[183, 184]		●						●		◑	◑	
Xie et al. and Chakraborty et al.	[37, 38, 289]		●					●			◑	◑	
monodisciplinary fields													
compressor simulation	[28]		●						●				●
high pass filter control design	[55]		●						●				●
evidence-based aircraft structural design	[22, 23, 287]			●						●	●		
aircraft taxi time prediction	[41]				●					●			●
fuzzy factory analysis	[100, 214]				●					●			●
fuzzy aircraft structural design	[45, 216]				●					●			●

*The ◑ symbol indicates that these publications mention the different uncertainty natures but do not treat them differently and consequently do not quantify their nuanced impact on the overall result.

This brief review of UQ in aeronautics underscores the prevalence of probability theory, especially in multidisciplinary and integrative contexts. In contrast, fuzzy set theory is mainly applied in monodisciplinary studies to manage linguistic uncertainties and model development via FIS, but it finds limited use in wider assessment frameworks. Although evidence theory is recognized for addressing epistemic uncertainties, its application is infrequent and limited to specific areas such as single-discipline studies, complex aircraft design, and technology portfolio management. Despite its robust mathematical base and noted potential, evidence theory's integration into TEA frameworks is almost entirely absent. Applied techniques for UQ within probability theory are MCS as well as local and global sensitivity analysis. The latter is found more frequently in the monodisciplinary domain, whereas tools and frameworks from the TEA area and TPS field tend to rely on LSA. The discussed literature tends to either overlook the specific nature of uncertainties - epistemic versus aleatory - or concentrate on one type only, without adequately exploring their distinct impact on the results and potential future uncertainty reduction efforts.

Building on these insights, the following barriers to a structured inclusion of input UQ in TEAs can be formulated:

1. UQ, an active research area, offers a variety of mathematical and engineering methods. While monodisciplinary studies often integrate non-probabilistic approaches, the techno-economic sector predominantly utilizes probability theory, MCS, and LSA. This reliance on traditional methods might stem from the greater familiarity of both practitioners and recipients of TEAs with probabilistic concepts and UQ metrics, such as the mean value ± a standard deviation [220, p. 9]. The preference for conventional approaches may also

be attributed to the scarcity of practical tools and interpretive aids for alternative non-probabilistic theories and techniques.

2. UQ efforts are typically demanding in terms of time, budget, and computational resources. Consequently, it is impractical and often unnecessary to address all uncertainties with equal depth. Sensitivity analyses, particularly GSA, help in breaking down output uncertainty into its input parameter components, pinpointing key parameters, and efficiently directing efforts to reduce future uncertainties. Despite their significant potential in uncertainty addressing TEAs, the adoption of GSA in such frameworks remains limited. This may be due to the wide variety of GSA methods available, which differ substantially in their approaches, compounded by the lack of discussion among researchers about their specific choice of GSA technique.

3. Recognizing the nature of uncertainty and its influence on outputs is pivotal in TEAs. When epistemic uncertainties dominate the output, it suggests that further research and development could effectively reduce uncertainties. In contrast, a predominance of aleatory uncertainties often calls for more robust design solutions, if viable. Despite its importance, this systematic differentiation between epistemic and aleatory uncertainties is rarely addressed in assessment literature. This oversight may stem from a general lack of awareness and challenges associated with combining and propagating these different types of uncertainties in a way that is both manageable and interpretable.

With these barriers in mind, the following Section presents the overall research objective as well as the guiding research questions and corresponding hypotheses.

1.3 Research Objective, Research Questions, and Hypotheses

The structure of this thesis, including the design of the conducted studies, is driven by an overarching research objective and three guiding research questions, each with its own hypothesis. The former can be formulated as:

> *The objective of this research is to enhance the transparency of prospective TEAs in aeronautics by overcoming key barriers that impede the implementation of an uncertainty management methodology that is discerning of uncertainty nature, efficient in application for practitioners, and comprehensible for recipients.*

In this context, the term *transparency* is central, particularly in light of the aforementioned reproducibility crisis and its potential to undermine the credibility of prospective TEAs. Enhancing transparency is crucial for addressing concerns about the reliability of the assessment studies, especially given the challenges in validating these assessments, particularly when predictive modeling and long-term forecasts are involved. The phrase *discerning uncertainty nature* underscores the critical importance of recognizing and differentiating the nuanced impacts of epistemic and aleatory uncertainties on the interpretation of outputs. *Efficient in application* encompasses both uncertainty modeling efforts and considerations of computational requirements. Finally, the phrase *comprehensible for recipients* ensures that the results and implications of the assessments can be readily understood and utilized by decision-makers, thus facilitating the adoption of the methodology.

The Research Questions (RQs), which aim to guide the research and overcome the aforementioned barriers, are thus formulated as follows:

> **RQ1** *How can the adoption of global sensitivity analysis in uncertainty-addressing TEAs be facilitated or promoted, given the variety of available techniques and the lack of discussion among researchers about their specific choice?*

> **RQ2** *How can non-probabilistic theories, such as evidence theory, be more widely adopted in TEAs, given the challenges in interpreting their concepts and metrics and the lack of comparative studies with prevalent probabilistic methods?*
>
> **RQ3** *How can the systematic differentiation between epistemic and aleatory uncertainties be effectively addressed in TEAs, considering the challenges associated with combining and propagating these different types of uncertainties in a manageable and interpretable manner?*

The first RQ is dedicated to the improvement of UQ efficiency and subsequent uncertainty reduction efforts by addressing current challenges with the wide-spread adoption of GSA methods. The corresponding hypothesis specifies the approach further and can be formulated as follows:

> **HYP1** *Through a comprehensive evaluation of the capabilities, interpretability of sensitivity, and computational efficiency of various GSA techniques on an example problem, followed by the demonstration of the selected method's effectiveness in a case study, it is anticipated that the insights gained will contribute to facilitating the adoption of GSA in uncertainty-addressing TEAs, informing future research directions and methodological choices in the field.*

The second RQ guides investigations in the domain of non-probabilistic theories, more specifically, evidence theory. Thereby, it addresses the limitations of probability theory to adequately model epistemic uncertainties, especially in data scarce situations. Practical aspects hindering a more frequent adoption of evidence theory are further specified in the corresponding hypothesis:

> **HYP2** *The development of publicly accessible and user-friendly tools for non-probabilistic UQ, such as for evidence theory, is expected to aid in the adoption of these methods in TEAs. Through demonstration, comparison with prevalent probabilistic methods, and analysis of key aspects such as interpretability, traceability, and computational efficiency, such tools have the potential to enhance decision-making processes in practical applications.*

The third and final RQ addresses a critical aspect of UQ in TEAs and subsequent decision-making processes. As highlighted in the discussion of the third key barrier, the nature of uncertainty can play a decisive role in the formulation of future recommendations. More specifically, if decision-makers are capable of assessing not only the overall uncertainty in the outputs but also of distinguishing between the influences of epistemic and aleatory uncertainties, their subsequent actions can be influenced in the following ways:

- When high epistemic uncertainties dominate a TEA, they indicate a significant potential for enhancing internal knowledge. Decision-makers should consider allocating more resources to research and development to deepen understanding of the epistemic input parameters, thereby effectively reducing the output uncertainty. This strategy is advisable only if the assessment's overall output - reflected, for instance, in its mean value - is sufficiently positive to justify the potential returns on such investments. Conversely, if the overall outcome is low or negative, further investments are dispensable.

- In scenarios dominated by high aleatory uncertainties, which represent inherent and product-specific risks, additional investment in knowledge acquisition does not typically mitigate these uncertainties. If the assessment's overall outcome is positive yet accompanied by considerable output uncertainty, the strategic response should pivot towards developing more robust product designs that mitigate these uncertainties. This approach is warranted only when the economic stakes are significant. However, if the overall output is low or the uncertainty negligible, pursuing such efforts would not be cost-effective as they are unlikely to deliver proportional benefits.

The corresponding hypothesis can be formulated as follows:

> **HYP3** *By systematically differentiating between epistemic and aleatory uncertainties in TEA using evidence-theoretic methods and nested Monte Carlo simulations, it is anticipated that enhanced interpretability and actionable insights will enable more informed resource allocation and risk mitigation strategies. Specifically, this approach is expected to identify dominant uncertainty types and guide strategic decisions on research investments or robust product designs.*

With the research questions and hypotheses formulated, the next Section presents the research plan, which shows what studies and analyses are used in this thesis to meet the research objective and overcome the challenges that researchers may face when incorporating the uncertainty management methodology.

1.4 Research Plan and Thesis Structure

The research plan for this dissertation is depicted in Figure 1.6, which outlines the foundational Chapters 2 and 3, followed by the three main analytical chapters, Chapters 4 through 6. A description of the research plan is provided next.

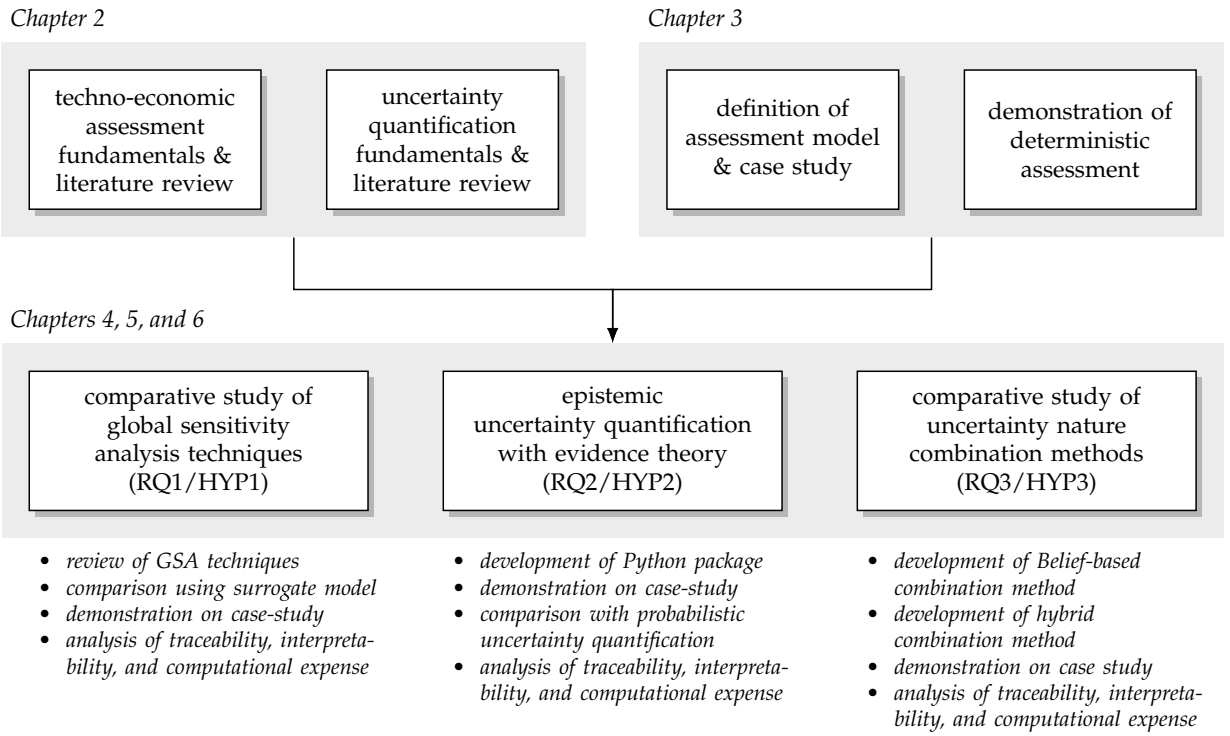


Figure 1.6 Overview of the research plan and thesis structure.

Chapter 2 comprehensively examines the fundamental theories and relevant literature associated with UQ and TEAs within aeronautic practice. This chapter establishes a solid foundation for subsequent analyses by detailing key mathematical concepts and their applications in engineering, focusing on techniques such as ProT and evidence theory. Additionally, it explores the basic principles of TEAs to enhance comprehension and define the scope of the evaluation.

Chapter 3 presents the use case of Hybrid Laminar Flow Control (HLFC)[12], including a description of the newly developed assessment framework, the aircraft under investigation, and

[12] It should be noted that the majority of the data for this use case were obtained from two Clean Sky 2 projects: LPA WP1.4.1 and LPA WP1.4.4. The goal of these projects was to mature the HLFC technology for the next generation of large passenger aircraft. Supplemental information about these projects can be found in Appendix A.3.

how the technology's impact is modeled. This chapter also details the boundary conditions and assumptions. It concludes with a deterministic assessment of HLFC, serving both as a traditional technology evaluation and as a baseline for subsequent uncertainty analyses.

Chapter 4 is dedicated to GSA, addressing the first research question and hypothesis. It explores various GSA methods, evaluating them in terms of their capabilities (e.g., (in)dependence of inputs), convergence behavior, and ease of interpretation. The most suitable GSA method is then applied to the case study, serving as a practical demonstration of its effectiveness.

Chapter 5 advances the discussion to the second RQ and hypothesis. More specifically, a Python package for evidence theory has been developed, tested, and verified on the HLFC example. This Chapter further examines aspects such as convergence behavior and the difficulty of interpretation. Additionally, a verification analysis is conducted, which includes hypothetical uncertainty reduction (e.g., due to new information becoming available), to discuss the traceability of the evidence-theoretic UQ method.

Chapter 6 addresses the third and final research question and hypothesis by comparing two techniques for nested MCSs. The first technique is based on evidence theory, while the second, though incorporating elements of evidence theory, stays within the probabilistic framework. The comparison focuses on their interpretability, convergence behavior, and utility in scenarios where new uncertainty-reducing information becomes available.

The thesis concludes in Chapter 7 with a comprehensive summary of conclusions and a reflection of the limitations of this work, alongside an outlook for further research in this field.

2 Fundamentals and Literature Review

This Chapter discusses the topics of TEA and UQ, addressing both their theoretical foundations and practical applications. It aims to provide readers from each discipline with essential knowledge of the other. The exploration begins with Section 2.1, which classifies various scopes and techniques of TEA and introduces the key terminology. Following this, Section 2.2 presents the mathematical theories and practical techniques prevalent in the UQ field. Section 2.3 offers a thorough literature review that complements the initial overview provided earlier. This Chapter concludes with a summary in Section 2.4.

2.1 Techno-Economic Assessment Fundamentals

TEA is used to evaluate the technical performance and economic feasibility of new technologies, particularly in the context of industrial applications. This methodology integrates engineering analysis with financial modeling to determine the viability of technologies under consideration. By examining both technical metrics, such as an improvement in operating performance, and economic factors, such as direct operating cost, TEA provides a holistic view of both the operational and economic implications of technology deployment. The insights gained from TEA are crucial for stakeholders in making informed decisions about technology investment, development, and implementation, ensuring that both technical and financial requirements are addressed effectively.

In the IPDP, early TEAs typically rely on qualitative techniques due to limited initial knowledge. As the process advances, more quantitative approaches become feasible. Consequently, there is no universal TEA process or metric that fits all purposes [77, p. 45]. Fig. 2.1 illustrates a widely recognized classification of cost estimation techniques [15, p. 3], distinguishing between intuition-based and analogy-based approaches for qualitative methods, and parametric and detailed analysis-based techniques for quantitative methods.

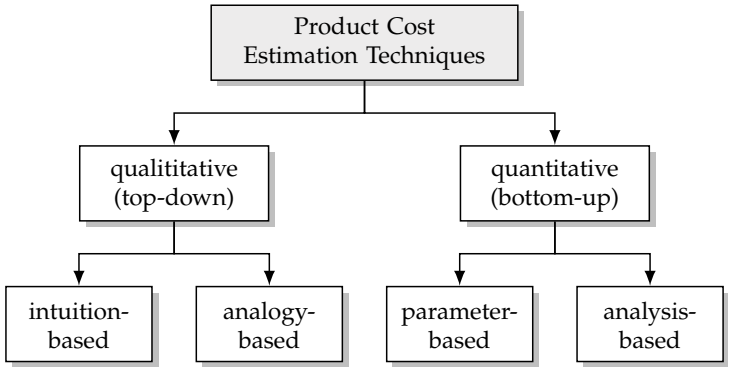


Figure 2.1 Classification of product cost estimation techniques in the general engineering domain, based on Niazi et al. [177, p. 569].

Qualitative, or top-down, techniques are characterized by their minimal information requirements but often lack precision due to their reliance on historical data and expert judgment. It is essential to evaluate their relevance carefully in each specific situation. On the other hand, quantitative, or bottom-up, methods involve more complex modeling, ranging from regression-based CERs to discrete event and/or agent-based models. While requiring expert knowledge, these methods generally provide a broader scope of evaluation.

In the field of aeronautical engineering, practitioners and researchers utilize a broad spectrum of cost estimation methods with a variety of temporal and content scope. For nomenclature clarity purposes, a classification based on three overall groups is suggested and shown in

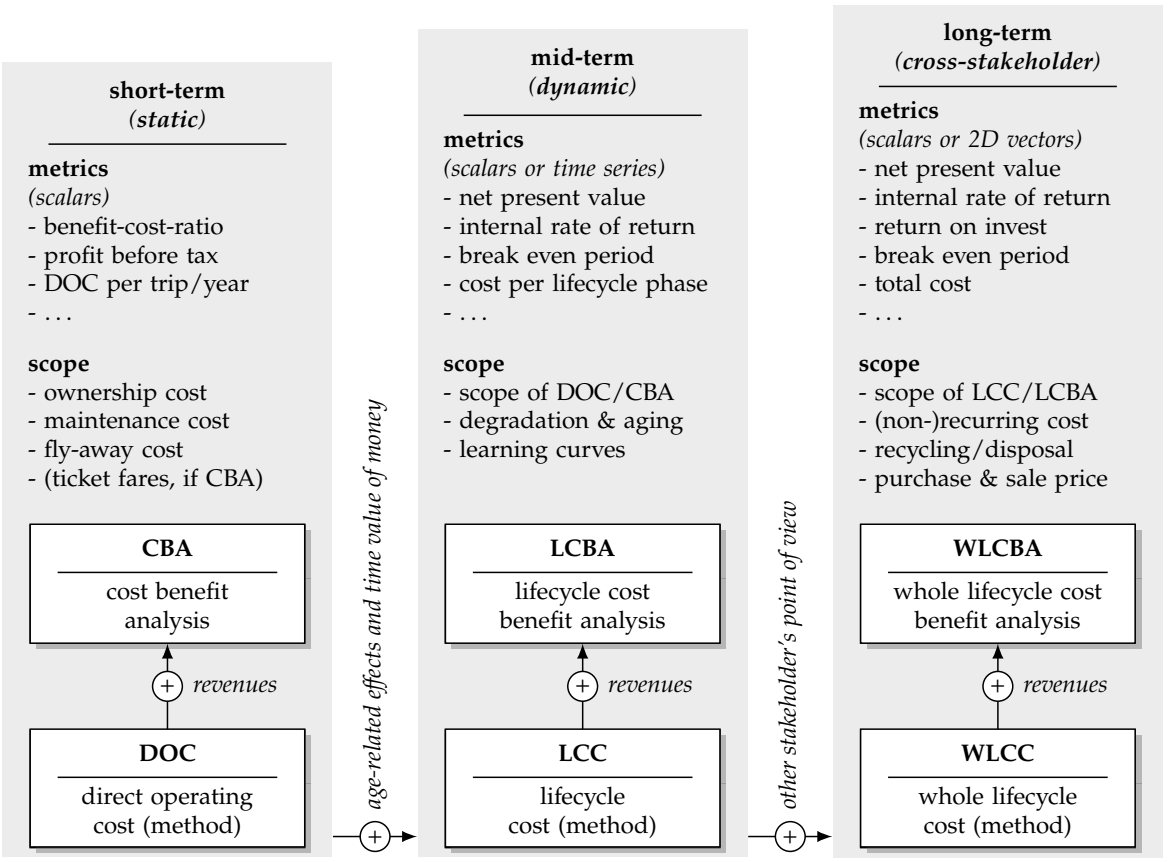


Figure 2.2 Categorization of TEA practices in aeronautics based on temporal and content-specific scope.

Fig. 2.2. Here, existing assessment approaches are categorized into: (a) short-term and static; (b) mid-term and dynamic; or (c) long-term and cross-stakeholder methods. The former two focus on estimating the economic impact for the customer, *i.e.*, airlines, whereas the latter encompasses the perspective of the aircraft manufacturer. This classification, though not exhaustive, seeks to represent the TEA application practice where customer integration is focused on. It furthermore provides the structure of the subsequent discussion, where techniques within each TEA category are described further.

2.1.1 Short-Term and Static Methods

The first category of TEA methods comprises DOC and Cost Benefit Analysis (CBA) approaches. They are named *short-term* since they typically consider shorter periods of time, such as one representative year or one representative mission. The *static* nature combines two aspects. The first has an economic background and addresses the fact that the temporal effect of the payments is not accounted for. The second aspect deals with the methods themselves. Many DOC and CBA methods tend to have a fixed parametric nature, which prohibits flexibility in terms of considering alternative, e.g., technology-specific parameters in the equation.

Direct Operating Cost

DOCs revolve around cost elements that are directly affected by the aircraft and can be broken down into Cash Operating Cost (COC) (comprising, for instance, fuel, maintenance, and operating various fees) and Cost of Ownership (COO) (including, but not limited to, insurance cost and interest charges). Together with Indirect Operating Cost (IOC), DOC represent the Total Operating Cost (TOC) of an airline, as shown in Fig. 2.3. As it is common for operators to compare different aircraft based on their associated DOCs, parametric techniques have been developed that predict these costs based on early accessible information about an aircraft [241, 275]. These techniques are known as DOC methods and are the predominant approach in

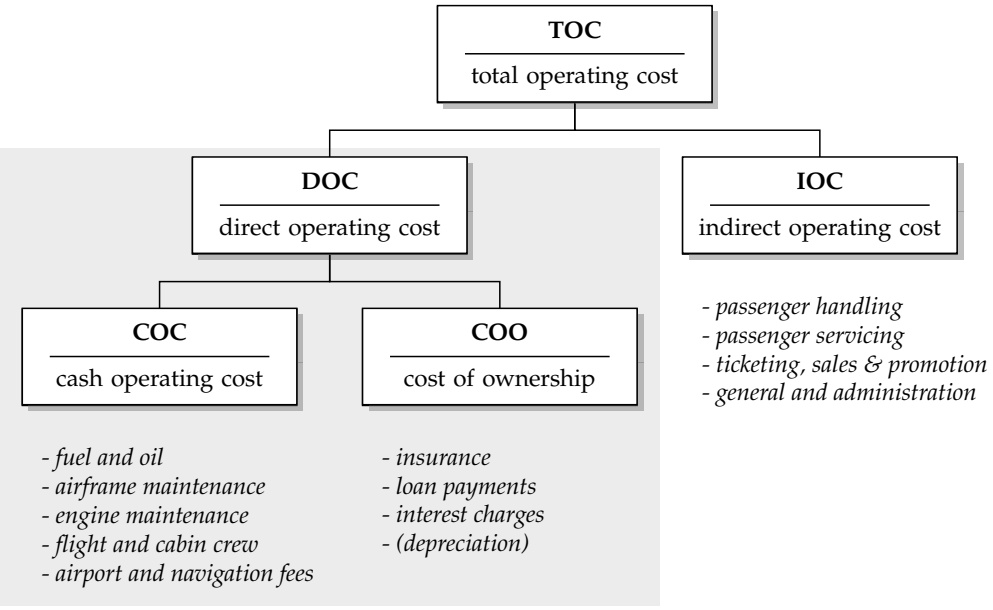


Figure 2.3 Common breakdown of total operating cost, based on Clark [46, p.251]

aeronautical practice for the evaluation of aircraft designs, technologies, and sometimes operational procedures [199]. Thus, the term DOC can be interpreted as both a cost element and a cost estimation method.

As a cost estimation technique, DOC methods utilize regression-based CERs, resulting in a set of equations that take, for instance, the aircraft's weight to calculate the expected average maintenance cost. The process of creating these was first developed by the U.S. Department of Defense [270] and consists of the following steps:

1. Collect input and output data from company databases, libraries, or contractors.
2. Evaluate and normalize the data for comparison.
3. Select a subset of easily accessible input parameters, especially useful in early stages.
4. Qualitatively analyze correlations, e.g.,, using data plots or correlation matrices.
5. Identify potential equations for regression and curve fitting.
6. Test predictions and determine the best-fitting equations.

Known for their simplicity and quick computational evaluation, DOC methods are frequently incorporated into programming environments, e.g., for Design of Experiences (DOEs). Despite their practicality, however, DOC methods have several limitations that require careful consideration [84, p. 428]. Because different DOC methods typically have different sets of input parameters, changes in the Object of Interest (OoI) may or may not be reflected by the CERs of the selected method. It is imperative to remember that "*. . . hypotheses, common sense, and engineering knowledge should precede, and then the association should be verified with statistical analysis*" (Roy et al. [225, p. 3]). Equally significant is that the same input parameters can yield disparate outcomes based on the chosen DOC method, as shown in Langhans [145] and Pohya et al. [199], which highlights the importance of the selection process. Therefore, it is suggested to confirm the selection and discuss the limitations of the selected DOC method with the customer or recipient of the study.

As outlined in Section 1.1.1, DOC methods typically fall short in terms of UQ as they often fail to provide the underlying dataset and the goodness of fit measures of the regression. This practice could potentially cloud the accuracy of the prediction, thereby compromising the transparency of these methods. Moreover, the inherently static and short-term nature of DOC methods places temporal and context-specific constraints on their usage, restricting factors such as the period

under consideration, the operational content evaluated, the object of interest, and the level of temporal detail[1]. Consequently, while DOC methods offer a seemingly streamlined approach to cost estimation in aeronautical engineering, their successful implementation requires the understanding and navigation of these constraints.

Cost Benefit Analysis: Extension of DOC by Revenues

The previously described DOC methods are capable of indirectly capturing the revenue potential if the outcome is divided by the number of seats and/or the traveled distance, assuming that the passenger load factor is identical among the alternatives. With the help of a revenue model, it is possible to enhance the TEA by estimating how many tickets are sold for which price while considering the seating class (i.e., economy, premium economy, business, and first) in conjunction with the flown distance and other factors (e.g., the airline's service level and reputation). This explicit benefit quantification allows for the calculation of basic metrics of static investment budgeting, such as the Benefit Cost Ratio (BCR) and the Profit Before Tax (PBT). However, CBAs of this kind are virtually nonexistent in TEA, presumably due to the OoI impact being sufficiently described by DOC methods only (often using the aforementioned normalization to the number of seats or traveled distance). If, however, the product under investigation affects the utilization of the aircraft (e.g., due to a shorter turnaround time or increased cruise speed), an explicit inclusion of revenues allows for quantifying the revenue *loss* or *gain* expected by employing said product. This, in turn, enables a more comprehensive picture of the IPA.

2.1.2 Mid-Term and Dynamic Methods

In the second TEA category, the focus shifts to the entire operational phase of aircraft. Here, the methods typically encompass a broader period and a more detailed environmental context for the product, employing a combination of parametric-based and analysis-based approaches. Consequently, they tend to offer a more flexible input space compared to the static DOC methods. Notable techniques in this mid-term and dynamic cluster include the Lifecycle Cost (LCC) and Lifecycle Cost Benefit Analysis (LCBA), which are described next.

Lifecycle Costing

Lifecycle costing is crucial in the aeronautic sector for evaluating the economic viability of new technologies adopted by airlines. It encompasses various dynamic elements that influence the overall economic performance of the airline, from initial acquisition to disposal. LCC for airlines go beyond the TOC. They include various dynamic elements that can significantly impact the overall economic performance of the airline. Transitioning from the initial acquisition to disposal, these costs evolve, influenced by factors such as learning curves and aging processes.

Learning curves depict the improvement of task efficiency over time as experience accumulates. The increasing expertise of technicians and maintenance personnel leads to reduced maintenance costs per unit over time. For example, while initial costs may rise with the adoption of new materials such as novel composites, they tend to decrease as best practices emerge. Aging processes, as illustrated by the "bathtub curve", impact operational costs, initially showing higher failure rates, stabilizing, and eventually increasing. Aircraft structures, for instance, are susceptible to risks such as fatigue and corrosion towards the end of their operational lifespan.

Depending on the technology, seasonal and geographical effects can significantly influence performance and, consequently, cost-effectiveness. The operational environment in which an aircraft operates throughout the year can subject, for instance, its engines to varying levels of wear and tear. In harsh, cold conditions, engines may encounter challenges associated with

[1]Supplemental information about the limitations of DOC methods can be found in Appendix B.1.

cold starts, increased icing risks, and heightened thermal stresses. Conversely, in contrasting environments such as hot and dusty climates, engines face challenges related to abrasion and contamination from airborne particles. A technology aimed at improving engine efficiency may exhibit its own susceptibility to seasonal and climatic factors. This intricate interplay of seasonal, geographical, and technological factors in aircraft operations underscores the importance of comprehensive LCC frameworks. By accounting for environmental variables and temporal effects, such frameworks facilitate a more accurate assessment of a technology's impact on performance and cost-effectiveness. Nevertheless, it's crucial to acknowledge the considerable effort required to capture and integrate these factors into analyses, especially when their influence on the study's outcome is limited.

LCBA: Extension of LCC by Revenues

Similar to short-term CBA, LCBA is relatively uncommon in aeronautic TEA. It necessitates an explicit model of the aircraft's revenue streams, allowing for the direct capture of the cost of missed opportunities. Consequently, LCBA facilitates the calculation of additional economic metrics through discounted cash flow analysis, a technique commonly employed in investment decision-making. These metrics encompass the Net Present Value (NPV), Internal Rate of Return (IRR), and Break-Even Period (BEP) [89]. The subsequent paragraphs introduce these economic metrics alongside the concept of the time value of money. For further insights, consider McCrary [167, pp. 57-69] for an economist's perspective or Clark [46, pp. 303-313] for an aviation perspective.

The Time Value of Money concept describes the tie between the value of money and the time it is spent or earned, as evident from the commonly known phrase "*a dollar today is worth more than a dollar tomorrow*". The two underlying principles of this concept are earnings potential and inflation. The former essentially describes the opportunity cost of not being able to invest today but at a later point in time. Similar to the earnings potential, the impact of inflation increases with longer time periods. The underlying formula to calculate the Future Value (FV) of an investment made today, i.e., the Present Value (PV), is:

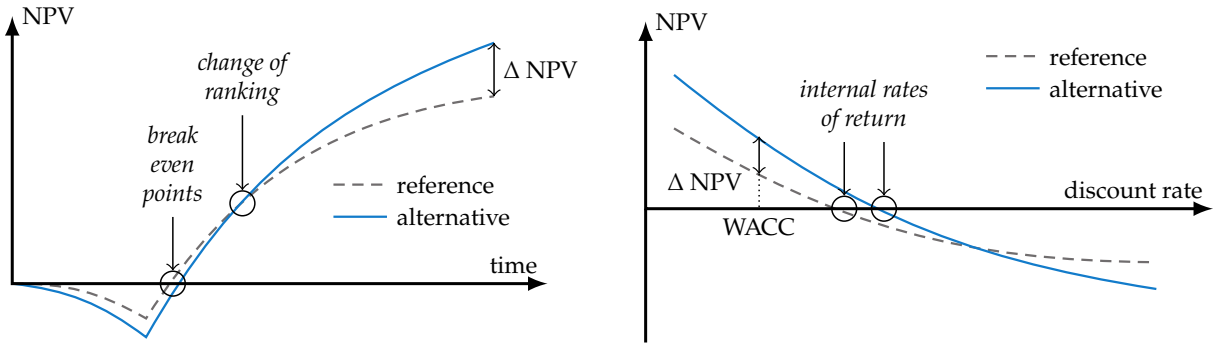
$$\mathrm{FV} = \mathrm{PV} \cdot (1 + r_{\mathrm{real}})^t \quad \text{with} \quad r_{\mathrm{real}} = \frac{1 + r_{\mathrm{interest}}}{1 + r_{\mathrm{inflation}}} - 1 \tag{2.1}$$

where r_{real} is the real interest rate, and t is the number of years considered. The real interest rate combines the expected inflation rate $r_{\mathrm{inflation}}$ and the nominal interest rate r_{interest}. In other words, a higher interest rate leads to a higher FV given the same PV, whereas a higher inflation rate decreases the FV for a given PV.

The Net Present Value is a measurement of the profitability of a potential investment project, i.e., it indicates how much value can be added to the investor. NPV is part of the discounted cash flow analysis and is a standard method for using the time value of money to evaluate long-term projects [32] and is commonly used in aircraft fleet planning [46]. To calculate the NPV, all (future) cash flows related to the aircraft have to be discounted to a reference year. Equation 2.2 can be directly inferred from Eqn. (2.1) and discerns whether the future cash flows C_t are inflation-corrected or not.

$$\mathrm{NPV} = C_0 + \sum_{t=1}^{T} \frac{C_t}{(1+r)^t} \quad \text{with} \quad \begin{cases} r = r_{\mathrm{real}} & \text{if } C_t \text{ excludes inflation} \\ r = r_{\mathrm{inflation}} & \text{if } C_t \text{ includes inflation} \end{cases} \tag{2.2}$$

Based on a reference case and a similar risky alternative, the difference in NPV represents the economic advantage, i.e., if $\Delta\,\mathrm{NPV} > \$0$, the investment project (e.g., aircraft with new technology) leads to an overall economic benefit, whereas $\Delta\,\mathrm{NPV} < \$0$ indicates an economic disadvantage (see Fig. 2.4 (a)). The used discount rate is a central element in the NPV calculation



(a) reference and alternative NPV over project time.

(b) reference and alternative NPV over discount rate.

Figure 2.4 Schematics of the NPV, IRR, and BEP.

and represents the rate of return to which the investment project is compared to. Since investors' discount rates are usually unknown, a common procedure is to use the Weighted Average Cost of Capital (WACC), generally published in the respective yearly reports [46, 167].

The Internal Rate of Return is directly derivable from the NPV. It is equivalent to the discount rate r for which the NPV equals zero and represents the annual percentage gain or loss of the investment project (see Fig. 2.4 (b)). When multiple investments are compared, the *Incremental Internal Rate of Return* (IRR_{incr}) can be used to determine the superiority of an alternative [44]. The IRR_{incr} is computed in a similar way to the standard IRR, but instead of using the cash flow of each project, the difference in cash flow (i.e., the cash flow of the standard investment subtracted by the cash flow of the proposed project) is used. To assess the economic advantage, it is common to compare the IRR_{incr} with the WACC of the investor, i.e., if $\text{IRR}_{\text{incr}} > \text{WACC}$, the proposed investment is favorable, whereas $\text{IRR}_{\text{incr}} < \text{WACC}$ indicates an economic loss. IRR and NPV can be used in a complementary way to capture a more comprehensive outlook on the investment potential.

The Break-Even Period describes the time period (e.g., month or year) where the cumulated and discounted revenues exceed the cumulated and discounted costs, as shown in Fig. 2.4 (a). The supporting rationale is that the earlier the BEP occurs, the better it is for the investor (due to the aforementioned time value of money). A frequently mentioned disadvantage of the BEP is that cash flows after the BEP are not considered at all.

2.1.3 Long-Term and Cross-Stakeholder Methods

The third and last category considers cash flow elements throughout the aircraft's whole lifecycle. To do so, the different stakeholder perspectives have to be brought together. Such analyses are usually conducted by the manufacturer, who incorporates the airline's operating costs in order to optimize the pricing strategy and/or design a cost-effective and competitive product.

WLCC: The Manufacturers' Point of View

The whole lifecycle cost approach was first introduced by the US Department of Defense after realizing that the majority of product costs incurred during operations and maintenance are defined earlier in the design phase [94]. WLCC are typically composed of CERs due to their quick evaluation while attempting to incorporate relevant aspects and cost drivers of each lifecycle phase from "*cradle to grave*". Depending on the sector, the "*cradle*" begins with an idea (e.g., for general studies), a design (e.g., for Design for X studies), or the collection of raw materials (e.g., for environmental Lifecycle Analysis (LCA)). A common lifecycle breakdown from an economics perspective considers four phases: (1) design, (2) production, (3) usage,

and (4) end of life [19, 77]. Figure 2.5 lists the specific work for each respective phase and gives an indication of the share of WLCC spent.

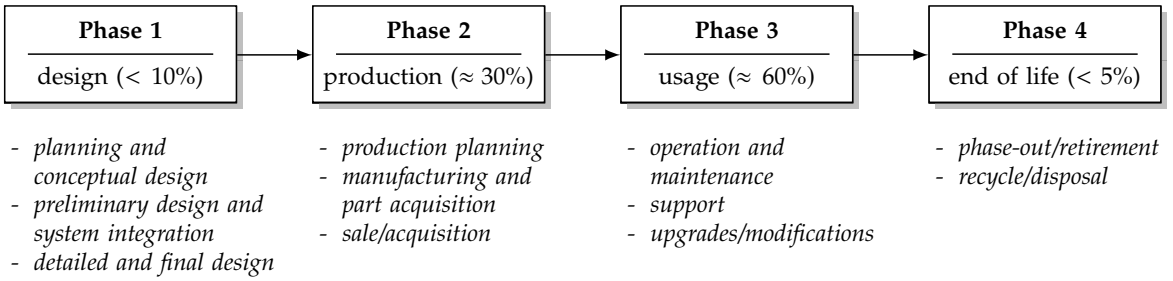


Figure 2.5 Phases of an aircraft's lifecycle (based on Fabrycky [77, p. 13]), their corresponding cost elements and an indication of their share (based on Raymer [211, p. 504] and US Department of Defense [274, p. 5]).

Phase 1, Design: This phase initiates with research to define product needs and future aircraft capacity, progressing through conceptual, preliminary, and detailed design stages. Ground and flight tests are conducted throughout to validate design choices. Non-recurring costs, predominantly Research, Development, Test, and Evaluation (RDTE), occur here, constituting less than 10 % of the WLCC, despite shaping 70 to 80 % of it [211].

Phase 2, Production: Once the design is finalized, the production phase commences. This phase entails meticulous planning and manufacturing or acquisition of inventories, materials, tooling, and other production-related elements. Some elements are recurring (e.g., materials), while others are non-recurring (e.g., manufacturing machines).

Phase 3, Usage: The third phase begins with the sale of the product to the operator. It primarily involves operating, maintaining, and, if applicable, modifying the product to meet evolving requirements. For aircraft, this phase can extend over 30+ years and constitutes the majority of the WLCC.

Phase 4, End of Life: Eventually, the physical life of the product concludes, marking the beginning of the end-of-life phase. For aircraft, this phase typically entails engaging a third party for disassembly, (partial) recycling, and disposal. Despite its environmental significance, this phase often holds negligible economic relevance.

WLCBA: Extension of WLCC by Revenues

With the inclusion of stakeholder-specific revenues, the WLCC becomes a WLCBA, enabling a direct and simultaneous profitability assessment for the manufacturer and operator. Here, the main link between both stakeholders is the aircraft price. If it is too high, the customer is unlikely to buy the product. If it is too low, the manufacturer will likely not offer the product, as the internal efforts outweigh the revenue. This simultaneous consideration of both parties represents the essential added value of WLCBA, allowing for a product design that is feasible for both the manufacturer and operator. While WLCBAs provide interesting and useful insights, they require an integrated model of each stakeholder, which entails a thorough level of understanding of all relevant processes and interconnections. Some manufacturers, such as Airbus, are known to simulate fleet operations and feed the insights back into the design [288]. However, studies of this kind are rare in aeronautic research.

2.1.4 Assumptions, Bias, and Cost Benchmarking

After reviewing the available methods in TEA and their purposes, some practical aspects are discussed next. These include the distinction between assumptions and uncertainties as well as a description of the general TEA process, including the cost benchmarking approach.

Assumptions and Uncertainties include parameters, scenarios, circumstances, or modeling details and occur naturally in all types of simulations as they imitate a real-world system, which is impossible to model perfectly. In integrative and complex analyses, their number, reason, limitation, and potential impact can quickly become confusing and seriously limit correct interpretation. Depending on which question the analysis is supposed to answer, the choice of assumptions may or may not be appropriate and justified. The distinction between an assumption and an uncertainty depends strongly on the definition of the assessment's system boundary. Thus, a specific parameter may be an assumption in one study but treated as an uncertainty in another. Deterministic IPAs are those that consider only assumptions, whereas a perfect probabilistic analysis would perform an uncertainty analysis on all assumptions. This is, however, prohibitive in terms of human and computational effort. Therefore, probabilistic IPAs have both assumptions and uncertainties. To differentiate between them, the following guidelines are given:

Consider an aspect X that is initially classified as an assumption. If, to the best of one's knowledge, a change in X does not affect the design or performance of the OoI, X can likely be treated as an assumption, i.e., assigned a deterministic value. An example would be the airframe maintenance cost if an engine-specific technology is being assessed. If, however, the OoI has some form of physical or other connection with X, it is more appropriate to treat it as an uncertainty, e.g., the fuel cost development for a fuel saving technology. A subsequent sensitivity analysis can then prove whether this assignment is appropriate or not. In practice, this distinction is not always easy to make, especially when dealing with highly complex models.

Intentional and Unintentional Bias add another layer of complexity when dealing with assumptions and uncertainties. Intentional biases, where individuals might prefer certain alternatives over others, can affect the choice of assumptions and skew the analysis in favor of the individual's interests. While this type of bias is generally easier to identify and subsequently mitigate, unintentional bias is more difficult to detect [107, pp. 2-3]. Thus, in both deterministic and probabilistic IPA, it is crucial not only to differentiate between assumptions and uncertainties but also to account for and mitigate the effects of biases (e.g., by following a well-documented and transparent process as well as verifying and validating the models) to ensure the reliability of the analysis.

Cost Benchmarking describes a method where a prospective TEA, which may be difficult or impossible to validate directly, can still yield meaningful results by comparing outputs to a reference case. For instance, consider the economic valuation of an aircraft with a newly developed technology. The absolute NPV depends on various factors such as company policies, political dynamics, and global economic conditions, making nearly every input parameter uncertain and reducing the practical value of the TEA. By simulating a reference aircraft under identical boundary conditions, the effects of technology-independent uncertainties offset each other. This not only enhances the robustness of the resulting outputs but also reduces the number of elements requiring assessment for assumptions or uncertainties.

2.2 Uncertainty Quantification Fundamentals

In Chapter 1, UQ was briefly introduced, laying the initial groundwork by outlining foundational concepts and introductory methodologies. However, the complexity inherent in this subject necessitates a more thorough investigation, particularly regarding the multifaceted nature of uncertainty in engineering endeavors. Scholars have grappled with defining and characterizing uncertainty, using terms such as "*ignorance*" (Dempster [60, p. 206]), "*that what is not precisely known*" (Booker et al. [30, p. 670]), and the "*intrinsic absence of necessary knowledge*" (Lohweg et al. [151, p. 193]). This diversity of terminology underscores the various sources from

which uncertainty arises, including flawed assumptions, lack of knowledge about input parameters, measurement errors, and the inherent variability of physical events. Building upon the introductory discussions, this Section aims to expand and enrich our understanding of uncertainty by delving deeper into its definition, exploring potential sources, and examining mathematical representations and engineering techniques for modeling it.

2.2.1 Definitions and Distinctions

In the context of this thesis, the following definition of uncertainty from Walker et al. [281] is used:

> ***Uncertainty*** *is any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system.*

In other words, everything that prevents a situation where one single value *fully* represents the truth is considered an uncertainty. This definition is fairly general, as it is based on the distinction from deterministic knowledge. For further categorization purposes, the three dimensions proposed by Walker et al. [281] are explained next, i.e., the *nature*, *location*, and *level* of uncertainty.

Nature of Uncertainty

As briefly described in Chapter 1, a common classification of uncertainties involves a differentiation between aleatory and epistemic uncertainty. The class of *deep* uncertainties completes this breakdown, as shown in Fig. 2.6.

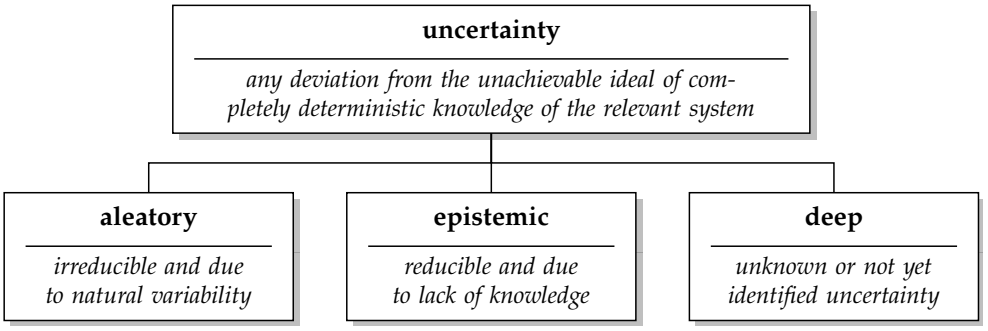


Figure 2.6 The nature of uncertainty according to Walker et al. [281].

Aleatory Uncertainty (also called irreducible, stochastic, or variability uncertainty) is used to categorize uncertainties that exist due to their inherent variability and are in principle not reducible with more effort [249]. For example, the frequency of how often aircraft surfaces experience rain can be obtained by statistics to a certain degree, but it will never be possible to reduce this forecast to a single, invariably true value, e.g., every three days. Another, more general example is a manufacturing process producing a part with a nominal length of 0.5 m. Due to the imperfection in the process, the actual length of any given part varies, which is known as manufacturing tolerances Roy et al. [225, p. 2132]. With a sufficiently large number of measured parts, a frequency distribution can be obtained, describing this uncertainty. For a given set of process settings, this uncertainty is considered irreducible, i.e., aleatory. Aleatory uncertainties are typically unbiased and more naturally represented in probability theory [249].

Epistemic Uncertainty (also called reducible, ignorance-based, or knowledge uncertainty) is used to categorize uncertainties that exist due to a *lack of knowledge*, i.e., epistemic uncertainty may be reduced by more research and empirical efforts. For instance, the impact of a fuel-saving technology may be estimated using a low-level method such as the Breguet Range Equation.

This estimation can then be improved upon with respect to its uncertainty by using a more sophisticated tool that is ideally validated with real aircraft performance data. Alternatively, in the above mentioned example of a part with a nominal length of 0.5 m, epistemic uncertainty would exist if only one or two parts were measured. Additional effort (without changing the process settings) in terms of measuring more parts would reduce this uncertainty. Epistemic uncertainties can be biased and are less naturally represented in probability theory [249].

Deep Uncertainty describes the situation where, for instance, a parameter not only has a deviation from the ideal deterministic true value but where the analyst "*does not even know what he does not know*" (Walker et al. [281, p. 13]) about that parameter. An example is coding mistakes in a simulation (which are usually corrected once identified). It should be noted that the efforts spent to reduce epistemic uncertainties may reveal deep uncertainties, potentially increasing total uncertainty.

It should be noted that this distinction is not mutually exclusive, i.e., uncertainties may be purely aleatory, purely epistemic, purely deep, or a mixture of any of these.

Location of Uncertainty

An overview of potential sources of uncertainty can help identify which parameters, attributes, assumptions, or, more generally, which pieces of information are prone to uncertainty. Literature provides many different classifications depending on the context of the analysis. If, for instance, measurement processes are the focus, the consideration of systematic and random errors in the test setup may prove to be sufficient. In other cases, a specific list of sources with respect to the involved process is compiled, as shown in Wen et al. [283]. A less structured but more comprehensive approach from Booker et al. [30] lists "*observation, measurement, recording, poorly understood initial conditions, random effects, uncontrollable effects, unknown effects, incomplete information, lack of knowledge, vagueness, ambiguity, (imperfect) physical, mathematical, statistical, and computational models, (imperfect) decisions, interpretations, extrapolation, interpolation, prediction, inferences being made, and conflicts among data, models, tests, and experiments*" as potential sources. The following categorization is loosely based on Ewing et al. [76] and Walker et al. [281] and aims to fit the nature of integrative TEA without being restrictive towards other applications. It should be noted that these clusters are not necessarily mutually exclusive either, as one uncertainty can be located in multiple places.

Model Inputs can be either (a) the parameters fed to the model, e.g., information about the product such as its price or the number of units sold, or (b) other submodels that are connected to the TEA tool. These submodels are either introduced *as is* (e.g., if the TEA framework has a modular structure) or via surrogate models (e.g., via response surfaces, lookup tables, regression functions, etc.). The uncertainties within these models have their own classes, i.e., model context and model form (see below).

Assumed Scenarios and Predictions involve a variety of parameters. In aircraft IPA, for instance, the fuel price's often plays a major role. Predicting the fuel price development over the course of the next few decades is inevitably uncertain. Another example is the simplification of boundary conditions, such as the aircraft's utilization or operating environment.

Model Context represents the scope of the chosen model and whether (and to what extent) it is suitable to capture all relevant aspects of the real world that it intends to simulate. This is comprised of the inclusion of all relevant stakeholders as well as the temporal and spatial dimensions of the model. An ill-chosen model context may falsely neglect relevant parameters and effects and hence require well-documented and accepted simplifications.

Model Form deals with the structure of the model and whether the level of detail in the system mapping is sufficient for the analysis. A lack of interrelationship modeling or system behavior representation may ultimately lead to ill-advised decision-makers and is hardly compensated by mentioning the simplifications made. Compared to the model context domain, the model form focuses on interaction modeling, whereas the context deals with the completeness of the model input space.

Technical Implementation of the Model refers to the accuracy of software and hardware, e.g., discretization errors, floating point errors, and sensor accuracy. Depending on the scale of the uncertainty and its propagation, the effect on the output uncertainty can range from being negligible to significantly altering the result.

Level of Uncertainty

The level dimension essentially describes *how high* the uncertainty is. Various classifications exist, aiming to aid researchers in qualitatively categorizing the current state of uncertainty. Therefore, they do not intend to prescribe an exact methodology to assign a specific state but rather promote effective communication by handling and expressing uncertainty in a transparent way. The following description is loosely based on the definition from Walker et al. [281] and its extension by Kwakkel et al. [142].

Level 1 (shallow uncertainty): The lowest level of uncertainty is present if data is plentiful and probability distributions can be assigned with high confidence. Uncertainties at this level are usually dealt with using probability theory. An example would be a sensor recording of thousands of landings, out of which a PDF can be created and a probability for the next flight to experience a harsh landing can be calculated.

Level 2 (medium uncertainty): Parameters can be classified as having medium uncertainty if their probabilities cannot be expressed with a continuous function but can only be ranked, e.g., from *unlikely* to *likely* to *very likely*. This is often used in scenario-based analyses and is best quantified with non-probabilistic theories, although probability theory provides some methods for modeling these to some degree.

Level 3 (high uncertainty): When uncertainties cannot be ranked but are only listed, their level is considered to be high. This often occurs when the understanding of the problem is very limited [205]. These uncertainties are hard, if not impossible, to quantify. Instead, researchers facing high uncertainties are advised to analyze the sensitivity of the model to these parameters and, if necessary and possible, acquire new knowledge to reduce the level of uncertainty to medium.

Level 4 (recognized ignorance): The highest level of uncertainty exists if no parameter alternatives can be listed. This classification is valid when researchers know that uncertainty is present but cannot describe it further. An indication for this would be if a future change in this parameter is not considered a surprise. Uncertainties at this level cannot be quantified without a transfer to lower levels, e.g., through basic research. Basic research refers to scientific inquiry aimed at understanding fundamental principles and phenomena without immediate application in mind. In the context of uncertainty, basic research may involve exploring the underlying causes of uncertainty, developing new theories or methodologies for characterizing uncertainty, or conducting experiments to gain deeper insights into the factors contributing to uncertainty.

Depending on the comprehensiveness of the study, some authors include *deep uncertainty* as a final level. As these cannot be listed and thus do not serve the intended purpose of communication and transparency, they have been omitted from this list.

2.2.2 Theories and Techniques

Uncertainty can be expressed in various ways. In everyday life, uncertain information is expressed linguistically and often (but not necessarily) accompanied by indicator words such as *about* or *roughly*. In more technical areas, a numerical representation is often found, typically using a nominal value and an interval, e.g., 88 kg±1.8 kg. Mechanical engineers are familiar with asymmetric tolerances such as $420^{+3.5}_{-2}$mm, and typical scientific uncertainty representation involves statistical parameters such as the mean and standard deviation. This Section discusses some fundamentals of mathematical theories and techniques dealing with uncertainty, covering conventional probabilistic approaches as well as non-probabilistic ones. It should be noted right away that these presented theories are not necessarily competitive but (depending on the subject) complementary [223].

Probability Theory: Theoretical Aspects

There is no doubt that the most frequently used theory for coping with uncertainties in science is ProT [302]. This is not surprising, as it is the one with the longest history and the only one being taught in most educational systems.

In ProT, the uncertainty of an event A is represented by a single value measuring the probability of its occurrence, $P(A)$. Thereby, $P(A)$ can be interpreted in different ways. The two main ones, causing somewhat of a philosophical debate [223, pp. 57-63], are: (a) the relative Frequentist perspective and (b) the subjective Bayesian perspective. Their fundamentals are introduced next, followed by a discussion of common techniques stemming from ProT.

The Frequentist Interpretation defines the probability $P(A)$ as the fraction of times that the event A occurs if the situation is repeated an infinite number of times [151, 302], i.e.,

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n} \quad . \tag{2.3}$$

Consider, for example, the punctuality of a flight, which represents hypothesis A. In the Frequentist perspective, past data would show that $n(A) = 87$ out of $n = 100$ flights were on time, resulting in an on-time probability for the next flight of $P(A) = 87\,\%$. $P(A)$ is interpreted as a *frequency* generated by some process and is considered "*. . . a measure of an empiciral, objective, and physical fact of the external world, independent of human attitudes, opinions, models, and simulations*" (Ross [223, p. 58]). Predictions of the future state of an event are made using only the data from the current experiment. Transferred to the types of uncertainty, the Frequentist interpretation is generally linked to the aleatory type (given that a representative sample size is available [151]). The Central Limit Theorem (CLT) and LLN are pivotal to this interpretation. Although highly interrelated and often used interchangeably, these are two different theorems. The CLT states that the distribution of the mean of a sample tends to be *normal* as the sample size increases, regardless of the variable's distribution in the population. This is commonly accepted for sample sizes greater than 30 [222]. The LLN also describes the sample behavior with increasing size, i.e., it states that the mean value of the sample approximates the mean value of the population. Both theorems are utterly important for the inference of information obtained by a sample onto the population. Recalling the information-sparse environment of economic studies in aeronautics, the sample size requirements of the CLT and LLN are not always met, which has to be considered when choosing appropriate UQ metrics [243].

The Bayesian Interpretation, known as the subjective approach, defines the probability $P(A)$ as the degree of belief regarding the occurrence or truth of event A. [25]. Hence, it is usually associated with epistemic uncertainties. The subjective assignment is based on the available

background knowledge and is be updated as soon as new evidence becomes available. The basis of this interpretation is the Bayes' theorem:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)} \tag{2.4}$$

where

A	is the event or hypothesis in question,
B	is additional information,
$P(A)$	is the *prior*, i.e., how likely the event A is, regardless of evidence,
$P(B)$	is the *evidence*, or how likely the B is, regardless of A,
$P(B \mid A)$	is the *likelihood*, i.e., assuming A occurred, how likely is B, and
$P(A \mid B)$	is the *posterior*, i.e., the probability of A occurring given that B has occurred.

Using the same flight punctuality example, suppose the prior probability $P(A) = 87\,\%$. With new information B, such as unfavorable weather, this belief can be dynamically updated. Assuming that past weather data analysis results in $P(B) = 10\,\%$ and further flight data analysis shows that among the on-time flights, $5\,\%$ experienced severe weather, i.e., $P(B \mid A) = 5\,\%$. This results in a posterior of

$$P(A \mid B) = \frac{0.05 \cdot 0.87}{0.1} = 0.435 \quad . \tag{2.5}$$

In other words, the updated belief of an on-time flight, given the new information of severe weather, is now 43.5 %.

It should be noted that neither interpretation is inherently superior to the other; each has its contexts where it is more applicable. The Frequentist approach is favored when data is plentiful, allowing for reliable frequency-based inferences, and is generally easier to implement, which makes it better understood by decision-makers and analysts [220, p. 7]. However, it tends to produce less conservative results and may give a false sense of exactness, particularly when underlying assumptions are not met [22, 101]. Conversely, the Bayesian perspective is suggested when data are sparse [91, p. 32], allowing for the integration of prior knowledge and continual updating of beliefs as new evidence becomes available. This adaptability can make Bayesian methods appear more subjective but provides significant advantages in contexts where prior information is crucial or when data acquisition is challenging. Ultimately, the choice between these approaches should be guided by the specific needs of the analysis, the availability of prior information, and the analytical context.

Probability Theory: Techniques for UQ

ProT provides a plethora of techniques to deal with uncertainties. The following descriptions focus on those that are either most commonly used or are fundamental for understanding the later analyses.

Expressing Uncertainty in ProT is, in its simplest form, an error term ϵ [78, 207], i.e.,

$$x = X + \epsilon \tag{2.6}$$

where x is the observed value and X is the true value [129, p. 171]. Thereby, ϵ is often broken down to systematic (bias) error and random effects. While this approach is useful for many applications such as laboratory experiments and measurement uncertainty, it cannot capture the diversity of uncertainties in more complex engineering applications. More elaborate alternatives include visual ones such as histograms or kernel density estimations as well as quantitative measures such as the standard deviation or variance. The latter is a measure

of the average degree to which each point (e.g., measurement) differs from the mean. More specifically, it is calculated via

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \quad \Rightarrow \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \quad . \tag{2.7}$$

Here, σ^2 is the variance, σ the standard deviation, x_i represent the observed values, μ is the mean of the population, and n is the total number of observations[2].

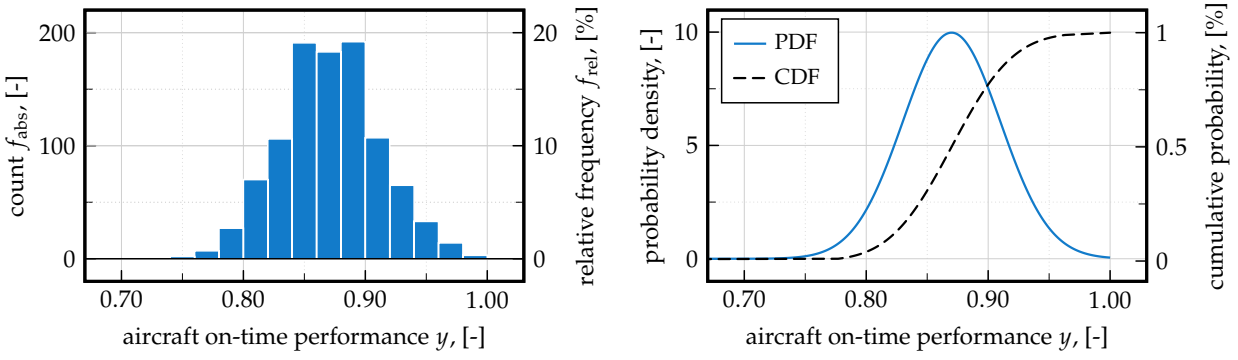
Histograms and statistical distribution functions allow for a qualitative and quantitative interpretation of the results and the underlying uncertainty. Histograms, for example, are used to visualize the outputs y of an experiment that is repeated n times, where the abscissa represents the range of y, divided into bins of the width Δy, and the ordinate shows the number of results falling within each bin, i.e.,

$$f_{\text{abs}}(i) = n(y \in [y_i, y_i + \mathrm{d}y]) \quad \text{and} \quad f_{\text{rel}}(i) = f_{\text{abs}}(i)/n \quad . \tag{2.8}$$

Here, f_{abs} and f_{rel} describe the absolute and relative frequency, respectively. The shape of the histogram allows for an estimation of the spread, symmetry, and outliers of the data. An example is shown in Fig. 2.7 (a). Provided that the data is continuous in nature, a more enhanced view on the distribution is accessible via PDFs. PDFs describe the likelihood of a random variable taking on a particular value or being in a particular interval of interest. The latter is represented by the area under the PDF curve, i.e.,

$$P(y \in [y_{\text{lower}}, y_{\text{upper}}]) = \int_{y_{\text{lower}}}^{y_{\text{upper}}} f(y)\mathrm{d}y \quad . \tag{2.9}$$

The CDF, which is essentially the integral of the PDF, allows for a quantification of the probability of the output being greater or less than a threshold value. An example PDF and CDF are shown in Fig. 2.7 (b).



(a) absolute and relative histogram.

(b) PDF and CDF.

Figure 2.7 Histogram and statistical distribution functions for a fictive dataset on aircraft on-time performance.

Conventional Sensitivity Analysis, also called local sensitivity analysis, is used to gain insights regarding a model's response when one or a few input parameters deviate from a nominal value. Consider a (mathematical or numerical) model

$$y = f(x_1, x_2, \ldots, x_n), \tag{2.10}$$

[2]Supplemental concepts and definitions from ProT can be found in Appendix B.2.

where one or more input parameters are uncertain, e.g., may take any value within a particular range,

$$x_i \in \left[x_{i,\min}, x_{i,\max}\right]. \tag{2.11}$$

In LSA, the sensitivity of the model output is calculated by sampling an array of n equally spaced values for x_i and evaluating the model n times, once for each sample. The sensitivity of the model towards a change in x_i can then be calculated using the derivative, typically evaluated at a nominal value x_i^* (e.g., its mean):

$$\left.\frac{\delta f}{\delta x_i}\right|_{x_i=x_i^*} \approx \frac{\Delta y}{\Delta x_i} = \frac{f(x_{i,\text{upper}}) - f(x_{i,\text{lower}})}{x_{i,\text{upper}} - x_{i,\text{lower}}} \tag{2.12}$$

For more than one uncertain parameter, LSA typically follows a one-at-a-time approach, i.e., it varies one input while keeping the others fixed. This is often accompanied by an Δoutput vs. Δinput visualization, as illustrated in Fig. 2.8.

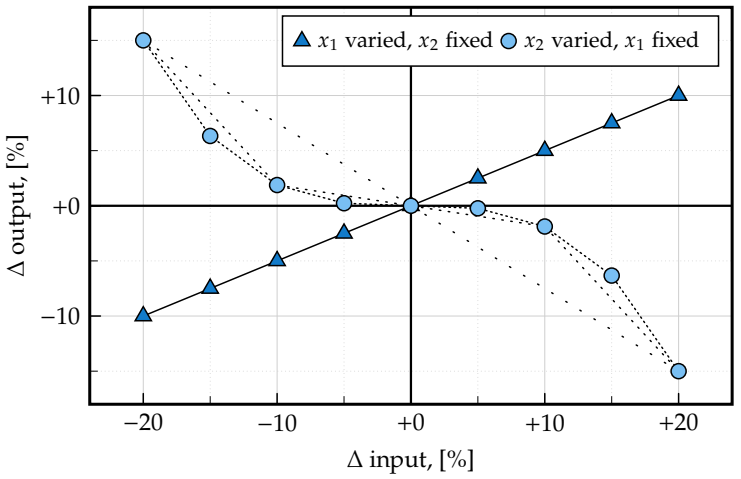


Figure 2.8 Example of LSA of a two parameter model.

While LSA are highly computationally efficient, requiring only a minimal number of model evaluations, they present several shortcomings. The sensitivity measure is significantly dependent on the choice of the nominal value x_i^*. For instance, as demonstrated in Fig. 2.8, the sensitivity for x_1, whose effect on the output is linear, is straightforward. Conversely, the sensitivity for x_2 may be nearly negligible if x_2^* is chosen, but considerable for other nominal values. Another limitation concerns the bounds for input variations, typically set symmetrically, e.g., ±15 %. This symmetric approach does not account for the likelihood that parameters will actually vary within these bounds, as some ranges may be more plausible than others, or even unrealistic. This limitation underscores why the analysis is termed *local*: it only considers variations near a nominal value, thereby restricting its applicability to this specific region [232, p. 1509]. Furthermore, LSA is restricted to analyzing additive effects when multiple uncertain variables are involved. By keeping other parameters at their nominal values, it is impossible to examine interactions between high and low values across different uncertainties, thus missing out on potential non-additive interactions. This is especially problematic when considering variables like x_2, where the output's response is nonlinear. The sensitivity analysis might differ substantially if only the extremes are considered compared to a more granular analysis within the interval [−20 %, +20 %], as illustrated by the varying dotted lines in Fig. 2.8. Despite these flaws, LSA remain widely used, often without recognition of their limitations. As Saltelli et al. suggest, this widespread but uninformed usage can lead to underestimating uncertainties and inaccurately estimating sensitivities, illustrating a misapplication of sensitivity analysis [232, 238].

Global Sensitivity Analysis can be described as "*how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input*" (Saltelli [239]). As mentioned by Helton et al. [102], GSA "*should be a fundamental part of any analysis that involves the assessment and propagation of uncertainty*". In contrast to the one-at-a-time approach of LSA, GSA considers all uncertain parameters simultaneously. Furthermore, instead of sampling equally distanced vectors, the probability distributions of the input parameters are taken

into account in GSA. This renders them more realistic, as likely samples of x_i occur more frequently than unlikely ones. Lastly, it should be mentioned that GSA does not assume linearity in the output vs. input behavior. Depending on the method, either the entire distribution or selected statistical moments (e.g., variance) are used.

The goal of GSA is to break down the total output uncertainty into its constituents, i.e., to quantify what share of the uncertainty is caused by each parameter x_1, x_2, their interaction (x_1, x_2), and so on. This is shown illustratively in Fig. 2.9. Provided sensitivity measures in GSA are therefore interpreted as *how critical the uncertainty per variable is*, whereas LSA sensitivities quantify *how sensitive the output reacts to a predefined change in the inputs*.

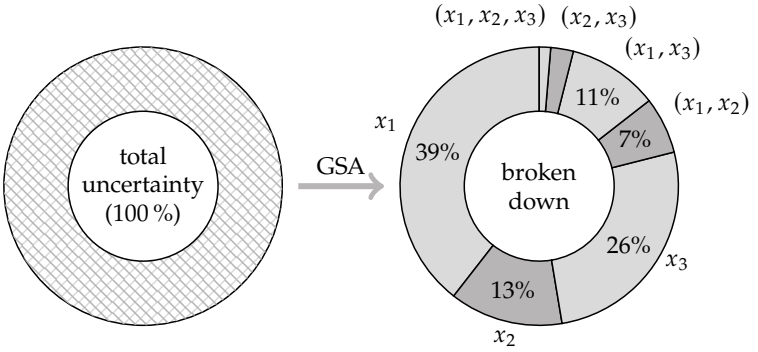


Figure 2.9 Example of a breakdown of total uncertainty into its first, second and third order effects for a function with three uncertain parameters.

A more in-depth investigation into the actively researched field of GSA reveals that there are a wide variety of different techniques, some of which are developed based on differing theoretical definitions of sensitivity. As a consequence, using different methods on the same problem may lead to significantly different, even conflicting, sensitivity results [64, p. 2]. A number of studies have been devoted to discuss GSA methods from different perspectives, e.g., Pianosi et al. [193], Razavi et al. [212, 213], and Saltelli et al. [236]. Considering that GSA have a yet untapped potential [213, p. 6], Chapter 4 will deal with this topic in the context of uncertainty-enabled TEA in more detail, which is why further details are omitted at this point.

Monte Carlo Simulations (MCSs) MCSs are widely used to propagate uncertainties through a system [243]. They are classified as non-intrusive methods, i.e., they do not interfere with the inner part of a model as they are defined on the outer level, making them particularly popular in black-box systems [220]. The idea is to draw random input samples from probability distributions and feed them to the model at hand in a sequential manner, as schematically shown in Fig. 2.10. Consider a general function or model $y = f(x_1, x_2, \ldots, x_n)$, where two parameters x_i and x_j are considered to be uncertain. As a first step, the distributions of these uncertainties need to be defined. For simplicity reasons, assume that x_i and x_j follow a uniform and normal distribution, respectively, i.e.,

$$x_i \sim \mathcal{U}\left(x_{i,\min}, x_{i,\max}\right) \quad \text{and} \quad x_j \sim \mathcal{N}\left(\mu_{x_j}, \sigma_{x_j}\right). \tag{2.13}$$

In step two, scalar samples are drawn from these distributions and fed to the model. This ensures that the model itself is, in a figurative sense, unaware that some variables are uncertain, i.e., it only receives deterministic values. The model execution is repeated n times in step three (with a new sample being each time) until the output converges. Here, convergence is defined by the analyst, as the true value is often unknown and can only be calculated if n approaches infinity. In practice, n is rarely chosen systematically. Instead, it is chosen based on the available computational budget. Once (assumed to be) converged, the output distribution is analyzed in terms of the mean value, standard deviation, and occasionally additional statistical moments in step four. The output analysis is typically visualized by a histogram and an empirical cumulative distribution function.

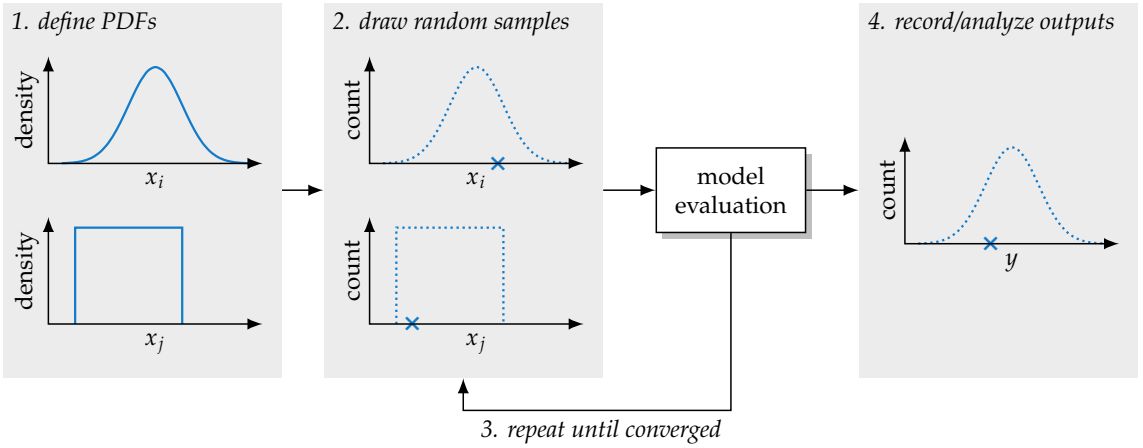


Figure 2.10 Schematics of a monte-carlo simulation.

Modifications of the Monte Carlo Method include, but are not limited to, quasi-random sampling techniques, nested MCSs, second-order probabilities, and the Markov Chain Monte Carlo (MCMC) method. The former aims to lessen the computational expense of MCS since the CLT and LLN require a large number of samples to be drawn and propagated through the model in order to achieve statistically reliable results [220]. Quasi-random sampling techniques such as Latin Hypercube Sampling (LHS) can improve the efficiency and convergence of the MCS by distributing the drawn samples more evenly across the input space compared to purely random sampling. This can reduce the overall number of required samples and, consequently, execution time. Nested MCS, on the other hand, are useful for solving problems with multiple layers of uncertainty. This makes them invaluable in hierarchical systems or multistage problems, enhancing the representation of system behaviors and the interaction of risk variables. Second-order probabilities can be seen as an extension of MCS as they deal with uncertainty in the PDFs, e.g., when the mean and variance of a normally distributed input variable are uncertain themselves. This process, essentially a "simulation within a simulation", yields a distribution of outcomes for each iteration, providing a deeper understanding of variability. MCMC simulations use a process where a Markov chain is utilized to sample effectively from complex probability distributions, with the samples becoming representative of the distribution after many chain transitions. This approach is particularly useful in high-dimensional spaces, commonly seen in Bayesian statistics and machine learning Richardson [217].

Evidence Theory: Theoretical Aspects

The first non-probabilistic theory discussed here is evidence theory, which was first introduced by Dempster [61] and later extended to a mathematical theory by Shafer [245], which is why it is often referred to as Dempster-Shafer Theory of Evidence [135]. Compared to ProT, evidence theory is relatively new, and techniques using it are rare. Furthermore, it should be noted that evidence theory may be more difficult when communicating results with decision-makers due to its limited applications [30, 101, 180].

Essentially, DSTE is a generalization of ProT and aims to explicitly model the *lack of knowledge* [26, pp. 477-479]. It is related to the Bayesian interpretation of ProT in the sense that both deal with subjective beliefs [287] rather than objective frequencies. The theory's primary premise is that information (or evidences) for a given hypothesis can be split into those that support it and those that contradict it. The supporting evidence is termed *Belief* whereas all evidences that do not contradict the hypothesis are called *Plausibility*[3]. These two concepts form the theory's representation of uncertainty, as illustrated in Fig. 2.11.

[3]To differentiate the usage of the terms "*Belief*" and "*Plausibility*" as key concepts in DSTE from their use in common parlance, these terms are capitalized throughout this thesis.

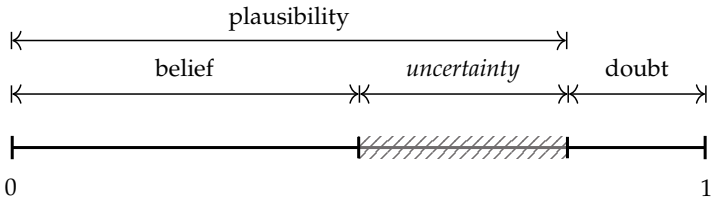


Figure 2.11 The Dempster-Shafer uncertainty interval, based on Worden [287].

In order to quantify Belief and Plausibility for a given hypothesis, one prior step needs to be taken. That is, all possible hypotheses are assigned a so-called *mass*. This mass, denoted as m, is subjective and can represent the belief, confidence, or trust that this hypothesis is true. Its assignment is a function formally defined on 2^Ω, in which Ω is the frame of discernment representing the universe of all possible and mutually exclusive events Θ_i. The power set 2^Ω describes all possible hypotheses that can be made with respect to Ω. Say, for example, Ω contains two elements Θ_1 = on and Θ_2 = off, its power set would be:

$$\Omega = \{\text{on}, \text{off}\} \tag{2.14}$$

$$\Rightarrow 2^\Omega = \{\emptyset, \{\text{on}\}, \{\text{off}\}, \{\text{on} \vee \text{off}\}\}\,. \tag{2.15}$$

Now, the function which assigns a value between 0 and 1 to each element A in 2^Ω, i.e.,

$$m : 2^\Omega \rightarrow [0, 1], \tag{2.16}$$

is called mass assignment function, provided that the following two requirements are met:

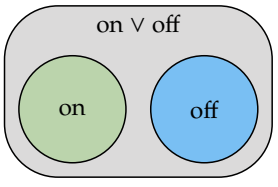
$$m(\emptyset) = 0 \quad \text{, i.e., the null hypothesis receives no mass, and} \tag{2.17}$$

$$\sum_{A \in 2^\Omega} m(A) = 1 \quad \text{, i.e., the sum of all masses equals one.} \tag{2.18}$$

If, for example, the overarching problem is to determine the state of a computer (matching the frame of discernment in Eq. (2.15)), one can assign masses to each element as specified in the second column of the table in Fig. 2.12. Note that the null hypothesis is given no mass by definition, and the last hypothesis, i.e., "on ∨ off", represents the "indeterminate" answer, i.e., the system could be either on or off. The Belief and Plausibility columns of the table in Fig. 2.12 are a result of this mass assignment, and the underlying concept is explained next.

hypothesis	mass	Bel	Pl
$\emptyset$	0	0	0
on	0.6	0.6	0.7
off	0.3	0.3	0.4
on ∨ off	0.1	1.0	1.0

(a) Mass assignment and resulting Belief and Plausibility



(b) Visualization of possible sets within Ω.

Figure 2.12 Tabular summary and visualization of the frame of discernment Ω as specified in Eq. (2.15).

The Belief Function represents the *support* for a hypothesis $A \in 2^\Omega$. It is calculated as the sum of all masses of $B \in 2^\Omega$ which are a subset or are equal to A, i.e.,:

$$\text{Bel}(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \tag{2.19}$$

Considering Fig. 2.12 (b), the hypothesis "on" has no other subsets within the elements of 2^Ω, so the Belief for it is simply $\text{Bel}(A = \text{on}) = m(A = \text{on}) = 0.6$. The same principle applies for the hypothesis "off". The Belief for the hypothesis "on $\vee$ off", however, does have subsets present in 2^Ω, namely "on" and "off" with the masses 0.6 and 0.3, respectively. Thus, the supporting evidence for this hypothesis is

$$\text{Bel}(A = \{\text{on} \vee \text{off}\}) = m(A = \text{on}) + m(A = \text{off}) + m(A = \{\text{on} \vee \text{off}\}) \tag{2.20}$$

$$= 0.6 + 0.3 + 0.1 = 1.0 \tag{2.21}$$

The Plausibility Function answers the question to what extent the evidence is *consistent* with the proposition. It is formally defined as:

$$\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{2.22}$$

Here, the sum is taken over all masses of $B \in \Omega$ which intersect the hypothesis in question, i.e., A. For the hypothesis "on", the Plausibility is thus calculated as the sum of masses of its own hypothesis (being 0.6) as well as of the hypothesis "on $\vee$ off" (being 0.1). The same principle is applied for the Plausibility of the hypothesis "off". For the Plausibility of the last hypothesis, all three masses need to be added, as all sets intersect "on $\vee$ off" (see Fig. 2.12 (b)).

In a way, Belief and Plausibility form the lower and upper bounds of the conventional probability for a hypothesis [152, p. 250f]. Uncertainty in this probability exists, if Belief and Plausibility differ, as shown in Fig. 2.11. In other words, if there is no (knowledge) uncertainty, DSTE converges to ProT[4]

Evidence Theory: Techniques for UQ

Although evidence theory's use in engineering is relatively rare [243], there are some noteworthy techniques that are described briefly next.

Combination Rules are a useful tool within DSTE. Often times, multiple pieces of evidence have to be considered to evaluate a proposition. For example, multiple experts can rate their Beliefs on how valid an assumption is. To do so, different combination rules can be used. The most common one is formally defined as:

$$m_{12}(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C)} \tag{2.23}$$

and yields the combined evidence of two experts (1 and 2) that support A, which is composed of the intersection of B and C [107]. While this general combination rule is useful for simple cases, it tends to ignore conflict existing within the evidence due to the normalization in the denominator [3, p. 286]. To overcome this, additional combination rules have been developed, e.g., the one from Hester [107], which specifically deals with the credibility of different experts. A comprehensive guide on combination rules, including advantages and disadvantages, can be found in Sentz et al. [244].

[4]Supplemental interpretation aids for Belief and Plausibility are provided in Appendix B.3.

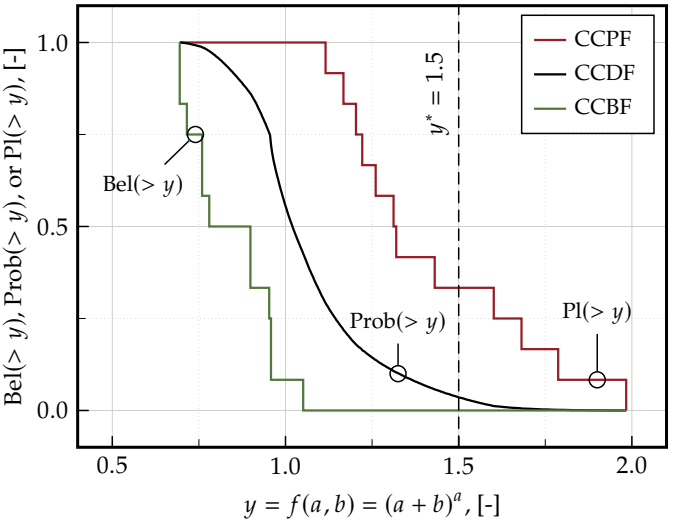


Figure 2.13 Calculated CCXFs of an exemplary function with evidences taken from Helton et al. [101, p. 52].

Complementary Cumulative Belief and Plausibility Functions represent DSTE's main technique for UQ and can be seen as an alternative to sampling input values from PDF, passing them to the model, and analyzing the output with respect to its statistical moments. The DSTE approach requires a modified sampling mechanism where, instead of generating PDFs from data, expert elicitation is used to create Belief spaces, which are then used to sample from. Additionally, a more complex post-processing analysis foresees using these Belief spaces, the drawn samples, and the results of the repeated model execution to calculate so-called Complementary Cumulative Belief Functions (CCBFs) and Complementary Cumulative Plausibility Functions (CCPFs) alongside the Complementary Cumulative Distribution Functions (CCDFs), as exemplified in Fig. 2.13. These functions, hereafter abbreviated as CCXFs, provide uncertainty bounds over the entire range of outputs and allow for statements that are threshold-driven, e.g., quantifying the probability as well as a lower and upper bound (represented by the Belief and Plausibility measure) of the output being greater than a threshold value y^*. For the shown example threshold, one interpretation of the CCXFs would be that the probability of exceeding $y^* = 1.5$ is 3.5 %, bounded on the lower and upper ends by a Belief of 0 % and a Plausibility of 33 %, respectively. While the computational expense has been mentioned as a limiting factor [101, 104], this DSTE-based UQ approach has been mentioned to be helpful when data is scarce but expert knowledge is available. As this can be the case for uncertainty-enabled IPA efforts, Chapter 5 is dedicated to this technique, aiming to understand its potential and added value in comparison to the ProT-based approach.

Possibility Theory and Fuzzy Set Theory: Fundamentals

Possibility Theory (PossT) is an alternative to ProT and was first introduced in 1978 by Zadeh [297]. Contrary to ProT, where uncertainty is rather interpreted as "*randomness*", PossT treats uncertainty as "*vagueness*" and is used when information is incomplete, imprecise, or inconsistent, which makes it difficult to assign a precise probability to an event.

Similarly to evidence theory, PossT uses two measures for uncertainty: *possibility* and *necessity*. The possibility measure describes the degree of feasibility of event occurrence, whereas the necessity measure describes to what degree the available evidence necessitates event occurrence. Both measures are explained below.

The Possibility Measure of a A (a subset of the universe of discourse[5] X) is defined as:

$$\Pi(A) = \sup\{\pi(x) : x \in A\} \tag{2.24}$$

where sup is the supremum operator, x is an element in A, and $\pi(x)$ is the possibility distribution $\pi : X \rightarrow [0, 1]$. In other words, the possibility of A is equal to the greatest possibility assigned to any element within A. Two extreme possibility states can be differentiated:

[5] The universe of discourse X is conceptually similar to the frame of discernment Ω used in DSTE. One fundamental difference is that the elements in X are not necessarily mutually exclusive.

- Complete knowledge: If the possibility of one element x_0 within A is one while the possibilities of all other elements within A are zero, i.e., $\pi(x_0) = 1$ and $\pi(x) = 0 \quad \forall x \neq x_0$, no uncertainty exists. In other words: If only one outcome is possible, it will happen.
- Complete ignorance: If every element x in A is possible, i.e., $\pi(x) = 1 \quad \forall x \in A$, there is a complete lack of knowledge regarding the outcome.

The Necessity Measure of $A \subseteq X$ is defined as:

$$N(A) = \inf \{1 - \pi(x) : x \in X \setminus A\}, \tag{2.25}$$

where inf is the infimum operator and $x \in X \setminus A$ describes those set elements that are part of the universe of discourse but are not part of the subset A. As an example, consider the universe of discourse $X = \{\text{sunny, cloudy, rainy}\}$ describing the future weather and a subset $A = \{\text{sunny, cloudy}\}$ representing a "*not rainy*" event. Now suppose a hypothetical possibility distribution of

$$\pi(\text{sunny}) = 1.0, \ \pi(\text{cloudy}) = 0.6, \text{ and } \ \pi(\text{rainy}) = 0.2. \tag{2.26}$$

These values are typically assigned subjectively, which is usually accepted in light of data scarcity and the consequential inability to use ProT. Note that this possibility distribution is normalized, meaning that there is at least one element with a Possibility of 1, which is typically done to reflect the idea that at least one event must occur, aligning the approach with our intuitive understanding of possibility. With this Possibility distribution, the possibility and necessity measures of A can be calculated as

$$\Pi(A) = \sup \{1.0, 0.6\} = 1.0 \quad \text{and} \quad N(A) = \inf \{1 - 0.2\} = 0.8 \quad . \tag{2.27}$$

Therefore, there is a high possibility of a non-rainy day. Because of the low possibility value for $x = \text{rainy}$, the necessity of having a non-rainy day is high as well. If the possibility of rainy weather had been higher, e.g., $\pi(x = \text{rainy}) = 0.6$, the Necessity of A (non-rainy day) would intuitively decrease to $N(A) = 1 - 0.6 = 0.4$.

While PossT provides a framework for handling vague and imprecise information, its application in the engineering domain is relatively scarce. Instead, the closely related fuzzy set theory is used[6]. Fuzzy set theory, first introduced by Zadeh [298] in 1965, is a mathematical framework that extends classical set theory to handle cases where information is incomplete, uncertain, or imprecise by using the concept of partial memberships. Fuzzy set theory addresses uncertainties that do not originate from randomness but from the "*absence of sharply defined criteria of class membership*" (Zadeh [298]).

Fuzzy Sets sets can be interpreted as an extension of the classical boolean sets (sometimes called *crisp* sets), which have the following characteristic or membership function μ_C:

$$\mu_C : \Omega \to \{0, 1\} \quad \text{so that} \quad \mu_C(\omega) = \begin{cases} 1 & \text{for } \omega \in A \\ 0 & \text{for } \omega \notin A \end{cases} \tag{2.28}$$

or in other words, an element ω out of a frame of discernment Ω does or does not belong to the set A, i.e., there is no inbetween. The membership μ_C of ω to A can therefore only have a value of either 0 or 1. In fuzzy sets, this membership function is allows values *between* 0 and 1, i.e.,:

$$\mu_F : \Omega \to [0, 1] \quad . \tag{2.29}$$

[6]For further information on possibility theory, consider the work from Dubois et al. [66–69].

To illustrate this difference, consider the following example. Someone is asked to provide the total number of adults in a region. Considering the classical set theory, the legal age can be used to determine whether a particular person is an adult or not, i.e.,

$$\mu_C(a) = \begin{cases} 1 & \text{for } a \geq 18 \\ 0 & \text{for } a < 18 \end{cases} \tag{2.30}$$

This is visualized in Fig. 2.14 as a step function. However, if the question is answered from a less legal and more societal or biological standpoint, the fuzzy set approach proves to be a more a useful concept. One potential and fuzzy membership function is:

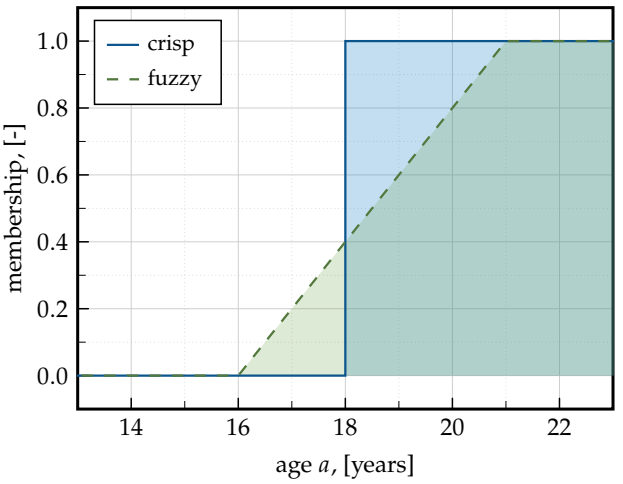


Figure 2.14 Crisp and fuzzy memberships of age to the set of adulthood

$$\mu_F(a) = \begin{cases} 1 & \text{for } a > 21 \\ (a-16)/5 & \text{for } 16 \leq a \leq 21 \\ 0 & \text{for } a < 16 \end{cases} . \tag{2.31}$$

So while a 17 year old individual is not an adult in the crisp set, the fuzzy set states that this person is an adult to a degree of 0.2 (or 20 %). This is illustrated as a ramp function in Fig. 2.14. The shape of the fuzzy membership function in this example was chosen to be a simple ramp. Depending on the user and problem at hand, other shapes such as triangular, gaussian, trapezoidal, or piecewise linear may be more suitable.

Possibility Theory and Fuzzy Set Theory: Techniques for UQ

Following the structure of Sections 2.2.2 and 2.2.2, the techniques for UQ related to PossT and fuzzy set theory are discussed briefly here.

Complementary Cumulative Necessity and Plausibility Functions are the equivalent of evidence theory's CCXFs and have been explored in Helton et al. [101] as an alternative and PossT-based approach for UQ. Similarly to the DSTE technique, this method foresees a modification of the sampling space as well as a post-processing step to calculate the Complementary Cumulative Necessityy Function (CCNecF) and Complementary Cumulative Possibility Function (CCPosF) alongside the CCDF. Fig. 2.15 illustrates the CCNecF, CCPosF, and CCDF using the same inputs and function as Fig. 2.13. While it is mentioned that under specific conditions this approach can yield the same uncertainty results as the DSTE counterpart [101, p. 66], there are some drawbacks of the PossT-based technique. If, for instance, the input sets are not consistent, the possibility distribution functions need to be scaled to meet the definition of possibility spaces. This can be done in either an

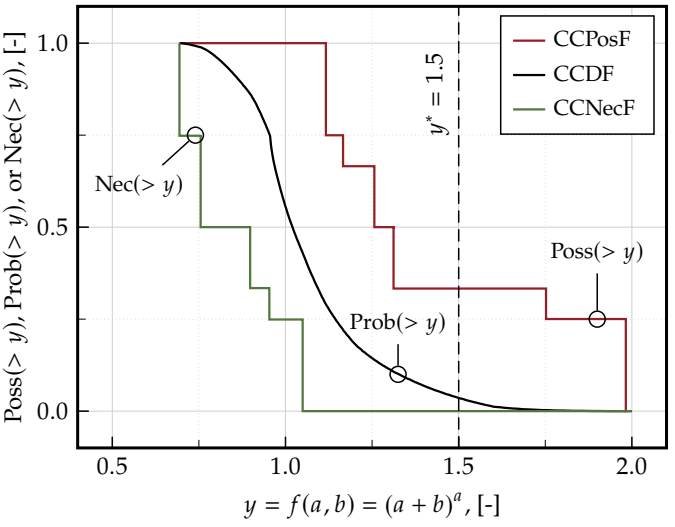


Figure 2.15 Possibilistic measures of an exemplary function with values from Helton et al. [101, p. 55].

additive or multiplicative manner, leading to different results. Additionally, the uncertainty definition using Possibility and Necessity is more diffuse and cannot be directly interpreted as bounds of probability [101]. These limitations and aspects may be the reason why there is virtually no application of this method in the literature.

Fuzzy Inference Systems are an approach to solving various problems in the fields of control, data classification, or decision-making. The main idea is to make use of expert knowledge and experience to replace (almost) exact but highly complex mathematical models. A FIS is a system that takes crisp inputs, processes them based on specified rules using fuzzy measures, and produces crisp outputs. Figure 2.16 illustrates the basic steps of one of the most common approaches for FIS, called the Mamdani FIS[7].

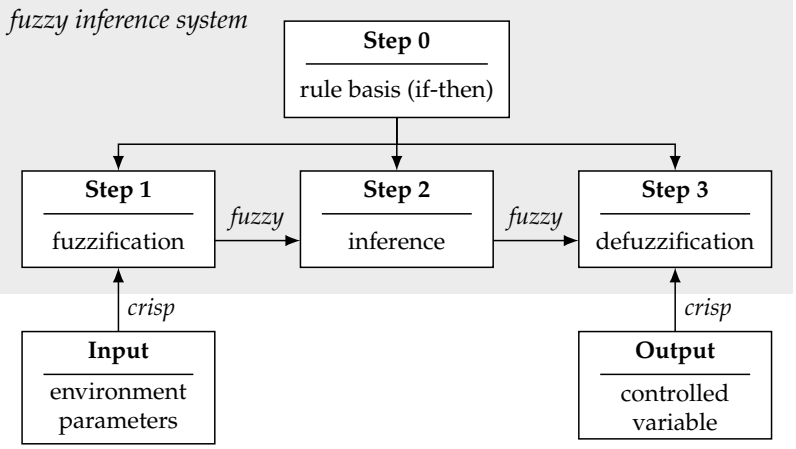


Figure 2.16 Steps of mamdani type Fuzzy Inference System

An example often given when introducing FIS is the restaurant tipping problem, where a question is asked about how much tip should be given based on the food and service quality. The core idea here is that human behavior can be modeled in a reproducible manner. Here, the input parameters (which are rated between 1 and 10) are fuzzified with a membership function (*poor*, *good*, and *excellent* for the service quality and *rancid* and *delicious* for the food). In step 1, the memberships for the inputs are calculated. The previously defined rule basis then combines these memberships with AND or OR operators and results in a degree of fulfillment of each rule. After the combination of each rule's fulfillment, a fuzzy output in terms of *cheap*, *average*, or *generous* is created. In the final defuzzification step, the fuzzy output is converted to a crisp value, e.g., 16 % tip. For a detailed description of this example, consider the Mathworks tutorials [265].

To conclude this Section, Table 2.1 summarizes the uncertainty measures, preferred utilization conditions, advantages, and disadvantages of the four discussed theories for UQ.

[7] Named after Ebrahim Mamdani who introduced this FIS type as a control system for a steam engine [155].

Table 2.1 Comparison of discussed mathematical theories for UQ.

Probability Theory *Frequentist*	*Bayesian*	**Evidence Theory**	**Possibility Theory & Fuzzy Set Theory**
Uncertainty Metrics			
$P(E) = \lim_{n\to\infty} \frac{n(E)}{n}$	$P(A\|B) = \frac{P(B\|A) \cdot P(A)}{P(B)}$	$\mathrm{Bel}(A)$ and $\mathrm{Pl}(A)$	$\Pi(A)$ and $N(A)$
Uncertainty is determined by the frequency of outcomes across many identical trials, focusing on the long-term behavior of an event under repeated experimentation	Uncertainty is viewed as a subjective degree of belief, which is updated by incorporating new evidence through Bayesian inference	Uncertainty is managed by distributing belief over sets of possibilities, using measures of belief and plausibility that reflect the strength and completeness of the available evidence	Uncertainty is addressed by describing events with varying degrees of membership in fuzzy sets and assessing the plausibility of events given this fuzzy information
When to Use			
Best used when there is a large amount of empirical data and the situation can be modeled by repeating identical trials under similar conditions to estimate probabilities directly from observed frequencies	Ideal for situations where prior knowledge is available, and the model parameters can be updated dynamically as new data becomes available, making it well-suited for iterative learning and decision-making processes	Particularly useful when dealing with incomplete, uncertain, or conflicting information, as it allows for an explicit representation of ignorance and the accumulation of evidence without requiring full probabilistic models	Optimal for scenarios where the information is imprecise, vague, or linguistically described, such as in expert systems and decision-making processes involving human-like reasoning under uncertainty
Advantages			
Offers clear interpretations in many real-world situations. Robust to prior beliefs. Ideal for reproducible and frequent events. Easy to implement [220] and best understood by decision-makers and analysts [30, 220]. Preferred choice when data is plentiful [42]	Can provide detailed posterior distributions, ideal for prediction. Works well in small sample situations. Incorporates prior beliefs. Suitable for aleatory and epistemic uncertainty [22, 293]. Allows modeling techniques considering a dynamic knowledge state using Bayesian inference	More versatile than traditional ProT. Suitable with scarce information [3, 103] and can deal with a wide range of uncertainty levels [180]. Useful when conflicting evidence is present [172, 220] and does not require assumptions from the analyst [220]. With increasing level of information, results approach those from ProT [103, 171, 220]	More suitable for describing rare events compared to probability theory [30]. Useful for epistemic uncertainties with subjective vagueness [293]. For control systems, FIS are easier and faster to implement and give better results than conventional models [139]
Disadvantages			
Produces least conservative results [22, 220] and may give false appearance of exactness [101]. Questionable applicability for epistemic uncertainty [3]	Results can vary significantly based on the choice of prior. Can be computationally expensive and is less understood by decision-makers	Assigning belief functions can be subjective. Decision-making is difficult when uncertainty intervals are high [220, 254]. When data is plentiful, evidence theory is less suitable for decision-making than ProT [254]. Propagation can be computationally expensive [103]	More qualitative than quantitative, making it less precise than other methods. Difficult to directly use with conventional statistical methods. Less understood by decision-makers [30] and lacks operational definition [49]. Can be computationally more expensive than conventional modeling techniques.

2.3 Literature Review

With the fundamentals of TEA in aeronautics and UQ explained in the previous Sections, this Section deals with more recent literature from both fields. This list does not claim to be exhaustive, but informative and expedient. The aim is to explore emerging research avenues to ensure that (a) the current research is novel and (b) recently developed methods and models from either domain are recognized and incorporated into this thesis when relevant. The literature review begins with publications focusing on economic evaluation in Section 2.3.1, continues with more specific UQ literature in Section 2.3.2, and finally includes those that combine both fields in Section 2.3.3.

2.3.1 Economic and Operational Assessment Studies

The following literature primarily focuses on the aerospace sector, highlighting integrative assessments, whether technological, economic, or both in nature. Although these publications often neglect or oversimplify uncertainties, their inclusion is essential. It provides insight into the prevailing status quo and identifies boundary conditions that might challenge the uncertainty management methodology introduced earlier. The literature is clustered into:

- *Valuation-Centered Publications*, which represent advances in assessment methods and evaluation frameworks,
- *Integrated Methods in Multidisciplinary Design Optimization*, where assessment methods are incorporated in product development frameworks, and
- *Operations-Focused Studies*, where assessments are performed on operational KPIs while taking an increased level of complexity into account.

Valuation-Centered Publications

Publications centered on valuation indicate that DOC methods are still commonly used in aeronautics, especially when evaluating aircraft designs resembling traditional models. For example, Lee et al. [148] use a DOC method to assess and compare costs for various widebody aircraft types. Similarly, Ali et al. [12] test and compare results from different DOC methods on 16 transport aircraft. Xu et al. [292] apply a DOC method for analyzing formation flights in airline operations, while Elham et al. [72] use it for winglet shape optimization. Iwaniuk et al. [117] adopted the operating cost model from Roskam [221] for cost-optimized CS-23 aircraft designs. Taking a more comprehensive approach, Isikveren [116] optimized flight mission parameters, such as cruise speed schedules, incorporating not just traditional DOC elements but also IOC and revenues. For newer DOC methods, reference can be made to Hong et al. [111], grounded on the Form 41 database [271], and Oliveira [182], who compiled a new method using various sources.

Given that CERs draw from historical data, they are primarily tailored for conventional aircraft or designs that resemble traditional models. However, some studies overlook this constraint. For example, Martinez-Val et al. [158] combine multiple DOC methods to assess a Blended Wing Body (BWB) aircraft. This design differs considerably from conventional models, and inferences from DOC methods regarding airframe and engine maintenance cost are questionable. Another instance can be found in Cuerno-Rejado et al. [51]. In this research, an unconventional aircraft with joined wings is designed and economically compared to a conventional counterpart. Notably, their joined-wing model neither increased maintenance nor ownership costs[8]. Remarkably, only the fuel cost differed, showcasing an improvement of nearly 6 % — a result the authors themselves found "*surprising*".

[8]The former is usually influenced by weight, consistent across both designs. The latter was simply omitted in the study, though qualitative impacts were mentioned.

Integrated Methods in Multidisciplinary Design Optimization

The benefits of CERs were touched upon earlier, emphasizing their straightforward implementation and quick evaluation, which renders them favorable for MDO frameworks. For instance, both the Preliminary Aircraft Design and Optimization Program (PrADO) from TU-Braunschweig [96] and the Multidisciplinary Integrated Conceptual Aircraft Design Environment (MICADO) from RWTH-Aachen [219] have integrated CER-based DOC and LCC methods into the environments.

Another MDO incorporating DOC methods is detailed by Colmenares et al. [48] from Cranfield University, focusing on aircraft engines. Their cost calculation, which seemingly covers only engine-associated DOC, uses a so-called "*Economics Module*". This module projects the DOC over three decades to determine the Net Present Cost (NPC) (which represents the NPV without revenues). A similar framework was used for intercooled engines by Kyprianidis et al. [143] and Saatlou et al. [227]. The same framework is used by Goldberg et al. [87], where the focus shifts from aircraft engines to overall aircraft designs. The assessment itself was detailed in a more systematic manner, factoring in metrics such as the NPV, IRR, and their version of the NPC[9]. The authors also draw attention to the existing gaps in comprehensive aircraft technology evaluation, stressing that "*It is therefore vital to establish the overall economic benefits and commercial viability of new technologies and policies, a type of study rather rarely seen.*".

Curran et al. provide additional examples of DOC method integration into multidisciplinary structures. In Ref. [57], the authors investigate the link between the manufacturing tolerances of an engine nacelle and increases in DOC. These increases are attributed to penalties in fuel consumption resulting from variations in aerodynamic efficiency. In this study, the DOC model, though not detailed in its methodology, is seamlessly integrated with a Computational Fluid Dynamics (CFD) framework. The same theme was subsequently explored by the authors in other works, each with its own unique central topic [36, 56, 208].

Operations-Focused Studies

Aircraft operations research is a field with substantial relevance to the subject at hand. This expansive domain encompasses studies that tackle a diverse range of subtopics, from airline fleet planning to management strategies. For instance, Santos et al. have contributed several insightful papers to this domain, addressing areas like fleet planning under demand uncertainty [215, 226], challenges of flight scheduling [240, 279, 280], strategies for maintenance schedule clustering [62], and end-of-life optimization techniques [175]. While these works detail various frameworks and methodologies tailored to resolve these problems, their central emphasis revolves around enhancing the resilience of planning algorithms to cope with unforeseen disruptions or events. Their focus is more on tackling these issues than evaluating specific aircraft or technologies. Nevertheless, the submodels introduced emphasize a critical observation: regression-based predictions fall short in accurately representing the intricate nature of the air transportation system. Instead, methodologies such as DES or agent-based modeling seem better suited to address this complexity.

Conclusions and Implications

The state-of-practice in aeronautical economic valuation is still significantly dominated by legacy DOC methods, with a few exceptions that have developed updated CERs. Applications range from aircraft designs to specific technologies up to operational procedures. Some of these seem to step out of the area of CER applicability, which, considering the impact of the implied financial decisions, can be highly critical for aircraft manufacturers and operators. With the recent stream of airline operations optimization, the need for a more detailed modeling

[9] The authors do not call the economic metric "*NPC*" but the "*difference in aircraft purchase cost*" or just "ΔX". However, their definition is equal to that of the NPC.

approach becomes apparent. The integration of costing methods in MDO environments emphasizes their *integrative* nature, implying the need for a comprehensive and flexible valuation tool.

2.3.2 Uncertainty Quantification Literature

This Section provides a review of more recent publications in the UQ domain. The sectors from which these studies originate are not restricted to aerospace but extend to all domains such as mathematics, computer engineering, and civil engineering. Their categorization is based on the following clusters:

- *Definition and Categorization Focused Publications*, where studies aim to aid in the differentiation of uncertainties,
- *Reviews and Method-Based Studies*, where relevant literature is reviewed and methods and metrics are discussed,
- *Visualization and Recipient-Considering Publications*, where the proper representation of uncertainties is dealt with.

In order to focus on general advances and theoretical developments, specific applications of UQ have mostly been filtered out in from this list. They can be found in Section 2.3.3.

Definition and Categorization Focused Publications

The seminal work from Walker et al. [281], which was the foundation of the uncertainty definition and classification discussed in Section 2.2.1, laid the groundwork for a number of uncertainty-addressing studies, especially in the environmental domain, e.g., Petr et al. [189] and Walling et al. [282]. Building on the three uncertainty dimensions from Walker et al., Janssen et al. [119] introduced two additional dimensions: the qualification of the knowledge base and the value-ladeness of choices. According to Van der Sluijs et al. [276], the first dimension pertains to the depth and support of the information, including data, theories, models, methods, and argumentation used in problem assessment. This dimension underscores methodological acceptability and the robustness of the adopted methods, knowledge, and information, offering insights into their reliability. The second dimension, value-ladenness of choices, highlights the influence of values and biases in various decision-making processes. This encompasses framing scientific questions, selecting and interpreting data, designing methodologies and models, and formulating explanations and conclusions.

Booker et al. [30] offer a comprehensive review of the quantification and assessment of uncertainty, highlighting some essential steps aiding in uncertainty management. Additionally, they introduce a novel classification of uncertainties, i.e., those rooted in inference. This includes, for example, uncertainties that arise from forecasts that infer future values based on past data or from attributes inferred about an entire population from a sample. While this extension of inference-based uncertainties provides a valuable perspective for identification, the study did not furnish any mathematical or engineering methodologies to address them.

Reviews and Method-Focused Studies

Schwabe et al. [243] present an extensive and excellent review of the metrics of UQ with respect to their potential use in aerospace innovation. They analyzed the frequency of different metric usage throughout the time, highlighting that classical metrics and procedures such as general statistics and MCSs are slowly being replaced or at least accompanied by non-probabilistic metrics. With respect to aerospace innovations, where the number of actual observations of the inputs or outputs afflicted with uncertainties is relatively low, the authors explain how the classical approach of using MCS and relying on the CLT is not always appropriate. Another key message of this publication is that UQ is generally a snapshot of a dynamic situation, i.e., the same OoI may be dealt with different uncertainty metrics over time. The authors also

highlighted that these changing uncertainty metrics should be mathematically congruent for a consistent UQ.

In Grenyer et al. [90], a systematic literature review of multivariate UQ for uncertainty quantification was conducted. In this context, the term "*multivariate*" refers to the combination of quantitative and qualitative UQ, where the former describes situations where data is available and probability theory is suitable and the latter refers to situations where data is scarce and expert-driven knowledge or experience is utilized instead. The authors claim that the combination of both will increase confidence and rigor in determining the impact of uncertainty. Their review highlighted the predominant use of ProT, Sensitivity Analysis (SA), and MCS. Furthermore, they conclude that evidence theory and interval analysis are suitable alternatives for more qualitative UQ. Fuzzy set theory was mentioned as being well suited for qualitative reasoning but not for estimating quantitative uncertainty. The study results in two research gaps: (a) a lack of frameworks to aggregate multivariate uncertainty and (b) limited approaches for forecasting uncertainty with limited data.

Hawer et al. [100] provide a brief literature review where they pick up the work from Walker et al. [281] and Janssen et al. [119]. Furthermore, they present a flowchart-based guideline for practitioners who wish to decide which theory or metric is appropriate for modeling the uncertainty of a reference object. This rather complex flowchart, however, focuses on the choice of probabilistic distribution functions and fuzzy set theory and does not cover evidence-theoretic approaches, nor does it provide guidelines for uncertainty propagation.

Visualization and Recipient-Considering Publications

The initial characterization from Walker et al. [281] was picked up by Viavattene et al. [278] and applied to uncertainties regarding flood risk management. In this study, the authors provide a qualitative method aiding the researcher in determining whether enough information is available to conduct a viable lifecycle cost evaluation, where they utilize the uncertainty dimensions from Walker et al. The outcome is an Uncertainty Inventory (UI), based on the *level* and *nature*[10] dimension of the uncertainties of cost elements. The authors provide a visualization plotting the cost magnitude over the UI, which may then be divided into regions of feasibility for the LCC evaluation. It should be noted that Viavattene et al. acknowledge the limitations of their methodology as being subjective in multiple facets and improvable as some experience is required.

Spiegelhalter et al. [255] specifically discuss the visualization of uncertainty about the future in the *science* journal, leading with the statement that "*probabilities are notoriously difficult to communicate effectively to lay audiences*" ([255, p. 1393]) and that previous reviews have shown that there is a lack of best practices in uncertainty visualization. They mention that communicating uncertainty can be done through tabulated summary statistics, but a greater impact is achieved by a graphical visualization, ideally when tailored for the audience. They furthermore address the fact that a suboptimal uncertainty visualization may lead to confused, suspicious, and more risk-averse recipients [95, 204]. The authors conclude their paper with a number of recommendations for uncertainty visualization, with the most important one being to assess the needs and capabilities of the audience to design a visualization format suitable to inform the recipients effectively and objectively.

Conclusions and Implications

The general categorization of uncertainties into epistemic and aleatory remains consistent throughout most publications. Additional categories and characteristics, such as the inference type and expert credibility, have been proposed but rarely picked up in other studies. In fact, there is a vast spread of metrics for UQ among researchers, ranging from those provided by

[10] The authors do not interpret the nature dimension as covering the epistemic/aleatory attributes but the variability throughout time, i.e., the dynamic properties of the uncertainties - a concept mentioned but not addressed by Walker et al.

classical probability theory over non-probabilistic approaches up to data science and machine learning-driven ones. From a comparative perspective, it can be seen that the underlying theories are often alternatives to each other but are occasionally used in a complementary way. The most frequently mentioned theories are (a) probability theory, (b) fuzzy set theory, and (c) DSTE. Fuzzy set theory was, however, mentioned as being less suitable for UQ but rather applicable for modeling systems with subjective inputs. As a final note, recipient-focused publications are relatively rare, although several authors highlight the need for a proper approach to communicating the complex UQ efforts to potentially lay audiences.

2.3.3 Uncertainty Addressing Design and Assessment Studies

After discussing the conventional assessment studies and general uncertainty quantification literature, this Section is dedicated to publications in-between, i.e., design and assessment studies that incorporate some form of UQ. The literature discussed here extends the overview given in Section 1.2 and is not limited to the aerospace sector. The clusters for this review are:

- *Monte Carlo Simulation and Surrogate Modeling Techniques*, where the latter aims to reduce the computational expense of the former,
- *Global Sensitivity Analyses*, for apportioning uncertainty and separating strongly contributing uncertainties from negligible ones, and
- *Non-Probabilistic Techniques*, where methods from DSTE and fuzzy set theory are used.

Monte Carlo Simulation and Surrogate Modeling Techniques

A large number of design and assessment studies deal with uncertainties via MCS [90, 243]. The approach is usually as follows: After uncertain parameters have been identified, they are assigned lower and upper bounds. Then, a uniform distribution is defined within these bounds, from which samples are drawn and propagated through the model for output distribution analysis. Applications vary and include aircraft CO_2 emission assessments [99], preliminary aircraft design [191, 192], certification-driven aircraft design [289], aircraft noise assessments [27, 125], economic assessments of biodiesel production [261], hydrogen production [147], and nuclear energy systems [82]. A noteworthy drawback of this approach is that, due to the lack of data, the assigned bounds are subjective in nature, and oftentimes there is no support for the assumption of a uniform distribution.

A noteworthy modification of the MCS in the field of Prognostic Health Management (PHM) can be found in Dewey et al. [63]. Here, epistemic and aleatory uncertainties are treated separately using a nested MCS architecture. In the inner loop, the parameters of aleatory uncertainties are sampled and analyzed once converged. The outer loop samples the epistemic uncertainties, which are usually of uniform distribution, which are then fixed for the inner loop, where the aleatory uncertainties are sampled. The output is visualized using sets of CDFs for the probability of failure over time. The authors highlight the computational expense of this method as a significant drawback. In addition, the same limitations with respect to the assignment of uniform distributions apply. A similar approach can be found in West et al. [284] for the UQ of supersonic aircraft configurations. Here, the authors acknowledge that uniform distributions "*can lead to inaccurate predictions in the amount of uncertainty in a system*" and therefore suggest using intervals, which, however, still suffer from subjectivity in the assignment of lower and upper bounds. To improve the computational runtime, a PCE-based surrogate model was used, and outputs were visualized using bounded CDFs.

Other applications of surrogate models in the MCS context can be found in DeGennaro et al. [59], where the authors investigate airfoil icing situations under aleatory uncertainty. Another example of PCE-based UQ was presented by Cheema et al. [40], who investigated the flutter behavior of an aircraft T-tail. Both studies present and discuss the results using histograms and statistical moments.

Allaire et al. [13] from Massachusetts Institute of Technology (MIT) provide a new methodology for developing a surrogate model. The authors apply their surrogate to an environmental aircraft performance tool called AEDT. This work is then continued in Allaire et al. [14], where their environmental tool has been further developed. Another update of this methodology was presented in Amaral et al. [16], where the application moved from the AEDT to a combined tool that incorporates aircraft design capabilities.

Global Sensitivity Analyses

An increasing number of studies acknowledge the potential of GSA methods to quantify and rank the criticality of input uncertainties. For instance, Opgenoord et al. [184] use the so-called Sobol' method to efficiently decide which input uncertainty should be reduced first. The authors discuss an aircraft sizing problem and introduce uncertainty budgets to solve an optimization problem. The same GSA method was also used by Sibdari et al. [246], where the authors reviewed datasets on fuel cost, passenger demand, and unemployment rate to quantify the impact on airline capacity (i.e., flight frequency, aircraft size, and load factor). Further applications of the Sobol' method can be found in Gong et al. [88], where the aerodynamic performance of a morphing-wing aircraft was optimized under geometric uncertainty of wing sweep, span, and chord length, or in Raj et al. [209], who quantified the sensitivity of uncertain icing shape and icing mass parameters on physical and modeling parameters of airfoils.

An alternative GSA method, namely the Fourier Amplitude Sensitivity Test (FAST), was used by Park et al. [186], who optimize the aero-structural design of an aircraft wing. A modified version of the FAST technique called FAST with Random Balanced Design (FAST-RBD) was utilized and extended to non-uniform distributions in Gao et al. [81], where the authors performed a stress analysis of aerospace rocket nozzles. Further (non-aerospace-related) applications of the FAST-RBD method can be found in Bui et al. [33], Gaspar et al. [83], and Hong-Qi [112].

Generally speaking, the research field of GSA methods is quite active, resulting in several newly developed techniques in the past years, some of which are tool-specific [13, 14], whereas others are generic and treat the underlying model as a black-box [149, 194]. Considering that some GSA methods have fundamentally different approaches to apportioning uncertainty and provide different types of sensitivity measures, researchers should choose the GSA method according to their needs. However, the selection is rarely explained in literature.

Non-Probabilistic Techniques

While non-probabilistic techniques are far less frequent in UQ-driven design and assessment studies, there are some noteworthy publications. Hester [107], for example, discusses the use of evidence theory and focuses on combining expert knowledge while accounting for varying expert credibility. They briefly demonstrate the potential using a case study where maintenance intervals are quantified using Belief and Plausibility measures. A more general discussion of the potential of evidence theory in MDO is presented by Agarwal et al. [3], who apply the methodology to an optimization problem for aircraft sizing and compute the CCXFs over the output domain. A similar approach can be found in Bae et al. [22] on a structural design problem for aircraft wings. Note that these publications highlight the potential of the evidence-theoretic approach and its superiority over ProT when data is scarce. However, they apply the techniques to highly simplified problems since the computational expense increases exponentially with the number of uncertain inputs. Furthermore, no distinction is made between aleatory and epistemic uncertainties.

Fuzzy set theory is mostly used for incorporating qualitative uncertainties in design and assessment efforts. Odedele et al. [181], for instance, compare the traditional deterministic NPV approach within oil and gas investment decisions to one that incorporates uncertainties. Instead of MCSs for uncertain inputs, they make use of their own predictions via "*fuzzy support vector machines*". For their final statement regarding economic viability, they incorporate a Fuzzy Inference System (FIS) with rules based on different investment metrics. Further examples

where fuzzy theory replaces MCSs can be found in Chowdhury et al. [45] and Rezaei et al. [216], both dealing with aircraft flutter investigations. A quantitative attempt at dealing with uncertainties is provided by Chen et al. [41], who predict aircraft taxi times using a mamdani-type FIS. The authors seem to use the goodness of the defuzzification step as a means to quantify the output uncertainty. Without providing details on the methodology, this approach suffers from a lack of reproducibility and transparency.

Conclusions and Implications

The applied UQ studies highlight the importance of distinguishing epistemic from aleatory uncertainties, as they have different consequences for decision-makers. Nested MCS have been proposed, in which epistemic uncertainties are modeled wit huniform distributions or intervals, which are subjective and may lead to inaccurate representations of uncertainty. Several authors mentioned the difficulty of computational runtimes, which led to the development of surrogate modeling techniques such as PCE. GSA methods seem to receive increasing attention for guiding uncertainty reduction efforts. The Sobol' method is a popular GSA technique, followed by FAST-RBD and FAST. However, the choice of GSA method is usually not explained, although different techniques provide different capabilities. In the non-probabilistic domain, the potential of evidence-theoretic techniques was identified but mostly applied using simplified examples due to the high computational expense. Furthermore, there seems to be a lack of studies combining DSTE and ProT for the quantitative inclusion of uncertainties. As a final note, fuzzy set theory is mostly used for qualitative uncertainty considerations, substantiating the theoretic discussions from Sections 2.2.2 and 2.3.2.

2.4 Summary

This Chapter provided the basics of TEA, UQ, and the interconnection of both. To do so, the theoretical foundations were outlined in Sections 2.1 and 2.2, complemented by a literature review in Section 2.3, which augmented the state-of-the-art discussion from Section 1.2. By analyzing conventional economic evaluation practices, general UQ literature, and uncertainty-addressing design and assessment studies, the barriers that led to the research questions and hypotheses presented in Chapter 1 were substantiated.

It is evident that, compared to the non-probabilistic domain, conventional ProT - encompassing both the Frequentist and Bayesian perspective and their associated techniques - is widely understood and utilized. Given that this thesis seeks to address challenges obstructing the effective implementation of the proposed uncertainty management methodology, the subsequent Chapters will focus on the following aspects:

- The potential of GSA in attributing output uncertainty to input parameters has garnered increasing interest, yet the rationale behind the selection of specific techniques often remains unexplained. It is recognized that different GSA methods can vary significantly in their theoretical foundation, applicability, and added value, depending on the problem at hand. Consequently, Chapter 4 is dedicated to comparing the efficacy of various GSA methods when applied to TEA, assessing their strengths and weaknesses, and providing guidance for users. The most suitable method is then applied in the HLFC case study.

- Quantitative non-probabilistic techniques for uncertainty, particularly those based on evidence theory, have shown potential to outperform traditional probabilistic methods, especially when data is scarce and expert input is available. Despite their promise, these methods are seldom explored extensively in existing literature. Furthermore, a detailed comparison of evidence-theoretic methods with traditional probabilistic approaches within the context of uncertainty quantification has yet to be conducted. Therefore, Chapter 5 provides a systematic evaluation of the effectiveness of evidence-theoretic methods against

conventional probabilistic techniques, achieved through a novel programming package developed specifically for this analysis.

- Some scholars emphasize the importance of differentiating between epistemic and aleatory uncertainties due to their unique impacts on decision-making. Nested MCS have been proposed to manage these uncertainties separately, often employing probability boxes for visualization. This method typically uses either uniform distributions or subjectively determined intervals to represent epistemic uncertainties, which can affect the accuracy of UQ results. The advantages of using evidence theory to create expert-driven belief spaces for epistemic uncertainties have not yet been integrated with conventional probabilistic techniques for aleatory uncertainties. Therefore, Chapter 6 investigates the integration of these methods within nested MCS, evaluating their effectiveness and clarity for decision-makers compared to alternative approaches.

To set the stage for the above, Chapter 3 introduces the newly developed LCBA framework as well as the HLFC case study in more detail.

3 Assessment Framework and Use Case Definition

This Chapter introduces the assessment model crafted for this thesis, aligning with the initial step of the uncertainty management methodology detailed in Chapter 1. Subsequent sections explore the use case of HLFC on a long-range aircraft, namely the Airbus A330-343, providing an overview of the technology and its relevant boundary conditions. From these conditions, assumptions and uncertainties arise, which will be discussed and tabulated for the later UQ efforts. The Chapter concludes with an evaluation of the use case within the assessment framework, executed in a deterministic context.

3.1 Discrete Event Simulation Framework LYFE

The framework named Lifecycle Cash Flow Environment (LYFE) aims to provide a generic (i.e., not project- or case-specific) environment for the LCBA of various products in aeronautics. This Section briefly describes the program structure and some relevant capabilities. It should be noted that, due to the multi-purpose nature of the LYFE, the framework will not be discussed in its entirety here. For a more detailed description, consider Pohya et al. [201] (on which this Section is based on) or LYFE's extensive documentation in Pohya et al. [203].

3.1.1 Program Structure

LYFE is a Python-based framework[1] that uses a modular program structure to simulate discrete events spanning an aircraft's lifecycle. These include, but are not limited to, the acquisition of the aircraft, all performed flights, scheduled and unscheduled line and base maintenance with corresponding downtimes, various non-recurring and recurring payment events (such as insurance, interest, or leasing rate payments) as well as the eventual resale, recycling, or disposal of the aircraft. These events are cataloged in what is called an "*event calendar*". Distinct modules (which include *acquisition*, *operation*, *maintenance*, *end-of-life*, and a *custom* one that can be specified by the user) autonomously create, define, trigger, and append events to this calendar, as shown in the overall program structure in Fig. 3.1.

The process initiates with LYFE processing user inputs and loading essential databases, such as airport data and aircraft performance response surfaces. Next, the individual modules are accessed. Each module typically comprises an initialization function `init` and a core function `main`. The `init` function executes preparatory tasks to speed up the overall runtime where possible. The `main` function, executed repeatedly in a loop, first determines if an event that this module can create is imminent. Event triggers may, for instance, include the operational schedule (for flight events), the aircraft's health status (for maintenance events), or the simulation's current timestamp (for recurring payments). If no event is due, the loop progresses to the next module in line. Conversely, if an event is due, it is created and filled with relevant attributes such as the duration, associated costs and revenues, and other specifics such as the calculated fuel burn. Once the designated lifetime is reached, a sequence of post-processing measures is triggered. This part not only computes operational statistics (e.g., the average annual utilization) but also economic metrics (e.g., the NPV, see Eqn. (2.2)). The results are consolidated, tabulated, and visually represented in a comprehensive report.

[1] Note that the term *framework* indicates that LYFE can be enriched by custom modules, allowing its application across a variety of assessment scenarios.

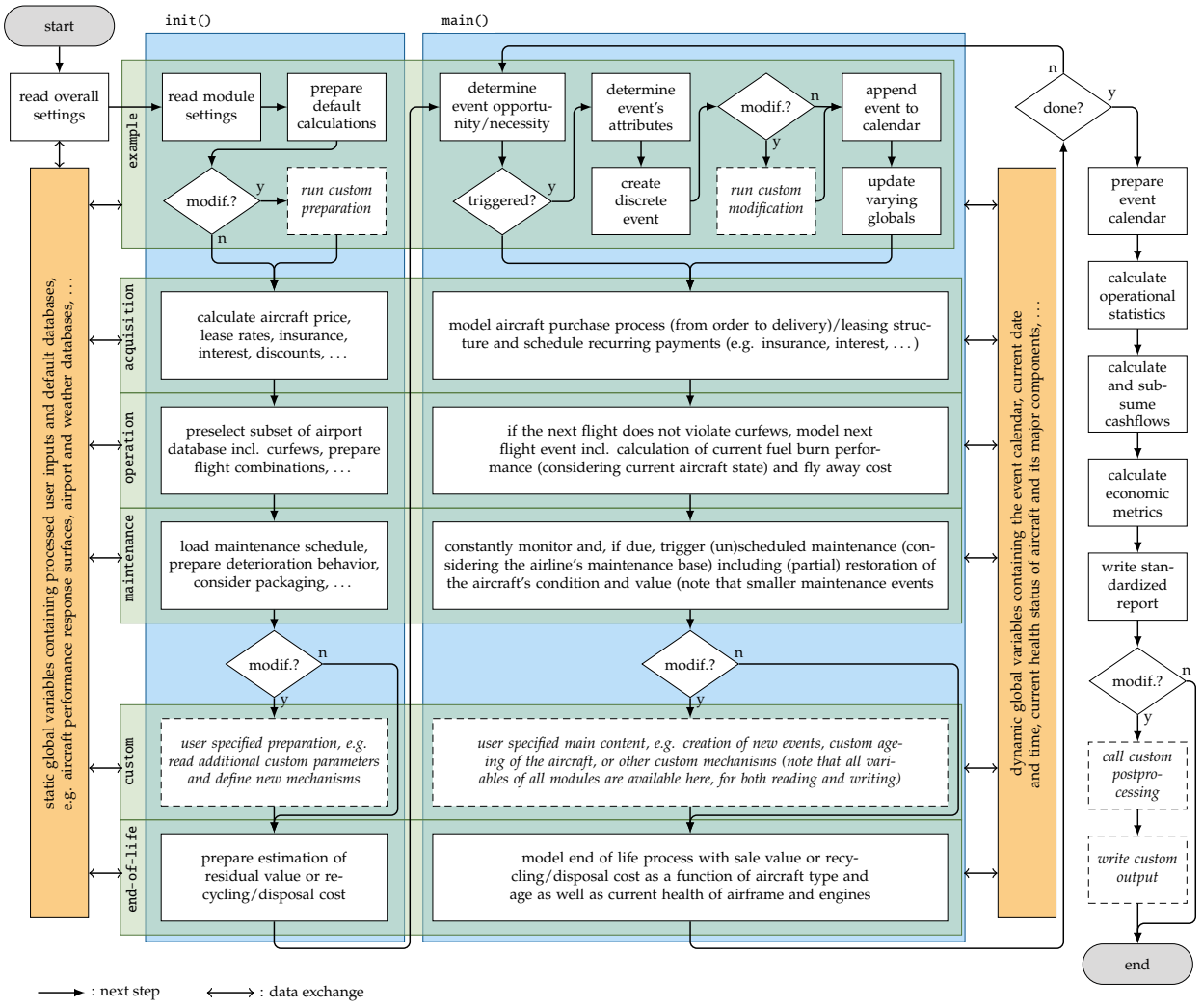


Figure 3.1 Modular program structure of LYFE with individual modules in green, their initialization and main functions in blue, databases in orange, and optional procedures in dashed boxes.

The adoption of DES aims to capture both primary and secondary effects throughout the lifecycle. Primary effects are those that directly influence the OoI or a stakeholder. On the other hand, secondary effects emerge either later or due to more complex interactions. For instance, a particular technology might decrease fuel consumption, leading to reduced fuel costs (a primary effect). However, when maintenance for this technology extends downtime, it can disrupt the flight schedule (a secondary effect). A thorough simulation in LYFE can consist of 50k to 100k events, varying with the use case and lifecycle duration. Consequently, a single execution of LYFE can take 1 to 3 minutes. Given that the cost benchmarking approach (refer to Section 2.1.4) demands two simulations for a full assessment, the overall runtime for one evaluation typically falls between 2 and 6 minutes.

3.1.2 Aircraft Performance Database

While LYFE makes use of a variety of databases, one of them is of particular relevance to the later-introduced use case, namely the aircraft performance database. It contains data for 35 different Airbus and Boeing aircraft, derived from a commercially available mission calculation tool called Piano-X [247]. For this, the tool was executed repeatedly with input values (using a full factorial design) for the flown distance r, carried payload p, as well as three technology factors, i.e., drag t_d, mass t_m (as a change in Operating Empty Weight (OEW)), and Specific Fuel Consumption (SFC) t_s. The outputs in terms of burned fuel, required time, and emissions (such as CO_2 and NO_x) were recorded and used to create response surfaces. As for the input

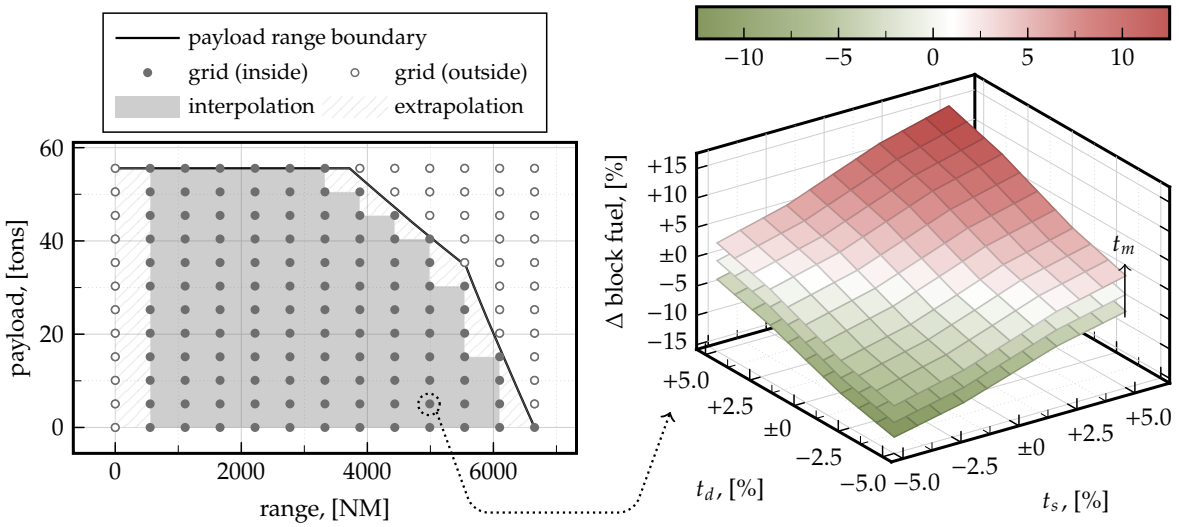


Figure 3.2 Aircraft performance response surfaces with a payload-range variation (left) and the impact of the three technology factors drag t_d, SFC t_s, and mass t_m (right).

space, the distance and payload were mapped from their lowest to highest values, whereas the technology factors were covered from 0.95 to 1.05. It should be noted that some of the input combinations cannot be reached by the aircraft because they exceed the payload-range boundary. This is illustrated in the left plot of Fig. 3.2. The right plot shows the fuel burn change for different values of the technology factors for a given grid point. With these response surfaces, LYFE calculates the fuel burn, flight time, and emissions with a 5-dimensional linear interpolation, i.e.,

$$m_{\text{fuel}} = f\,(r, p, t_d, t_m, t_s) : \mathbb{R}^5 \rightarrow \mathbb{R} \quad . \tag{3.1}$$

This performance model, which is attached to each aircraft object, allows users to quickly and easily simulate the impact of a variety of technologies. For the use case of HLFC, however, the performance model for the aircraft under investigation needed to be recalculated. That is, the lower bound of the drag factor was decreased to 0.9 in order to capture the anticipated drag reduction potential. Additionally, since the HLFC is sensitive to different speeds (explained later in Section 3.2.2), one additional input dimension was needed, namely the cruise mach number t_c. This modification resulted in a new performance model with six instead of five inputs to interpolate from.

Each of the fuel burn values represents the result of a detailed mission calculation, the results of which are shown for the (later introduced) reference aircraft in Fig. 3.3 on the design and an off-design mission. As illustrated in the upper plot, the mission simulation encompasses the entire flight profile, starting from takeoff, continuing through cruise with several step climbs, followed by descent, and concluding with touchdown. The resulting fuel flow is depicted in the lower plot, showing the peaks during takeoff and each step climb. The final amount of burned fuel for the design and off-design missions is 65.8 and 34.7 tons, respectively. The takeoff weight for each mission includes the required contingency fuels[2]. For this aircraft, the resulting passenger-specific fuel consumption is:

$$\gamma_{\text{fuel}} = \frac{m_{\text{fuel}}}{n_{\text{pax}} \cdot 100\,\text{km}} = \begin{cases} 2.69\,\text{kg}/(\text{pax} \cdot 100\,\text{km}) & \text{for design mission} \\ 3.09\,\text{kg}/(\text{pax} \cdot 100\,\text{km}) & \text{for off-design mission} \end{cases} \quad . \tag{3.2}$$

A comment on the validity of the mission calculation tool is appropriate. Acquiring real data on fuel consumption for commercial aircraft is challenging. Such data are not only scarce

[2]This includes 5 % of trip fuel, a diversion to an alternate airport which is 200 NM away from the destination airport, and allowances for taxiing, takeoff, approach

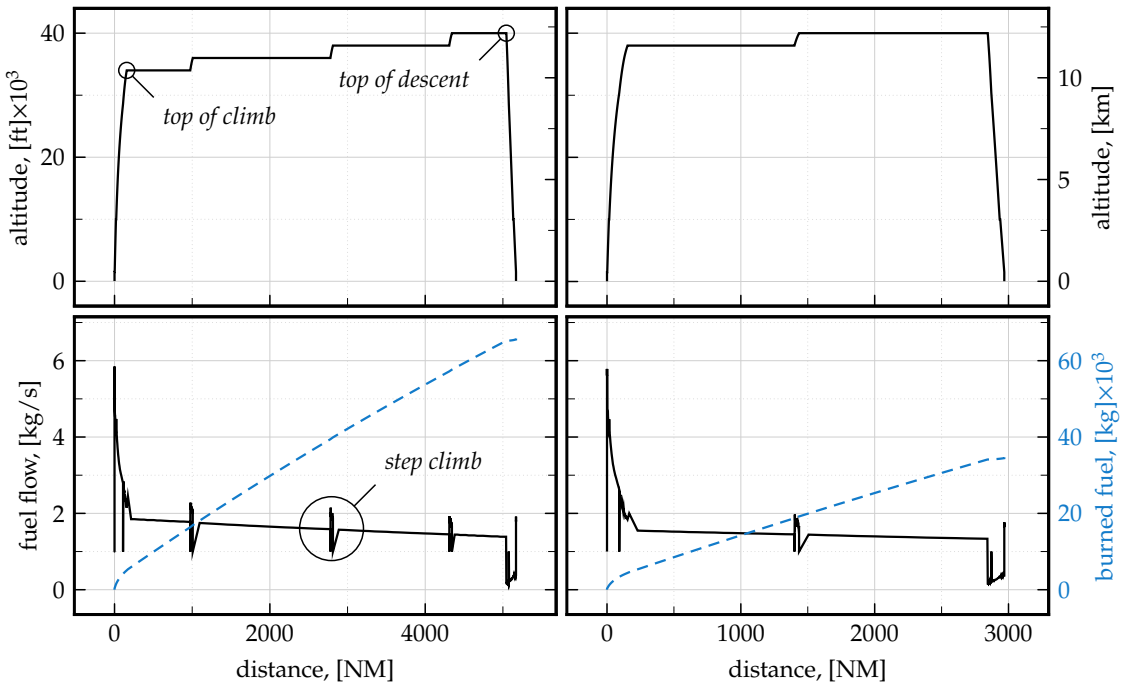


Figure 3.3 Performance of the reference aircraft for the design mission (left, 5169 NM and 100 % load factor) and off-design mission (right, 2968 NM and 80 % load factor)

but are also influenced by various factors, often undisclosed, e.g., prevailing wind conditions. Consequently, a comprehensive validation has not been conducted. Nonetheless, a partial validation was undertaken by comparing results with a publication by Aircraft Commerce [5], showing good accuracy with errors well below 0.5 %[3]. Therefore, the fuel calculation within LYFE, especially for the A330-343 (the reference aircraft used in subsequent analyses), is deemed sufficiently accurate to exclude it as a source of uncertainty.

3.2 Hybrid Laminar Flow Control

Due to its multidisciplinary and complex character, as well as its susceptibility towards environmental boundary conditions and lack of operational knowledge, the evaluation of HLFC is chosen as a use case for this thesis. This Section begins with a brief introduction to the working principles of HLFC for readers who are unfamiliar with the technology. Afterwards, the boundary conditions are dealt with in more detail, describing how the overall efficacy is affected by elements from different domains such as operations or the environment. It should be noted that the descriptions in this Section are of a qualitative rather than quantitative nature.

3.2.1 Working Principles

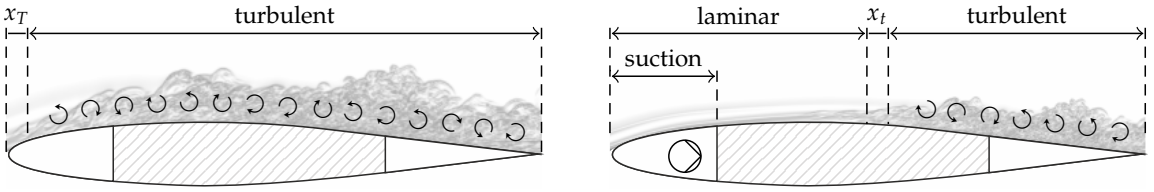
While HLFC is a primarily aerodynamic technology, its overall success depends on multiple disciplines such as systems, structures, materials, operability, and maintainability. It can be applied to components such as the wing, Horizontal Tail Plane (HTP), Vertical Tail Plane (VTP), and engine nacelles. The primary aim of the technology is to reduce the aircraft's drag, and consequently, its fuel burn by delaying the laminar-turbulent transition x_t of the boundary layer in chordwise direction as much as possible. In conventional aircraft, x_t occurs relatively close to the leading edge, whereas studies and early flight tests of aircraft with HLFC achieve transition lengths of 40 to 60 % of chord, see Fig. 3.4. As a *hybrid* technology, HLFC employs two general concepts[4]:

[3] Supplemental information on this topic can be found in Appendix C.1.

[4] For a more comprehensive description of the technology, consider Joslin [123], Young [296], or Risse [218].

(a) An active compressor-based suction of the boundary layer in the forward region of the airfoil (typically up to the front spar), through a microperforated or otherwise porous outer skin, and

(b) an airfoil geometry with specifically designed pressure distribution properties to dampen premature turbulence-inducing mechanisms, enabling a laminar-turbulent transition x_t beyond the suction area.



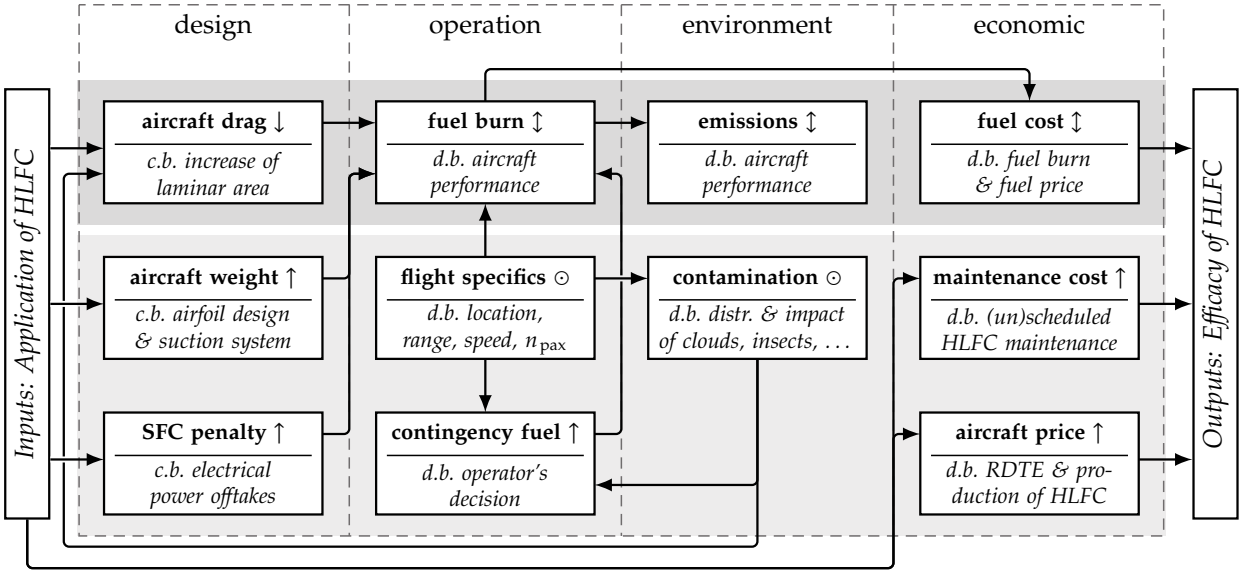
(a) conventional airfoil with transition near the leading edge. **(b)** HLFC airfoil with transition at 50 % of chord.

Figure 3.4 Illustration of turbulent and laminarized airfoils with background image from Ref. [21].

In essence, the ideal chain of effect of HLFC is as follows: The delayed transition position increases the laminarized area, which decreases aircraft drag for as long as the technology is activated. This, in turn, improves fuel consumption (and reduces emissions), which reduces fuel cost and ultimately operating costs, increasing the economic value for the operator. The aircraft manufacturer benefits from either an advantage in competitiveness or an increase in revenue (through an increase in aircraft price). The actual operation of HLFC is, however, characterized by a number of boundary conditions that affect different aspects of the described chain of effects. These are explained next.

3.2.2 Boundary Conditions

As several studies have pointed out, the overall efficacy of HLFC is influenced by a number of factors (e.g., Meifarth et al. [168] or Schrauf [242] for an airline's and manufacturer's perspective, respectively). Fig. 3.5 shows a diagram of effects for overview purposes, categorizing the influencing factors into the four domains of design, operation, environment, and economics. Before explaining their origin and repercussions on the HLFC's performance in more detail, some further considerations of the general domain are described.



↑: increase, ↓: decrease, ↕: increase or decrease, ⊙: causing element, c.b.: caused by, d.b.: defined by

Figure 3.5 Diagram of effects of HLFC.

Domain 0: General Aspects

When dealing with HLFC assessments, the first question that often arises is: *Which aircraft is it applied to?* While Natural Laminar Flow (NLF) is typically investigated for short-range aircraft, HLFC is usually chosen for medium- and long-range aircraft. This is primarily due to aerodynamic reasons. Achieving laminarity with NLF is difficult, if not impossible, for higher leading edge sweep angles, higher Mach numbers, and higher Reynolds numbers, but possible with HLFC[124, 218, 242]. Additionally, the relative portion of cruise time affects fuel burn reduction since most studies assume HLFC is activated only during cruise[296, p. 28]. This relative time in cruise increases with distance, making long-range more operations attractive for the technology. Another question is: *On which components is the technology applied?* Potential areas include the wings, tail planes, and engine nacelles. Several aerodynamic restrictions need to be considered. Areas close to the fuselage are difficult to laminarize due to challenging aerodynamic conditions. On the wing, HLFC application faces further difficulties. While the upper and lower sides of the wing can be laminarized, many studies consider only the upper side, often with a Kruger high-lift device. This is because the conventional high-lift system introduces steps and gaps that must meet very low manufacturing tolerances, a problem not yet sufficiently solved [196]. The Kruger flap can also shield the sensitive laminar area near the leading edge from insect impacts. The next question is about the *degree of freedom* with which the aircraft and airfoils are (re)designed. While airfoil geometries can be changed for HLFC design, the planform geometry may or may not be fixed. Furthermore, a final Overall Aircraft Design (OAD) optimization may or may not be part of the study. Such optimization includes snowball effects, e.g., reduced fuel burn allows for smaller fuel tanks and thinner airfoils, which leads to weight decrease and further fuel burn reduction, until the design converges. However, any decrease in airfoil thickness must be carefully managed due to its significant impact on structural integrity. The bending stiffness of the wing is proportional to the cube of the airfoil's height, meaning that even minor reductions in thickness can substantially weaken the wing's load-bearing capacity. Therefore, a balance between aerodynamic efficiency and structural safety must be achieved. The final question is: *Who is the customer?* Different airlines operate aircraft in different regions with varying maintenance strategies, resulting in the benefit of HLFC being different for each airline and operation. This is closely linked to the operational and environmental boundary conditions explained later. These questions need to be answered transparently to make a statement about the overall efficacy of HLFC.

Domain 1: Design Aspects

The design domain defines a variety of parameters. These include the system components and their layout (e.g., one or a few centralized compressors or multiple distributed ones), the changed structure in the leading edge region, the type of perforated skin, the resulting steps and gaps, and many others. The knowledge about these aspects varies with the stage of product development and, unless a large-scale prototype or demonstrator is built, remains incomplete. At an early stage, the assumptions in the design domain can be broken down into (a) the achieved laminar area, (b) the mass increase due to the installed systems and materials, and (c) the SFC penalty due to the power offtake of the compressors, inverters, and control computers. While (a) has a desired impact and is often maximized in optimization loops, (b) and (c) worsen the efficacy of HLFC. This trade-off can be explained with the Breguet-Range Equation:

$$R = \frac{V}{g} \cdot \underbrace{\frac{L}{D}}_{\text{(a)}} \cdot \underbrace{\frac{1}{\mathrm{SFC}}}_{\text{(c)}} \cdot \underbrace{\ln\left(\frac{m_0}{m_0 - m_t}\right)}_{\text{(b)}} \tag{3.3}$$

where V is the (constant) speed, g is the gravitational acceleration, L and D are the lift and drag, respectively, m_0 is the initial mass of the aircraft, and m_t is the mass of the burned fuel. The extended laminar flow of HLFC increases L/D, increasing the range, while the SFC penalty

and mass increase have the opposite effect. In order to reduce fuel burn for a given range, (a) therefore has to outweigh the combined negative impacts from (b) and (c), especially when further degradation from other domains is taken into account. As with all assumptions, these three design impacts can be estimated with a wide range of detail. Early methods may use analogies and expert knowledge, whereas more advanced procedures such as CFD are used in later stages. Full certainty can only be obtained with large-scale flight tests and the respective instrumentation for drag, mass, and power measurement.

Domain 2: Operational Aspects

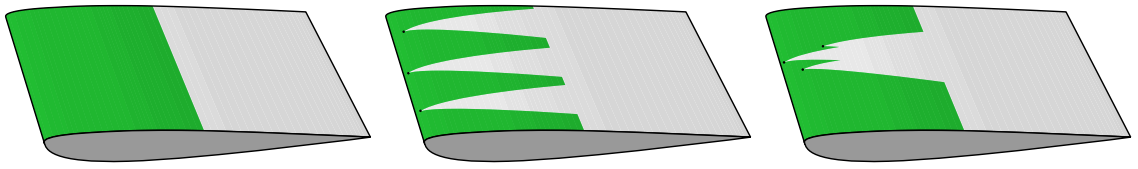
Another domain for the assumptions (later treated as uncertainties) stems from the operation of an HLFC aircraft. The previously mentioned design parameters typically assume specific flow conditions, e.g., specified cruise speeds and cruise altitudes. Airlines, however, operate their aircraft at a wide range of speeds and distances with different passenger loading conditions, creating off-design circumstances that influence the overall HLFC efficiency. Deviations from the design speed may affect the transition position, depending on the robustness of the concept. Flown distances and carried (passenger) loads also affect the overall savings, as HLFC is more efficient at longer routes and higher load factors. Additionally, operators of HLFC aircraft may choose to carry additional contingency fuel to account for potential but unforeseen losses of laminarity. This additional weight also influences the efficacy of the technology, and a trade-off between reliability (i.e., likelihood reaching the destination without the need to refuel) and fuel consumption needs to be made.

Domain 3: Environmental Aspects

A particular difficulty with laminar aircraft is their susceptibility to environmental boundary conditions [168]. Contamination of the laminar surfaces (e.g., through insect accumulation) decreases the laminar effectivity η_{L}, which is defined as follows [73]:

$$\eta_{\mathrm{L}} = \frac{A_{\mathrm{lam,1}}}{A_{\mathrm{lam,0}}} = \frac{A_{\mathrm{lam,0}} - A_{\mathrm{lam,loss}}}{A_{\mathrm{lam,0}}}, \tag{3.4}$$

where $A_{\mathrm{lam,0}}$ is the laminar area without any contamination, $A_{\mathrm{lam,1}}$ is the laminar area in operation, and $A_{\mathrm{lam,loss}}$ is the lost laminar area due to contamination or other effects. For example, insect contamination can induce turbulent wedges, creating a loss of laminar area $A_{\mathrm{lam,loss}}$. This loss generally increases with the number of insects but also depends on the accumulation pattern, as agglomerated insects share portions of their turbulent wedges and therefore have a smaller impact on η_{L} than evenly distributed ones [285].



(a) undisturbed. **(b)** distributed. **(c)** agglomerated.

Figure 3.6 IIllustration of insect contamination on laminarity.

Another degrading environmental element is the cloud encounter. At high altitudes, the cloud particles (primarily ice crystals in cirrus clouds) move through the boundary layer and disturb laminar flow. As flight tests in the 1970s have shown [188], laminarity decreases when flying through and in the vicinity of clouds and is immediately restored after emergence. Thus, the time in cloud has been selected in many studies as a key parameter to describe the degradation [121]. Other degradation mechanisms occur due to the presence of ash, sand, dust,

general air pollution, and other particles such as pollen, which may clog the microperforated sheet or lead to a faster deterioration of the compressors, e.g., through fouling.

Domain 4: Economic Aspects

The last domain covered here deals with economic aspects. As stated earlier, the ultimate goal of the technology is to reduce operating costs, which include cost of ownership, maintenance costs and fly-away costs (see Fig. 2.3). The economic superiority of HLFC aircraft therefore depends on its impact on all of these elements. For instance, an increase in fuel costs is beneficial for technology development, while an expected increase in maintenance costs due to the introduced systems and components decreases the value. These may consist of HLFC-specific efforts in line maintenance (such as inspections or the frequent cleaning of surfaces), base maintenance (mostly for the mechanical systems), and unscheduled maintenance (primarily driven by the electronic components). Additionally to the maintenance aspect, it is conceivable that HLFC aircraft are more expensive to procure than their conventional counterparts, due to the prior made RDTE effort of the manufacturer, the increased production cost due to higher requirements regarding tolerances in steps and gaps, as well as the acquisition of HLFC components and materials.

3.3 Deterministic Assessment

The first assessment performed in this thesis is of a deterministic nature, i.e., it does not attempt to quantify any uncertainties. It is presented here to get a more practical understanding of how the assessment framework LYFE works and also serves as a comparison point in later uncertainty-addressing studies. First, the goal of the study is formulated, followed by a documentation of the assumptions. Afterwards, the inputs for LYFE are described, before the results of the deterministic analysis are presented.

3.3.1 Goal of the Study

A clear and agreed-upon goal of an analysis, study, or project is considered standard practice and aids in the selection of appropriate models and justified assumptions. Therefore, it is formulated first. Here, it is distinguished between the goal of the assessment itself and the goal of the uncertainty consideration.

> *The goal of the assessment is to quantify the techno-economic performance of an HLFC aircraft compared to a conventional* (*i.e., turbulent*) *Airbus A*330-300 *from an operator's point of view. The result shall aid in a fictive investment decision-making process where an airline is confronted with choosing which aircraft types to buy in the future* (*or to declare sincere interest in*). *Thus, the assessment is prospective, predicting the economic performance of both aircraft under identical boundary conditions.*

Complementary to this, the deterministic setting has the following intention:

> *The goal of the deterministic setting is to obtain two scalar outputs*: $\Delta\gamma_{fuel}$ *and* ΔNPV. *These outputs represent the changes in technological and economic performance, respectively, between the HLFC aircraft and the reference aircraft. Recognizing that some input parameters are uncertain, point-values are used to treat these parameters as fixed assumptions. This choice, while potentially affecting overall outcomes and decision-making, is made to facilitate a clearer initial discussion of KPIs. This approach makes it easier to understand subsequent UQ and demonstrates a process commonly used in many TEAs.*

The change in fuel consumption, $\Delta\gamma_{\text{fuel}}$, is defined as

$$\Delta\gamma_{\text{fuel}} = \frac{\gamma_{\text{fuel,ref}} - \gamma_{\text{fuel,hlfc}}}{\gamma_{\text{fuel,ref}}} \quad . \tag{3.5}$$

Negative values of $\Delta\gamma_{\text{fuel}}$ indicate a technological success of the HLFC aircraft, as they reflect a reduction in fuel consumption compared to the reference aircraft. More negative values signify a greater reduction, demonstrating a stronger performance advantage of the HLFC aircraft.

In terms of economic performance, the KPI Δ NPV is defined as

$$\Delta\,\text{NPV} = \text{NPV}_{\text{hlfc}} - \text{NPV}_{\text{ref}} \quad . \tag{3.6}$$

This metric measures the difference in NPV (see Eqn. (2.2)) between the HLFC aircraft and the reference aircraft. Higher values of Δ NPV reflect greater economic benefits derived from the HLFC technology.

Assessing both $\Delta\gamma_{\text{fuel}}$ and Δ NPV provides a comprehensive evaluation of the technological and economic impacts of implementing HLFC on an A330-300 type of aircraft.

3.3.2 Assumptions

The explicit documentation of assumptions is vital for transparency and reproducibility [213]. The following descriptions follow the same structure outlined in Section 3.2.2, starting with general considerations, followed by aspects from the domains of design, operation, environment, and economics.

Domain 0: General Assumptions

The following considerations answer the four questions raised in the general domain of Section 3.2.2, namely:

1. Which aircraft type is HLFC applied on?
2. Which components of the aircraft are equipped with HLFC?
3. What is the degree of freedom with respect to the HLFC aircraft design?
4. Who is the customer?

Which aircraft type is HLFC applied to? This question has partially been answered in the goal of the assessment. The reference aircraft is a conventional (hence turbulent) Airbus A330-300, and the HLFC aircraft will be a close derivative redesign of that. As several A330-300 variants are in operation, an exploratory analysis was performed to identify a representative (or, in other words, most common) version. The result of this analysis is the Airbus A330-343, which is equipped with two Rolls-Royce Trent 772-B-60EP engines and has 255 seats, divided into three classes. Key parameters are summarized in Table 3.1.

Table 3.1 Key parameters of the A330-343.

Parameter	Value
Weights	
Maximum Takeoff Weight	233,000 kg
Operating Empty Weight	119,430 kg
Maximum Zero Fuel Weight	175,000 kg
Maximum Landing Weight	187,000 kg
Maximum (Structural) Payload	55,570 kg
Seating Layout	
Business Class	42
Premium Economy Class	28
Economy Class	185
Total No. of Seats	255

It is important to note that this study does not employ a hypothetical baseline aircraft design estimated to enter service alongside the HLFC model, primarily due to the scope of the research projects that underpin this thesis. Moreover, HLFC technology, having been researched for decades, is assumed to sufficiently mature to suggest that its incorporation into aircraft design could proceed rapidly, potentially outpacing other significant developments such as engine efficiency improvements. Consequently, both the reference and the HLFC aircraft were assumed to benefit equally from other technological advancements, thus allowing the comparative simulation to focus primarily on the impacts of HLFC itself. This approach mitigates the influence of other variables and aims to isolate the specific contributions of HLFC technology to aircraft performance and operational costs, as discussed in the cost benchmarking discussion of Section 2.1.4.

Which components of the aircraft are equipped with HLFC? The aircraft implements HLFC on the upper surfaces of the wings, extending from the engine positions outward. HLFC is also applied on both the upper and lower surfaces of the HTPs and on both the left and right sides of the VTPs. The initiation of laminar flow on the HTPs and VTPs does not begin directly at their junction with the fuselage but starts with a slight offset, addressing the complex aerodynamic interactions near the aft part of the fuselage. The HLFC system includes a suction setup comprised of a microperforated titanium sheet that replaces the conventional aluminum skin from the leading edge to the front spar, supported by an array of compressors, converters, control units, and sensors powered by the aircraft's Integrated Drive Generator (IDG), operational only during cruise. On the wing's lower side, a Krueger flap, positioned slightly ahead and below the leading edge, functions as a high-lift device during takeoff and landing, while also serving as an insect barrier to protect the laminar areas from debris, thus preserving the efficacy of the HLFC system throughout the flight.

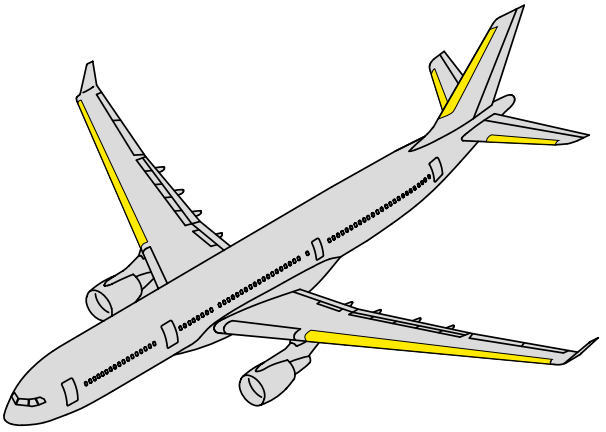


Figure 3.7 Isometric view of the HLFC aircraft design, based on Airbus [4].

What is the degree of freedom with respect to the HLFC aircraft design? As mentioned before, the degree of freedom in the HLFC development can range from a modification (without airfoil geometry changes) up to entire, HLFC-optimized aircraft that only share the Top Level Aircraft Requirements (TLAR) with their conventional counterparts. In the present case, it is assumed that the degree of freedom is somewhere in between. More specifically, the airfoil geometries of all components with HLFC are free to change, whereas the planform is kept fixed. The reason not to go further with the degree of freedom is as follows: An OAD-based optimization typically assumes ideal operating conditions for the laminar flow, and then employ expected fuel savings and their impact on the aircraft design as snowball effects. As the actual operation of HLFC is characterized by a number of degrading elements, such an optimized design is assessed to be too progressive[5], especially considering that operators are already concerned about the reliability of the technology [168]. The second reason is more of a practical nature: the OAD work is not tackled by any of the research questions or hypotheses and hence is out of scope for the present work.

[5]This is substantiated by several conversations with current and former Airbus employees responsible for HLFC related research and development projects.

Who is the customer? While manufacturers are more likely to think of groups of customers and aim to develop a one-fits-all aircraft, operators are only interested in the performance of a new aircraft on their specific route network. The assessments performed here take the operators' perspective rather than the manufacturers'. The assessment results can then be fed back to the manufacturer, as it is done in both supporting research projects. The airline (or customer) in this study is chosen to be Lufthansa, which operates 13 Frankfurt-based A330-343. Fig. 3.8 shows the route network of this fleet, with destinations in North America, Africa, the Middle East, and Central Asia. With an average flight time of a little above 7 hours, the A330s in the Lufthansa fleet are considered mid- to long-range utilized.

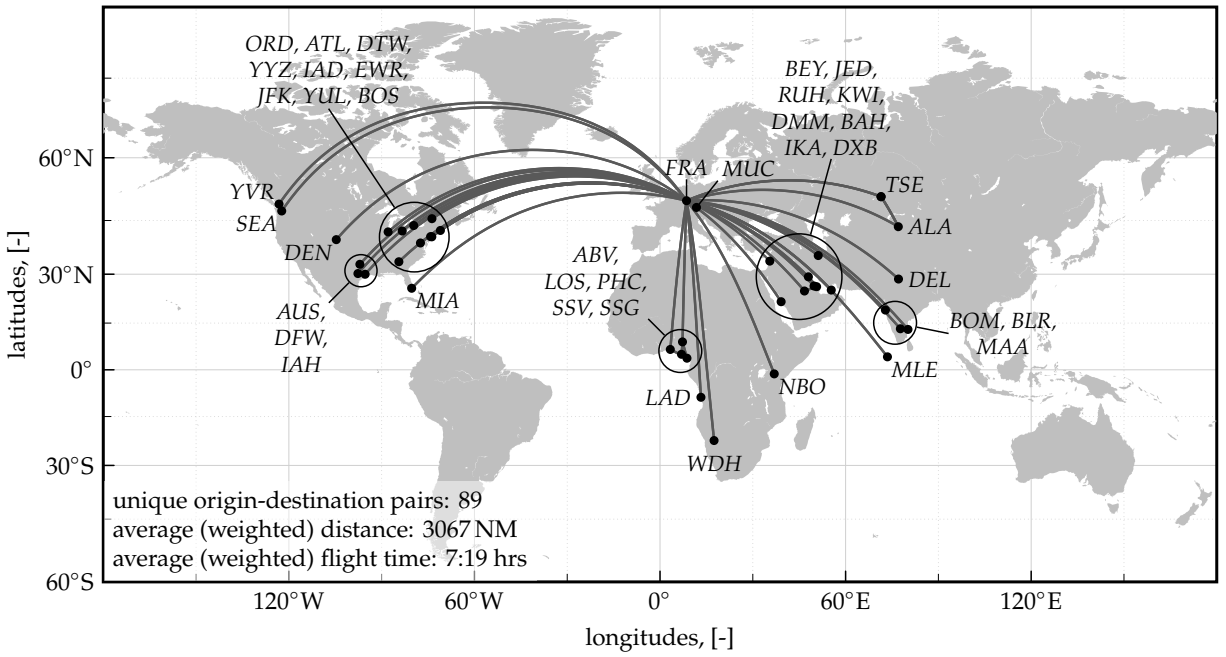


Figure 3.8 Route network of Lufthansa's A330-343 fleet as of 2020 shown as great circles, data taken from Sabre Airline Solutions [228].

Additional Considerations. Apart from the aforementioned aspects, it should be noted that this assessment will simulate only one aircraft, i.e., not a fleet. Generally speaking, simulating fleets with LYFE is suggested when there is indication that an OoI's performance affects or is affected by other aircraft in the fleet. It is assumed that this is not the case for HLFC. With respect to the TEA itself, it is assumed that it is situated at an early stage of product development. That is, detailed information regarding, for instance, the suction chamber designs, system layouts, and component sizing issues are assumed not to be available and have to be estimated, introducing uncertainties in the process. Due to the deterministic setting of this study, these uncertainties are treated as assumptions, i.e., are given point values. As a final note, the simulation time needs to be specified. It is assumed that the entry into service of the HLFC aircraft (as well as of the conventional A330) is in the year 2030 and will operated for 20 years.

Domain 1: Design Assumptions

The assumptions in the design domain comprise the drag reduction potential (without any degradation), the mass increase due to the HLFC-related components and materials, and the SFC penalty due to power requirements of the electrical components.

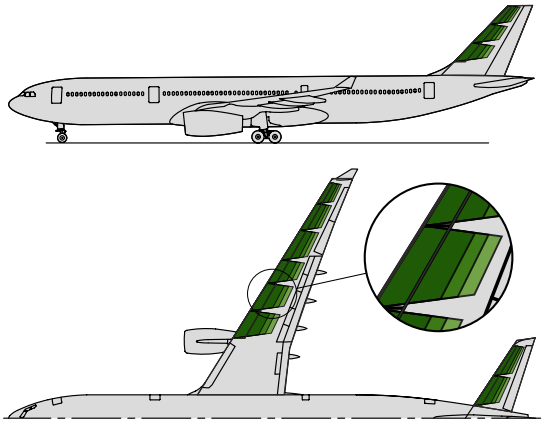
For this deterministic setting, the drag reduction potential is estimated using a two-dimensional approximation based on early valuation practices from Airbus and DLR [131]. It consists of the following steps:

1. Given 2D technical drawings, measure the base area A_{base} of each of the components that are supposed to be laminarized. For the wing, for instance, A_{base} represents a polygon that is bounded by the leading and trailing edges as well as the wing tip and position of the engine bay.

2. Given an assumed relative transition position $x_t/c = 0.5$, create a trapezoid that represents the idealized laminar area. In the wing example, this results in a polygon bounded by the wing tip, engine bay position, leading edge, and a line that is halfway between the leading and trailing edge. For a more realistic approximation, incorporate turbulent wedges between the leading edge segments with an opening degree of $\phi = 15°$ [131, p. 6] and subtract them from the idealized laminar area. Measure the resulting expected laminar area A_{lam}.

3. Given a drag breakdown from Schrauf [242], (including the assumption of the wing upper side making up two thirds of the wing's friction drag [218]), an estimate of the average drag coefficient during cruise of 260 dc, and an assumption of laminarized areas having 90 % less friction drag, estimate the impact of HLFC on the overall aircraft level.

Although simplified, this procedure has provided reliable results when compared to later, more sophisticated CFD-based studies. Fig. 3.9 visualizes the trapezoidal areas and tabulates the results broken down per component. On an overall aircraft level, this approximation results in a drag reduction potential of 19.3 dc (or 7.4 %), out of which the majority stems from the application on the wing upper side (13.9 dc), followed by the HTP and VTP (3.5 dc and 1.9 dc, respectively).



(a) Laminar areas on the HLFC aircraft.

	Unit	Wing	HTP	VTP	Sum
trapezoidal areas					
total area	[m^2]	304.7	118.1	90.3	513.2
suction area	[m^2]	37.3	18.8	12.6	68.7
laminar area	[m^2]	150.0	43.5	24.0	217.5
friction drag breakdown					
from [242]	[%]	12.0	4.0	3.0	16.0
savings potential					
Δdrag (absolute)	[dc]	13.9	3.5	1.9	19.3
Δdrag (relative)	[%]	5.3	1.3	0.7	7.4

(b) Estimated aerodynamic impact of HLFC.

Figure 3.9 Visualization and summary of HLFC application in terms of drag reduction potential, assuming a relative transition position of 50 %.

The total suction area can be used to estimate the two remaining design parameters, i.e., mass increase and suction power. The underlying assumption is that both parameters scale linearly with the suction area. For this, previous studies of HLFC designs have been consulted, namely Refs. [18, 29, 70, 218, 296], in order to plot the mentioned mass increases and suction powers over the suction area of each aircraft design. The results are shown in Fig. 3.10. The mass increase due to HLFC in the selected literature (top plot) ranges from 5.7 to 16 kg per m^2 of suction area. With the 68.7 m^2 suction area of the present aircraft design, this translates to a mass increase between 384 and 1104 kg. Due to the deterministic nature of this study, the intermediate value of 744 kg is chosen. The same approach yields 174 kW for the suction power P_{suc} (bottom plot). Note that this refers to the aerodynamic power needed to stabilize the boundary layer. To calculate the power taken from the engine P_{HLFC}, the efficiency losses

of the compressors, inverters, power lines, and generators need to be considered. For these, recent literature has been reviewed, resulting in the ranges shown in Table 3.2.

Table 3.2 Efficiencies of electric components for the HLFC operation.

Component	Symbol	Range	Average	References
Compressors	η_{comp}	0.92 . . . 0.97	0.95	[110, p. 874] [144, p. 7]
Inverters	η_{inv}	0.93 . . . 0.99	0.96	[17, p. 3] [173, p. 4399]
Generators	η_{gen}	0.81 . . . 0.87	0.84	[264, p. 1011]
Power Lines	η_{PL}	0.95	0.95	[137, p. 10]

As the present study is deterministic, i.e., does not aim to quantify input or output uncertainties, the averaged values are used, resulting in:

$$P_{\text{HLFC}} = \frac{P_{\text{suc}}}{\eta_{\text{comp}}\eta_{\text{inv}}\eta_{\text{gen}}\eta_{\text{PL}}} = \frac{174\,\text{kW}}{0.95 \cdot 0.96 \cdot 0.84 \cdot 0.95} = 240\,\text{kW} \tag{3.7}$$

For the final translation to SFC penalty, an estimation from a performance model of an A330 type of engine from Young [296, p. 89] is used. Here, a slope of 0.5 % SFC penalty per 100 kW of power offtake is mentioned, resulting in an SFC penalty of 1.18 % (during cruise).

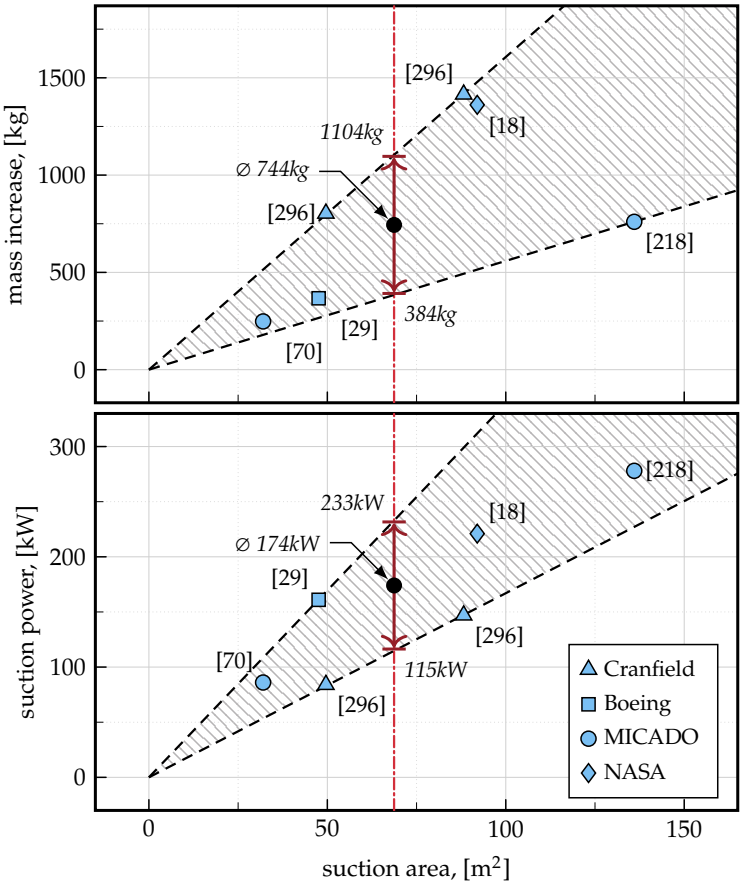


Figure 3.10 Estimation of mass increase and suction power based on HLFC studies from Young [296], Boeing [29], and the OAD environments MICADO [70, 218] and FLOPS [18].

As a last note on the design considerations: In this and all upcoming analyses, it is assumed that the replacement of the conventional high lift system of the A330 with a Kruger device does not introduce a change in mass, maintenance cost, or purchase price.

Domain 2: Operational Assumptions

The first operational assumption addresses the cruise speed. For this assessment, it is assumed that the airline operates the HLFC aircraft as well as its turbulent counterpart with a constant cruise Mach number of 0.82. Load factors are also assumed to be constant, regardless of the route or season, at 80 % of the maximum passenger capacity, resulting in 204 carried passengers for each flight. The passenger payload, which includes the total weight of all passengers aboard, is assumed to be 20,400 kg, calculated with an average of 100 kg per passenger, encompassing personal belongings and standard carry-on luggage. It is further assumed that the airline does not carry additional cargo, focusing solely on passenger service and associated luggage for revenue generation.

The next assumption revolves around contingency fuels that are carried specifically for the HLFC operation. Consider, for example, an upcoming flight of 3000 NM and 20,400 kg of payload. The turbulent reference aircraft burns m_0 = 32,750 kg of fuel while carrying only the reserve fuels that are required by law, i.e., 5 % of the trip fuel plus reserves for flying to an alternate airport as well as for holding patterns. These reserve fuels sum up to 6400 kg for this mission. If the same mission is performed by the HLFC aircraft (without any additional HLFC-specific reserve fuels), two cases can be distinguished using the state vector

$$\varsigma = \left[\Delta\text{drag}, \Delta\text{mass}, \Delta\text{SFC}\right] \quad . \tag{3.8}$$

The two cases are:

1. HLFC is activated and working as expected, corresponding to a state vector of $\varsigma_1 = \left[-7.4\,\%, 744\,\text{kg}, 1.18\,\%\right]$, which represents an operation under ideal conditions. This leads to a fuel burn of m_1 = 31,000 kg, which is 5 % less than the turbulent aircraft. Using the same reserve fuel rules, this aircraft carries 6200 kg of (unburned) fuel reserves for this mission.

2. The HLFC system is known to have a malfunction and hence turned off. The state vector here is $\varsigma_2 = \left[0\,\%, 744\,\text{kg}, 0\,\%\right]$, i.e., there is no drag reduction and no SFC penalty (as the system is turned off), but the mass of the system is still taken into account. This leads to a fuel burn of m_2 = 32,800 kg, which is almost the same as the HLFC aircraft and hence 5 % (or 1800 kg) more than the HLFC aircraft in case 1.

As the HLFC aircraft's malfunction in case 2 may occur after take-off, the unexpected increase in burned fuel has to be taken from the reserves. As these reserves have their specific purpose and have to be carried by law, operators may wish or need to carry a HLFC-specific contingency fuel. If, for example, an additional 5 % of the trip fuel is carried in order to capture any unforeseen HLFC-related effects, the fuel burn in the above mentioned cases changes to:

1. If no failure occurs, the additional HLFC contingency fuel is never burned. Its weight leads to a increase of the fuel burn to $m_1^{'}$ = 31,150 kg, slightly reducing the fuel consumption improvement.

2. If a failure occurs, the additional HLFC contingency fuel is burned. As the take-off has increased, the new fuel burn is $m_2^{'}$ = 33,200 kg, which represents an even worse fuel consumption at the benefit of not using the conventional reserve fuels.

This example highlights the dependencies between the additional weight of HLFC-related reserve fuels and the expected fuel burns. If no malfunctions occur, carrying no HLFC-specific reserves is better for the expected fuel consumption, as the additional weight will not be burned. Considering HLFC-specific reserves increase the takeoff weight of the aircraft and hence its fuel burn, but in case of a malfunction, this approach may be safer as the conventional reserves remain largely untouched. For this assessment, a HLFC-specific contingency fuel of 2.5 % of the trip fuel is carried before each mission, which is on top of the other reserve fuels. Note that neither this nor any other study in this thesis simulate the occurrence of a malfunction due to the lack of information on failure times.

Domain 3: Environmental Assumptions

While many HLFC design and assessment studies neglect the degrading effects of the environment in which the aircraft operates (e.g., Hartmann et al. [98]), those that do incorporate them have shown that their effect can be significant [294]. To date, there has been no comprehensive analysis of the environmental conditions, as long-term experience is not available. Out of the many aspects mentioned in Section 3.2.2, the insect contamination and cloud encounter effects are incorporated into this study. The reason for this is that their impact on the HLFC's efficacy

has been somewhat studied in the past, whereas the effects of other environmental occurrences such as ash, sand, and dust have only been dealt with qualitatively.

The insect contamination can be modeled in a wide range of detail. In fact, this field is actively researched, significantly driven by Young et al. [295], with foci revolving around anti-contamination devices, different insect types and the protein stains they leave behind, and angles of impact. Their influence on laminarity has also been studied in Elsenaar et al. [73], where an average laminar efficacy of 65 % has been reported. While seasonal variation was observed, with peak values in summer, the assumption for this analysis is a year-round fixed value, which is set to $\eta_{\mathrm{L,ins}} = 65\,\%$. As the Kruger device is assumed to shield successfully the upper side of the wing from any insect contamination,[6] this insect-induced loss of laminar efficacy only affects the laminarity of the HTP and VTP.

The degradation of laminarity due to cloud encounter has also received some attention, e.g., in Young [296] as well as Pohya et al. [197, 198, 202]. The modeling detail can vary from mission-specific analyses where the fuel flow of the HLFC aircraft reacts to the current presence of clouds to a more top-level consideration where the average time in cloud is translated to an average loss of laminarity. The deterministic setting pursues the latter, using an average time in cloud of 10 % and a resulting average loss of laminarity (due to clouds) of $\eta_{\mathrm{L,cld}} = 93\,\%$. This degradation affects all laminarized surfaces. The overall laminar efficacy at aircraft level is composed of the individually degraded drag savings, scaled by their contribution mentioned in Fig. 3.9 (b), which stated that the wing made up 72 % of the total drag savings, whereas the tails contribute 28 %. The laminar efficacy at aircraft level is thus:

$$\eta_{\mathrm{L}} = \underbrace{0.72 \cdot \eta_{\mathrm{L,cld}}}_{\text{wings}} + \underbrace{0.28 \cdot \eta_{\mathrm{L,cld}} \cdot \eta_{\mathrm{L,ins}}}_{\text{tails}} \tag{3.9}$$

$$\approx 0.72 \cdot 0.93 + 0.28 \cdot 0.93 \cdot 0.62 \tag{3.10}$$

$$\approx 0.82 \tag{3.11}$$

Domain 4: Economic Assumptions

The last set of assumptions discussed here are of the monetary nature. They comprise the fuel price development as well as the increase of maintenance cost and aircraft price due to the application of HLFC. The fuel price development is already modeled in LYFE as a time series with three different scenarios based on forecasts provided by the U.S. Department of Transportation [272]. For the present analysis, the so-called "*base*" scenario is chosen, which starts at \$ 0.79/kg in 2030 and increases to \$ 0.94/kg in 2050, as shown in Fig. 3.11.

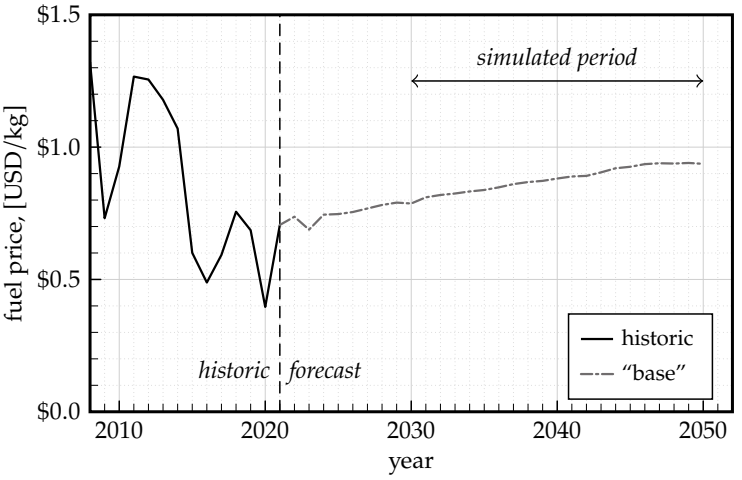


Figure 3.11 Historic and projected kerosene price under the scenario "*base*", taken from U.S. Department of Transportation [272].

[6] The Kruger device specifically shields the upper side of the wing, where laminarity is primarily targeted and maintained, as previously mentioned.

The increase in scheduled maintenance costs due to HLFC is estimated by evaluating multiple DOC methods using the mass increase of 744 kg derived earlier. Results are shown in Table 3.3, ranging from $12k to $22k annually. After averaging these, the assumed value for the scheduled maintenance cost is $16.8k per year. Unscheduled maintenance cost are dealt with using LYFE's default method, which uses a ratio-based approach and stems from Suwondo [258]. The aircraft price increase is estimated using the similar CER as before (including one from LYFE itself), resulting in a range between $1.2M and $1.5M, see rightmost column in Table 3.3. The average of $1.3M is chosen for the aircraft price increase (before discounts). It should be noted that all monetary values are discounted to the year 2023.

Table 3.3 Estimation of annual scheduled maintenance cost and aircraft price increase due to the HLFC system using DOC methods.

source	maintenance	price
AEA [20]	$15,132	*not available*
TU-Berlin [268]	$22,145	$1,172,222
NASA [97]	$12,144	$1,462,834
NASA [150]	$17,708	*not available*
LYFE [203]	*not available*	$1,462,834
average	$16,782	$1,334,064

Further Assumptions when using the LYFE framework

Apart from the technology-specific assumptions described previously, the assessment framework requires additional parameters for each of the modules to perform the DES, which mostly concern the reference aircraft or HLFC unrelated elements. They can be categorized into the acquisition process and end of life behavior, as well as the flight and maintenance schedule. These will be documented for transparency purposes next.

Acquisition Process and End of Life It is assumed that the aircraft are purchased, not leased. The acquisition process model is based on Marx et al. [159], where an initial downpayment of 10 %, a final payment of 70 %, and several distributed payments accounting for 20 % of the aircraft price are foreseen. The time between the first order and delivery is assumed to be 5 years for both aircraft. The price for the reference A330 is determined using the list price of $264.2M (as of 2018). As airlines are known to rarely pay the list price [46, p. 1], a discount of 30 % is assumed, resulting in a purchase price of $185.5M for the reference aircraft. This also lessens the price increase due to HLFC from the aforementioned $1.3M to $930k, resulting in a purchase price for the HLFC aircraft of slightly over $186.4M. Other aircraft price driven cost elements such as insurance or depreciation use LYFE's default process described in Pohya et al. [201]. At the end of the simulation, the aircraft is assumed to be sold, which generates the final cash inflow. The aircraft's value at the end of the simulation is estimated to be 10 % of its purchase price (without discounts).

Flight and Maintenance Schedule The flight schedule is a table-based input for LYFE, specifying the origin, destination, flown distances, and factors for passenger and cargo loads. Note that dedicated departure and arrival times are not required, as these are simulated by LYFE. The overall network, which is the basis for the flight schedule, was already introduced in Fig. 3.8, where 89 unique origin-destination pairs were identified. The database from which this network was obtained provides information about the scheduled flights of an airline and/or aircraft type, whereby the number of departures per queried time period is given. With some minor adjustments (such as the information about the load factors), this data can be used directly within LYFE. These routes are set to be flown in a random sequence, whereby origin and destination compatibility is checked and local curfews[7] are considered automatically[8].

The next set of inputs pertains to the aircraft's maintenance schedule. A comprehensive tabular maintenance schedule, based on the works of Aircraft Commerce [5, 6], is utilized.

[7] Strictly speaking, the curfew consideration in LYFE considers *typical operational times* of the airports rather than actual curfews. These were obtained through an extensive analysis of the Sabre Airline Solutions [228] database between 2010 and 2019.

[8] Supplemental information on the reference aircraft including the flight schedule can be found in Appendix C.2.

This schedule includes events for line and base maintenance, heavy component maintenance, and dedicated shop visits. Maintenance intervals are specified in Flight Hourss (FHs), Flight Cycless (FCs), or time. The maintenance module in LYFE automatically triggers each respective maintenance event, taking into account downtime as well as the required man-hours (MH) and material costs.

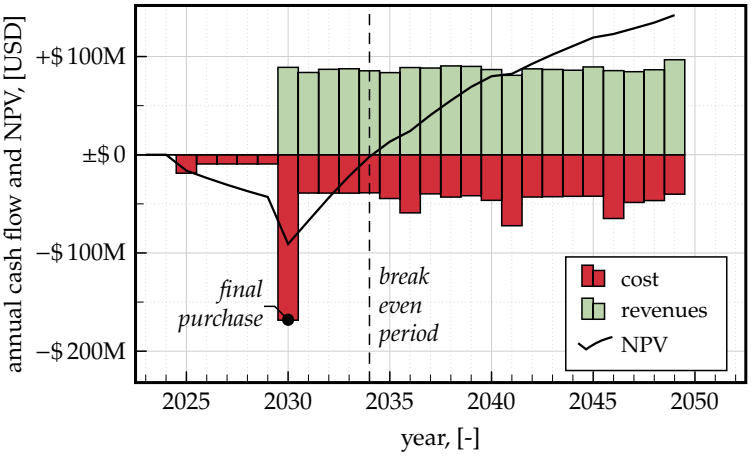
For an overview of assumptions for this deterministic study, see Appendix C.3.

3.3.3 Results

The results of the discrete event simulation are discussed in this Section. First, the reference aircraft is analyzed with respect to its overall economic value and a more detailed look at its operating costs. As mentioned in the Section 2.1.4, meaningful results are only obtained in a comparative setting, these first results only serve overview and comprehension purposes. The actual assessment is performed afterwards, where the simulation results of the HLFC aircraft are compared to those of the reference. A local sensitivity analysis of selected parameters completes this Section.

Reference Case

The first set of results represents the overall economic value of operating the conventional (turbulent) Airbus A330-343. These results are best introduced by the annual breakdown of costs, revenues, and the development of the NPV, as shown in Fig. 3.12 (a). The years from 2025 through 2029 are non-operational, i.e., the aircraft has been ordered and the first payments have been made, but the airline has not received the aircraft, yet. The final purchasing payment in 2030 heralds the operational phase. Here, annual revenues (from ticket and ancillary sales) remain fairly constant at around $87M except for a slightly higher value in the last year, when the aircraft was sold for its residual value of $12.6M. The annual costs in the operational phase range from $39M to $72M. Peaks represent expensive maintenance events (which are discussed in more detail later). The NPV curve starts decreasing between 2025 and 2030 as no revenues have been generated here yet. As soon as the aircraft begins to operate, the economic value begins to increase, crossing the $0 axis after 4 years of service, which marks the BEP. It continues to rise to its final value of $142.2M in 2049, when the simulation ends after 20 years of operation. Some complementary KPIs are summarized in the table in Fig. 3.12 (a) including the IRR and Cumulative Direct Operating Cost (CDOC) as well as some operational statistics. Note that the lifetime averaged γ_{fuel} value is slightly higher than the one calculated for the average mission in Eqn. (3.2). This is caused by the shorter routes of the underlying



(a) annual cost, revenues, and NPV.

parameter	value
NPV	$142.2M
BEP	2035-02
CDOC	$1098M
$\varnothing\gamma_{\text{fuel}}$	3.18 kg/(pax · 100km)
ØFH/FC	6.9 h
Ødistance	2992 NM
utilization	639 year
utilization	4400 h/year

(b) key economic and operational metrics.

Figure 3.12 Overall results of the (turbulent) reference aircraft in the deterministic assessment setting.

flight schedule, where the fuel consumption of the A330 is generally worse compared to longer routes.

Next, a closer look is taken at the operating costs and their development throughout the simulation. These are shown in Fig. 3.13. The overall cost breakdown in the top right plot reveals that about a third of the CDOC are spent on fuel, followed by about a fifth on the ownership, maintenance, and crew costs. The smallest contribution is allocated to navigation and airport-related fees. As the other plots show, the cost elements (and thus their share in CDOC) vary significantly with time. The top left plot, which depicts the ownership-related cost, shows the acquisition process with the initial down payment, the three distributed payments, and the final delivery payment. After that, the ownership cost reduced to interest and insurance, out of which the expenses for the former became zero after 15 years of operation. From here on, the aircraft is considered to be paid off. The flyaway costs in the bottom left plot are mostly constant. Only a slight increase in fuel costs over time is observable, which originates from the chosen fuel price scenario discussed earlier. The final plot in the bottom right shows the highly irregular expenses for maintenance. Especially the costs for base maintenance and heavy components (particularly engine shop visits) lead to significant peaks.

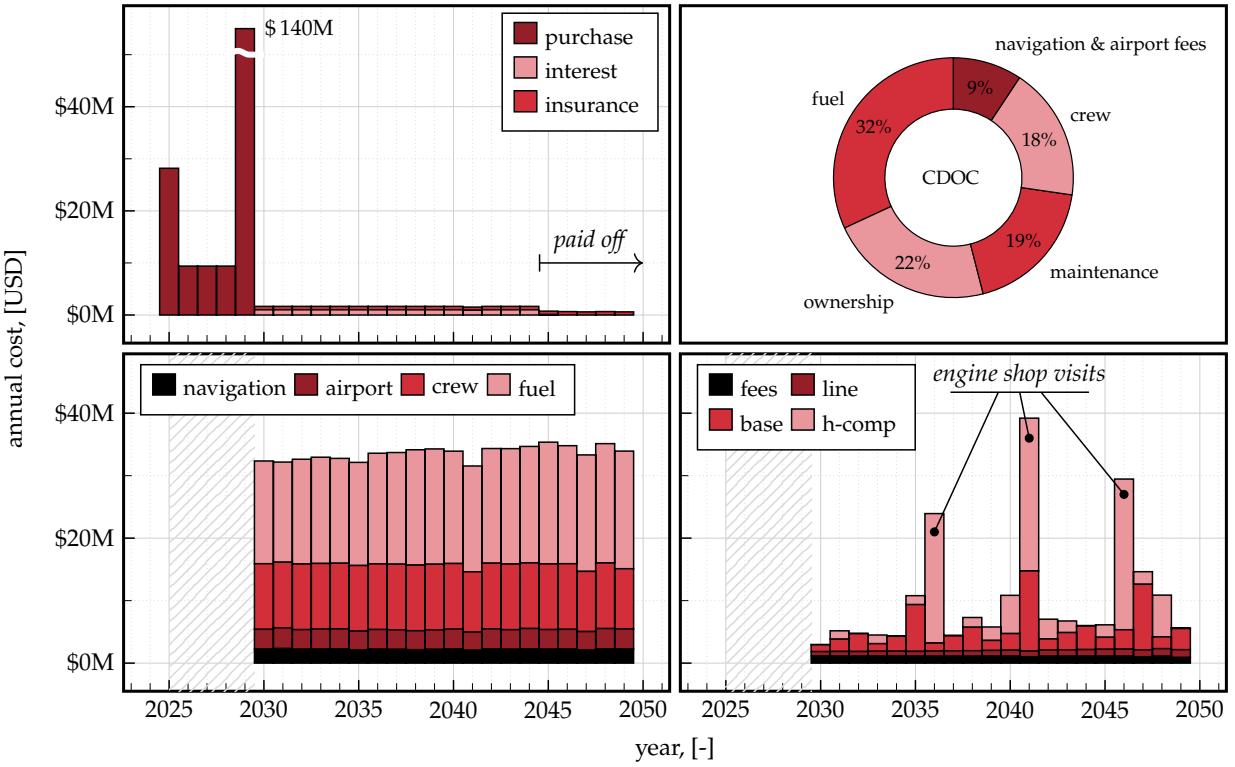
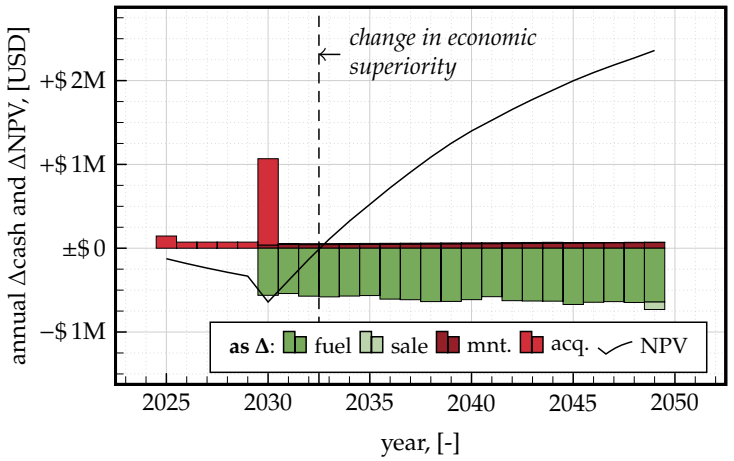


Figure 3.13 Breakdown of annual cost of ownership (top left), fly-away (bottom left), maintenance (bottom right) and overall share (top right) of the conventional (turbulent) reference aircraft in the deterministic assessment setting.

Comparison between HLFC Aircraft and Turbulent Reference

Fig. 3.14 shows the comparison between the two aircraft under investigation. In subfigure (a), the difference in annual cost and revenues is depicted, alongside the ΔNPV. The initial increase in acquisition cost represents the impact of the aircraft price increase, leading to a total increase of ownership cost of $1.58M. From 2030 onward, the annual HLFC maintenance cost start at $36k in 2030 and increase to $68k in 2049. This increase is due to the aforementioned unscheduled maintenance cost portion, which is implemented as an age-dependent function in LYFE. As the Table in Fig. 3.14 (b) shows, the total amount of HLFC maintenance cost is $1M. Both the maintenance cost increase and ownership cost increase are outweighed by the savings in fuel cost, which range from $540k to $671k annually. The variation stems from slightly

different annual utilizations as well as the fuel price model shown in Fig. 3.11. In total, more than $ 12M are saved due to the application of HLFC, which results in an increase of $ 2.46M in NPV (due to the time value of money). As the progression of ΔNPV in subfigure (a) shows, the HLFC aircraft becomes economically superior to its turbulent counterpart after three years of operation. Note how there is no change in BEP between the two simulations due to the HLFC technology requiring an initial investment before it leads to fuel cost savings.



(a) annual cost, revenues, and NPV.

parameter	value
ΔNPV	+$ 2.46M
ΔBEP	none
ΔCDOC	$ 9.5M
Δfuel	$ 12.2M
Δmaintenance	$ 1M
Δownership	$ 1.58M
$\Delta\varnothing\gamma_{\text{fuel}}$	3.06

(b) key economic and operational metrics.

Figure 3.14 Comparison of overall results between HLFC aircraft and conventional (turbulent) reference aircraft in the deterministic assessment setting.

As the latter stem from the decrease in fuel burn, it is worth discussing the impact of HLFC on the overall fuel consumption in more detail. As more than 12,000 flights per case are simulated, boxplots are chosen to analyze the overall impact, see Fig. 3.15. The upper plot shows that the median fuel consumption of the reference aircraft is 2.97 kg/(pax 100 km) and the IQR ranges from 2.95 to 3.03 kg/(pax 100 km). The HLFC aircraft has a noticeably better fuel consumption with an IQR and median of $[2.86, 2.92]$ and 2.87 kg/(pax · 100 km), respectively. This represents a median improvement of fuel consumption of 3.4 %, as the lower plot shows. The IQR of the fuel consumption improvement is $[-3.5, -3.2]$%. Since the payload has been fixed and no wind effects have been considered, the spread in $\Delta\gamma_{\text{fuel}}$ values ranging from −3.7 % to −2.8 % is solely caused by the flown distances[9]. Note that these values consider a fixed loss of laminarity due to cloud encounter and insect contamination, see Eqn. (3.11).

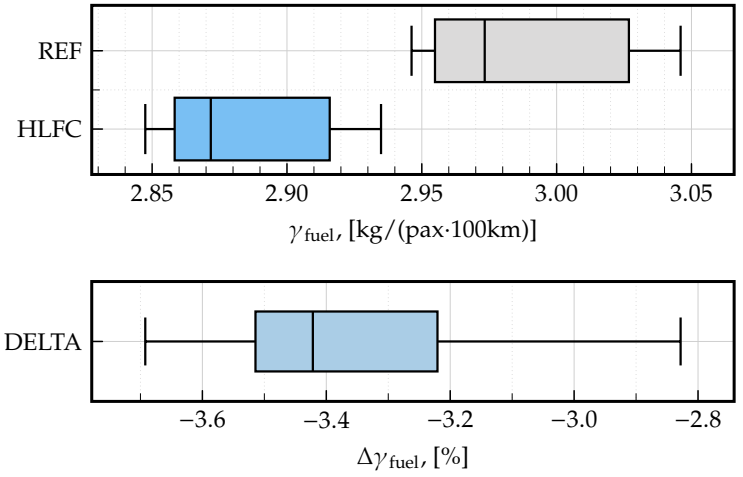


Figure 3.15 Boxplots of fuel efficiencies of the HLFC and reference aircraft (top) and the difference (bottom) in the deterministic assessment setting.

As a final step, a LSA is performed on the three design assumptions: the drag reduction potential, the mass increase, and the SFC penalty. Each of these parameters have been varied by ±25 % from their nominal value, while the others are kept fixed, resulting in 12 additional

[9] Refer to the discussion on p. 54f.

simulations. The results are depicted in Fig. 3.16, discerning the impact on the fuel consumption improvement $\Delta\gamma_{\text{fuel}}$ and economic superiority ΔNPV.

The sensitivity of both metrics is highest to the drag reduction potential, followed by the SFC penalty and mass increase. As expected, an improvement of the drag reduction potential (i.e., an increase by 25 %) leads to an overall better performing HLFC aircraft, reducing the average $\Delta\gamma_{\text{fuel}}$ by 1.1 %-points (or 38 % in relative terms). A +25 % change in SFC penalty decreases the fuel consumption improvement slightly, whereas the mass increase sensitivity is only minor, if not negligible. This behavior seems to be mirrored on the lower end of the parameter values, where a lower drag reduction potential decreases the fuel consumption improvement by 38 %, while the lower values of mass increase and SFC penalty lead to a (minor) improvement of $\Delta\gamma_{\text{fuel}}$. A similar picture is drawn for the economic superiority, as both are interconnected. Here, the increase of drag reduction potential by 25 % adds an economic value of $1.5M over the lifetime of the aircraft, whereas the same increase in SFC penalty increase subtracts less than $500k. The impact of the changes in mass increase is again minor, if not negligible.

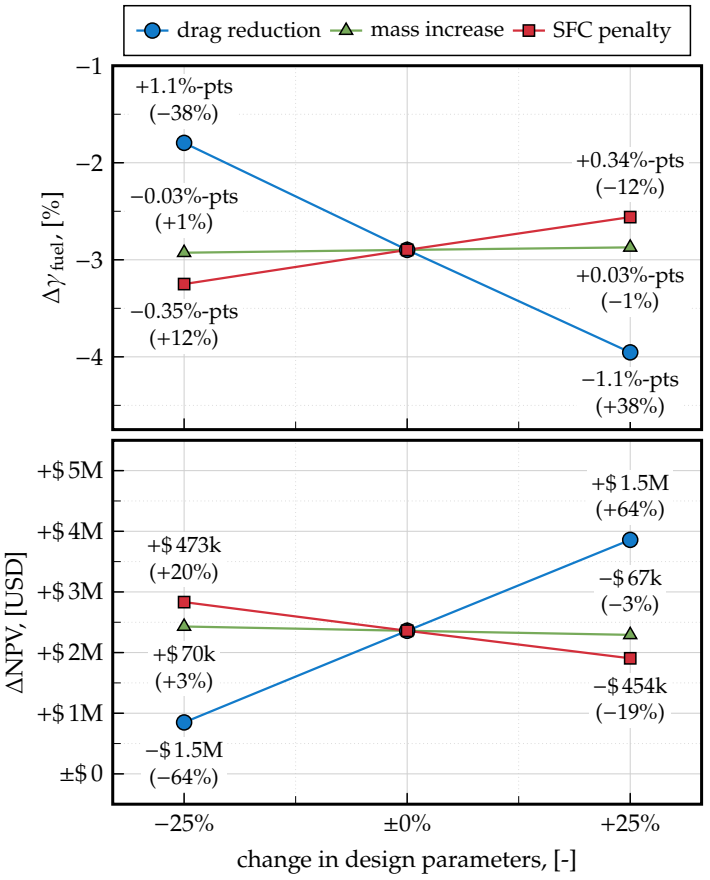


Figure 3.16 Results of the LSA in the deterministic setting (annotations represent the difference to the nominal value)

3.4 Conclusions

In this Chapter, the LYFE assessment framework and its integrated aircraft performance models were introduced. Subsequently, the HLFC use-case was detailed, emphasizing the technology's impact on and interplay with domains such as design and environment. Assumptions for each domain were discussed, and average values were derived from literature and other references. Using these average values, two lifecycle simulations were conducted, one for the conventional reference aircraft and one the HLFC aircraft. Results highlighted variations in NPV, CDOC, and other metrics, revealing an economic advantage of $2.46M for the laminar aircraft. When compared with the reference NPV of $142M, this superiority can be described as modest. A LSA on design parameters showed a significant sensitivity to the drag reduction potential. Given that this parameter, along with several others detailed in Section 3.3.2, was determined using broad estimations, the inherent uncertainties in this deterministic study are of significant concern and should be addressed systematically. As highlighted in Chapter 2, the LSA itself, while serving as a preliminary verification and a sanity check of the coding implementation, is not comprehensive enough for UQ. Therefore, the next Chapter will explore GSA methods, discussing the advantages and disadvantages of various techniques.

4 Global Sensitivity Analysis

This Chapter is driven by the first research question and hypothesis, which seeks to evaluate the capabilities, interpretability, and computational efficiency of various GSA techniques, aiming to facilitate their adoption in uncertainty-addressing TEAs. Specifically, the first hypothesis posited that through a comprehensive evaluation of these aspects on an example problem, followed by a case study demonstration, the insights gained will help promote the adoption of GSA in TEAs, informing future research directions and methodological choices in the field. Therefore, the practice of GSA is investigated in detail. According to Razavi et al. [213] and Saltelli et al. [230], GSA serves the following purposes:

- *Getting scientific insights* of the content in question, e.g., identifying and understanding parameter causalities and their interaction [93],
- *reducing complexity* by fixing or removing non-influential factors from a system [252], and
- *decision support*-based on the uncertainty measures, e.g., to direct future uncertainty reduction efforts or system design improvements [92, 263].

While all three are highly useful, the uncertainty inclusion efforts benefit most from the last aspect, especially considering the IPDP and the limited budget in RDTE. It should be recalled from the discussion in Section 2.2.2 that a number of different GSA methods exist, each with its own strengths and weaknesses. To investigate these and address the first hypothesis, multiple analyses have been conducted. The first set of studies are comparative of nature. Here, multiple GSA methods are applied on a mechanistic surrogate of LYFE that uses the same set of inputs and outputs as the original while being significantly less expensive in computational cost. The questions discussed in this comparative investigation are threefold:

(a) which insights are gained by each method,

(b) how are the different sensitivity measures interpreted, and

(c) how fast does the calculation of the sensitivity measures converge?

The second study in this Chapter then incorporates the most suitable GSA technique into LYFE and applies it to the use case of HLFC. Thereby, the focus shifts from the technique itself to the results, i.e., the sensitivity of the output uncertainty towards each input uncertainty. The outcome of this analysis is then used in the subsequent Chapters to identify which aspects of the TEA should receive more attention.

4.1 Investigated Methods

The comparative study performed first builds on an open source library for sensitivity analysis that is written in Python and named SALib [105]. The available methods can be classified into variance-based and non-variance-based ones. An overview of their characteristics is given in Table 4.1. The column N_{tot} represents the total number of samples that are created with each method, which is a result of u and N. The former (u) is the number of uncertain inputs, while the latter (N) describes the number of sample points (per variable) that are required to estimate the mean of the model response [230, p. 89][1]. The next column lists the provided sensitivity

[1] In other words: If one were to run a MCS with one parameter being varied, N would be the number of samples at which the mean of the simulation is considered to be converged.

Table 4.1 Feature comparison of GSA methods as implemented in SALib.

method	N_{tot}	sensitivity measures	supports** given-data	correlation	groups***
Variance-Based					
FAST [54, 229]	$N \cdot u$	S_I, S_T	○	○	○
FAST-RBD [194, 262]	N	S_I	○	○	○
Sobol'* [231, 237, 251]	$N(k \cdot u + 2)$	$S_I, (S_{II}), S_T$	○	○	●
SCSA [149]	N	S_a, S_b, S_c, S_T	●	●	○
Not Variance-Based					
Morris [170]	$N(u + 1)$	μ, μ^*, σ	○	○	●
DMIM [31, 194]	N	S_I, δ	●	●	○

* *here, calculating S_{II} is optional. If S_{II} is of interest, then $k = 2$, otherwise $k = 1$.*

** ●*: capability is mentioned to be supported,* ○*: capability is not mentioned to be supported.*

*** *refers to SALib's out-of-the-box capability as of version 1.4.7*

measures. Some methods only calculate one overall measure, whereas others calculate an entire set where each element covers a different effect such as interaction (denoted as S_{II}) or correlation (denoted as S_b). The last three columns indicate further capabilities, which are explained next.

Given-Data Capability is a crucial feature of certain GSA techniques, enabling them to work with existing data, such as observations or measurements, without the need for the input sampling sequence that generated them. Typically, GSA methods dissect the output uncertainty into its components by mapping specific samples to their corresponding outputs, which usually necessitates a defined sampling technique or sequence. This implies that the sampling process is inherently part of the GSA. However, methods with given-data capability bypass this requirement, allowing for the use of samples generated outside the GSA process. This capability is particularly valuable in practical applications where the input sampling sequence is unavailable or when using real-world observational data. A common alternative for methods without given-data capability is to fit a PDF to the observations or measurements, then use method-specific sampling techniques to generate samples from this distribution. This approach allows users to circumvent the method's limitation of not supporting given-data. However, this may not always be feasible or accurate, especially if the data does not conform well to a specific PDF, leading to potential inaccuracies in the GSA results. Therefore, GSA methods with given-data capability offer a significant advantage by providing greater flexibility and applicability in real-world scenarios.

Correlation Support is another important capability to consider when selecting an appropriate GSA method. Most GSA methods assume statistically independent[2] input parameters, an assumption that holds true in many cases [43, p. 876]. However, when this assumption is violated (e.g., due to the presence of correlation) the GSA results may be biased [64]. To address this, it is recommended to use a GSA method that explicitly accounts for correlated inputs or, depending on the researcher's knowledge and experience, techniques such as combining copulas and the previously mentioned given-data capability. This approach is adopted, for example, in Kucherenko et al. [140]. For additional information on GSA and correlation support, Jacques et al. [118], Xu et al. [290], and Chastaing et al. [39] are recommended readings.

Grouping Capability is an invaluable feature when the computational cost of a GSA method becomes prohibitive, for example, if the model itself is computationally demanding and/or the number of uncertain input parameters is high, e.g., greater than 100 [230, p. 92]. Grouping enables the combination of uncertain input parameters and the calculation of a single sensitivity measure per group. Under certain circumstances, it is possible to use the grouping feature

[2]It is important to note that correlation and dependence are not synonymous. Correlation is one of several measures of statistical dependence. Statistical independence implies a correlation of zero, but the reverse is not necessarily true.

to create "*supersaturated sampling designs*", where the number of samples N is smaller than the number of uncertain parameters u. Despite its utility, grouping is not fully established in GSA and is not well understood by practitioners [230, p.96]. Lastly, it is worth noting that the grouping capability indicated in Table 4.1 refers to whether the SALib package natively supports grouped variables. Technically, every GSA method can be extended, albeit with considerable effort, to handle groups as long as it treats the model as a black box.

While the aforementioned features are helpful for narrowing down the selection of suitable methods, a comprehensive understanding of each method's theory, sampling strategy (if applicable), and other characteristics is necessary for making an informed decision. Sections 4.1.1 and 4.1.2 provide a concise discussion of the variance-based and non-variance-based methods. This knowledge will equip readers with the necessary theoretical background to fully comprehend and critically assess the analyses performed in Sections 4.2 and 4.3.

4.1.1 Variance-Based Methods

The variance-based methods use, as their name suggests, the (co)variance[3] of the output to apportion the total uncertainty into its causing input constituents. Since this is one of the most commonly mentioned techniques, the underlying concept, i.e., the High Dimensional Model Reduction (HDMR), will be briefly described first.

Consider a square integrable function f on Ω^u, which is the u-dimensional unit hypercube

$$\Omega^u = (x \mid 0 \leq x_i \leq 1;\ i = 1, \ldots, u)\,. \tag{4.1}$$

Assuming the inputs of the function are independently and uniformly distributed within Ω^u, it follows that f can be described as with a finite expansion into terms of increasing dimensions, each of which is also square integrable:

$$\begin{aligned} f(x_1, x_2, \ldots, x_d) = f_0 + \sum_{i=1}^{u} f_i(x_i) + \sum_{i<j}^{u} f_{i,j}(x_i, x_j) \\ + \ldots + f_{1,2,\ldots,u}(x_1, x_2, \ldots, x_u) \end{aligned} \tag{4.2}$$

In here, f_0 is the mean response of the model, whereas $f_i(x_i)$ is the first-order univariate function and represents the independent contribution due to the individual parameters. The bivariate functions $f_{ij}(x_i, x_j)$ quantify the interactions between x_i and x_j, with similar interpretations for higher-order interaction terms [249, p. 289]. For a function with $u = 3$ input variables x_1, x_2 and x_3, Eqn. (4.2) becomes:

$$\begin{aligned} f(x_1, x_2, x_3) = f_0 + \underbrace{f_1(x_1) + f_2(x_2) + f_3(x_3))}_{\text{first order functions}} + \underbrace{f_{12}(x_1, x_2) + f_{13}(x_1, x_3) + f_{23}(x_2, x_3)}_{\text{second order functions}} \\ + \underbrace{f_{123}(x_1, x_2, x_3)}_{\substack{\text{third order function} \\ \text{(usually neglected)}}} \end{aligned} \tag{4.3}$$

In practice, the expansion is typically approximated up to the second order functions, based on the assumption that higher order interaction terms have no significant effect on the response [249, p. 290]. As Sobol' proved, the terms of Eqn. (4.2) are orthogonal if each term has

[3]The structural and correlative sensitivity analysis presented by Li et al. [149] uses the covariance decomposition instead of the variance decomposition as discussed later.

a zero mean. Provided this, the terms of the functional decomposition can be expressed using conditional expectations, i.e.:

$$f_0 = E\left[f(x)\right] \tag{4.4}$$

$$f_i(x_i) = E\left[f(x) \mid x_i\right] - f_0 \tag{4.5}$$

$$f_{ij}(x_i, x_j) = E\left[f(x) \mid x_i, x_j\right] - f_0 - f_i - f_j \tag{4.6}$$

With this perspective, it becomes clear that first order function is equal to the effect of varying x_i only, whereas the second order functions describe the *additional* effect due to the simultaneous variation of x_i and x_j. After squaring and integrating each term of Eqn. (4.2) over Ω^u, the variance decomposition can be obtained, i.e.,

$$\int f^2(x)\mathrm{d}x - f_0^2 = \sum_{s=1}^{u} \sum_{i_1<\ldots<i_s}^{u} \int f_{i_1\ldots i_s}^2 \mathrm{d}x_{i_1} \ldots \mathrm{d}x_{i_s} \tag{4.7}$$

$$\Leftrightarrow V(f(x)) = \sum_{i=1}^{u} V_i + \sum_{i<j}^{u} V_{ij} + \ldots + V_{12\ldots u} \tag{4.8}$$

where the left and right hand side represent the total and partial variances, respectively. Similarly to Eqs. (4.5-4.6), these partial variances can be rewritten using conditional variances, i.e.:

$$V_i = V(E\left(f(x) \mid x_i\right)) \tag{4.9}$$

$$V_{ij} = V(E\left(f(x) \mid x_i, x_j\right)) - V_i - V_j. \tag{4.10}$$

For the three parameter example, the full variance decomposition reads:

$$\begin{aligned} V(f(x)) = &\underbrace{\int f_1^2(x_1)\mathrm{d}x_1}_{V_1} + \underbrace{\int f_2^2(x_2)\mathrm{d}x_2}_{V_2} + \underbrace{\int f_3^2(x_3)\mathrm{d}x_3}_{V_3} \\ &+ \underbrace{\iint f_{12}^2(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2}_{V_{12}} + \underbrace{\iint f_{13}^2(x_1, x_3)\mathrm{d}x_1\mathrm{d}x_3}_{V_{13}} + \underbrace{\iint f_{23}^2(x_2, x_3)\mathrm{d}x_2\mathrm{d}x_3}_{V_{23}} \\ &+ \underbrace{\iiint f_{123}^2(x_1, x_2, x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3}_{V_{123}} \end{aligned} \tag{4.11}$$

The Sensitivity Indices (SIs) are then obtained by dividing both sides of the variance decomposition by the total variance $V(f(x))$. This leads to the first order SI:

$$S_i = \frac{V_i}{V(f(x))} = \frac{V(E\left(f(x) \mid x_i\right))}{V(f(x))} \quad , \tag{4.12}$$

which describes the main effect of the i-th parameter, and the second order SI

$$S_{ij} = \frac{V_{ij}}{V(f(x))} = \frac{V(E\left(f(x) \mid x_i, x_j\right)) - V_i - V_j}{V(f(x))} \quad , \tag{4.13}$$

captures the interaction effect between the i-th and j-th parameters. Beyond the first and second order SIs, the total sensitivity index S_T is essential. It measures the overall sensitivity of the model to a particular parameter, including all interaction effects with other parameters. For instance, the total sensitivity index for x_1 is:

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123} \quad . \tag{4.14}$$

This means that the total sensitivity index includes the main effect of x_1 and its interactions with every combination of other parameters. Consequently, while the sum of all first order and higher order SIs equals 1, each total SI can be greater than 1 due to these interaction effects.

While all four variance-based GSA techniques mentioned in Table 4.1 use the HDMR, they differ in their approach to sampling the input space and (numerically) computing the SIs.

Fourier Amplitude Sensitivity Test (FAST)

FAST was first introduced by Cukier et al. [52–54] and has become one of the most popular GSA techniques [291]. Compared to other methods, FAST requires fewer samples (N_{tot}) to calculate the SIs, using the ergodic theorem to simplify the computation of multi-dimensional integrals[4]. In FAST, the input space is populated with N samples distributed over a periodic curve, each assigned a distinct frequency $\{\omega_1, \omega_2, \ldots, \omega_n\}$. These samples are sequentially fed into the model, and the Fourier transformation of the output is performed to decompose its variance into partial variances, as shown in Eqn. (4.7). First-order sensitivity indices are calculated using Eqn. (4.12). An extension by Saltelli et al. [229] allows FAST to calculate the total sensitivity index for a parameter. However, FAST does not support the calculation of parameter-specific higher-order SIs. Additionally, transforming the input and output space into the frequency domain precludes the use of given-data, and the method does not support correlated inputs. FAST may become unstable and biased when the number of inputs exceeds 10 [114, 269]. FAST works by transforming the problem into the frequency domain, where the influence of each input parameter on the output can be studied separately. By assigning unique frequencies to each input parameter, FAST can efficiently decompose the output variance into contributions from each input and their interactions, reducing the need for a large number of samples. However, it cannot directly use given-data and assumes uncorrelated inputs, which may not be practical in all scenarios. When the number of input parameters exceeds 10, the method's performance can degrade, leading to instability and bias.

FAST with Random Balanced Design (FAST-RBD)

FAST-RBD was introduced by Tarantola et al. [262] as a modification of FAST, where the input space is explored using only one frequency ω. This method leverages a LHS-based random permutation to ensure sufficient input space coverage, significantly reducing the number of required samples by a factor of u, the number of uncertain parameters. This reduction in sample size decreases computational effort while maintaining the efficiency of the sensitivity analysis. Despite these advantages, FAST-RBD has limitations: it cannot calculate total SIs or interaction effects, handle correlated input parameters, or utilize given-data situations. By combining the strengths of the FAST method with the computational efficiency of random balance design, FAST-RBD provides a more efficient approach to sensitivity analysis, especially in complex models with many input parameters, but it sacrifices some of the versatility offered by the original FAST method.

[4]The ergodic theorem states that, under certain conditions, the time average of a function along the trajectories of a dynamical system will converge to the ensemble average. In FAST, this means that the average behavior observed over time (or samples) can be used to estimate the overall behavior of the system, simplifying the computation of multi-dimensional integrals.

Sobol' Indices

The Sobol' method is an extension of FAST that employs a Monte Carlo-based approach [230, p. 160], approximating the partial variances directly through a large number of model evaluations [251]. To achieve this, the input space is sampled using a low-discrepancy approach, meaning it is filled more evenly than pseudo-random sampling, utilizing the Sobol' sequence. This high number of samples is necessary to calculate not only the first but also the second-order SIs, as well as total effects. The computation of higher-order SIs is optional and can be omitted if deemed unnecessary by the user. Several studies [176, 233] have demonstrated that the Sobol' method yields more accurate results than the FAST technique. However, it is important to note that, similar to the FAST and FAST-RBD methods, the Sobol' method does not support given-data and input parameter correlation support in its list of capabilities.

Structural and Correlative Sensitivity Analysis

The SCSA method is one of the more comprehensive ones in terms of capabilities and provided information. Introduced by Li et al. [149], it is able to handle input distributions that are independent and/or correlated. To do so, the approach foresees a decomposition of the covariance of the unconditional variance of the output. This process is essentially a generalization of the variance decomposition and simplifies to the previously described variance decomposition if the input variables are independent. Consequently, the provided sensitivity measures can also be interpreted as an extension of the independence-assuming SIs in Eqn. (4.12) and (4.13). More particularly, in SCSA, the first order, second order, and total SI distinguish three different measures, each. The first describes the *non-correlative* contribution and is equivalent to Eqn. (4.12) and (4.13) and is denoted with an a, i.e., S_I^a, S_{II}^a, and S_T^a. The second one describes the *correlative* contribution and is analogously denoted with b. The third sensitivity measure is termed *total* in the SCSA method but will be referred to as *combined* SI to circumvent confusions with S_T. It is the sum of the non-correlative and correlative SI, e.g., $S_I^c = S_I^a + S_I^b$. A specialty of this method is that it does not require a specific sampling strategy, as the algorithm identifies the patterns between the input space and output space on its own. Therefore, the SCSA method is the only variance-based method that supports given-data situations.

4.1.2 Non-Variance-Based Methods

The following methods do not use the variance to calculate the sensitivity of the output towards its inputs. Instead, either a derivative-based or distribution-based approach is taken. Both ideas are explained below, followed by a description of the most common GSA method in each category.

Derivative-Based Approaches and the Method of Morris

As their name suggests, derivative-based methods use the derivative (or an approximation of it) of a model's output with respect to its input to calculate sensitivity measures. Their calculation is more simple and less computationally expensive compared to the SIs of the variance-based methods but often provide less information. Their use is frequently recommended in situations where the number of uncertain inputs is too large to perform other GSA methods [141, 230]. A well-known representative of the screening techniques is the method of Morris [170] and its modification by Campolongo et al. [34].

The method of Morris provides two sensitivity measures per investigated input parameter: μ and σ [170]. To interpret their values, the concept of the "*elementary effect*" needs to be understood first. The elementary effect is similar to the finite difference approach in numerical differentiation where the derivative of a mathematical function or computational model is approximated. For a model Y with u independent inputs X_i (where $i = 1, \ldots, u$) defined on

the u-dimensional unit hypercube Ω^u, the elementary effect EE of the i-th parameter is defined as follows:

$$EE_i = \frac{Y(X_1, X_2, \ldots, X_{i-1}, X_i + \Delta, X_{i+1}, \ldots, X_u) - Y(X_1, X_2, \ldots, X_u)}{\Delta} \tag{4.15}$$

Thereby, Δ denotes a small increment, which has to be chosen in a way that $X_i + \Delta$ is still within Ω. By randomly sampling different $\mathbf{X}$ from Ω, one can obtain the distribution of elementary effects that are associated with the i-th input factor. This distribution is named F_i, i.e., $EE_i \sim F_i$. The sensitivity measures μ and σ describe the mean and standard deviation of F_i, respectively. A high μ indicates a high influence of the parameter on the output. A high σ indicates high nonlinear and/or interaction effects. A modification of Morris' μ was introduced by Campolongo et al. [34], where the modified mean μ^* uses the absolute values of the elementary effects, i.e., $|EE_i|$. This tweak mitigates an issue with the original μ when F_i contains both positive and negative values which partially cancel each other out when estimating the mean value. Additionally, μ^* has been mentioned to have similarities with the total sensitivity index S_T as it allows for parameter ranking, although no formal proof of the link was given [34]. As Saltelli et al. [230, p. 111-112] suggest, it may be insightful to calculate all sensitivity measures, i.e., μ, μ^*, and σ, especially considering the fact that the modification does not require additional computations.

The method of Morris is typically used with a specific sampling approach which requires several parameters to be understood and set, which may impede the usage of this method: The approach is based on the creation of a number of trajectories r in the u-dimensional input space. A trajectory represents a sequence where each element contains one sampled value per input dimension u. Every subsequent element in the trajectory differs from the previous one by an increment Δ in one randomly selected dimension[5] $i \in \{1, 2, \ldots, u\}$. The value of Δ depends on the underlying grid resolution of Ω^u. which is called *level* and is usually denoted with p. While n_r, u, Δ, and p evidently influence the final samples, their values seem to be set based on (potentially non-transferable) experience rather than calculated systematically [34, 230][6].

Distribution-Based Approaches and the Delta Moment-Independent Measure

In this category, the sensitivity measures are computed using more comprehensive information given by the output distribution as opposed to only using its variance. This is especially useful if the output has a multi-modal, highly skewed, and/or fat-tailed distribution [206] and the variance does not sufficiently represent the statistical characteristics completely [253]. That is also why methods from this category are often referred to as being (statistical) moment-independent. The DMIM, introduced by Borgonovo [31], is the most popular distribution-based method to date [253]. It "*measures the difference between the unconditional distribution of the output and its conditional counterparts*" (Razavi et al. [213, p. 3]). The DMIM provides the so-called delta-index, which is defined as

$$\delta_i = E_{X_i}\left[s\left(X_i\right)\right] = \frac{1}{2}\int\left[\int\left|f_Y(y) - f_{Y|X_i}(y)\right|\mathrm{d}y\right] f_{X_i}\left(x_i\right)\mathrm{d}x_i \quad , \tag{4.16}$$

where s is the shift (i.e., difference) between the unconditional PDF $f_Y(y)$ and the conditional PDF $f_{Y|X_i}$ of the output Y. With this approach, the method is capable of dealing with correlated inputs, although a quantification of the actual correlative effect is not foreseen. With the update

[5] Which is why this sampling technique is considered to be a one-at-a-time. However, this is not to be confused with one-at-a-time LSA. In the Morris' sampling strategy, every parameter is varied one-at-a-time with each sequential step to obtain one sensitivity measure, whereas in LSA the sensitivity measure of one parameter assumes all others to be fixed throughout the sequence.

[6] Supplemental information and illustrations on this topic can be found in Appendix D.1.

presented by Plischke et al. [195], the DMIM is capable of treating given data and was made independent of the sampling method.

With the theory and characteristics of the investigated GSA methods outlined, it is evident that these techniques differ fundamentally in their theoretical foundations, the information they use to define sensitivity, and the types of sensitivity measures they provide. Additionally, their capabilities vary widely, highlighting that theoretical knowledge alone is not enough to select the most appropriate GSA method. Given that this has been identified as a key barrier for a successful adoption of the uncertainty management methodology introduced in this thesis, the next Section focuses on a comparative analysis that will help in making an informed decision on the most suitable method for a specific application. This analysis is aligned with the first research hypothesis, which anticipates that through a comprehensive evaluation of these methods, valuable insights will be gained to facilitate the use of GSA in uncertainty-addressing TEAs.

4.2 Comparative Study of GSA Methods

The structure and setup of the analyses that make up the comparative study are presented in Subsection 4.2.1. The results, which cover aspects regarding output consistency as well as computational efficiency, are then discussed in Subsections 4.2.2 and 4.2.3.

4.2.1 Methodology and Use Case

The first and foremost goal of the comparative study is to gain practical knowledge about the different GSA methods. For this, two separate analyses are conducted, which both follow the methodology outlined in Fig. 4.1.

The process begins with input UQ, generating a PDF for each uncertain input parameter. A GSA method is selected, and the initial sample size is set. Depending on the GSA method, samples are drawn either randomly or using a specific sequence to define the input space.

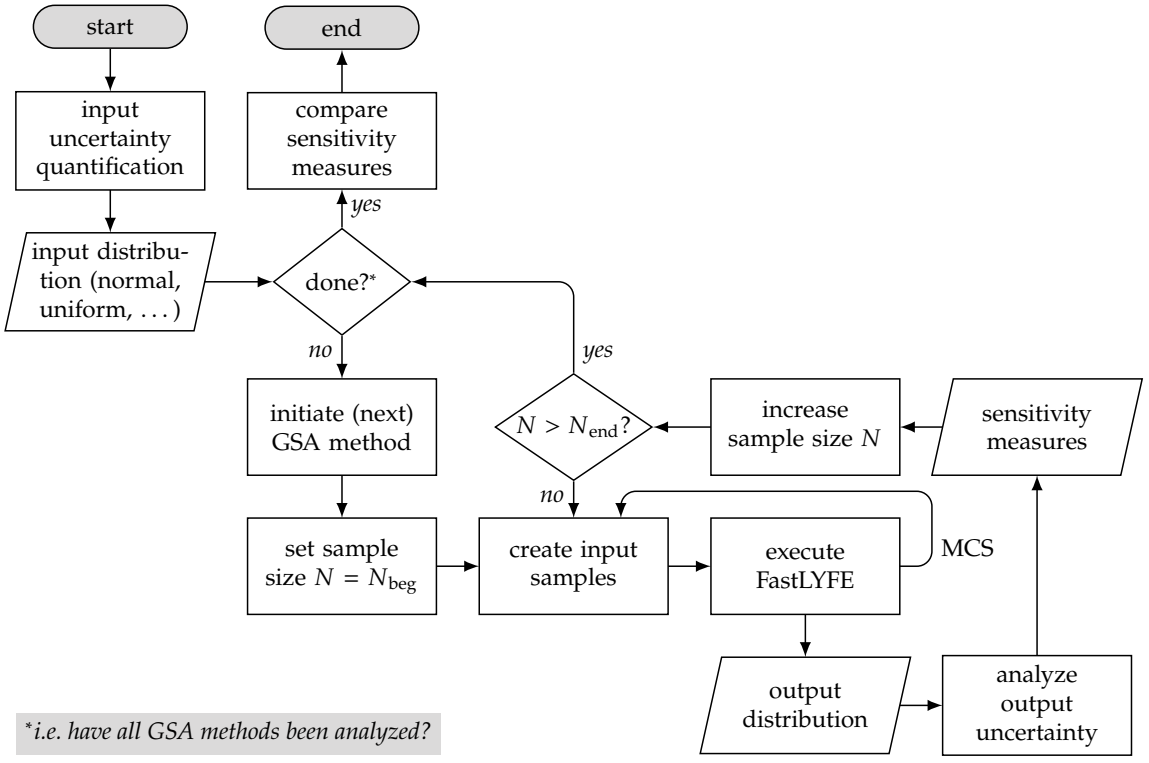


Figure 4.1 Methodology behind the comparative GSA study.

These samples are then processed through a MCS using the a surrogate model to calculate the KPIs for the TEA. This step is repeated across all sample sets, resulting in a distribution of KPIs, which is then analyzed to derive method-specific sensitivity measures.

In the first analysis of this study, the focus is on the interpretation and consistency of these sensitivity measures across different methods under identical conditions. To simplify the comparison, the effect of increasing sample sizes on convergence behavior is investigated separately in the second analysis, which also addresses computational expense.[7].

Fast Lifecycle Cashflow Environment (FastLYFE), which is used throughout this comparative GSA study, is a mechanistic surrogate of LYFE. It is used to approximate the estimated economic superiority of one technology over another with significantly less computational expense compared to LYFE[8]. Contrary to the DES approach of LYFE, FastLYFE is based on a set of regression-based CERs (which are based on LYFE's own methods). Therefore, instead of triggering specific flight and maintenance events when they are due, the surrogate model uses a flight hour-based estimation of the annual flight cycles and a weight-based maintenance cost relationship.

To further reduce complexity, the use case for the comparative GSA is not HLFC, but a fictive engine technology. This technology is assumed to reduce fuel consumption (modeled by applying an SFC factor $t_s < 1$ on the fuel consumption), but the exact value of t_s is uncertain. Similarly, the price of the technology is yet to be determined, representing the second of five investigated input uncertainties. The remaining three are the flown distance, carried payload, and fuel price. Table 4.2 shows the bounds of the uniformly assumed distribution for each parameter.

Table 4.2 Summary of input distributions for the comparative GSA.

	SFC factor	**flown distance**	**carried payload**	**fuel price**	**technology price**
Abbr.	t_s	d	m_{PL}	p_{fuel}	p_{tech}
Unit	[-]	[NM]	[tons]	[USD/kg]	[USD]
Distr.	$\mathcal{U}(0.95, 0.99)$	$\mathcal{U}(500, 4000)$	$\mathcal{U}(15.3, 23.0)$	$\mathcal{U}(\$0.5, \$1.8)$	$\mathcal{U}(\$0, \$10\text{M})$

Further assumptions for this comparative study are:

- The aircraft under investigation is the same A330-343 used in Chapter 3.
- The SFC factor is applied throughout the entire mission.
- The assessment is performed using FastLYFE with results expressed as ΔNPV.
- The lifetime is set to 20 years.
- The maintenance cost are estimated using the DOC method from TU Berlin [268].
- The discount factor for the NPV calculation is set to 10 %.
- All monetary values are inflation corrected to the fiscal year 2023.

4.2.2 Comparison of Sensitivity Measures

The results of the first analysis are categorized into variance-based and non-variance-based methods and include a discussion on the ranking of each parameter by order of criticality, as well as a comparison of the sensitivity values assigned to each parameter by the different methods.

[7] A third and complementary analysis can be found in Appendix D.2, where the effect of correlation is investigated.

[8] The deterministic study discussed in Chapter 3 was executed in 3 minutes with LYFE. Using FastLYFE, the computational time reduces to 25 seconds.

Variance-Based Methods

Fig. 4.2 depicts the first order SIs (solid bars) and total indices (solid plus hatched bars) using the ΔNPV as an output. This particular representation form explicitly shows the interaction effects one variable has with all others (which are represented by the height of the hatched bars).

Upon first inspection, the variance-based techniques produce consistent sensitivity measures. The average difference among the first order indices is 1 %-point and there is no indication that one method is generally more conservative or progressive than another. The consensus ranking is as follows[9], being (1) p_{tech}, (2) p_{fuel}, (3) t_s, (4) d, and, (5) m_{PL}. The latter appears to have virtually no effect, which is plausible from use case standpoint. Both fictitious aircraft (i.e., with and without the SFC improving technology) carry the same payload in each mission, and the effect of carrying more or less weight weight on the technology's relative fuel savings is minimal[10].

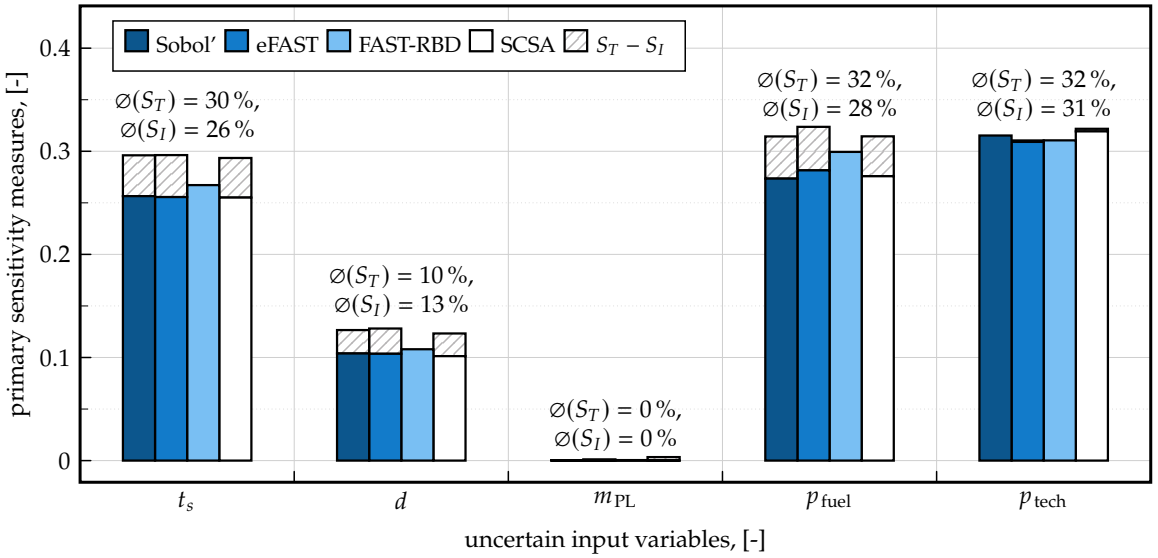


Figure 4.2 Comparison of first order and total sensitivity measures provided by the variance-based GSA methods.

Three out of the four variance-based methods quantify the total sensitivity S_T, thereby providing a measure of non-linear interaction effects represented by $S_T - S_I$. For p_{tech} and m_{PL}, these interaction effects are negligible. However, for t_s, d, and p_{fuel}, interaction and higher-order effects quantified by $S_T - S_I$ range from 3 % to 4 %. This is expected because the efficacy of the engine technology, represented by $t_s < 1$, the flown distance d, and the fuel price p_{fuel} are known to have a non-linear interactive effect on the cost savings, which in turn affects the discussed output ΔNPV. More specifically, the reduction in SFC leads to a reduction in fuel burn, but the absolute fuel savings have a non-linear dependency on the flown distance[11]. The resulting fuel burn of each aircraft (with and without the investigated technology) is then multiplied with the uncertain fuel price, which further amplifies the non-linear effects. m_{PL} showing no interaction effects is self-explanatory considering that its first order sensitivity index was also virtually non-existent.

[9] For this ranking, the first order sensitivity measure S_I was used.

[10] Note that this does not mean that the fuel consumption of either aircraft, with or without the fictive technology, is not sensitive to the carried payload. The sensitivity measures calculated here use the ΔNPV, in which the relative change in fuel consumption between both aircraft are considered. Therefore, the statement made here is concerned with the relative change in economic superiority and whether it is sensitive towards the carried payload.

[11] Less fuel burn leads to less fuel carried, which in turn further reduces fuel burn, further reducing the carried fuel (and hence weight), and so on. This effect is known to be stronger with longer routes, leading to the non-linear interaction effects between t_s and d.

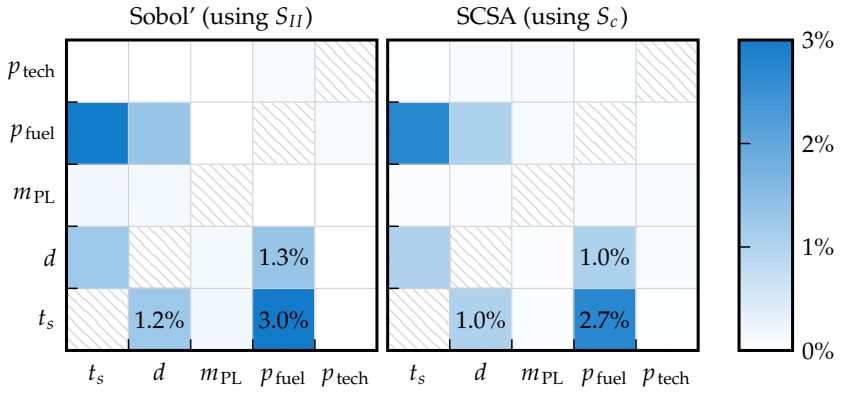
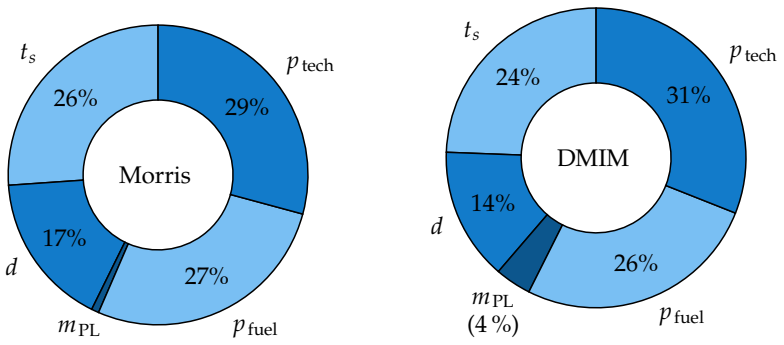


Figure 4.3 Comparison of second order sensitivity measures provided by the variance-based GSA methods using a matrix plot.

While these repercussions are relatively easy to explain in this example, understanding interaction effects in more complex situations might not be as straightforward. In such cases, the detailed and pairwise calculated second order SIs can be more useful. Among the variance-based methods, only the Sobol' and the SCSA method provide such. These are shown in Fig. 4.3 as a matrix plot. These plots are diagonally symmetric (as the interaction between two parameters has no direction, i.e., the interaction effect of p_{fuel} and d is the same as between d and p_{fuel}). It can be seen that the highest interaction (about 3 %) is in fact between p_{fuel} and t_s, while d and t_s as well as d and p_{fuel} make up the rest (about 1 % each). However, it should be clarified that the magnitude of the second order SIs are negligible when compared to the first order SIs. Furthermore, it can be concluded that there is no significant difference between the Sobol' and SCSA approach. The only rather minor mentionable aspect would be that SCSA tends to be slightly more conservative with the interaction sensitivity.

Non-Variance-Based Methods

In contrast to the variance-based methods, the derivative-based and moment-independent GSA techniques provide conceptually different sensitivity measures, as discussed in Chapter 4.1.2. After some postprocessing of the results[12], the ranking and sensitivity quantification of both methods can be compared. Fig. 4.4 shows the relative share of parameter sensitivity of each technique.



(a) relative share of μ^*. **(b)** relative share of δ.

Figure 4.4 Comparison of primary sensitivity measures provided by the non-variance-based GSA methods.

The parameter ranking is not only consistent between the Morris method and the DMIM, but also with that of the variance-based methods. Differences in the quantification are insignificant for the three leading parameters p_{tech}, p_{fuel}, and t_s. The lower two, however, show slightly higher deviations from one another. The sensitivity towards the distance parameter, for example, was consistently quantified by the variance-based methods at 10 %, whereas DMIM and the Morris method yield 14 and 17 %, respectively. DMIM also shows a non-negligible sensitivity towards the carried payload, albeit a rather small one. Three possible explanations

[12] Neither the Morris method nor the DMIM calculate a relative share of the sensitivity. Thus, the shown percentage values were calculated by dividing the parameter specific sensitivity measures by the sum.

include (a) numerical error, e.g., due to insufficient sample sizes and (b) the sensitivity of d and m_{PL} does not sufficiently manifest in the output variance for the other methods to properly quantify it.

As these deviations do not significantly alter the implications of the GSA with respect to decision support, complexity reduction, or gathering of scientific insight, a deeper investigation towards them is omitted at this point. Instead, the practitioner's perspective is taken next.

The derivative-based Morris method provides its results in absolute terms as summarized in the left side of Table 4.3 with μ being the sign-specific mean, μ^* the modified and sign-ignoring mean, and σ the standard deviation of the distribution of the elementary effects (recall Eqn. (4.15)). For t_s and p_{tech}, μ is negative. This can be explained by the fact that higher values of these parameters lead to a decrease in the output metric being the economic superiority of the technology equipped aircraft. The relationship is reversed for the other parameters, e.g., a higher fuel price makes the fuel savings technology even more desirable. Thus, μ can serve as a sanity check regarding the implementation. Apart from this, the absolute numbers obtained from the Morris method are of little value to users. Therefore, the relative (i.e., scaled so that the sum equals one) numbers are computed and shown in the right side of Table 4.3. The relative μ is not meaningful and should not be interpreted. The relative modified mean μ^*, however, is very consistent with SIs provided by the variance-based methods. In addition, the relative σ seems to capture the interaction effects correctly. That is, the highest interaction is assigned to t_s, p_{fuel} and d. It should be kept in mind that this σ is more of an *indicator* than it is a quantification, i.e., the values of σ are not directly comparable to its counterpart (being $S_T - S_I$) in the variance-based methods, but the ranking is. As for DMIM, the absolute δ is principally sufficient for interpretation as is. The relative counterpart was only computed for comprehensiveness reasons.

From a user's perspective, it should be noted that the Morris method, as described earlier, may require some level of engagement the sampling strategy and what the trajectories do in order to determine an appropriate value for the number of levels p, as the latter has to be supplied to the method when using SALib. In contrast, the usage of DMIM is very straightforward, as it does not require a method specific sampler and the underlying concept is comparatively easy to understand.

The following points conclude this first analysis:

1. The majority of investigated methods show consistent results, especially with respect to parameter ranking.
2. Regarding the first order SIs, it makes little difference which variance-based technique is selected. Users should consider the degree of information required to select the most suited approach. FAST-RBD, for example, only provides the first order sensitivity, whereas FAST additionally calculates the total sensitivity per parameter and Sobol' and SCSA complement the analysis by quantifying pairwise interaction sensitivity.

Table 4.3 Sensitivity measures provided by the Morris method in absolute numbers (left) complemented by the calculated relative values (right).

	absolute				***relative***			
	Morris			DMIM	Morris			DMIM
	μ	μ^*	σ	δ	μ	μ^*	σ	δ
t_s	−6,728,126	6,728,126	2,684,720	0.180	2.442	0.261	0.342	0.244
d	4,269,141	4,269,141	2,263,430	0.105	−1.550	0.166	0.289	0.143
m_{PL}	199,040	211,615	180,337	0.029	−0.072	0.008	0.023	0.039
p_{fuel}	7,034,501	7,034,501	2,715,588	0.195	−2.553	0.273	0.346	0.264
p_{tech}	−7,529,412	7,529,412	0	0.229	2.733	0.292	0	0.310
Σ	−2,754,855	25,772,795	7,844,075	0.737	1.000	1.000	1.000	1.000

3. The analyzed non-variance-based methods seem to be a reasonable alternative to the variance-based techniques. The Morris method shows some difficulty in practice as it is based on a specific trajectory-based input sampling strategy. Outputs of this method also require some postprocessing to be fully interpreted. The DMIM is easier to use but provides only one sensitivity measure.

4.2.3 Analysis of Convergence

The second analysis deals with the convergence of the investigated techniques. Here, a GSA is considered to be converged if the SIs no longer vary substantially above a certain sample size. While this is determined primarily on a visual inspection basis, the confidence intervals, which are calculated for each method and sample size and correspond to a confidence value of 95 %, serve as a numerical indication of convergence. More specifically, for each sample size N, the parameter averaged confidence interval is compared to a threshold (in this case 0.01), i.e.,:

$$\frac{1}{u}\sum_{i=1}^{u}(\mathrm{CI}_i(N)) \begin{cases} \leq 0.01 & \Rightarrow \text{indication for convergence,} \\ > 0.01 & \Rightarrow \text{no indication for convergence} \end{cases} . \tag{4.17}$$

Aside from these, the computational time[13] to execute the code is recorded with each sampling size so recommendations for practitioners can be formulated.

Input Uncertainties and Assumptions

The inputs for the convergence analysis are identical to those of the previous analysis (see Table 4.2). That is, all distributions are considered to be uniform and all parameters are sampled independently from one another. The sample size variation used here is as follows[14].

$$N = 2^x \quad \text{with} \quad x \in [9, 10, \ldots, 15] \tag{4.18}$$

$$= 512, 1024, \ldots, 32786. \tag{4.19}$$

Depending on the GSA method, very different N_{tot} are created, ranging from 512 to 393,216 (see the N_{tot} column in Table 4.1), which is expected to significantly affect both the convergence and computational expense.

Results

The results of the convergence analysis for each method are depicted in Fig. 4.5. The upper four subplots reflect the four variance-based methods, where the first order sensitivity index S_I is depicted on the left ordinate. The lower two show the modified mean μ^* and the δ of the Morris method and DMIM, respectively. Akin to the previous analysis, both μ^* and δ are quantified in relative terms for comparison purposes. Apart from these primary (relative) sensitivity measures, the confidence intervals are depicted as error bars. Finally, the dashed line, which the right ordinate refers to, shows the total time to compute the results. Note that, for ease of comparison, both ordinates and the abscissa are identical throughout all subplots.

The distribution of S_I in the Sobol' method (top left subplot) remains fairly constant as N increases, suggesting an early convergence. As indicated by the magnitude of the confidence intervals, this may be a mere coincidence, at least for small sample sizes. For $N \geq 2^{14}$ (which equates to $N_{\mathrm{tot}} \geq$ 393,216), the averaged confidence intervals begin to fall below the threshold of 0.01. Thus, the total execution time till convergence for the Sobol' method is 14 min in this analysis. Generally speaking, this technique is computationally heavy, especially when the number of uncertain parameters is high. However, it should be recalled that the Sobol' method

[13] Executed on a laptop with a 3 GHz i7 processor, run in serial (i.e., not parallelized) mode.

[14] While the values of x have been chosen heuristically, using a power of two is suggested by SALib [105] for the Sobol' sequence.

provides not only the sensitivity measures of first order, but also those of second order as well as total sensitivities.

The FAST method (top right subplot) seems to converge relatively early. According to the confidence intervals, a sample size of $N = 2^{10} = 1024$ (which corresponds to $N_{\text{tot}} = 5120$) is already sufficient. With an execution time of 22 sec at this sample size, this technique is one of the fastest. However, recall that the computation of higher order SIs is not computed by this method.

Also appearing to converge rather quickly is the SCSA (center left subplot), which provides the most comprehensive set of sensitivity measures (including correlation). The sample size of $N = 2^{12} = 4096 = N_{\text{tot}}$ is where the confidence intervals begin to dip below 1 %. Taking only about 20 sec to compute (at this N), this method is exceptionally efficient, especially due the fact that N_{tot} is always equal to N, regardless of the number of uncertain input parameters.

The last variance-based method (center right subplot), being FAST-RBD, shows relatively large confidence intervals for most N. The distribution seems to have settled in only at the highest sample size of $N = 2^{15} = 32{,}768 = N_{\text{tot}}$, which is also where the confidence intervals begin to drop below the threshold. However, with a computational time of 136 s at this sample size,

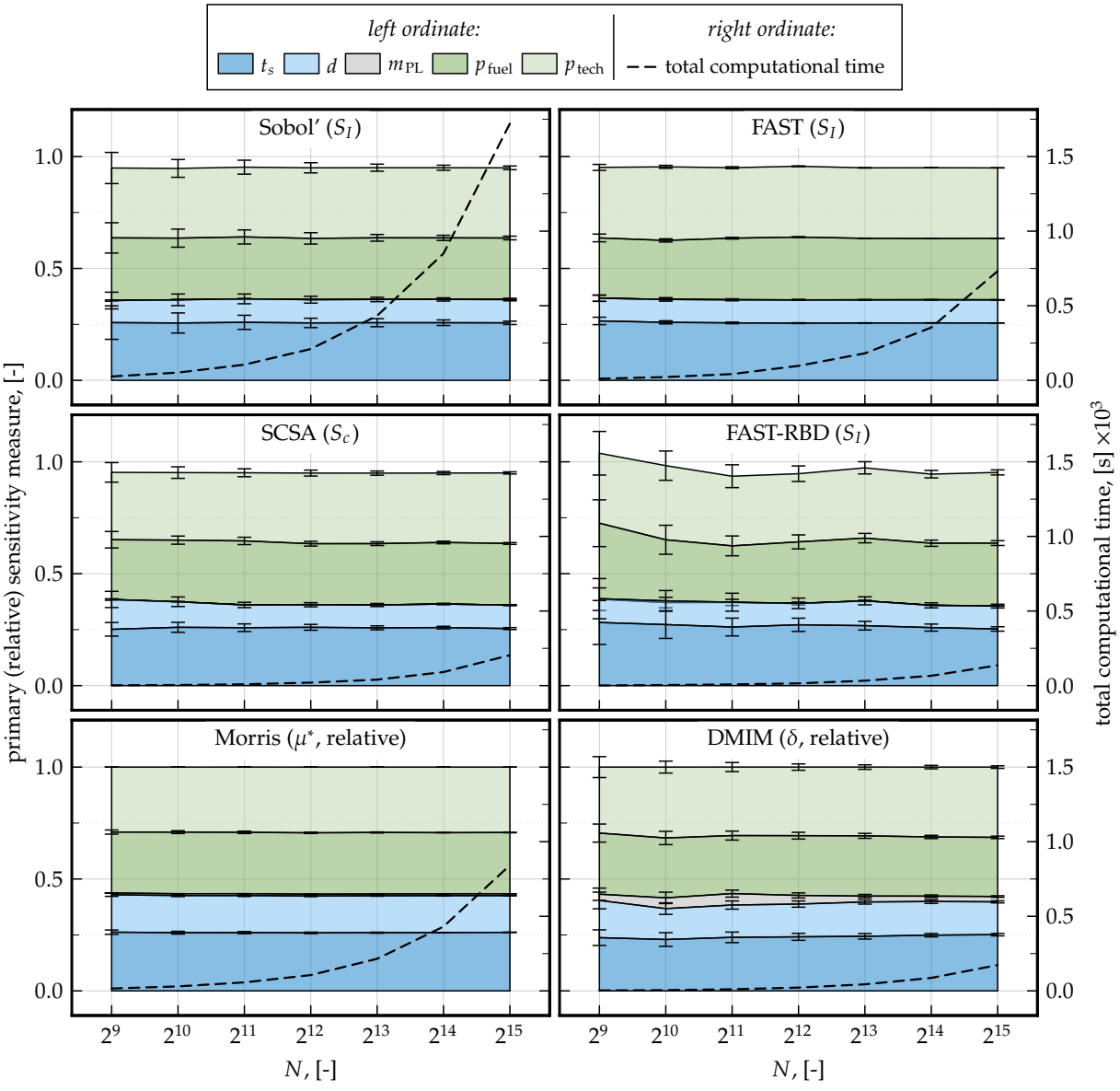


Figure 4.5 Analysis of convergence of all investigated GSA methods.

this method is moderately quick to execute, although it should be recalled that FAST-RBD calculates the first order sensitivity index, only.

The Morris method (bottom left subplot) seems to converge with the lowest sample size. There does not appear to be any discernible change in the distribution right from the start. This is also indicated by the small confidence intervals, which are hardly visible for sample sizes above $N > 2^9$. With an execution time of less than 20 s, this method is the fastest method of all.

DMIM (bottom right subplot) appears to be slightly indecisive with the distribution at first, mostly for the parameters d and m_{PL}. Values seem to settle in at sample sizes of 2^{14} and above, for which the computational expense is still moderate, i.e., 1.5 min.

Table 4.4 summarizes the execution time and sample sizes for which each GSA method has been considered to be converged, showing that the Sobol' method is by far the slowest. The fastest three (with comparable times) are: Morris, FAST and SCSA. Taking the comprehensive set of provided sensitivity measures into account, the SCSA technique performed best in this analysis.

Table 4.4 Overview of converged sample sizes (per variable N and in total N_{tot}) for each GSA method.

	Unit	**Sobol'**	**FAST**	**FAST-RBD**	**SCSA**	**Morris**	**DMIM**
N	[-]	2^{14}	2^{10}	2^{15}	2^{12}	2^9	2^{14}
N_{tot}	[-]	196,608	5120	32,768	4096	3072	16,384
execution time	[s]	850	22	136	20	16	87

The convergence analysis of the GSA methods aimed to shed some light on their overall efficiency. The following aspects can be concluded:

1. There are large differences when it comes to convergence. The Sobol' method, for example, is more than 40 times slower than the SCSA, which provides more sensitivity measures.
2. The method specific translation from input sample size N to required sample size N_{tot} (shown in Table 4.1) can only be used as an indication for the computational expense. For FAST-RBD, for instance, N_{tot} is equal to N, which suggests it being one of the fastest, which is not the case.
3. When in doubt, the consideration of confidence intervals can be a good indicator for convergence. The threshold to define convergence, however, depends on the required accuracy and is ultimately a choice for the user to make.

4.2.4 Conclusions

The comparative study, comprising analyses, shed some light on the actively researched field of GSA and aimed to answer some fundamental, practice oriented questions. The following conclusions can be summarized:

- Different GSA methods build on different theories and definitions of sensitivity, rendering their comparison non-trivial. Recommendations heavily depend on what the users are interested in, as the different techniques quantify different aspects.
- Despite these conceptual differences, the ranking (i.e., the sequence of input parameter criticality) was fully consistent throughout all techniques for the assessment example analyzed here.
- The quantification of (primary) sensitivity measures and their relative share was mostly consistent throughout the methods, with minor differences between the variance-based ones and the non-variance-based techniques.

- The primary sensitivity measures of most variance-based methods are easy to interpret, i.e., quantifying the percentage of the output variance which is caused by the input uncertainty.
- Among the non-variance-based methods, the Morris' method requires a few postprocessing steps and a minimal knowledge of the technique's algorithm for the interpretation.
- Interaction-based sensitivity, which is only provided by the Sobol' method and SCSA, is consistent. However, there is a significant difference in computational expense between these two methods with a clear recommendation towards the latter.
- Convergence speeds vary greatly among the methods. FAST, SCSA, and the Morris method are the quickest to converge.

Finally, it should be said that these insights may be different in situations where the number of uncertain input parameters is significantly higher and/or the use case and model are fundamentally different. A general recommendation cannot be made, as the usefulness of each method depends on the requirements and boundary conditions of the analysis. For TEA, which are in the center of attention of this thesis, the SCSA method appears to be a good choice due to its comprehensive set of sensitivity measures and relatively quick convergence.

4.3 Global Sensitivity Analysis using the Original Model

The second part of the overall GSA investigation is subject of this Section. Building on the insights gained by the previous study, GSA is now applied to a more complex example, i.e., the HLFC use case. The goal of this study is to demonstrate a GSA application starting from the input UQ and finishing with the recommendations for the decision making regarding where to invest future efforts.

4.3.1 Input Uncertainty Identification

Revisiting the methodology presented in Chapter 1, the first steps include the model creation and uncertainty identification. Because LYFE has already been selected as the LCBA simulation framework, the model creation phase is complete. The identification of uncertainties, i.e., deciding which parameters to vary and calculate sensitivity measures for in the GSA, builds on the assumptions discussed in Chapter 3. Here, several aspects that may have an impact on the TEA of HLFC have been explored and grouped as general, design, operational, environmental, and economic issues. This results in the eleven identified uncertainties shown in Table 4.5.

Table 4.5 Identified input uncertainties in the HLFC case study.

Design Domain		**Environmental Domain**	
$\mathcal{D}_1$	drag reduction potential	$\mathcal{E}_1$	insect contamination
$\mathcal{D}_2$	mass increase	$\mathcal{E}_2$	cloud encounter
$\mathcal{D}_3$	SFC penalty		
Operational Domain		**Economic Domain**	
$\mathcal{O}_1$	cruise speed	$\mathcal{M}_1$	future fuel price
$\mathcal{O}_2$	extra contingency fuel	$\mathcal{M}_2$	HLFC maintenance cost
$\mathcal{O}_3$	carried payload	$\mathcal{M}_3$	aircraft price increase

4.3.2 Input Uncertainty Quantification

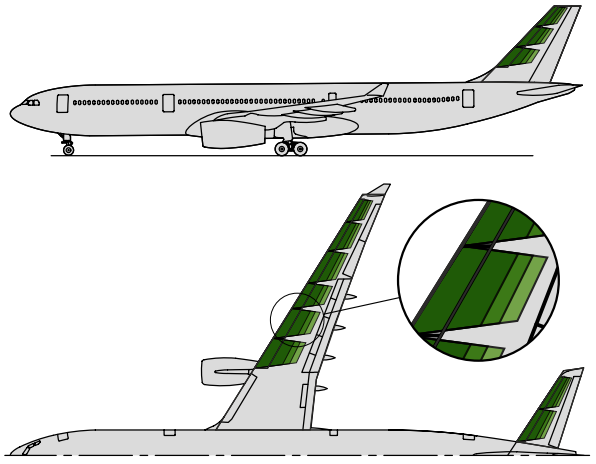
The next step of the methodology foresees the a quantification of the uncertain input parameters. In this analysis, which remains in the probabilistic realm, this entails that each uncertainty is assigned a PDF using either literature and past research, or general estimation methodologies. Since this can be done with a wide range of detail, it is helpful to recall the purpose of this study, which is to direct future research efforts more efficiently by distinguishing crucial from negligible uncertainties. Therefore, the input UQ performed here remains rather cursory, with each parameter receiving roughly equal attention.

For the assignment of PDFs when plentiful data is available, a Python-based distribution fitter was developed and used. This tool, which is aimed at practitioners, takes a dataset (e.g., a vector x where each entry represents one observation) and automatically fits all distributions that are available from the Python package `scipy`. For each, a goodness of fit test is performed, the residuals are calculated, and the fit itself is visualized alongside the original data. Users may then choose the most appropriate one based on their needs. If only few data are available, a uniform distribution ranging from the lowest to the highest value is assumed.

To avoid repetitions of the fundamentals discussed in Chapter 2 and the more parameter specific information introduced in Chapter 3, the following explanations put the uncertainty that is involved in the parameter estimation into focus.

Drag Reduction Potential

In Chapter 3, where an area-based estimation method was used to quantify the drag reduction potential, it was assumed that the transition position x_T/c is 0.5 throughout the span and for all laminarized areas. In this analysis, this assumption is treated as an uncertainty. As Fig. 4.6 shows, the relative transition position is now $x_T/c \in [0.4, 0.6]$, which, using the same estimation method, results in a total drag reduction between 6.0 and 8.6 %. With no additional detailed aerodynamic information available, the distribution of this uncertainty is assumed to be uniform. In alignment with all other uncertain parameters in this GSA, the drag reduction potential is sampled on a lifecycle simulation basis, i.e., is constant for one lifecycle simulation. Finally, it is considered to be of epistemic nature, as said information would arguably help to reduce this uncertainty.



(a) laminar areas at different transition positions.

	Unit	Wing	HTP	VTP	Sum
laminar area					
at $x_T/c = 0.4$	[m^2]	122.4	35.4	19.9	177.6
at $x_T/c = 0.5$	[m^2]	150.0	43.5	24.0	217.5
at $x_T/c = 0.6$	[m^2]	175.1	51.1	27.8	253.9
absolute drag savings potential					
at $x_T/c = 0.4$	[dc]	11.3	2.8	1.6	15.7
at $x_T/c = 0.5$	[dc]	13.9	3.5	1.9	19.2
at $x_T/c = 0.6$	[dc]	16.2	4.1	2.2	22.5
relative drag savings potential					
at $x_T/c = 0.4$	[%]	4.4	1.1	0.6	6.0
at $x_T/c = 0.5$	[%]	5.3	1.3	0.7	7.4
at $x_T/c = 0.6$	[%]	6.2	1.6	0.8	8.6

(b) aerodynamic impact of estimated laminar areas.

Figure 4.6 Visualization and summary of HLFC application with a transition position $x_T/c \in [0.4, 0.5, 0.6]$ resulting in a uniform distribution of the drag reduction potential $\mathcal{D}_1$.

Mass Increase and SFC Penalty

As discussed in Section 3.3.2, the total suction area can be used as to estimate the mass increase and required suction power. With the help of aircraft designs from other studies, a mass increase between 384 and 1104 kg and a suction power between 115 and 233 kW was estimated, see Fig. 3.10. While the previous study used the median value of these ranges, they are now treated as uniformly distributed and epistemic uncertainties. To translate the range for the suction power to a distribution of SFC penalties, which is shown in Fig. 4.7, the efficiencies of the electric components from Table 3.2 are used.

In order to keep the effort of UQ manageable, a manual grouping strategy is proposed with which the number of uncertain inputs is reduced. To do so, the SFC penalty distribution is obtained through a MCS, where the intervals for the power offtake as well as for the various electrical efficiencies are treated as lower and upper bounds of uniform distributions. Now, the "*model*" in Eqn. (4.20), which calculates the SFC penalty, is executed repeatedly with a new set of sampled values. The resulting values have been fitted with a general normal distribution for the later GSA. The data range from 0.73 to 1.73 % and show an average value of 1.2 %.

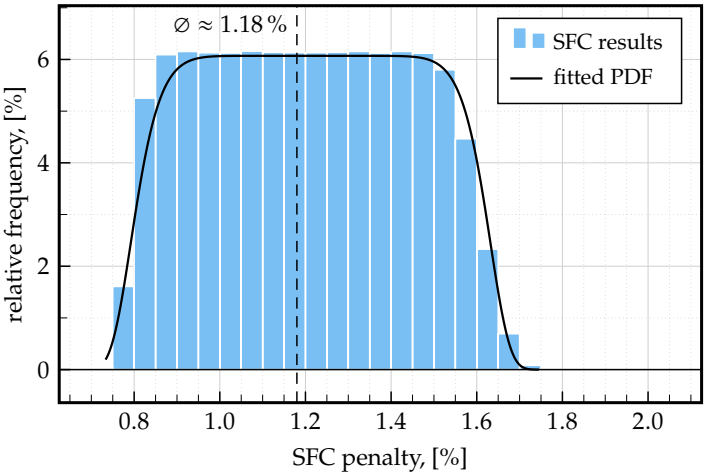


Figure 4.7 Histogram and fitted PDF for the SFC penalty uncertainty $\mathcal{D}_3$.

$$\mathcal{D}_3 : t_s = \frac{P_{\text{suc}}}{\eta_{\text{comp}} \cdot \eta_{\text{inv}} \cdot \eta_{\text{gen}} \cdot \eta_{\text{PL}}} \cdot \frac{0.5\,\%}{100\,\text{kW}} \tag{4.20}$$

Mach Number during Cruise

This parameter's integration takes two stages. First, the fluctuation of cruise speeds needs to be obtained. Second, the impact of this fluctuation on the laminarity has to be quantified. For the first stage, representative flight trajectories from FlightRadar24 [79] were analyzed. As these only provide the current time, altitude, latitude, and longitude (and hence the ground speed) of an aircraft, the cruise Mach number is not readily accessible. Therefore, these trajectories were coupled with the weather database from the European Centre for Medium-Range Weather Forecasts (ECMWF) [106] to calculate the present wind speed, temperature, pressure, and speed of sound that the aircraft experiences. The resulting Mach number variation is depicted in Fig. 4.8 (top), alongside a fitted PDF of the general normal distribution. From this distribution, it can be observed that the A330 flies with an average cruise Mach number close to 0.82. However, with a standard deviation of $\sigma = 0.013$ there is considerable variation from this value, which is expected to affect the laminarity noticeably. The non-hatched area between Ma= 0.79 and Ma= 0.85 represents the bounds of the sampled cruise speeds.

For the second step, i.e., the impact of different cruise speeds on the laminarity, a CFD simulation of a highly similar aircraft with HLFC on the wing upper side was used, where the span- and chordwise transition location was calculated with the DLR-TAU code [200]. For the laminarity on the tailplanes, it is assumed that the behavior is similar. These data were used to create a regression that translates the sampled cruise Mach number to a laminar efficacy of the form

$$\eta_{\text{L,mac}} = f(x) = -4833x^3 + 11{,}909x^2 - 9763x + 2664, \tag{4.21}$$

which is shown in the bottom left plot of Fig. 4.8. Note how deviations from the design point of Ma = 0.82 change the laminar effectivity depending on the direction of the deviation. Speeds below the design point decrease the efficacy by up to 38 % (i.e., $\eta_{\text{L,mac}} = 62\,\%$). Speeds above 0.82 positively affect the laminarity, increasing the effectivity up to 32 %. This effect is primarily caused by the effect of the transition position x_T/c, which has been evaluated throughout the span for each CFD simulation, as shown in the right plot of Fig. 4.8. As no information about the laminarity at a cruise Mach above 0.84 is available, $\eta_{\text{L,mac}}$ is capped to the highest observed value.

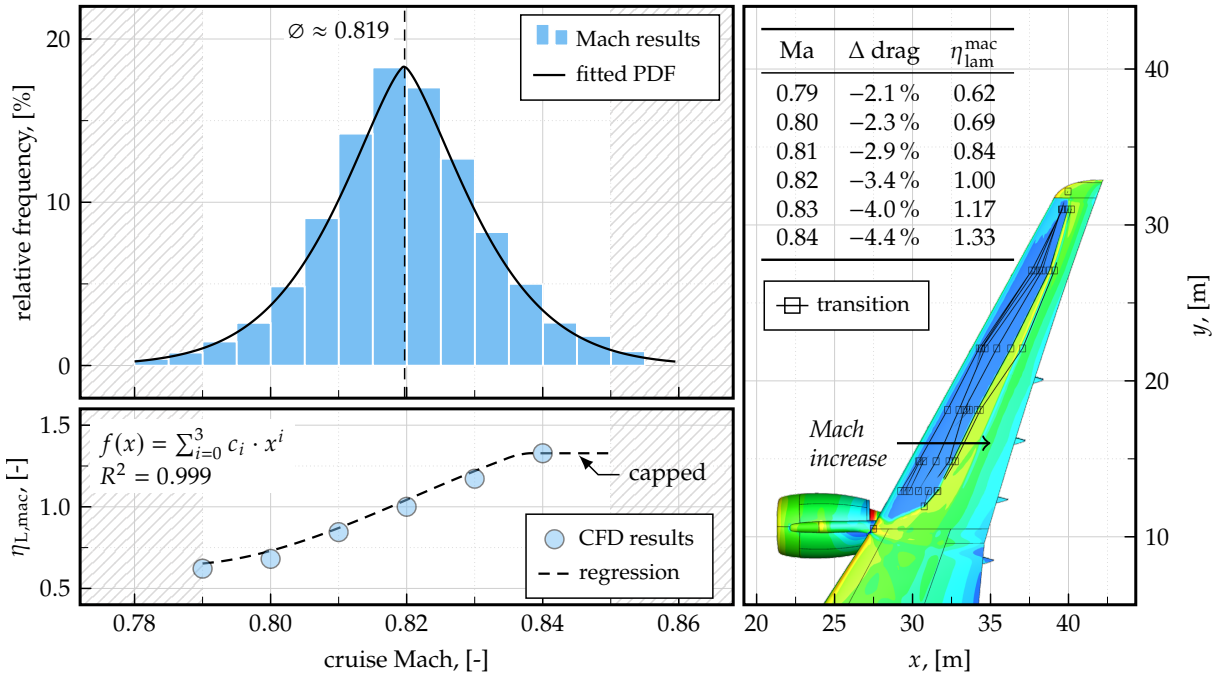


Figure 4.8 Histogram and fitted PDF of (calculated) cruise Mach number O_1 of an Airbus A330-343 (Registration D-AIKI) based on trajectories from FlightRadar24 [79] and weather data from ECMWF [106] (top left), CFD-Tau results of an A330 similar aircraft design with a laminarized upper wing, color coded by the skin friction coefficient [200] (right), and the regression model obtained from it (bottom left).

Although there are aspects where additional information could reduce the uncertainty of this parameter (such as directly measured Mach numbers obtained from Airlines), the influence of cruise speed on boundary conditions that cannot be changed is estimated to dominate this uncertainty. These include the weather conditions as well as tactical decisions of the airline. Therefore, this parameter is classified to be aleatory in this analysis.

Extra Contingency Fuel

As described in the operational considerations Section of of Chapter 3, operators of HLFC aircraft may wish to include some additional contingency fuel to account for unforeseeable losses of laminarity, e.g., due to system malfunctions. Depending on the operator's preferences, these additional fuel reserves can vary substantially from risk-affine to risk-averse. While the deterministic study in the previous Chapter used 2.5 % of the trip fuel as a HLFC-specific contingency fuel, this study uses a more detailed fuel planning approach. First, an interpolant $c \in [0, 1]$ is sampled on lifecycle simulation level. Before each mission, the expected fuel burn of the HLFC aircraft is estimated twice: one without any malfunctions m_1 and one with a malfunction directly after takeoff m_2, i.e., $m_2 > m_1$. The HLFC-related contingency fuel is then determined using

$$O_2 : m_{\text{cofu}} = c \cdot (m_1 - m_2)\,. \tag{4.22}$$

Thus, low values of the interpolant c represent a more risk-affine operator whereas high values indicate a risk-averse fuel planning strategy. The distribution of the c is assumed to be uniform and the uncertainty itself is classified as epistemic.

Carried (Passenger) Payload

The load factor is often averaged to 80 % but is known to vary with the airline, route, season, and other conditions. For the current study, the published data from the US Department of Transportation (DoT) is used [273], where the load factor is defined as the ratio of revenue passenger miles and available passenger miles. The distribution of data from 2002 to 2019[15] is shown in Fig. 4.9, showing a range of 66 to 90 %. The PDF fitted to the data is of the

[15] The year 2020 was excluded due to the impact that Covid-19 had on the air transportation system.

general logistics distribution type. With more data and more sophisticated models which take seasonal effects into account, the load factor prediction could theoretically be improved upon. Therefore, this parameter is treated as an epistemic uncertainty.

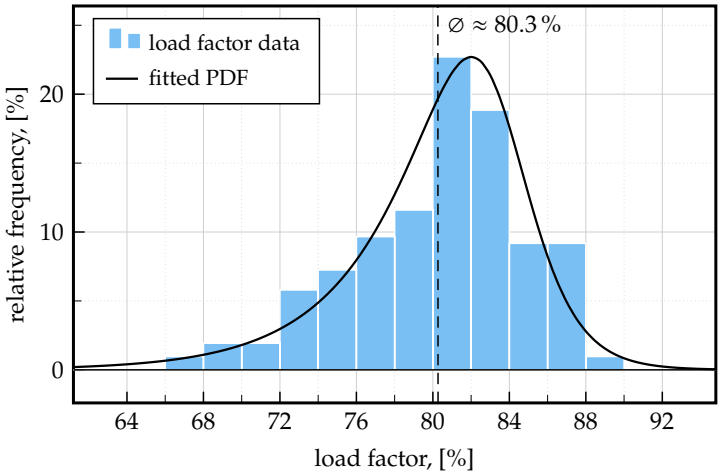


Figure 4.9 Histogram and fitted PDF of average load factors O_3 based on data from the US DoT [273].

Impact of Insect Contamination

Insect contamination can be modeled with a significant variation in level of detail. For instance, it is known that the insect population density is correlates with season, wind speed, temperature, precipitation, and altitude as well as local flora and fauna [50, 136]. And for a given insect population density, the impact on laminarity is not deterministic, either. It depends on the impact position, angle, mass, and type of the insect, which determine the pattern of lost laminar areas, resulting in a decreased laminar efficacy. For this study, the impact of insect contamination is simplified. The model builds on the devoted study of Elsenaar et al. [73], who counted the number of insects on several aircraft surfaces over the course of one year and used the data to calculate the loss of laminar efficacy. While the data shown in Fig. 4.10 shows a seasonal trend, there is a significant variation, which the authors trace back to potential natural cleaning events, i.e., rain.

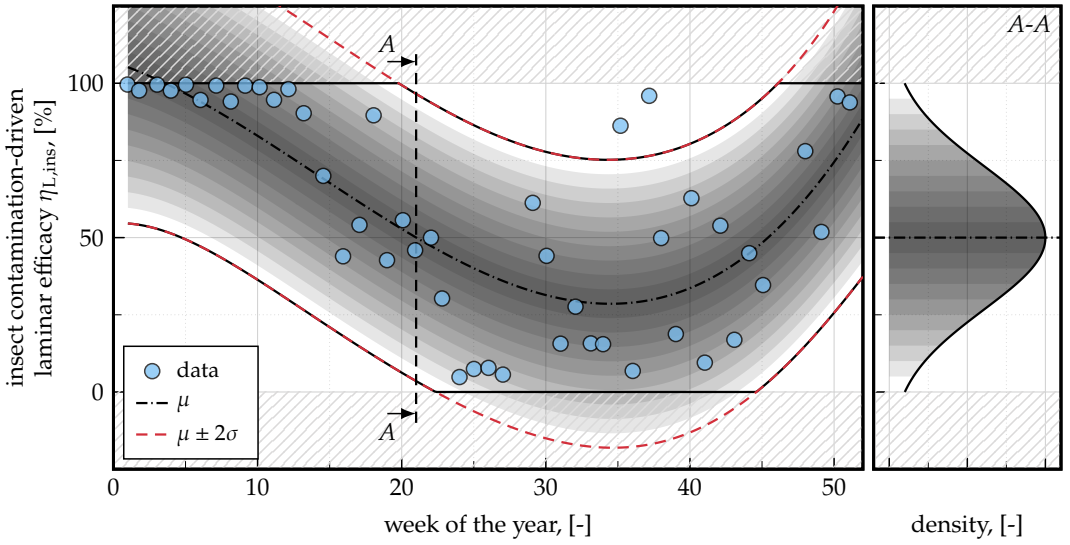


Figure 4.10 Weekly distribution of laminar efficacy from Elsenaar et al. [73] with equally spaced arrays between minimum and maximum fitted curves representing $\mathcal{E}_1$.

For modeling the insect contamination effect in this study, a Bayesian ridge regression of the third order was used, resulting in the predicted mean $\mu(t)$ as well as the standard deviation

$\sigma(t)$. The $\mu \pm 2\sigma$ curves serve as bounds the sampling process, which works as follows. Before each simulation, an interpolant $c \in [0, 1]$ is sampled and used to calculate a time series of $\eta_{\mathrm{L,ins}}(t)$ that is fixed for that simulation, i.e.,

$$\eta_{\mathrm{L,ins}}(c, t) = \eta_{\mathrm{L,ins}}(t)\big|_{\mathrm{worst}} + c \cdot \left(\eta_{\mathrm{L,ins}}(t)\big|_{\mathrm{best}} - \eta_{\mathrm{L,ins}}(t)\big|_{\mathrm{worst}} \right). \tag{4.23}$$

Here, the worst and best curves of the laminar efficacy are represented by $\mu - 2\sigma$ and $\mu + 2\sigma$, respectively. As some of the curves evidently exceed $\eta_{\mathrm{L,ins}} = 100\,\%$ as well as fall below of $\eta_{\mathrm{L,ins}} = 0\,\%$, which is not possible, the resulting arrays are capped for values above and below 100 % and 0 %, respectively. The distribution from which c is sampled from is assumed to be normal, as the A-A cut in the right plot of Fig. 4.10 shows, resulting in more samples being drawn in the vicinity of the predicted mean compared to the extremes.

It should be noted that this model is applied only on the laminarity of the tail planes, as the Kruger high lift device is assumed to successfully shield the wing upper side from insect contamination. To account for this, the relative contribution of component's HLFC application on the overall drag reduction potential is used. With two thirds of the total drag reduction caused by the application on the wing and one third by the HTP and VTP combined, the overall insect contamination-driven laminar efficacy is:

$$\eta_{\mathrm{L,ins,total}} = \underbrace{1 \cdot 66\,\%}_{\text{wing}} + \underbrace{\eta_{\mathrm{L,ins}} \cdot 22\,\%}_{\text{HTP}} + \underbrace{\eta_{\mathrm{L,ins}} \cdot 11\,\%}_{\text{VTP}}. \tag{4.24}$$

Due to the strong simplifications regarding seasonality and climate aspects in this model, this parameter is categorized as epistemic as well.

Impact of Cloud Encounter

For the second environmental uncertainty, cloud encounter at high altitudes, statistics of the time in cloud t_c obtained from flight test data [120] were used and fitted with an exponential distribution as shown in Fig. 4.11 (top). With a range from 0 to 60 % and an average value of 9.2 %, t_c is sampled before each simulation and fed to the linear translation function (shown in the bottom plot of Fig. 4.11). Note that this impact on laminar effectivity (i.e., the linear regression) is not treated as an uncertain parameter, but as a deterministic function, due to the negligible variation in the source data and lack of further data.

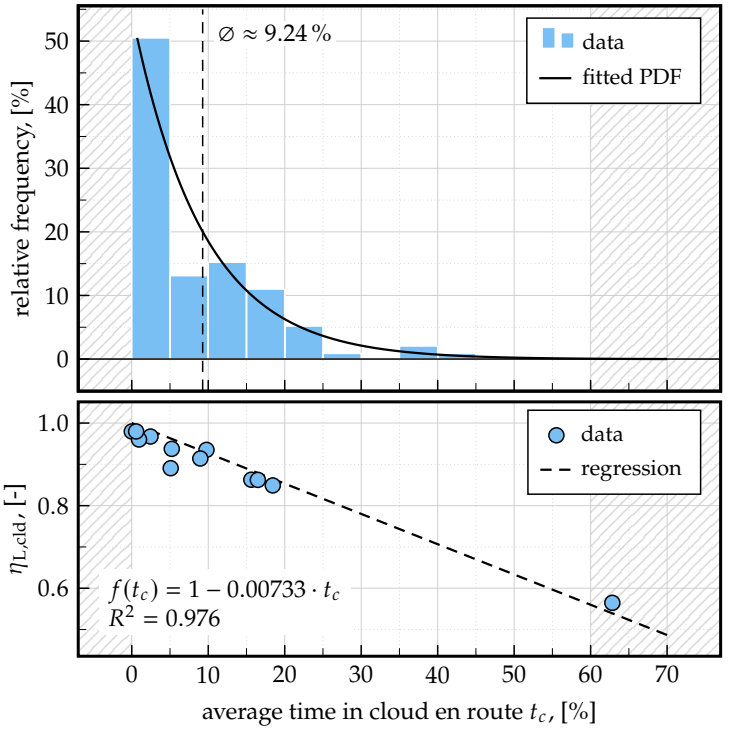


Figure 4.11 Histogram and fitted PDF for average time in cloud en route $\mathcal{E}_2$ (top, data from [120]) and its impact on laminar efficacy (bottom, data from [58]).

It should be noted that the combined impact of cloud encounter and insect contamination is modeled in an multiplicative nature. Consider, for example, an overall insect contamination-driven laminar efficacy of 85 % and a cloud encounter-driven laminar efficacy of 90 %, the resulting overall

laminar efficacy on aircraft level is

$$\eta_{\mathrm{L}} = 1 - (\eta_{\mathrm{L,ins,tot}} \cdot \eta_{\mathrm{L,cld}}) \tag{4.25}$$
$$= 1 - (0.85 \cdot 0.9) = 0.765. \tag{4.26}$$

Fuel Price

With HLFC being a technology aiming to reduce fuel burn and consequently fuel cost, the uncertainty regarding the fuel price development needs to be addressed appropriately. However, predictions of future price developments are inherently uncertain, especially considering the long periods of the prediction. The source of the fuel price developments was already introduced in the economic assumptions of Section 3.3.2, where the "*base*" scenario was chosen. For this analysis, a time-series sampling approach is chosen that is similar to that of the laminar efficacy in Eqn. (4.23). That is, an interpolant $c \in [0, 1]$ is sampled on lifecycle level and used to calculate the fuel price development array using the "*low*" and "*high*" as lower and upper bound, respectively:

$$p_{\mathrm{fuel}}(t) = p_{\mathrm{fuel}}(t)|_{\mathrm{low}} + c \cdot \left(p_{\mathrm{fuel}}(t)|_{\mathrm{high}} - p_{\mathrm{fuel}}(t)|_{\mathrm{low}} \right). \tag{4.27}$$

To shift the array sampling slightly towards the base scenario (which is considered to be the most likely development), the interpolant c is drawn from a skewed normal distribution with $\alpha = 2.6$, $\mu = 0.19$ and $\sigma = 0.39$. This distribution is shown in the *A*-*A* cut on the right plot of Fig. 4.12 and has been fitted so that the highest density matches the relative position of the base scenario within the interval.

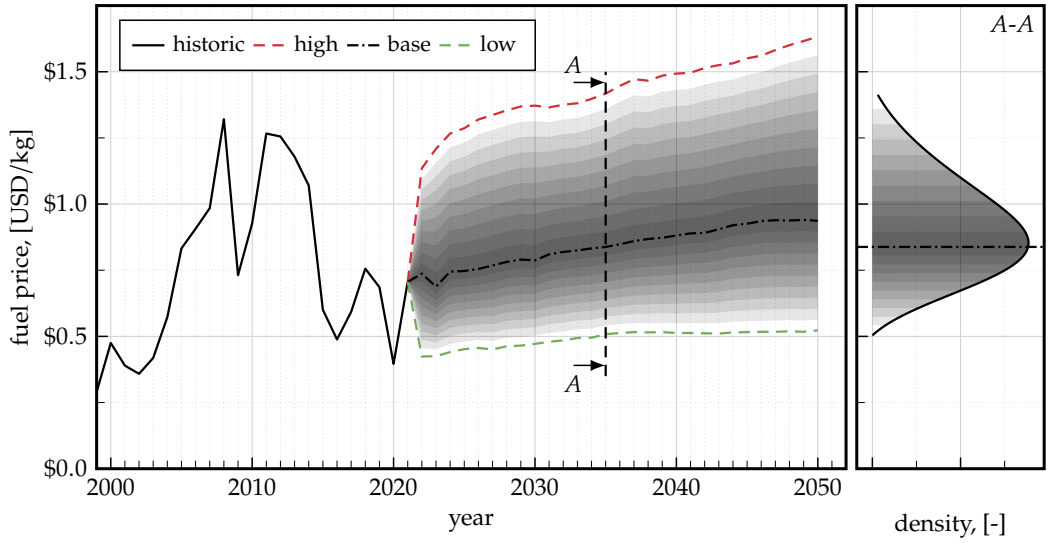


Figure 4.12 Historic and projected kerosene price under "*low*", "*base*" an "*high*" scenario (data taken from U.S. Department of Transportation [272]) with linearly spaced arrays between minimum and maximum prices representing $\mathcal{M}_1$.

HLFC Maintenance Cost

The expected increase in maintenance cost due to the HLFC system is treated as an epistemic uncertainty. It comprises a scheduled and an unscheduled portion. The scheduled HLFC maintenance cost uncertainty is derived in a similar manner to the method outlined in Table 3.3, utilizing various DOC methods with the previously established mass increase range of [384, 1104] kg. While this is a reasonable assumption, considering that heavier HLFC systems likely correlate with larger areas of application, resulting in more HLFC-specific maintenance

needs, this does not imply that the two parameters (mass increase and maintenance cost increase) are modeled in a correlative manner. In fact, for simplicity, all parameters in this study are sampled independently from one another, although it should be noted that the GSA method used is capable of quantifying the correlative sensitivity effect[16].

It should be noted that the maintenance increase uncertainty is considered to be epistemic, as more detailed approaches (e.g., a comprehensive analysis involving a parts list with reliability values, failure mode and effect analyses, etc) will likely improve the parameter's UQ. However, such information was not available during the course of the two underlying projects, which is why the previously described and simplified approach was used. The result of the input UQ, along with the fitted PDF of the Nagasaki type, is depicted in Fig. 4.13.

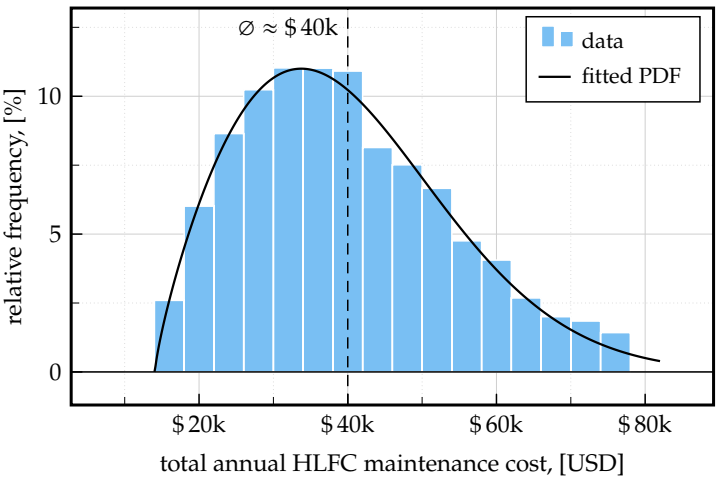


Figure 4.13 Histogram and fitted PDF of total HLFC maintenance cost $\mathcal{M}_2$ with data obtained from different CERs.

The data indicates an annual HLFC-related maintenance cost varying between $ 14k and $ 82k, with an average at $ 40k. It is important to note that this data combines both scheduled and unscheduled maintenance costs, the latter of which is integrated using the methodology proposed by Suwondo [258]. This method delineates the unscheduled to scheduled maintenance cost ratio as a time series bounded within $[1.24, 1.50]$ over a 20-year simulated operation period. Assuming a uniform distribution within this range, a Monte Carlo-based technique similar to that employed for the SFC penalty uncertainty was utilized to calculate the total HLFC-related maintenance cost.

HLFC Aircraft Price

The last uncertainty to quantify is the potential increase in aircraft price due to the development of the HLFC system. Analogous to the approach adopted for estimating the rise in maintenance costs, the aircraft price increase is projected by leveraging the mass increase uncertainty, coupled with the DOC methods outlined in Table 3.3. Given the complex nature of aircraft pricing, which encompasses potential deviations from the list price, the MCS-based methodology previously applied to the (un)scheduled maintenance cost uncertainty is employed here as well. Consequently, the aircraft price distribution illustrated in Fig. 4.14 not only reflects the CER-based outcomes but also accounts for the existing knowledge

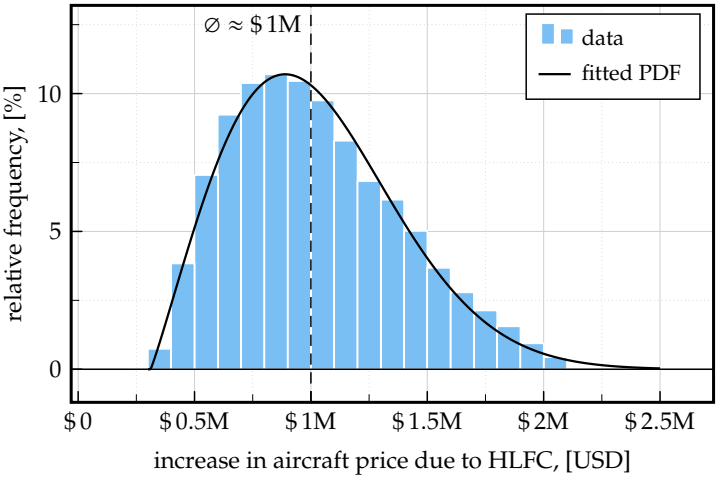


Figure 4.14 Histogram and fitted PDF of the aircraft price increase $\mathcal{M}_3$ due to HLFC with data obtained from different CERs.

[16] A detailed inclusion of input correlation in GSA requires the discernment of S_a, S_b, and S_c for each sensitivity measure (i.e., first order, second order, and total SI). This increases the complexity when discussing the results, especially for the higher-order effects, as the analysis in Appendix D.2 showed.

gap concerning the potential discount range, presumed to uniformly span the [0, 50]% interval. The data conforms to a Nagasaki type distribution, ranging between $ 300k and $ 2.1M, with a mean value of $ 1M.

Alongside a list of assumptions for this study, all uncertain parameters fed to the GSA are summarized in Appendix D.4 including their type (i.e., nature), fitted distribution, lower and upper bound, as well as the parameters for recreating the distribution with the Python package `scipy`. With eight epistemic and three aleatory elements, the input uncertainties are dominated by lack of knowledge rather than natural variability. Note that this is, however, not only a categorization that is somewhat subjective, but also one that can change when models are improved upon. The load factor, for instance, was modeled in a very simple manner, leaving significant room for improvement. If these improvements were to be implemented, the parameter's variation might decrease but is unlikely to be eliminated. Instead, the remaining variability would be exclusively aleatory. Another note worth to make deals with truncation. For those distributions that are not uniform, a modification of SALib was required, where the inverse probability integral transformation was used. It ensures proper probability density scaling as well as a more efficient sampling approach[17].

4.3.3 Analysis and Results

With all uncertain inputs described and assigned a distribution, the results of the GSA can now be discussed. To recall, the comparative study performed in Chapter 4.2 revealed the SCSA method to be both comprehensive (in the sense of provided sensitivity measures), quick to converge, and relatively easy to interpret. Therefore, this method is selected for the present GSA. The number of samples per variable, N, has been set to $2^{15} = 32{,}768$, leading to 65,536 lifecycle simulations with a total number of $n > 840$M individual flights[18]. The upcoming discussion of results are organized as follows. First, the overall output distribution of the two investigated metrics, namely the ΔNPV and the fuel consumption $\Delta\gamma_{\text{fuel}}$, are analyzed. Afterwards, the output uncertainty is broken down to its constituents, discerning the primary effects from the secondary and correlation induced effects.

Output Distribution

Like most GSA methods, the SCSA method requires repeated executions of the model to calculate the sensitivity measures. Given a sufficiently large N, these repeated executions are equivalent to a conventional MCS, which allow to calculate the statistical moments of the output distribution. This helps to understand how large the uncertainty in the overall study is and should be part of any GSA.

Fig. 4.15 depicts the overall distribution in four plots. The left two deal with the technological metric γ_{fuel}, whereas the right two show the NPV representing the economic KPI. The upper plots show the aircraft-specific results using boxplots, while the lower plots visualize the Δ values using histograms and the empirical CDF. Before diving into the discussion of the results, it should be noted that the fuel consumption metric in this GSA is calculated on a lifecycle simulation basis. That is, one lifecycle simulation (which comprises more than ten thousand flight events) yields one lifetime averaged value for γ_{fuel}. While the flight individual γ_{fuel} values are technically available, the GSA process requires one model execution of LYFE to result in one scalar value for the variance decomposition.

As the boxplots of γ_{fuel} in Fig. 4.15 show, the laminar (LAM) aircraft is more fuel efficient than the reference (REF) aircraft. The latter shows γ_{fuel} values ranging from 2.75 to

[17] As opposed to, for example, repeating the sampling process if a sample falls out of the bounds specified by the truncation. Supplemental information is provided in Appendix D.3.

[18] This campaign was run on a workstation with 60 logical cores in parallel and took 27 hours to complete.

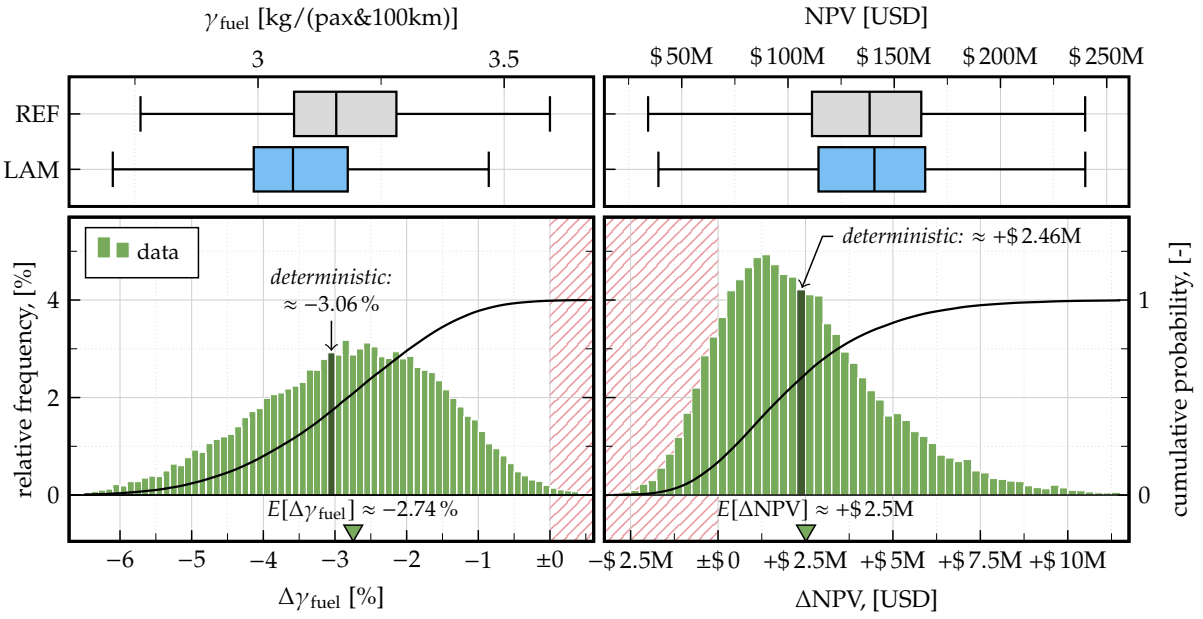


Figure 4.15 Monte-Carlo results of the GSA with fuel efficiencies on the left, NPVs on the right, aircraft design-specific results at the top, and histograms and empirical CDFs of Δ values at the bottom.

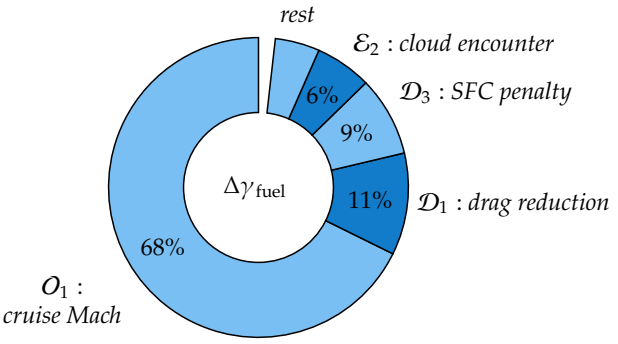
3.6 kg/pax · 100km with an IQR of [3.06, 3.29]. The median and average values for this turbulent aircraft are 3.16 and 3.19 kg/pax · 100km, respectively. The laminar aircraft's fuel consumption values range from 2.7 to 3.46 kg/pax · 100km and shows an IQR of [3.0, 3.18]. The median and average values are 3.07 and 3.10 kg/pax · 100km, respectively. This represents a fuel consumption improvement of 2.74 %, which is in line with the expected value of the histogram of $\Delta\gamma_{\text{fuel}}$ in the lower left plot. The deterministic value of 3.06 % shown in Fig. 4.15 refers to the analysis of Chapter 3. It being close to the expected value of 2.74 % indicates that (a) the mean values of the uncertain inputs in the GSA are close to the assumption values of the deterministic analysis, and (b) the sample size of the MCS part of the GSA is sufficiently large. This histogram shows the spread of fuel consumption changes in more detail, revealing $\Delta\gamma_{\text{fuel}}$ values as low as −6 %. With 95 % of data being within the interval of [−5.6, −0.9] and the standard deviation being 1.2 %, the histogram highlights that uncertainties are not negligible. One interesting aspect is that in a small number of simulations, the laminar aircraft turns out to be less fuel efficient than the turbulent reference (i.e., $\Delta\gamma_{\text{fuel}} > 0\,\%$). In these cases, the set of parameter samples is highly unfavorable for the technology, e.g., low cruise Mach, low drag reduction potential, high SFC penalty, and so on. It should be noted that such designs are not likely to be actually pursued in the product development process.

The economic metric (right plots) substantiates the overall superiority of the HLFC aircraft, albeit with a smaller lead. Considering the aircraft specific boxplots, the advantage of LAM over REF is existent, but difficult to discern. The ΔNPV depiction at the bottom right allows a more detailed view, showing a spread of values between −$ 3M to +$ 12M (and 95 % of data being in the interval $[700k, 800M]). The expected economic superiority is quantified with +$ 2.5M, which is fairly close to the deterministic result of the analysis in Chapter 3. The share of results where the laminar aircraft performs worse than the turbulent counterpart is noticeably higher when considering the ΔNPV metric over $\Delta\gamma_{\text{fuel}}$. The reason for this is that the calculation of the NPV encompasses more variables including the additional maintenance and aircraft price increase, which are relevant for a comprehensive assessment, but do introduce less favorable conditions. Overall speaking, the output distribution of either metric indicate a superiority of HLFC, but the involved uncertainty (measured as the spread or the standard deviation) is too high to accept a deterministic result, only. And with the uncertainty breakdown of a GSA, which is discussed next, it becomes possible to pinpoint what affects this output uncertainty to what degree.

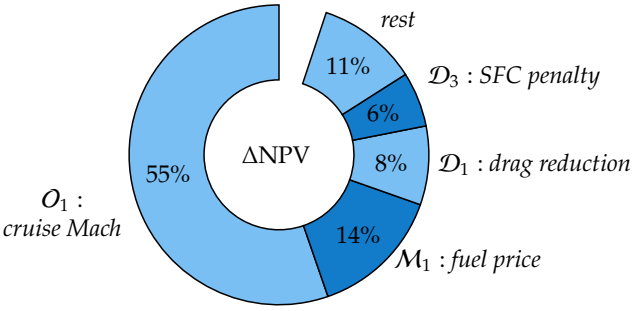
Uncertainty Breakdown

With the output uncertainty discussed, the next set of results focus on its breakdown. The overall goal is to understand which uncertain input parameter has the highest influence on the output (and with that deserves the most attention in subsequent uncertainty reduction efforts) and which ones can easily be fixed to a mean value without oversimplifying the model.

The primary effect, which is represented by the S_c sensitivity measure of the SCSA method, is shown in Fig. 4.16 (a) and (b) for $\Delta\gamma_{\text{fuel}}$ and ΔNPV, respectively. In both, the cruise Mach speed (O_1) is by far the most dominant parameter. As this is an aleatory uncertainty, there are limited options with respect to input uncertainty reduction and/or submodel improvement. It is, however, possible and even advisable to communicate these results back to the design, emphasizing the need for a more robust and less susceptible design. For $\Delta\gamma_{\text{fuel}}$, the other mentionable sensitivities are towards drag reduction potential uncertainty $\mathcal{D}_1$ and the time in cloud uncertainty $\mathcal{E}_2$ with $S_c = 10\,\%$ and 5 %, respectively. All other parameter sensitivities were less than 5 % and have thus been subsumed to the entry *rest*. The small portion which is left unfilled represents the amount of output uncertainty that cannot be explained by the primary effects, i.e., are caused by interaction or even higher order effects.



(a) Breakdown for $\Delta\gamma_{\text{fuel}}$.



(b) Breakdown for ΔNPV.

Figure 4.16 Uncertainty breakdown using S_c.

The breakdown of the ΔNPV is similar, albeit slightly more balanced. The cruise Mach speed sensitivity, which is at 55 %, is still the most sensitive parameter, followed by the future fuel price development with $S_c = 18\,\%$, which is aleatory in nature as well. This leads to the overall conclusion that the output uncertainties of both metrics are dominated by aleatory uncertainties, although only three out of the eleven investigated uncertainties were aleatory to begin with.

As the SCSA method provides more than just one sensitivity measure, a more detailed analysis allows for additional insights. Fig. 4.17 shows the direct effect, correlation, and interaction via S_a, $S_c = S_a + S_b$ and S_T, respectively, for each uncertain parameter. Note that correlation and interaction effects are not depicted explicitly, but can be measured as follows. If, for a given parameter X_i, the combined sensitivity measure S_c is noticeably higher or lower than the non-correlative sensitivity measure S_a, a noticeable positive or negative correlation has been identified by the method. The same concept applies for the interaction, where an increase in S_T (X_i) over S_c (X_i) indicates a noticeable interaction sensitivity of this parameter.

Starting with the correlative effect, there is no significant change of S_c over S_a observable for any of the uncertainties, as expected (since all parameters were provided to the model independently). The interaction effects do not seem to affect most parameters, either. The only two parameters for which the S_T is visibly larger than S_c are the cruise Mach speed O_1 and the fuel price $\mathcal{M}_1$ (for the ΔNPV metric), indicating that these two have a non-additive combined effect.

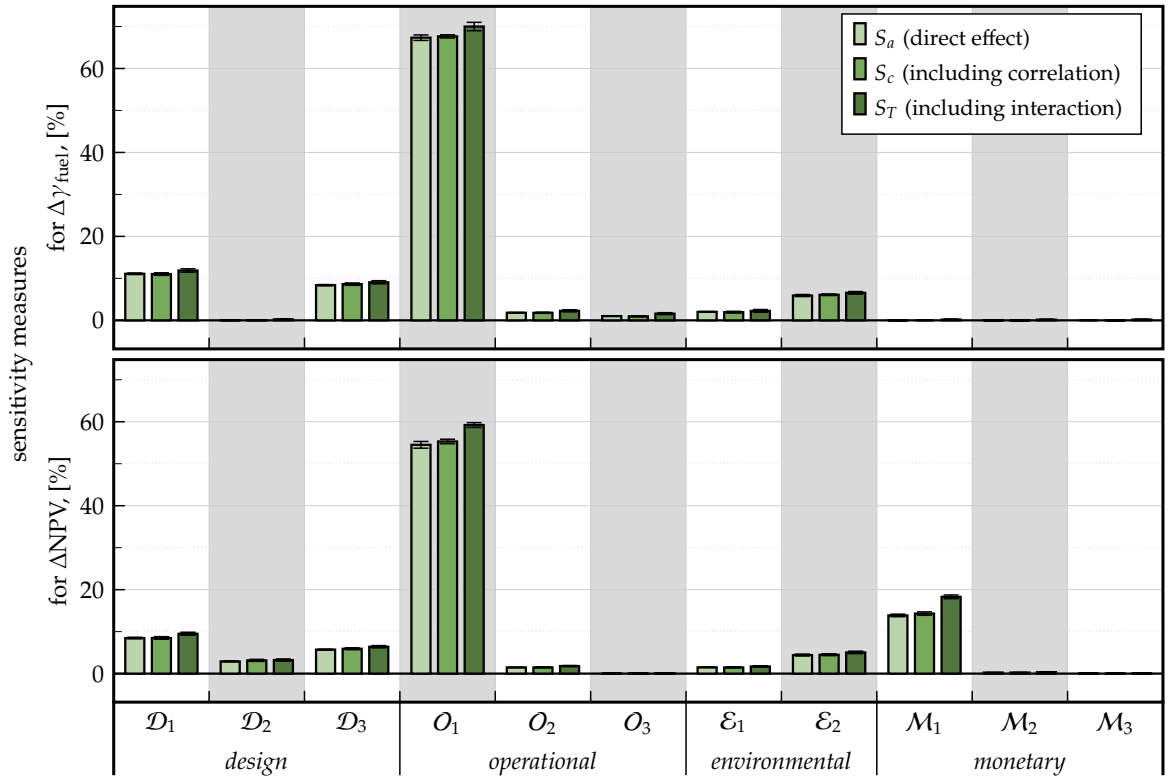


Figure 4.17 Breakdown of $\Delta\gamma_{\text{fuel}}$ (top) and ΔNPV (bottom) uncertainty using S_a, S_c, and S_T.

For completeness reasons, the interaction map illustration introduced in the comparative GSA is discussed next. Fig. 4.18 shows these for the two considered metrics. Here, the interaction effects of S_c were used. The highest value for S_c^{II} for $\Delta\gamma_{\text{fuel}}$ (left plot) is less than 0.5 %, rendering any interpretation efforts meaningless. For ΔNPV (right), the previous indication of $\mathcal{O}_1$ (cruise speed) and $\mathcal{M}_1$ (fuel price) is substantiated by this interaction map, although the interpretability of a maximum interaction effect of ≈3 % is questionable at best.

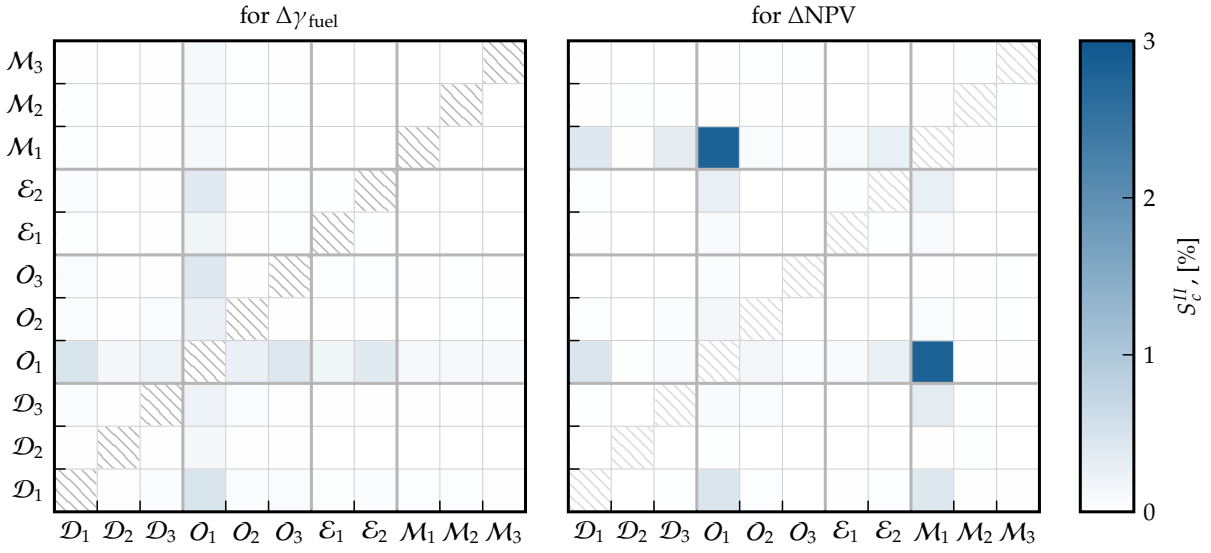


Figure 4.18 Second order sensitivities for $\Delta\gamma_{\text{fuel}}$ (left) and ΔNPV (right).

Improved Off-Design Behavior

The previous GSA has shown that the expected fuel consumption improvements as well as the economic superiority are dominated by the uncertainty in cruise Mach speeds. Due to its

aleatory nature, there is little room for uncertainty reduction. It is, however, possible to improve the HLFC design by making it less susceptible towards off design speeds. In this last analysis, a hypothetical improvement of this kind is introduced and, for conciseness reasons, discussed using only the economic metric. The goal is to demonstrate how the product development process could react to such insights provided by the GSA. The improvement is shown in the top left plot of Fig. 4.19, showing the (hypothetically) updated datapoints next to the original (and previously used) ones[19]. The updated regression uses the third degree polynomial

$$\eta_{\mathrm{L,mac}} = f(x) = -2417x^3 + 5954x^2 - 4882x + 1333 \tag{4.28}$$

which is, consistently with the previous implementation[20], capped at the $\eta_{\mathrm{L,mac}}$ value at the cruise speed of 0.84. The top right plot of Fig. 4.19 shows the Monte-Carlo results of the updated HLFC design (in gray) next to the original results (light green), alongside the empirical CDFs of both simulations. The updated results are visibly less spread, which is reflected by the reduction of the standard deviation from originally $\sigma_{\mathrm{NPV}} \approx \$\,2.3\mathrm{M}$ to $\sigma_{\mathrm{NPV}} \approx \$\,1.9\mathrm{M}$. The expected value of $E[\Delta\,\mathrm{NPV}] \approx \$\,2.5\mathrm{M}$ did not change significantly, which, due to the symmetrical improvement of $\eta_{\mathrm{L,mac}}$, is not surprising. The reduced spread, combined with the positive expected value, did slightly improve the probability of success from $P(\Delta\,\mathrm{NPV} > 0) = 83\,\%$ to $89\,\%$. The final comparison of the uncertainty breakdown of both simulations is depicted in the lower plot of Fig. 4.19 with the original S_c in green and the updated sensitivity

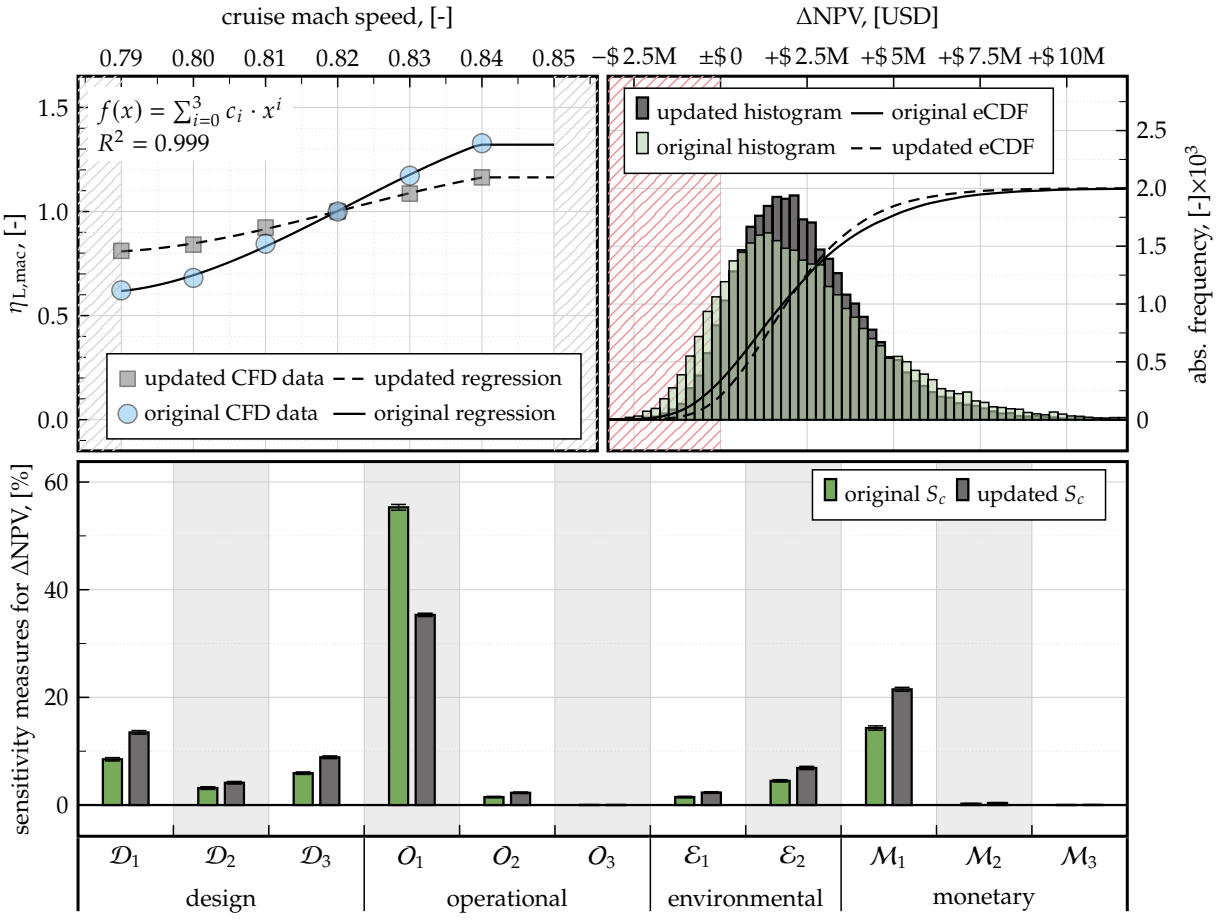


Figure 4.19 Effect of a hypothetical improvement of maintaining laminarity at off design speeds including the polynomial fit (top left), Monte-Carlo-based results (top right) and uncertainty breakdown using S_c (bottom).

[19] The updated data were created by taking half of the distance to $\eta_{\mathrm{L,mac}}$, resulting in a symmetrical behavior, i.e., an improved performance at lower Mach speeds, but also leads to a lower laminar efficacy at cruise Mach speeds above 0.82.

[20] Recall Fig. 4.8 on page 87.

measure in gray. The sensitivity towards the cruise Mach speed uncertainty reduced from previously 55 % to 35 %. This loss of 20 percentage points is now proportionally distributed over the remaining uncertainties. The sum of S_c over each parameter is > 95 %, indicating only insignificant interaction effects.

4.4 Conclusions

This Chapter revolved around the first RQ and hypothesis which deals with differentiation of relevant and irrelevant uncertainties using GSA. To do so, the theory and application of GSA were investigated in a two part approach. In the first part, a comparative study of multiple GSA methods was performed on a mechanistic surrogate of the LCBA framework called FastLYFE. Goals of this study included gaining practical knowledge regarding convergence, easiness of interpretation, and capability fulfillment. Considering the characteristics of the investigated techniques, the SCSA method from Li et al. [149] revealed itself to be most suitable for the underlying use case. This method was then applied in the second part of this Chapter, where the original LYFE model was executed on the HLFC use case. Here, literature was reviewed and simplified methods and models were developed and/or used to quantify the uncertainties of 11 input variables from 4 different domains. This input UQ is expressed as a PDFs for each parameter, from which the subsequent GSA repeatedly draws samples from. The output distribution, which is available as a byproduct when performing SCSA, was analyzed first. Whilst indicating superiority of the HLFC aircraft over the conventional counterpart when using averaged values, the distribution revealed a substantial spread in the economic and technological KPIs, which were the ΔNPV and $\Delta\gamma_{\text{fuel}}$, respectively. Thus, the necessity of a UQ enabled assessment approach was highlighted. Results of the GSA have shown the cruise speed to be the dominant input uncertainty as it is responsible for more than half of the output uncertainty. It is followed by the fuel price ($\approx$ 19 %) and the drag reduction potential ($\approx$7 to 10 %, depending on the KPI). Negligible uncertainties include the HLFC related maintenance cost, the increase in aircraft price, as well as the load factor. Correlation and interaction effects were also quantified by the SCSA method but are negligible for this use case. As a final analysis in this study, a hypothetical improvement of the HLFC aircraft's off-design performance was introduced, which served as a verification step. Results showed the expected decrease of the sensitivity towards O_1, which is the cruise Mach number, while the sensitivity measures of the other parameters increased proportionally due to the relative breakdown nature of the GSA.

5 Non-Probabilistic Uncertainty Quantification

This Chapter addresses the second research question, focusing on DSTE as a promising non-probabilistic method for quantifying epistemic uncertainty. The goal is to explore how DSTE's concepts, techniques, and metrics can complement or even replace traditional uncertainty quantification methods, such as those used in ProT. The research hypothesis guiding this investigation posits that a systematic demonstration of DSTE's capabilities will reveal its potential advantages. The Chapter begins with an introduction in Section 5.1, explaining how DSTE's metrics can be integrated into MCS-based uncertainty analyses using a newly developed Python package. Section 5.2 then outlines the methodology for conducting systematic investigations into DSTE-based uncertainty quantification, detailing the overall study design and the expert elicitation process. In Section 5.3, describes the selection, quantification, and preparation of HLFC-related epistemic input uncertainties for the simulations specified in the methodology. The results of these simulations are presented and compared to their probabilistic counterparts in Section 5.4. Note that all analyses in this Chapter have been performed with a PCE-based surrogate model of LYFE[1], as the computational expense of the DSTE-enabled MCS is too high to be performed with either the DES version or the mechanistic surrogate of LYFE.

5.1 Integration of DSTE in Monte-Carlo Simulations

As mentioned in Chapter 2, DSTE revolves around the concepts of Belief and Plausibility, which aim to complement the conventional concept of probability by providing a lower and upper bound. To compute these in MCSs, a new sampling process is needed, which is based on prior expert elicitation. In addition, an additional posterior calculation is required to obtain (complementary) cumulative functions of Belief and Plausibility over the output y in order to construct statements about the uncertainty of the results[2]. These elements are described in this Section, beginning with some theoretical aspects in Section 5.1.1, followed by practical guidelines discussed in Section 5.1.2, including a description of the developed Python package for DSTE-based UQ.

5.1.1 Theoretical Aspects

This Subsection extends the fundamentals around DSTE described in Chapter 2 and is based on the work from Helton et al. [101–104] and Oberkampf et al. [179, 180], complemented by notional examples and illustrations. The modified sampling strategy is explained first, followed by a description of the (complementary) cumulative function calculation.

Belief-Based Sampling Strategy

The Belief-based sampling strategy is a useful alternative when little to no data is available and PDFs other than a uniform distribution cannot be obtained. It is based on the Basic Belief Assignments (BBAs) provided by Subject Matter Experts (SMEs), for which an expert elicitation process is recommended. For demonstration purposes, consider a simple model

$$y = f(a, b) = 100(a^2 + b^2)^{a-b} \quad , \tag{5.1}$$

[1] Supplemental information on the PCE-surrogate development can be found in Appendix E.1.

[2] See the discussions on pp. 36.

where both inputs a and b are uncertain. Some information (either obtained through literature or expert elicitation) about a and b is available. Assume that one source provided three intervals for a being $a_1 = [0.5, 0.7]$, $a_2 = [0.6, 0.9]$, and $a_3 = [0.4, 1.0]$ and three for b being $b_1 = [2.0, 2.5]$, $b_2 = [2.2, 2.3]$ and $b_3 = [2.2, 2.4]$. This yields the following input sample spaces:

$$\mathcal{A} = \bigcup_{i=1}^{n_a} a_i = [0.4, 1.0] \quad \text{and} \quad \mathcal{B} = \bigcup_{i=1}^{n_b} b_i = [2.0, 2.5] \ . \tag{5.2}$$

Now, each of the evidences provided in $\mathcal{A}$ and $\mathcal{B}$ is assigned a BBA. It reflects the Belief of the expert in this interval. Say, for example, that the expert is more confident in the interval of a_1, whereas less Belief exists that the true value of a lies within the intervals of a_2 or a_3. A possible BBA assignment for a could then be $m(a_1) = 0.5$ and $m(a_2) = m(a_3) = 0.25$. For b it is assumed that no interval is weighted more than the others, yielding $m(b_1) = m(b_2) = m(b_3) = 1/3$. With these defined, the next step is to calculate the general density functions $D_a(a)$ and $D_b(b)$ defined on $\mathcal{A}$ and $\mathcal{B}$, respectively. These are

$$D_a(a) = \sum_{i=1}^{n_a} \frac{d_{a_i}(a)}{n_a} \quad \text{and} \quad D_b(b) = \sum_{i=1}^{n_b} \frac{d_{b_i}(b)}{n_b} \tag{5.3}$$

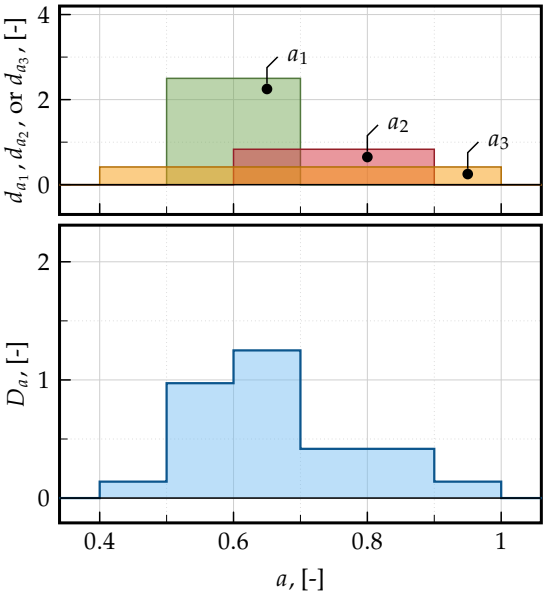
Hereby, d_{a_i} and d_{b_i} represent the evidence-specific density functions defined on $a \in \mathcal{A}$ and $b \in \mathcal{B}$. These can be calculated, for instance, using the width of the interval, i.e., to assign higher densities when small intervals are given and vice versa:

$$d_{a_i}(a) = \begin{cases} \dfrac{m(a_i)}{\max(a_i) - \min(a_i)} & \text{if } a \in a_i \\ 0 & \text{otherwise} \end{cases} \tag{5.4}$$

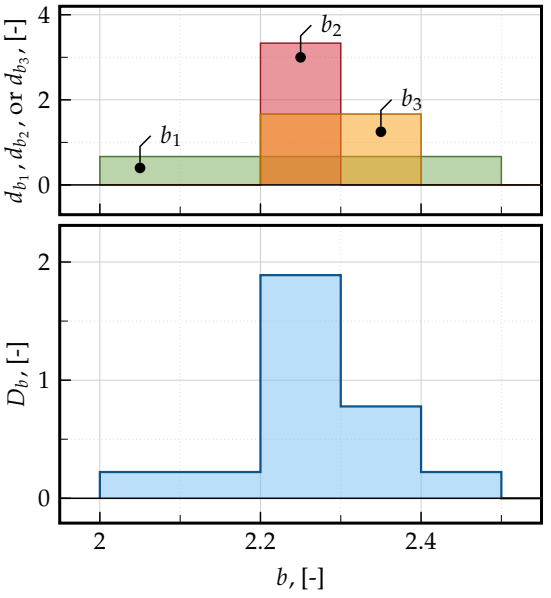
and accordingly for b. For a_1, this results in

$$\begin{aligned} d_{a_1}(a_1) &= \frac{m(a_1)}{\max(a_1) - \min(a_1)} \\ &= \frac{0.5}{0.7 - 0.5} = 2.5. \end{aligned} \tag{5.5}$$

Consequently, the other evidence-specific densities are $d_{a_2}(a_2) = 0.8\bar{3}$, $d_{a_3}(a_3) = 0.41\bar{6}$, $d_{b_1}(b_1) = 0.\bar{6}$, $d_{b_2}(b_2) = 3.\bar{3}$, and $d_{b_3}(b_3) = 1.\bar{6}$. With Eqn. (5.3), the density functions shown in Fig. 5.1 can be calculated. For each variable, the upper plot shows the individual densities and the bottom subplot depicts respective general density functions. Note how the interval $[0.6, 0.7]$ in a is the highest as all interval statements a_i contain it. The same applies to the interval $[2.2, 2.3]$ in b.



(a) uncertainty a.



(b) uncertainty b.

Figure 5.1 Density functions for a and b created for the Belief-based sampling process.

These density spaces can now be used to sample from, e.g., by using the inverse probability

integral transformation method described in Appendix D.3, and fed to the model through a MCS. The combined samples of a and b for this sampling strategy are shown in the top left subplot of Fig. 5.2, whereas the output of a conventional sampling strategy (with uniform distributions) is shown in the top right subplot. The bottom two subplots depict the results of the MCS for each strategy. The results show that the standard deviation (highlighted in light gray around the mean) for the Belief-based approach is 40 % lower due to the incorporation of weights instead of using uniform distributions. This highlights how the conventional approach of assuming uniform distributions can misrepresent the 'true' uncertainty that human knowledge provides, as elicited by the hypothetical SMEs and represented by DSTE.

These density spaces can now be used for sampling, e.g., by using the inverse probability integral transformation method described in Appendix D.3, and then fed into the model through a MCS. The combined samples of a and b for this strategy are shown in the top left subplot of Fig. 5.2, whereas the output of a conventional sampling strategy (with uniform distributions) is shown in the top right subplot. The bottom two subplots depict the results of the MCS for each strategy. The results show that the standard deviation (highlighted in light gray around the mean) for the Belief-based approach is 40 % lower due to the incorporation of weights instead of using uniform distributions. This highlights how the conventional approach of assuming uniform distributions can misrepresent the "*true*" uncertainty that human knowledge provides, as elicited by the hypothetical SMEs and represented by DSTE.

The strengths of the Belief-based sampling process lie in its ability to deal with multiple evidences, especially when given in intervals. Deterministic evidence (i.e., experts being sure that a parameter has exactly one value) can be considered as well by using a dirac-delta function to transform it to an (infinitesimal) interval. However, it should be noted that the foundations of DSTE and consequently of the sampling strategy revolve around intervals, and deterministic evidence is generally the exception (see examples discussed in Helton et al. [101]).

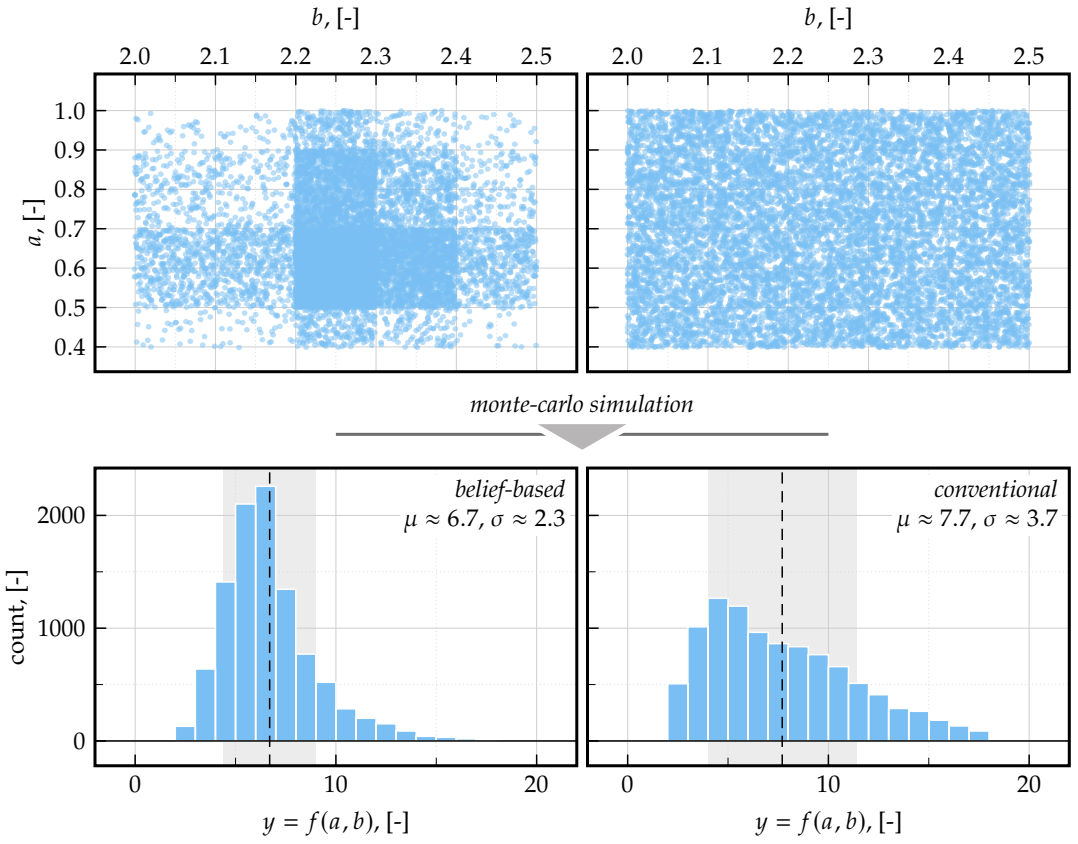


Figure 5.2 Samples (top) and monte carlo results (bottom) for the evidence-based sampling process example (left) and a simple uniform sampling (right).

Complementary Belief/Plausibility Functions

The concept of using CCBF, CCDF, and CCPF (which are referred to as CCXFs) for quantifying uncertainty has been briefly outlined in Section 2.2.2. To recall, they enrich the results of a Belief-based sampling-driven MCS by providing a distribution of Belief and Plausibility over the output range. From these, uncertainty-addressing statements can be constructed. While the CCDF is simply obtained by taking the complement of the conventional (empirical) CDF, the CCBF and CCPF calculations foresee retracing which input evidences lead to which values of the output distribution and then calculating the supporting and contradicting evidences over the output space. As such, it is performed after the results of the MCS have been gathered. As the process is quite complex, the following paragraphs dive deeper into the calculation of CCBFs and CCPFs while picking up the previous example model from Eqn. (5.1).

Starting with some preparation work, the first step is to pairwise combine the input evidences a_i and b_i to create a list of all k possible combinations of evidence intervals E_k. These are shown in the first few columns of Table 5.1. Next, a combined BBA need to be calculated for each of these combined intervals, which is the product of the individual BBAs. For $k = 2$, for instance, the individual BBAs are $m(a_1) = 1/2$ and $m(b_2) = 1/3$, resulting in a combined BBA of $m_{k=2} = 1/2 \cdot 1/3 = 1/6$. The next step requires creating a subset of the output space y_k for each k for which the input samples in a and b are in the corresponding evidence interval $E_k(a)$ and $E_k(b)$, respectively.

Table 5.1 Preparations for the calculation of the CCBF and CCPF.

Combination	Evidence Intervals $E_k(a)$	$E_k(b)$	BBA m_k	$y_k = (a \in E_k(a) \wedge b \in E_k(b))$ $\min(y_k)$	$\max(y_k)$
$k = 1 : a_1, b_1$	$[0.5, 0.7]$	$[2.0, 2.5]$	$1/2 \cdot 1/3 = 1/6$	2.376	14.133
$k = 2 : a_1, b_2$	$[0.5, 0.7]$	$[2.2, 2.3]$	$1/2 \cdot 1/3 = 1/6$	4.609	8.106
$k = 3 : a_1, b_3$	$[0.5, 0.7]$	$[2.2, 2.4]$	$1/2 \cdot 1/3 = 1/6$	3.320	8.106
$k = 4 : a_2, b_1$	$[0.6, 0.9]$	$[2.0, 2.5]$	$1/4 \cdot 1/3 = 1/12$	2.774	17.614
$k = 5 : a_2, b_2$	$[0.6, 0.9]$	$[2.2, 2.3]$	$1/4 \cdot 1/3 = 1/12$	5.270	10.485
$k = 6 : a_2, b_3$	$[0.6, 0.9]$	$[2.2, 2.4]$	$1/4 \cdot 1/3 = 1/12$	3.843	10.485
$k = 7 : a_3, b_1$	$[0.4, 1.0]$	$[2.0, 2.5]$	$1/4 \cdot 1/3 = 1/12$	2.050	19.769
$k = 8 : a_3, b_2$	$[0.4, 1.0]$	$[2.2, 2.3]$	$1/4 \cdot 1/3 = 1/12$	4.015	12.005
$k = 9 : a_3, b_3$	$[0.4, 1.0]$	$[2.2, 2.4]$	$1/4 \cdot 1/3 = 1/12$	2.884	12.005

Out of these subsets y_k, the minimum and maximum values need to be found, which are shown in the last two columns. In other words, for each of the possible evidence input combinations, the corresponding bounds in the output space needs to be calculated. This is illustrated in Fig. 5.3 for $k = 6$, where the highlighted sample space subset lies with $a_2 \in [0.6, 0.9]$ and $b_3 \in [2.2, 2.4]$ (left subplot) and the corresponding output space for these samples is $y_k \in [3.843, 10.485]$ (right subplot).

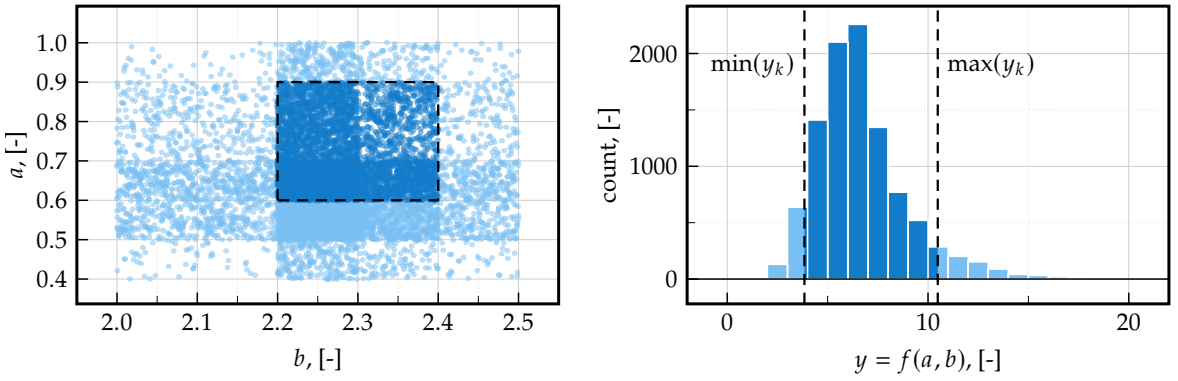


Figure 5.3 Highlighted DSTE sample space (left) and corresponding output space (right) for $k = 6$.

With the preparatory work done, the CCBF and CCPF can be calculated according to:

$$\text{CCBF} = \left\{ y_i, 1 - \sum m_k \right\} \quad \forall i \quad \text{for which} \quad \min(y_k) \leq y_i \quad \forall\, k \tag{5.6}$$

$$\text{CCPF} = \left\{ y_i, \sum m_k \right\} \quad \forall i \quad \text{for which} \quad \max(y_k) > y_i \quad \forall\, k \quad . \tag{5.7}$$

The first value in the braces y_i represents the abscissa vector, and the second is the ordinate vector. To evaluate this expression in practice, it is useful to sort the output values y, loop through them, evaluate the condition and calculate the abscissa and ordinate values for both functions. Consider, for example, the output value $y_i = 3.6$. Comparing this value with the last two columns of Table 5.1, the CCBF conditional statement $\min(y_k) \leq y_i$ is true for the rows $k = 1, 3, 4, 7$, and 9, leading to the CCBF of

$$\text{CCBF}(y_i = 3.6) = \left\{ y_i = 3.6, 1 - \sum m_k \right\} \quad \text{for} \quad k = \{1, 3, 4, 7, 9\} \tag{5.8}$$

$$= \{1.7, 1 - (2 \cdot 1/6 + 3 \cdot 1/12)\} = \left\{1.7, 0.41\bar{6}\right\} \quad . \tag{5.9}$$

For the CCPF, all rows fulfill the condition of $\max(y_k) > y_i = 3.6$, leading to:

$$\text{CCPF}(y_i = 3.6) = \left\{ y_i = 3.6, \sum m_k \right\} \quad \text{for} \quad k = \{1, 2, \ldots, 9\} \tag{5.10}$$

$$= \{3.6, 3 \cdot 1/6 + 6 \cdot 1/12\} = \{3.6, 1.0\} \quad . \tag{5.11}$$

By evaluating this across all values of y, it is possible to obtain the output space-spanning CCXFs, as depicted in the bottom part of Fig. 5.4. The ordinate represents the Belief, probability, or Plausibility of the output being greater than a particular y value, while the abscissa depicts the output y. Considering the previously discussed threshold value of $y_i = 3.6$ (dashed line), the CCDF (black curve) states that the probability of exceeding y_i is 0.95. This statement is entirely probabilistic, albeit derived from Belief-based sampling. The CCBF curve (in green) at this value shows that there is a Belief (i.e., quantification of supporting evidence) of ≈ 0.42 that the output is greater than $y = 3.6$ as calculated before. Meanwhile, the CCPF (in red) at this value is 1, indicating that the hypothesis ($y > 3.6$) is entirely plausible (i.e., there is no contradicting evidence). For further clarity, the upper plot of Fig. 5.4 illustrates the bounds for each k (i.e., row in Table 5.1). The evidence

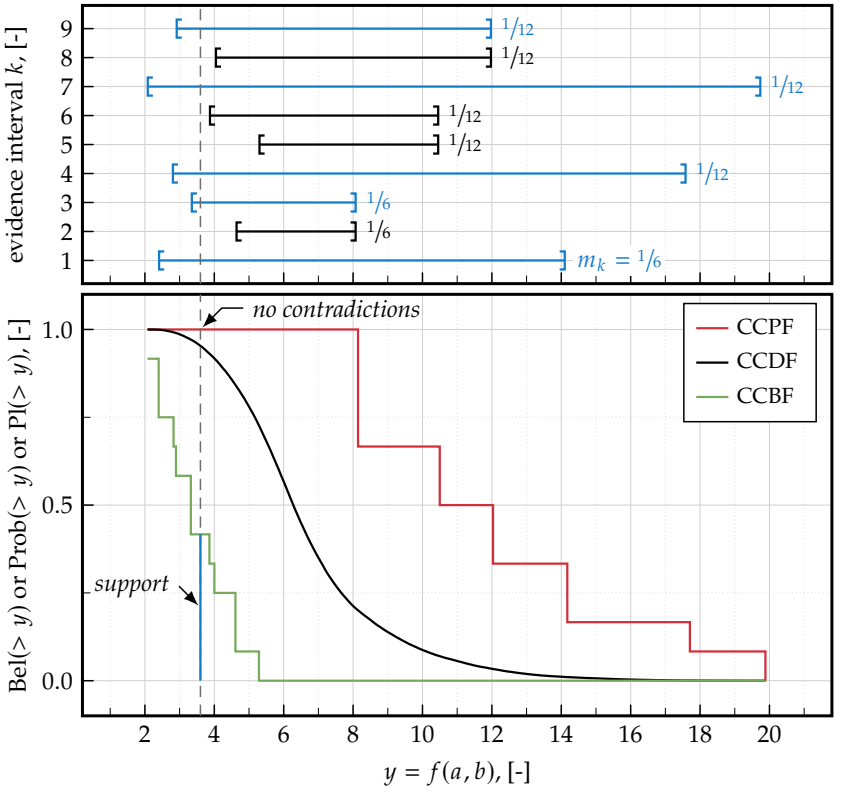


Figure 5.4 CCXF over the output space y (bottom) and bounds per evidence interval k (top).

intervals for which the current y_i value is inside are highlighted in blue. As y_i values increase, the model encounters more lower bounds, leading to a stepwise decrease in the CCBF. In a similar manner, exceeding more upper bounds results in a stepwise decrease in the CCPF. It is important to recognize that the number of evidence intervals increases exponentially with the number of input statements, contributing to the high computational cost of this method.

5.1.2 Application

Building on the second research hypothesis, a publicly available Python package named `dste` was developed to facilitate evidence-theoretic UQ studies. This package provides a comprehensive and practical tool for researchers and practitioners, aiming to make non-probabilistic UQ more accessible and user-friendly. To illustrate its usage, a notional evidence-based quantification of input uncertainties is presented. Code E.1 shows a minimal example of how to use `dste`, comprising four steps which are described in detail.

Code 5.1 Minimal example for using the DSTE package.

```
1  import dste  # available at PyPi (use >> pip install dste)
2  import numpy as np
3
4  # Step 1: Preparatory Work
5  # ------------------------
6  # Defining the model
7  def blackboxmodel(a, b):
8      return 100*(a**2 + b**2)**(a-b)
9
10 # Step 2: Define statements and combine them to evidences
11 # ---------------------------------------------------------
12
13 # Starting with a, for which the expert gave three statement.
14 st_a1 = dste.Statement(0.5, 0.7, 0.50)              # First statement for a
15 st_a2 = dste.Statement(0.6, 0.9, 0.25)              # Second statement for a
16 st_a3 = dste.Statement(0.4, 1.0, 0.25)              # Third statement for a
17 ev_a = dste.Evidence([st_a1, st_a2, st_a3])         # Evidence for a
18
19 # The second parameter b also received three statements in total
20 st_b1 = dste.Statement(2.0, 2.5, 1/3)               # First statement for a
21 st_b2 = dste.Statement(2.2, 2.3, 1/3)               # Second statement for a
22 st_b3 = dste.Statement(2.2, 2.4, 1/3)               # Third statement for a
23 ev_b = dste.Evidence([st_b1, st_b2, st_b3])         # Evidence for b
24
25 # Step 3: Create samples and run MCS
26 # ----------------------------------
27 n = 10000
28 evidences = [ev_a, ev_b]
29 samples = dste.advanced_sampling(evidences, n_samples=n, rule='latin_hypercube')
30
31 y = np.zeros(n)
32 for idx in range(n):
33     y[idx] = blackboxmodel(samples[idx, 0], samples[idx, 1])
34
35 # Step 4: Analyze output uncertainty
36 # -----------------------------------
37 X, CCDF, CCBF, CCPF, errcnt = dste.calc_evidence_metrics(y, evidences=evidences, samples=samples)
```

Step 1: Define the main model or make it accessible by importing it. The first preparatory step requires users to make the main model accessible to the code at hand. As it is treated as a black box, modifications to the main model are not necessary. In this example, the previous model from Eqn. (5.1) is used.

Step 2: Define statements and convert them to evidences. Based on the information gathered from SMEs in the form of intervals and BBAs, this step foresees the definition of statements and evidences. In this example, only one expert was interviewed, resulting in three statements (lines 14-16 and 20-22) and one evidence object per variable (lines 17 and 23). If, alternatively, multiple experts had provided statements, DSTE's combination rule would be utilized, for which a Python code example is provided in Code 5.2. Here, three experts provided one

Code 5.2 Adjustment of Step 2 (lines 13-23 in Code 5.1) for multiple experts.

```
1  # Individual statements of each expert for parameter a, each with full belief
2  st_a_e1 = dste.Statement(0.5, 0.7, 1.0)            # First expert's statement for a
3  st_a_e2 = dste.Statement(0.6, 0.9, 1.0)            # Second expert's statement for a
4  st_a_e3 = dste.Statement(0.4, 1.0, 1.0)            # Third expert's statement for a
5
6  # Transformation to evidence objects. Each statement receives its own evidence object.
7  ev_a_e1 = dste.Evidence([st_a_e1])
8  ev_a_e2 = dste.Evidence([st_a_e2])
9  ev_a_e3 = dste.Evidence([st_a_e3])
10
11 # Combination of evidences with the first expert receiving
12 # twice as much credibility than experts two and three.
13 ev_a = dste.combine_evidences([ev_a_e1, ev_a_e2, ev_a_e3], weights=[2, 1, 1])
14
15 # Individual statements from expert 1 for parameter b
16 st_b1_e1 = dste.Statement(2.0, 2.5, 0.5)
17 st_b2_e1 = dste.Statement(2.2, 2.3, 0.5)
18 ev_b_e1 = dste.Evidence([st_b1_e1, st_b2_e1])
19
20 # Individual statement from expert 2 for parameter b
21 st_b_e2 = dste.Statement(2.2, 2.4, 1.0)
22 ev_b_e2 = dste.Evidence([st_b_e2])
23
24 # Combination of evidences with equal weights
25 ev_b = dste.combine_evidences([ev_b_e1, ev_b_e2], weights=[1, 1])
```

interval each for a, with a BBA of 1. These are then combined, assuming that the first expert is twice as credible as the other two. For b, one expert provided two intervals, which are then transformed into one evidence object. A second expert stated only one interval for b. These statements are then combined to evidences assuming equal weights in the last line of Code 5.2. For visualization purposes, the package provides an evidence plotting function that depicts the so-called BBA space, which is constructed out of the general density functions described earlier. For the adjusted example (i.e., with multiple experts interviewed), this is shown in Fig. 5.5 for a (left) and b (right). The advantage of this depiction is that the individual statements are discernible, and their weights are scaled considering not only the individual BBAs of the statements but also the credibility of the experts when weighting them. Therefore, the height of the box is denoted with `st_a_e1` (first expert's statement) is twice as high as those of `st_a_e2` (second expert's statement) and `st_a_e3` (third expert's statement) due to the weighting in line 13 of Code 5.2, while all boxes for b have the same height due to equal expert credibility.

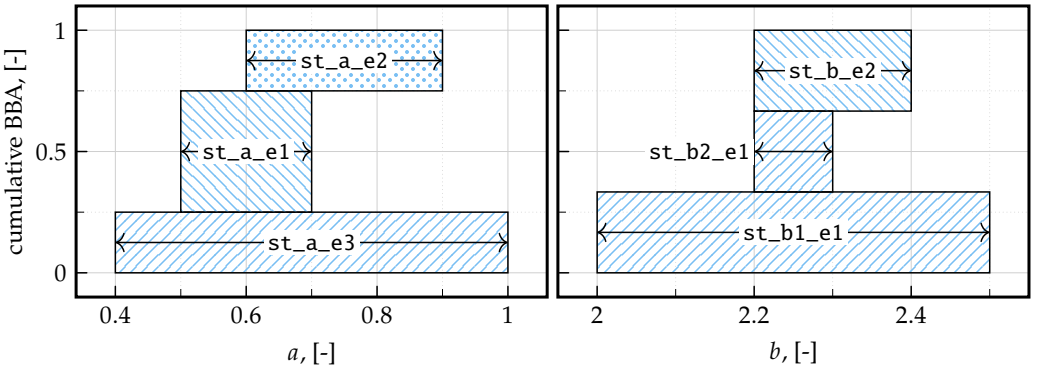


Figure 5.5 BBA space for a and b after combining statements to evidences.

Step 3: Create samples from the evidences and run a MCS. With the evidences for a and b created, the next step is to generate a number of samples and feed them to a MCS of the original model. The package allows random and quasi-random sampling schemes such as LHS or the Sobol' sequence. Here, 10,000 samples are generated using LHS (see line 27 of Code 5.1). Next, these samples are fed to the original model, recording its response `y` (in lines 31-33 of Code 5.1).

Step 4: Calculate evidence-theory-based metrics to analyze output uncertainty. In the last step, DSTE's key metrics, CCBF and CCPF, are calculated alongside with the CCDF. With the developed package, all calculations described in Section 5.1.1 reduce to one function, i.e., `calc_evidence_metrics` (line 37), which was used to create Fig. 5.4.

5.2 Methodology

This Section begins with an outline of the study design, which presents the overall sequence of steps for the upcoming analyses. Afterwards, the expert elicitation process is described in more detail.

5.2.1 Study Design

The steps performed to understand the potential of evidence-theoretic methods are illustrated in Fig. 5.6 and are divided into two areas: (a) steps specific to DSTE, shown at the top half, and (b) the conventional probabilistic approach, which serves as a comparison and is depicted in the lower portion. The first step foresees a selection of epistemic uncertainties for both domains. Criteria for this selection are, apart from the epistemic nature, data scarcity and the availability of SMEs that are able to provide Belief statements. The next step in the evidence-theoretic domain entails interviewing said SMEs. These interviews, which are conducted using a common questionnaire, produce intervals and Beliefs (i.e., BBAs) in them. In an effort to imitate such intervals and Beliefs for uncertainties for which no expert elicitation is available, the literature sources from which the probabilistic intervals were derived are revisited. The results from this step are used for the evidence-theory-based sampling strategy. A subsequent MCS is performed with these samples, resulting in an output distribution that can be compared to the conventional probabilistic counterpart. The next step in the evidence path is to calculate the Belief and Plausibility curves of the output, which provide information about the uncertainty in the evidence-theoretic domain. These are then compared to the conventional uncertainty measures of ProT.

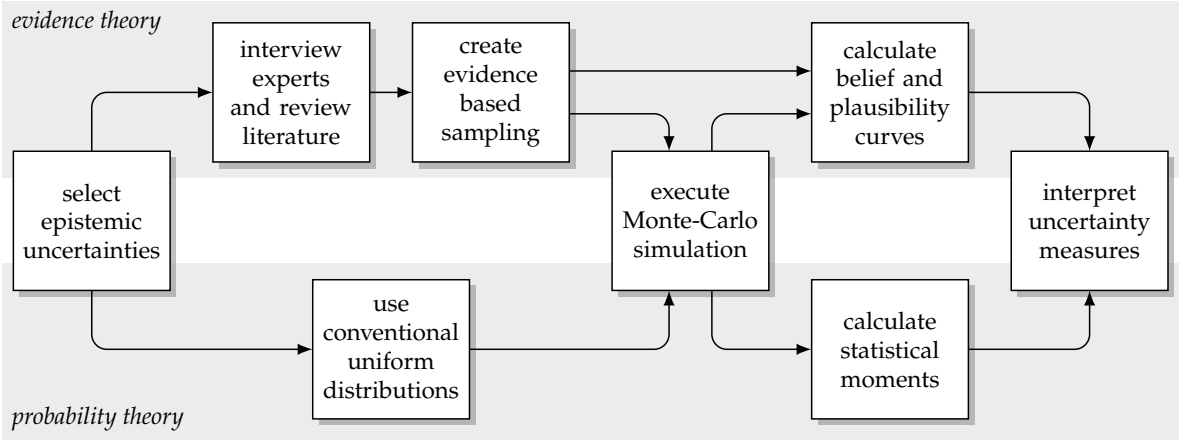


Figure 5.6 Methodology for the evidence-theoretic uncertainty quantification study.

5.2.2 Expert Elicitation Process

Expert judgment can be defined as the educated opinion of a SME, based on their knowledge and experience, in response to a technological problem [185]. It is essentially a snapshot of the knowledge of an individual at the time of the elicitation [130]. Its formal elicitation is a useful method to obtain evidences for uncertainties when data availability is too low to continue with conventional probabilistic methods. It draws from various fields such as decision analysis,

statistics, and knowledge acquisition. Advantages include improved rigor, defensibility of judgments, and the capacity to alter conclusions when new information becomes available [223, p. 105f]. The overall process can be divided into four phases, which are loosely based on the guidelines from Ross [223, Chapter 6] and have been adjusted to be used with DSTE.

Before discussing the recommendations for the expert elicitation process, a note on potential bias is appropriate. During elicitation, especially when conducted in an TEA setting, intentional bias can occur when the outcome of the study is thought to potentially affect the future research and development funding situation. Unintentional bias can occur when experts are not interviewed independently, potentially resulting in statements being altered to match those of other experts. While intentional bias is hard to detect and even harder to circumvent, unintentional bias can be easily mitigated by interviewing each expert individually and by not disclosing the answers from other interviews[3].

Phase 1: Select Appropriate Experts

In the first phase, available experts need to be identified. Apart from their general willingness to be interviewed, the individuals need to have some level of experience with the epistemic parameters in question. Hereby, different aspects, such as the number of years worked in this field, should be considered. However, as there is no formal definition of an expert, the final identification is up to the responsible research party. For this thesis, which draws from the Clean Sky 2 Projects HLFC-Win and ECHO, five individuals from industry as well as academia have been selected. They comprise the two advisory board members *Roger Taplin* and *Heinz Hansen*, the DLR aerodynamicist *Thomas Kilian* and systems engineer *Alexander Bismark*, as well as the HLFC researcher *Tim Effing* from RWTH Aachen University. Short biographies can be found in Appendix E.2.

Phase 2: Prepare the Interviews

In the second phase, the questions, answer structure, and interview process are designed. As it is well known that the way questions are asked can affect the answers, this step is the most critical with respect to mitigating bias and maximizing the usefulness of the outcomes. For this study, two options were investigated: (a) provide predefined intervals and ask for an assignment of basic Beliefs to these, or (b) ask the experts to provide their own intervals and BBAs. Due to the open nature of the questions in the latter option, it tends to be less subjected to bias (from the interviewer), but can lead to experts providing only one deterministic value (representing what is thought to be the *most likely* one), or one - often large - interval with no differentiation of BBAs within it[4]. In contrast, option (a) may introduce some subjective bias (from the interviewer), depending on how the provided intervals are derived, but it motivates the experts to think in multiple intervals and BBAs, which better fits the intended use in DSTE. Both options were tested in a pilot interview, resulting in more differentiated results with option (a), which is why intervals have been provided in the final interviews. These are of the form depicted in Fig. 5.7, which shows the response from one of the experts regarding the drag

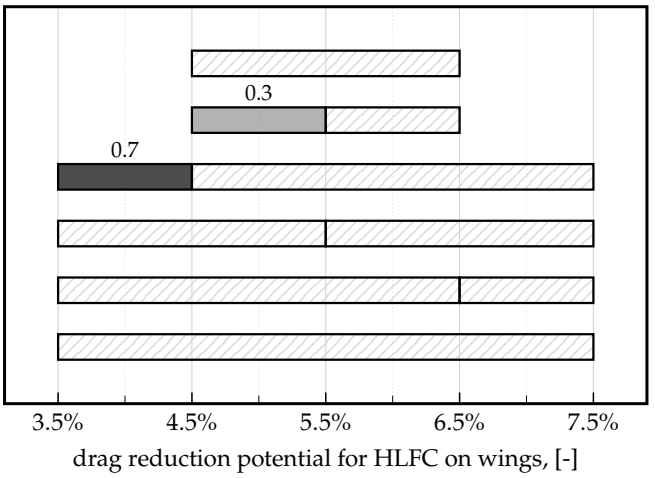


Figure 5.7 Visualization of interview intervals.

[3]For a more detailed classification of bias, consider Hester [107].

[4]This is due to two reasons: (1) typical interview situations, or even the day to day queries, tend to look for deterministic values, and (2) DSTE is, contrary to ProT, not widely taught, resulting in most people being more comfortable with providing single values or at most one interval with a subjective probability of the real value falling within that interval.

reduction potential of HLFC when applied to the upper wing portion. Here, the overall bounds of 3.5 and 7.5 % have been defined beforehand for each parameter based on literature values and previous studies. Next, a smallest interval width is chosen, which is 1 % in this example. These intervals are then combined successively with their neighbors, so that all possible intervals are shown as bars in one figure. At first, all intervals are *inactive* (shown as hatched bars). Experts are then asked to select an interval (e.g., between 4.5 and 5.5 %) and assign a BBA (e.g., 0.3), which activates the interval (e.g., the bar in the second row in Fig. 5.7). Figures like this are created for each uncertain parameter before the interviews. As this representation is not self-explanatory, the first part of the interview requires an introduction, which is part of the next phase.

Phase 3: Conduct the Interviews

In this phase, the interviews are conducted, beginning with a short introduction in which not only the motivation behind the interview is reiterated but also an example answer is discussed with the experts. Furthermore, each parameter requires a proper definition to avoid the typical "*well, it depends...*" answer. Once these aspects have been cleared up, the individuals are asked if they feel comfortable giving an expert judgment on this question. If not, the interview continues to the next question. It is important to allow this and not push the individual toward an answer in order to ensure the evidences are, in fact, based on knowledge and expertise. As HLFC is a highly interdisciplinary technology, and most of the interviewed individuals have expertise in one or a few specific disciplines, not every expert felt comfortable answering every question, which is something that should be expected from the start. It is self-explanatory that each interview should be documented properly.

Phase 4: Prepare the Outcomes

In the last phase, the outcomes of the interviews need to be prepared for their use in the evidence-based study, which may include some adjustments. Consider, for example, the expected maintenance cost of the HLFC system. While the input UQ in Chapter 4 used CERs with the system mass as a driving parameter, resulting in maintenance cost per kg, experts are more comfortable expressing maintenance cost per flight hour or flight cycle. Therefore, some transformation may be needed in order to feed this parameter to the lifecycle model. Similarly to the interviews themselves, the outcome modifications need to be well documented as well.

5.3 Evidence Based Input Uncertainty Quantification

As outlined in the methodology of this Chapter, the first step in the DSTE-based UQ is to select the epistemic uncertainties for which evidences are to be gathered. As per the analyses in Chapter 4, eight of the eleven parameters are epistemic: (1) drag reduction potential, (2) mass increase, (3) SFC penalty, (4) extra contingency fuel, (5) load factor, (6) insect contamination, (7) maintenance cost increase, and (8) aircraft price increase. The criteria for the selection are multifold. First and foremost, the intended acquisition of evidences should be considered. For this analysis, the main source will be expert elicitation through interviews. As an additional source, the literature from which the probabilistic bounds have been derived (see pp. 61f and 86f) is revisited. Therefore, uncertainties for which no experts are available and the literature does not seem to be sufficiently existent are not considered for the selection. This pertains to parameter (4), i.e., the extra contingency fuel, as no airline or operations specialist is available and literature on this topic is sparse. For similar reasons, parameter (6), i.e., insect contamination, is not selected either. Another criterion is concerned with the source of epistemic uncertainty. If the lack of knowledge originates from model implementation rather than input values, expert elicitation (in the form of asking for intervals and/or BBAs) is not particularly useful. This pertains to the Bayesian ridge regression used for the insect contamination as well as to the simplified load factor variation (i.e., parameter 5), albeit to a

Table 5.2 Overview of selected epistemic uncertainties for the DSTE-based UQ.

parameter	experts*						notes
	AB	HH	RT	TE	TK	LIT	
$\mathcal{D}_1$: drag reduction	○	●	○	●	●	○	split into wing, HTP, and VTP
$\mathcal{D}_2$: mass increase	○	○	○	○	○	●	see Fig. 5.10
$\mathcal{D}_3$: SFC penalty	○	○	○	○	○	●	see Fig. 5.10 and Table 3.2
$\mathcal{M}_2$: HLFC maintenance	●	◑	●	●	○	○	split into scheduled and unscheduled
$\mathcal{M}_3$: price increase	●	●	●	●	○	○	–

* AB: Alexander Bismark, HH: Heinz Hansen, RT: Roger Taplin, TE: Tim Effing, TK: Thomas Kilian, LIT: literature.

lesser degree. As the latter parameter had virtually no impact on the overall uncertainty (see GSA results in Section 4.3.3), it is left out as well. This leads to the five selected parameters shown in Table 5.2. For three out of the five parameters, expert statements were gathered from interviews, whereas for two, namely $\mathcal{D}_2$ (SFC penalty) and $\mathcal{D}_3$ (mass increase), available literature was revisited and treated as hypothetical experts. As the notes column shows, the drag reduction potential of HLFC was split into the three components wing, HTP and VTP, as experts were more comfortable providing these individually. The same applies to the HLFC maintenance, where *Heinz Hansen* provided a value only for the unscheduled part (hence the half-filled harvey ball).

Next, the evidences and resulting sample spaces for each of these inputs are discussed. Thereby, the expert interview-derived ones are presented first, followed by the two literature-based evidences.

5.3.1 Drag Reduction

The drag reduction potential was estimated by three SMEs with 20 statements in total. With experience in studies and industrial projects where HLFC was applied to individual components, the experts were more comfortable providing BBAs for the application on the upper wing, HTP, and VTP, individually. The drag reduction range for the upper wing application was limited in the interview process between 3.5 and 7.5 %, whereas the bounds for the HTP and VTP application was [1.0 %, 2.5 %] and [0.5 %, 2.0 %], respectively. The latter two were split into bins of 0.5 % and larger, while the smallest intervals of the former were 1 % wide. Fig. 5.8 shows the BBA space (top row) and the resulting sample histograms (middle row) for each component. Here, the BBA space reflects the already combined evidence[5] and each fill pattern represents the statements of a different expert. The statements for the drag reduction potential HLFC applied on the upper side of the wing (top left plot) span the entire provided range, with a slight favor towards the lower end. This is also discernible in the drawn samples, which have a mean value of 5.2 %. For the HTP, the experts provided identical intervals (with different BBAs) several times, despite being interviewed independently. The interval with the highest Belief assignment was [1.5, 2.0], resulting in samples with a mean of 1.8 %. The VTP sample space has a similar shape, albeit with a less prominent favorite interval. Here, the mean value computes to 1.2 %. As the overall model expects one drag factor as an input, the sample spaces of each component needed to be summed up[6]. The result is shown in the bottom plot of Fig. 5.8, with values ranging from 5 up to nearly 12 %. Compared to the previously derived uniform distribution between 6 and 8.6 % (as shown with the dark gray overlayed histogram in the bottom plot), this expert-based sampling process is expected to change the overall result of the MCS noticeably. The mean and standard deviation of the evidence-theoretic samples are 8.1 % and 1.2 %, respectively, which are both higher than the respective values when using the conventional uniform distribution, which had a mean value and standard deviation of 7.2 % and 0.69 %, respectively.

[5] As shown conceptually in Fig. 5.5.

[6] For this, the drag reduction values of individual components are assumed to be independent of one another.

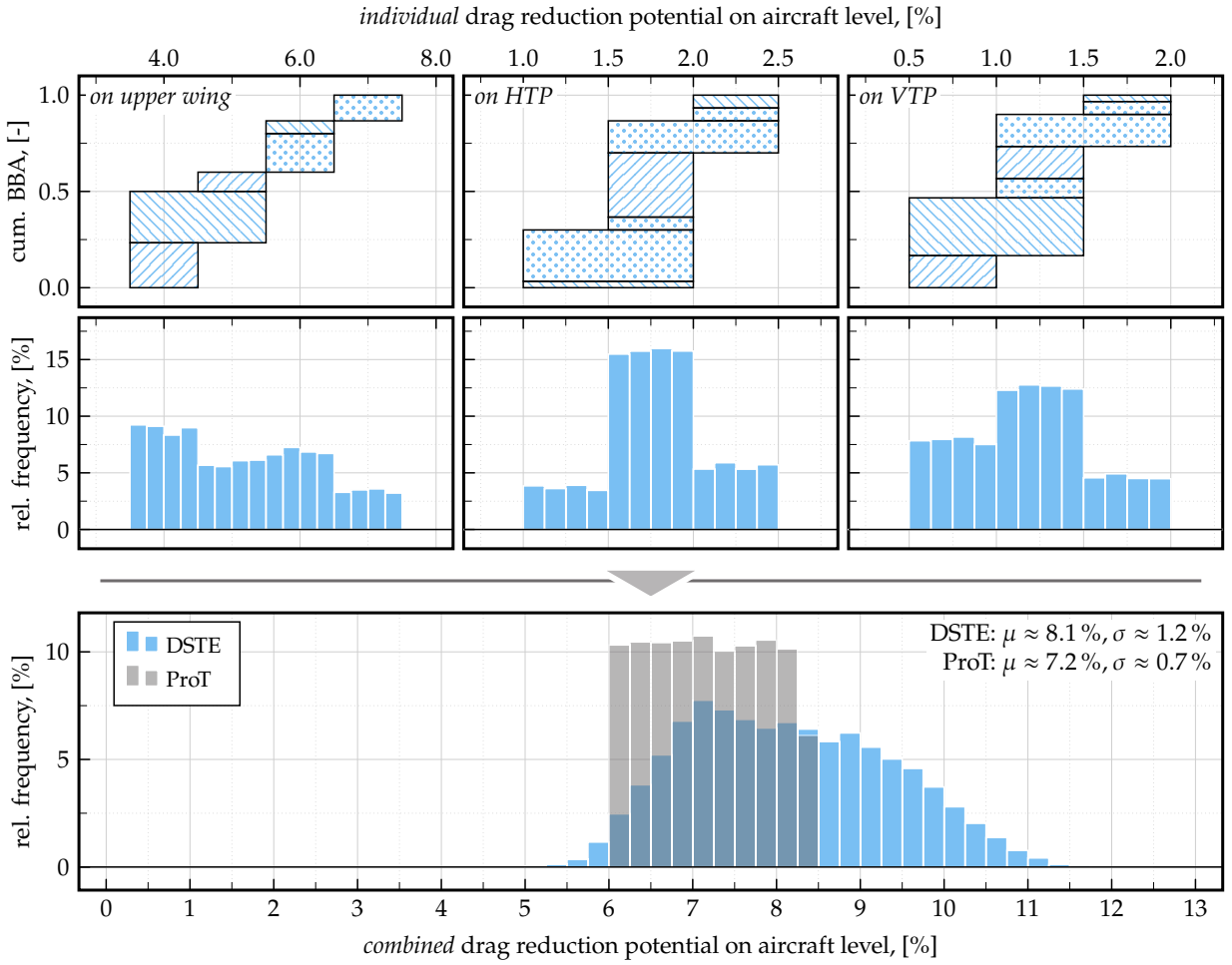


Figure 5.8 BBA space (top) and histograms (center and bottom) for the individual and combined drag reduction potentials with overlayed ProT-based histogram.

5.3.2 Increase in Maintenance Cost and Aircraft Price

The next epistemic uncertainty for which the experts gave evidences for is the expected increase in HLFC-related maintenance cost, which is split into a scheduled and an unscheduled portion. For the former, three of the five experts gave eight statements in total, whereas the latter has received five statements from three SMEs[7]. As opposed to the \$/kg representation used in Chapter 4, the pilot interview has revealed that the maintenance cost per flight hour is a more appropriate choice of representation. Bounds for this question were derived using literature values for component and system maintenance and were \$1.0/FH and \$11.5/FH, split into bins with a minimum width of \$2.5/FH. To further facilitate the elicitation, orientation values for scheduled maintenance cost were provided, which comprise tires (\$3.5/FH), APU (\$7/FH), thrust reversers (\$11/FH), and landing gears (\$19/FH) [5, p. 27]. The BBA spaces and sample histograms are depicted on the left side of Fig. 5.9. As the top left plot shows, the SMEs's statements cover the range from \$3.5 to \$11.5/FH but favor the centermost interval between \$6.5/FH and \$9/FH, resulting in an average value of \$7.5/FH. The unscheduled portion of HLFC maintenance was specified as a ratio of the scheduled cost, with provided intervals ranging from 80 to 240 % in bins of 40 %[8]. The most favored interval lies between 120 and 160 , leading to an expected ratio of 140 %. From these two elements, the total annual HLFC-related maintenance cost was calculated (using an expected annual utilization of 4400 FH/year; see Fig. 3.12 (a)), for which the histogram is shown in the bottom plot of Fig. 5.9. Values range from

[7] Note that there is no lower limit to the number of statements in evidence-driven expert elicitation. If only one statement is given, it receives full Belief by definition, rendering the sampling approach equal to the conventional uniform distribution.

[8] These values stem from the bounds provided by Suwondo [258].

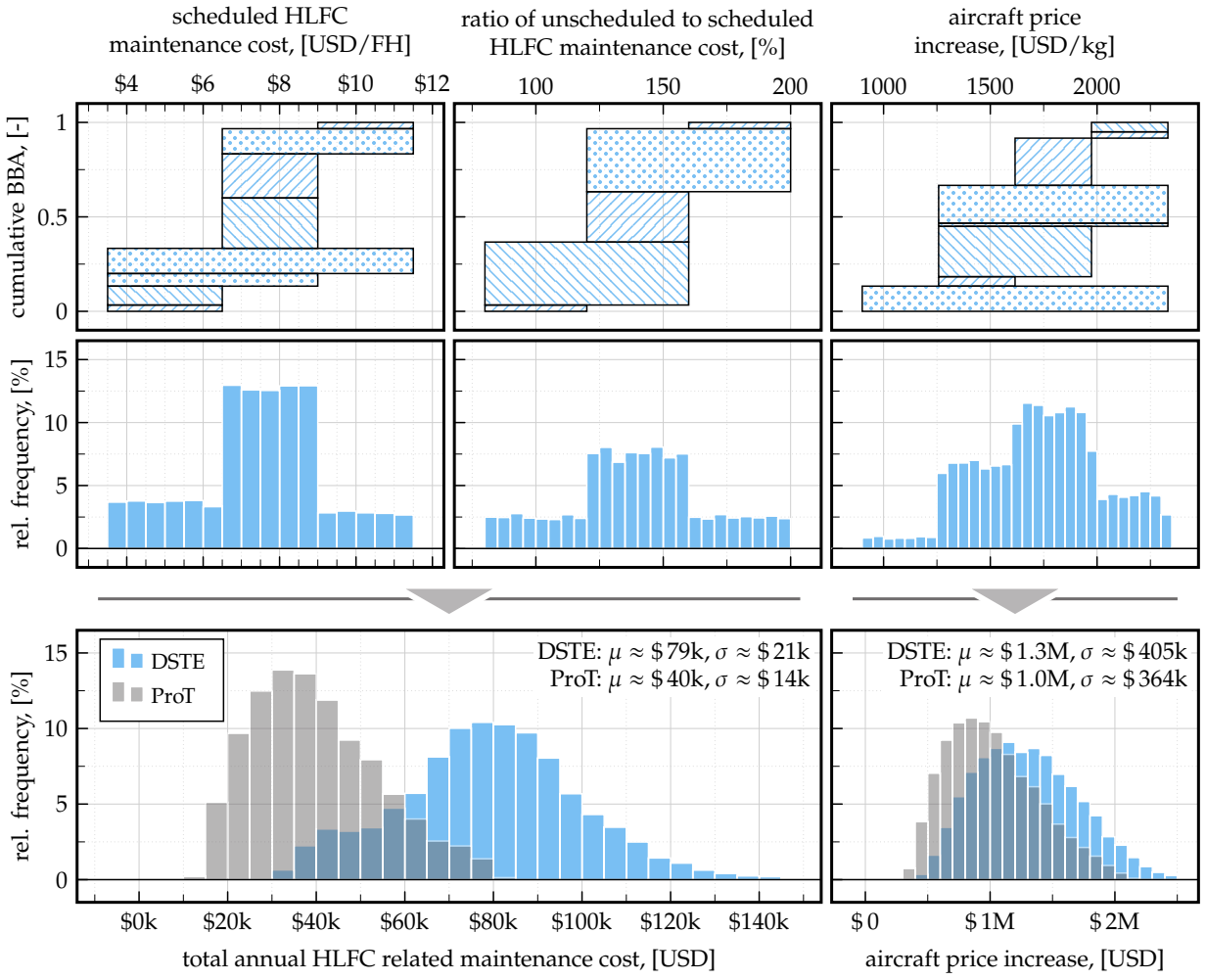


Figure 5.9 BBA space (top) and histograms (center and bottom) for the (un)scheduled and total HLFC maintenance cost (left) and aircraft price increase (right) with overlayed ProT-based histogram.

$30k to $150k and show a mean value and standard deviation of $79k and $21k, respectively. Similar to the drag reduction parameter, the UQ of the HLFC maintenance parameter differs significantly when using the DSTE-based technique compared to the conventional approach, which had a lower mean value of $40k and a lower standard deviation of $14k.

The third and last parameter specified by the interviewed experts is the aircraft price increase. For this, three experts have provided answers, resulting in a total of eight statements. Similarly to the approach for the scheduled maintenance cost, comparable recurring costs have been provided, including the landing gear ($830/kg), wing ($3.3k/kg), and systems[9] ($1.7k/kg) [157, p. 67]. Input intervals ranged from $542/kg to $2332/kg, with an interval width of multiples of $380/kg. These are shown in the third column of Fig. 5.9. The majority of samples range between $1260 and $2300 per kg, with only a few samples residing in the lowest interval between $900 and $1200 per kg. These were translated to absolute cost using the mass increase variation (which is explained in the next Section), resulting in the histogram in the bottom right plot of Fig. 5.9. The average value and standard deviation for this parameter are $1.3M and $405k, respectively. Compared to the conventional approach derived for the previous GSA (which had a mean value and standard deviation of $1M and $364k, respectively), it can be concluded that the expert-based sampling strategy is relatively close to the conventional technique.

[9] A category described to contain onboard systems including avionics, fuel, flight controls, hydraulics, and others. Note that the cost were inflation corrected from the year 2002 to 2023.

5.3.3 Mass Increase and SFC Penalty

For the last two epistemic uncertainties, an alternative to the interview process was needed, as no SME was able to provide BBAs for the mass increase and SFC penalty. Thus, the literature sources from which the bounds of the GSA had been obtained in Chapter 4 were revisited. As these sources only provided point values, a truly objective derivation of statements and evidence proved difficult. It is technically possible to create evidences with one or a few point-valued statements. However, creating evidences completely without intervals is contrary to DSTE's interval-based foundation, as described before. Therefore, a different approach to the literature-derived evidences was taken. Here, the six literature sources were regionally grouped into those stemming from the United States (i.e., references [18, 29], both from or for Boeing, shown in red) and those from Europe (i.e., references [70, 218, 296], shown in blue); see Fig. 5.10 (top row). Next, the lowest and highest values from these groups were treated as statement intervals with an individual BBA of 1. This resulted in the intervals $\left[384\,\text{kg}, 1103\,\text{kg}\right]$ (Europe) and $\left[530\,\text{kg}, 1016\,\text{kg}\right]$ (USA) for the mass increase. The suction power intervals are [115 kW, 184 kW] (Europe) and [165 kW, 242 kW] (USA). To further distinguish the groups, the average time since their publication was considered. With the publications from the US being more than three times older than those from Europe, a subjective credibility of three to one in favor of Europe was implemented. This results in the BBA spaces in the center row of Fig. 5.10, which in turn are used to draw the samples from, as depicted in the bottom row of Fig. 5.10. Note that differences from the probabilistic counterpart in Chapter 4 are marginal, as evident by the comparison of mean values and standard deviations in the plots in the bottom row.

In line with the process described in Section 3.3.2 (pp. 61f) and Section 4.3.2 (pp. 86), the

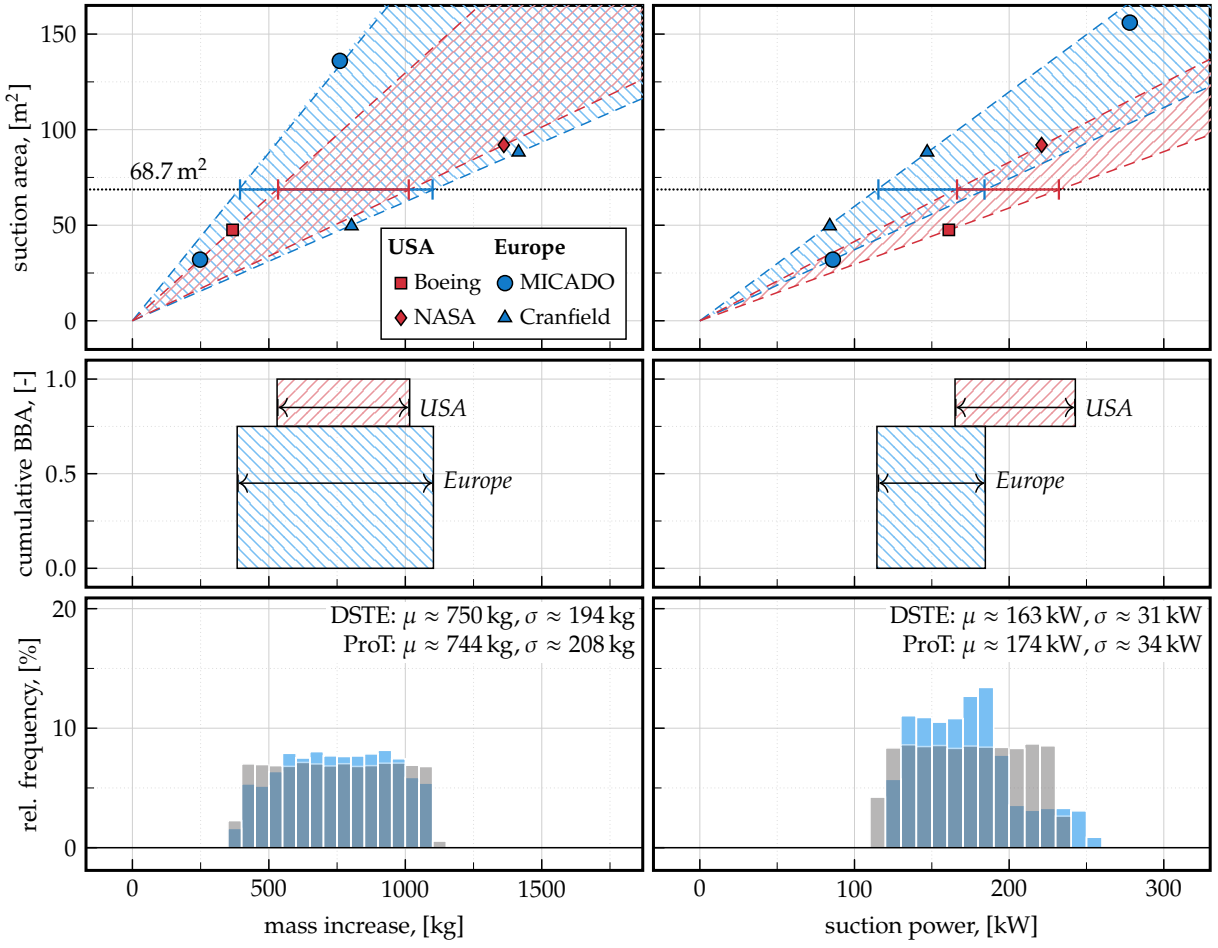


Figure 5.10 Revised literature data (top), BBA space (center), and histogram (bottom) for the uncertainties of mass increase and suction power with overlayed ProT-based histogram.

suction power needs to be translated to an SFC penalty using the efficiencies of the electrical components (including their uncertainties; see Table 3.2). Since estimating the suction power with the DSTE approach does not differ significantly from the conventional approach, the resulting SFC penalty shown in Fig. 5.11 is fairly close to the ProT-based sampling. The only observable difference is a minor prioritization of lower values (i.e., < 1.3 %) when using the evidence-based approach.

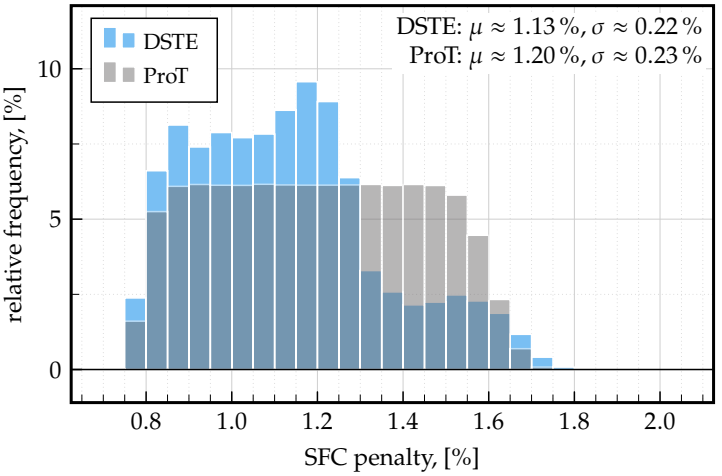


Figure 5.11 Histograms of resulting SFC penalty samples with the DSTE and conventional approach.

Before moving on to the next Section, it should be noted that there are various restrictions to consider when using DSTE to build samples from literature-provided values as it was done here. First, it is essential to note that this method is only one of many possible approaches, and alternatives such as Bayesian inference could be considered. Second, this method is subjective due to the arbitrary weighting of groups, the grouping itself, and the difficulties in merging these uncertainties with others, such as those pertaining to power losses due to the electrical efficiencies of HLFC-components. When evaluating the use of DSTE to create samples from literature-provided values, it is essential to bear these constraints in mind.

5.3.4 Parameter Overview

For overview purposes, the input parameter space of DSTE is compared to its probabilistic counterpart. Table 5.3 shows the bounds, the average value, and the standard deviation for both approaches. The aleatory parameters (namely $\mathcal{O}_1$, $\mathcal{E}_2$, and $\mathcal{M}_1$) as well as three epistemic parameters (namely $\mathcal{O}_2$, $\mathcal{O}_3$, and $\mathcal{E}_1$) have been fixed to their average value from Table D.1. By doing this, the focus of this study is set on the five epistemic parameters that have been previously identified as suitable for the evidence-theoretic sampling.

In the design domain, using DSTE for the UQ leads to a noticeable change in $\mathcal{D}_1$ (drag reduction) of +13 % (or 0.9 %-points), which is beneficial for the assessment of the HLFC aircraft. However, the DSTE-based approach led to a significant increase in the standard deviation of $\mathcal{D}_1$ of 71 % (or 0.5 %-points). The other two design uncertainties, $\mathcal{D}_2$ (mass increase) and $\mathcal{D}_3$ (SFC penalty), did not change noticeably, neither in their average value nor in the standard deviation. Since these parameters are the only varied ones that affect the fuel consumption improvement $\Delta\gamma_{\text{fuel}}$, it is expected that the evidence-theoretic approach will show a significant improvement over the probabilistic UQ while showing a larger spread in the output space (due to the higher standard deviation in $\mathcal{D}_1$). The economic superiority ΔNPV, however, is more difficult to predict. Here, the maintenance increase doubled in its average value (albeit with a lower standard deviation) and the aircraft price uncertainty increased by 30 %. Both of these DSTE-induced changes decrease the economic potential of the HLFC aircraft and may outweigh the benefits of the expected improvement in fuel consumption.

5.4 Results

With the input parameters fully defined, the results of the evidence-theoretic-driven MCS are now discussed. Section 5.4.1 begins with a comparison between the probabilistic outputs and

Table 5.3 Summary of uncertain inputs for the DSTE-based UQ.

domain			**probability-based**			**evidence-based**		
parameter	unit	type*	distribution	avg.	std.	source	avg	std.
design								
$\mathcal{D}_1$: drag reduction	[%]	E	uniform	7.2	0.7	experts	8.1	1.2
$\mathcal{D}_2$: mass increase	[kg]	E	uniform	744	208	literature	750	194
$\mathcal{D}_3$: SFC penalty	[%]	E	gen. normal	1.18	0.23	literature	1.13	0.22
operation								
$\mathcal{O}_1$: cruise mach	[-]	A	— *fixed to 0.82* —			— *fixed to 0.82* —		
$\mathcal{O}_2$: extra cont. fuel†	[-]	E	— *fixed to 0.50* —			— *fixed to 0.50* —		
$\mathcal{O}_3$: load factor	[-]	E	— *fixed to 0.80* —			— *fixed to 0.80* —		
environment								
$\mathcal{E}_1$: insect contamin.‡	[-]	E	— *fixed to 0.50* —			— *fixed to 0.50* —		
$\mathcal{E}_2$: cloud encounter	[-]	A	— *fixed to 0.10* —			— *fixed to 0.10* —		
economic								
$\mathcal{M}_1$: fuel price§	[-]	A	— *fixed to 0.46* —			— *fixed to 0.46* —		
$\mathcal{M}_2$: maint. increase	[USD/yr]	E	nagasaki	\$40k	\$14k	experts	\$79k	\$21k
$\mathcal{M}_3$: price increase	[USD]	E	nagasaki	\$1.0M	\$364k	experts	\$1.3M	\$405k

*E: epistemic, A: aleatory; †see Eqn. (4.22); ‡see Eqn. (4.23); §see Eqn. (4.27).

those stemming from the DSTE approach. Here, the overall distributions, including their mean values, standard deviations, and probabilities of success, are discussed. Section 5.4.2 presents the evidence-theoretic uncertainty measures in more detail and derives statements regarding the output using the CCXFs. For verification and sanity check purposes, the input space is modified while monitoring the impact on the DSTE metrics in Section 5.4.3. Finally, some insights into the convergence behavior of this approach are discussed in Section 5.4.4.

5.4.1 Comparison of the Output Distribution

The first set of results is depicted in Fig. 5.12, with the change in fuel consumption on the left side and economic superiority on the right. For both, the upper portion shows the results of the probabilistic sampling, whereas the lower part shows the results with the updated, evidence-based sampling strategy. These results have been obtained with a sample size of $N = 250{,}000$.

For $\Delta\gamma_{\text{fuel}}$, the probabilistic sampling leads to an average of $-2.1\,\%$ and a standard deviation of $0.6\,\%$ (visualized by the width of the light gray area around the mean). The overall shape resembles a triangular distribution that is bound in $[-3.4\,\%, -0.6\,\%]$. In contrast, the updated and evidence-based sampling strategy leads to a more tail-heavy distribution with wider bounds, i.e., $[-5.7\,\%, -1.1\,\%]$. The updated average and standard deviation are $-3.5\,\%$ and $0.8\,\%$, respectively. Thus, compared to the probabilistic approach, the expert elicitation process resulted in a $70\,\%$ better-performing HLFC aircraft in terms of fuel consumption at the cost of higher uncertainty.

The economic results, expressed in ΔNPV, lead to an average superiority of \$1.2M with a standard deviation of \$905k when using the conventional probabilistic sampling. Here, the share of samples where the HLFC aircraft is better than its counterpart (i.e., the probability of success) is $90\,\%$. The overall bounds are [−\$1.6M, +\$3.9M]. Similarly to the $\Delta\gamma_{\text{fuel}}$ results, the evidence-based sampling approach led to an improvement in the economic superiority of the HLFC aircraft. That is, the average ΔNPV doubled to \$2.5M and the probability of success has increased to $98\,\%$. The standard deviation as an indicator for the output uncertainty also increased to \$1.2M, which is most likely caused by the substantial increase of the input uncertainty of $\mathcal{M}_2$ and $\mathcal{M}_3$ in the evidence-based sampling (see Table 5.3). Evidently, their higher average values (compared to the probabilistic sampling), did not compensate for the effect of the improved average values of $\mathcal{D}_1$ and $\mathcal{D}_3$.

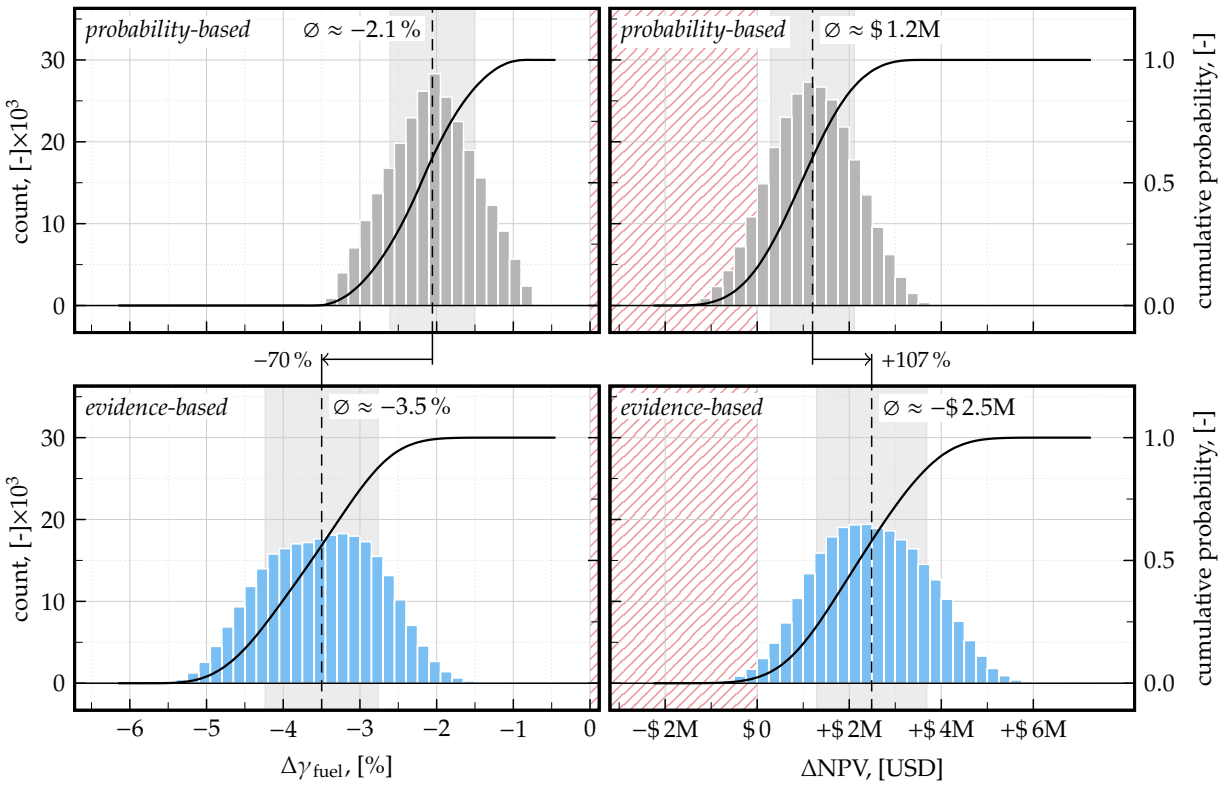


Figure 5.12 Histogram for $\Delta\gamma_{\text{fuel}}$ (left) and ΔNPV (right) using the probabilistic sampling (top) and the evidence-based sampling (bottom) including the average value (dashed line) ± the standard deviation (light gray).

In summary, an analysis of both the probability-based and evidence-based sampling methods' output distributions shows that expert elicitation leads to a more favorable assessment of the HLFC aircraft. However, this comes at the cost of heightened input parameter uncertainty, resulting in an increased standard deviation in both outputs $\Delta\gamma_{\text{fuel}}$ and ΔNPV. In the realm of probabilistic analysis, this marks the extent of the evaluation; no further metrics or processes are anticipated to quantify the output uncertainty. In contrast, DSTE offers additional avenues for insight into the MCS, which will be explored in the subsequent Section.

5.4.2 Evidence-Specific Metrics

Before moving on to the discussion of the CCXFs introduced in Section 5.1.1, it is important to first explain the necessary modification for $\Delta\gamma_{\text{fuel}}$. In general, CCXFs are useful for making statements about the output being higher than a certain threshold value. For example, if the CCDF of a ΔNPV value in this case at $y = \$0$ is 0.6, it indicates a 60 % chance that the HLFC aircraft is economically better. However, when discussing the change in fuel consumption, using the CDF is more logical, since lower (or more negative) $\Delta\gamma_{\text{fuel}}$ values are preferred over higher values, unlike the case with ΔNPV. Therefore, instead of using the CCXFs for $\Delta\gamma_{\text{fuel}}$, it is advisable to use the non-complementary versions[10]. To implement this, the following points should be considered: Firstly, the Belief in a statement $\mathcal{S}$ (e.g., $\mathcal{S} = y > y^*$) and the Plausibility of the opposite statement $\mathcal{S}_c$ (e.g., $\mathcal{S}_c = y \leq y^*$) should add up to one, as mentioned by Helton et al. [101, p. 42]. This principle is represented using the threshold notation, which is:

$$\underbrace{\text{Bel}(y > y^*)}_{\text{ordinate of CCBF}} + \underbrace{\text{Pl}(y \leq y^*)}_{\text{ordinate of CPF}} = 1 \quad \text{and} \quad \underbrace{\text{Bel}(y \leq y^*)}_{\text{ordinate of CBF}} + \underbrace{\text{Pl}(y > y^*)}_{\text{ordinate of CCPF}} = 1. \tag{5.12}$$

[10] Therefore, the references to the (complementary) cumulative functions is abbreviated as (C)CXF (when referring to all) or (C)CBF, (C)CDF, and (C)CPF (when referring to a particular cumulative function).

Using the definitions of the evidence-theoretic metrics (see Helton et al. [101, p. 43]), it becomes clear that the ordinates of CBF and CPF can be calculated a posteriori from the ordinates of the CCBF and CCPF[11]:

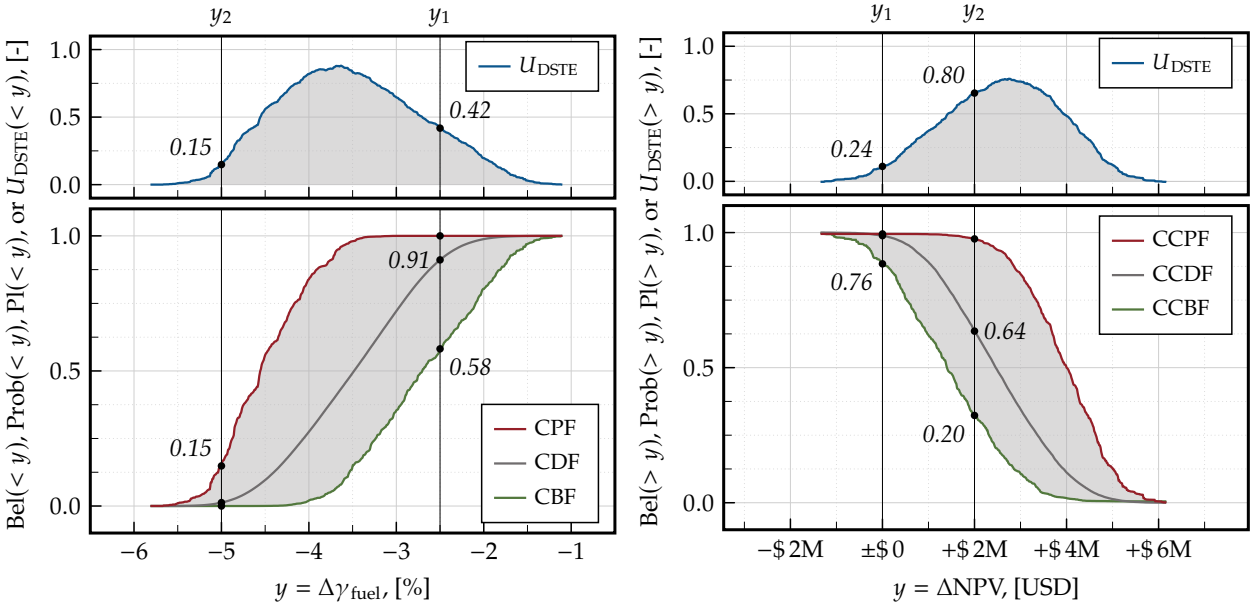
$$\text{CBF} = 1 - \text{CCPF} \quad \text{and} \quad \text{CPF} = 1 - \text{CCBF}. \tag{5.13}$$

With this modification for $\Delta\gamma_{\text{fuel}}$, the DSTE-based results shown in Fig. 5.13 can be discussed. Starting with ΔNPV, the CCBF for a given value for y^*, e.g., $y_1 = \$\,0$ (which represents the general threshold of success), is 76%, which can be interpreted as follows:

> *There is a 76% Belief in the statement that the HLFC aircraft will be economically superior to its counterpart.*

Simultaneously, the CCPF of y_1 is equal to 1, which states that:

> *There is no evidence that contradicts that the HLFC aircraft is economically superior to its counterpart.*



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 5.13 (C)CXFs for the DSTE-based campaign.

In other words, it is completely plausible for the HLFC aircraft to be economically superior. An easier and more recommended interpretation for the CCBF and CCPF uses them in combination with the CCDF, i.e.,:

> *The probability of the HLFC aircraft being economically superior to its turbulent counterpart is 100 % (taken from the CCDF). This probability has an upper bound of 100 % (taken from the CCPF) and a lower bound of 76 % (taken from the CCBF).*

The fact that the upper and lower bounds are relatively close indicates little epistemic-based uncertainty in this statement. If, for example, another threshold for y^* is chosen, e.g., $y_2 = \$\,2\text{M}$, the interpretation would be as follows: The probability of the HLFC aircraft being economically superior by at least $\$\,2\text{M}$ is 64 %. This probability is has a lower and upper bound of 20 % and 98 %. A statement for this threshold is thus highly uncertain (when using DSTE's interpretation of uncertainty, i.e., the difference between Plausibility and Belief). It follows that DSTE's metrics

[11] For brevity purposes, Eqn. (5.13) shows ordinate vectors, only. The abscissa remain the same, i.e., y_i in Eqn. (5.6) and (5.7).

assign an uncertainty to a probability, rather than treating the probability itself as the only measure of uncertainty. This evidence-theory-based uncertainty

$$U_{\mathrm{DSTE}}(> y) = \mathrm{CCPF}(y) - \mathrm{CCBF}(y) \tag{5.14}$$

can be plotted over the output space, as it is done in the upper plot of Fig. 5.13, showing a lower value of 0.24 for the statement regarding y_1 and a higher value of 0.80 for y_2. The overall progression shows that statements on the likelihood of exceeding thresholds between \$1.8M and \$4M are highly uncertain (i.e., have an uncertainty of $U_{\mathrm{DSTE}} > 0.5$) due to how the lack of knowledge in the inputs (i.e., expert statements) are distributed.

For the change in fuel consumption $\Delta\gamma_{\mathrm{fuel}}$, the modified CCXFs allow constructing similar output statements. That is, the likeliness of the HLFC aircraft's fuel consumption being $y_1 = -2.5\,\%$ or better (i.e., less) is 91 % (taken from the CCDF), with lower and upper bounds of 58 % and 100 %, respectively. The likelihood to meet the alternative threshold of $y_2 = -5\,\%$ is 0 % and is bounded by 0 % and 15 %. The upper plot shows that statements between $-4.5\,\%$ and $-2.7\,\%$ are significantly uncertain (i.e., $U_{\mathrm{DSTE}} > 0.5$).

5.4.3 Results with Reduced Uncertainty

In addition to the MCS with the DSTE-based input space described before, a second set of MCSs was performed. Here, the epistemic uncertainty in $\mathcal{D}_1$ (being the overall drag reduction potential) is reduced. This hypothetical uncertainty reduction could represent a more advanced state of knowledge, e.g., obtained through more sophisticated drag analyses. The intention of this MCS is to understand the effect of the uncertainty reduction on DSTE's uncertainty metrics. Therefore, it not only aids in the understanding of the DSTE approach, but also serves as a verification attempt.

Recall that $\mathcal{D}_1$ was constructed by obtaining intervals and BBAs from SMEs about how much HLFC could reduce the overall drag when applied to (a) the wing upper side, (b) both sides of both HTPs, and (c) both sides of the VTP. For the reduced uncertainty scenario, it is assumed that the knowledge about the drag reduction potential for the wing upper side (a) is significantly improved. More specifically, it is assumed that it is now fixed to 5.2 % (which was the average value when sampling from the Belief space in Fig. 5.8). Essentially, there are three ways to implement this change for the MCS:

Case 1) Keep the number of input statements constant. As it will be described in more detail when discussing the convergence behavior later, the number of statements affects the total number of evidence interval combinations. Furthermore, each provided statement affects the general density functions, as - by definition - the BBAs of each evidence have to sum to 1. Increasing or reducing the number of statements thus scales the other evidences. Therefore, keeping the number of statements fixed is assumed to facilitate comparability to previous analyses. To incorporate the uncertainty reduction in this approach, each of the six responses for $\mathcal{D}_1$ was hypothetically adjusted to the average value with an infinitesimal interval around it (using the dirac-delta function). The total number of responses, however, is not changed here.

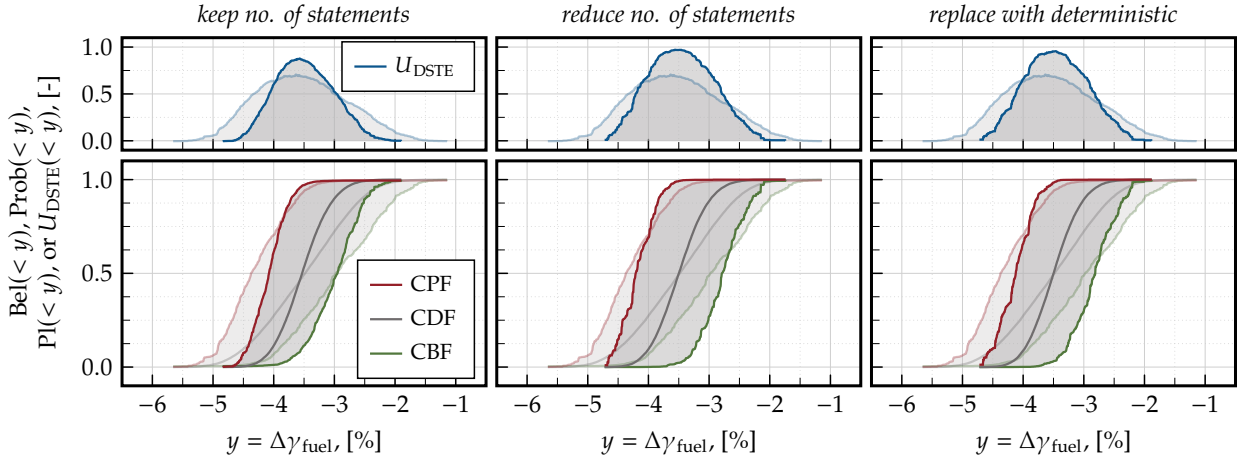
Case 2) Reducing the number of statements. This approach represents a more intuitive incorporation of uncertainty reduction. That is, only one statement with a single value of 5.2 % is used, instead of six as in case 1. This reduces the number of statements and consequently the number of evidence interval combinations drastically, but can alter the overall Belief and Plausibility distribution due to the aforementioned repercussions.

Case 3) Replacing the evidence treatment of the parameter with a deterministic value. Here, the drag reduction potential of the upper wing is not treated with DSTE anymore but is

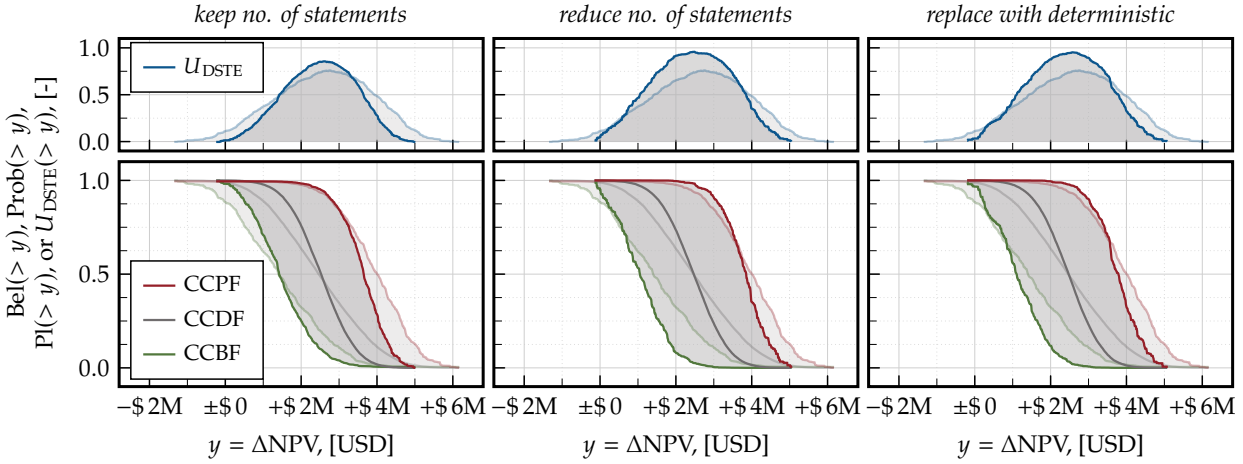
instead fixed to the single value of 5.2 %. This is similar to the previous approach but reduces the number of epistemic input parameters dealt with DSTE by one.

Each of these alternatives has been incorporated and tested. The results are shown for $\Delta\gamma_{\text{fuel}}$ and ΔNPV in Fig. 5.14 (a) and 5.14 (b), respectively. In each subplot, the results with the original input space from the previous Section are shown with lighter colors for comparison purposes.

While all three approaches lead to a reduction in the overall spread in y (shown by the slope of the (C)CDF), the shape and progression of the (C)CBFs and (C)CPFs seem to differ depending on how the uncertainty reduction is performed. Keeping the number of statements fixed (left column in both plots of Fig. 5.14) retains the shape of the (C)CBF and (C)CPF. That is, the width of the gray area in the reduced uncertainty case at low ordinate values is slightly smaller than the width at high values, which was also the case with the initial input space in Fig. 5.13. In contrast, the other two approaches now seem to have a more constant width of the gray area throughout the ordinate. This is, however, an arguably less important observation, as this width is not used to construct any statements about the results[12]. For a more sensible comparison, the change in uncertainty can be analyzed for each approach.



(a) change in fuel consumption.



(b) change in economic superiority.

Figure 5.14 (C)CXFs for the DSTE-based campaign with reduced epistemic uncertainty.

[12] Recall that the uncertainty metrics of DSTE use a particular threshold value y^* for constructing statements, which essentially means that a value of the abscissa is followed vertically until it meets the (C)CXF curves. Thus, the mentioned width of the gray area merely indicates which y values have the same Belief and Plausibility values, i.e., are not used for output statement construction.

Analyzing U_{DSTE}, one observes that all three alternatives seem to shrink the width of the curve, while increasing its peak near the average value of y. This is more visible for the $\Delta\gamma_{\mathrm{fuel}}$ results due to the fact that the reduction in $\mathcal{D}_1$ generally has a larger relative effect on the fuel consumption than on the ΔNPV. The increase of the peak of U_{DSTE} near the average values of y, which may appear counter-intuitive, is more pronounced for the second and third case as it is for the first. This is due to the repercussions of each uncertainty reduction approach has on the remaining BBAs. Cases 2 and 3 suffer from a scaling effect, since, by definition, the sum of all BBAs must equal 1, ensuring that the total belief is fully distributed across all possible subsets of the frame of discernment. As a consequence, the remaining evidence intervals E_k (see Fig. 5.4) receive a higher (combined) BBA throughout the entire output space y, which also leads to the shape change of the CCXFs. This leads to scaling-based increase of U_{DSTE}. In other words, when less input evidences are given, the remaining ones will receive more attention in the CCXF calculation. The peak increase in case 1, which is more concentrated with respect to y, occurs due to the fact that the evidence intervals E_k are now shifted (and positioned near the average value of y). As these have a higher combined BBA (compared to the original analysis) due to the dirac-delta function, the CCBF and CCPF curves become steeper, increasing U_{DSTE} further. The fact that cases 2 and 3 have a higher peak increase of U_{DSTE} indicates that the scaling effect is larger than the effect of the updated evidence interval position in case 1.

To quantify this effect, the numerical integral of U_{DSTE} for each approach was calculated, which serves as a quantification of said overall uncertainty. These values are summarized in Table 5.4. Only the first reduction case (where the number of statements is kept constant) leads to a reduction in the overall uncertainty (compared to the initial results). Therefore, it is recommended to not alter the number of statements when the comparability of the results is important.

Table 5.4 Numerical integral of U_{DSTE} for the reduced epistemic uncertainty campaign.

uncertainty reduction approach	$\Delta\eta_{\mathrm{fuel}}$	**ΔNPV**
initial input space	1.36 %	$ 2.48M
keep number of statements constant	1.08 %	$ 2.12M
reduce number of statements	1.41 %	$ 2.73M
replace evidence treatment with deterministic value	1.37 %	$ 2.61M

5.4.4 Insights into DSTE's Convergence Behavior

As mentioned before, calculating the DSTE metrics requires a large number of model executions in order to provide stable and meaningful results, i.e., to converge. For the DSTE-based UQ, the convergence behavior is twofold: a) the (relative) coverage of input evidence intervals and b) the behavior of approaching a value (in this case, an array), both of which are affected by increasing the sample size as well as other factors.

The (relative) coverage refers to combined evidence intervals described illustratively in Table 5.1 and shown in Fig. 5.4. To recall, the CCBF and CCPF calculations foresee looping over the output space and retracing which input samples (and more specifically, which evidence statements) led to which output values. Subsequently, Eqns. (5.6) and (5.7) are used to consider the individual BBAs. Note that the example shown in Fig. 5.4 was based on two parameters and three statements for each, leading to $k = 3^2 = 9$ evidence intervals to loop through. In the HLFC use case, the expert and literature elicitation led to a total number of 45 statements for a total of $m = 8$ parameters. The number of combined evidence intervals, which is equal to the product of the number of statements n_i per parameter (as per the Cartesian product), can

hence be calculated to[13]:

$$k = \prod_{i=1}^{m} n_i = 376{,}320. \tag{5.15}$$

As the algorithm loops through these k intervals, it looks for corresponding y values. For a given combined interval, it may happen that there are no y values to be found. In this case, the algorithm continues to the next combined interval. If this happens often enough, the CCXFs' informative value is reduced, as their shape changes slightly with every skipped interval. For convergence, it is therefore recommended to ensure high coverage of combined intervals. Three aspects affecting this coverage can be distinguished:

1) **The sample size**. With a larger sample size, the likelihood of drawing samples that cover every combined interval increases.
2) **The number of statements**. For a given sample size, the number of input statements will influence the number of samples per interval. This is typically difficult to influence, as it is a result of expert elicitation.
3) **The expert's BBA**. Considering a parameter for which multiple statements are given, a strong imbalance of the BBA can lead to a distribution in which the interval with a lower Belief is not sampled from at all.

To shed some light on the first aspect, a convergence analysis has been performed, where the sampling size was varied from $N = 10^4$ to $N = 10^6$. In addition, each sample size was evaluated with six different sampling schemes, as these may have a substantial effect on the representativeness of the drawn samples[14]. For each combination of sample size and sampling scheme, the number of skipped intervals was recorded and plotted in Fig. 5.15. All schemes show a quick reduction of skipped intervals with increasing sample size, reaching near-zero values at $N > 10^5$. For sample sizes smaller than this, the hammersley scheme seems to perform the best, followed by the halton method, whereas the conventional random sampling showed the weakest performance.

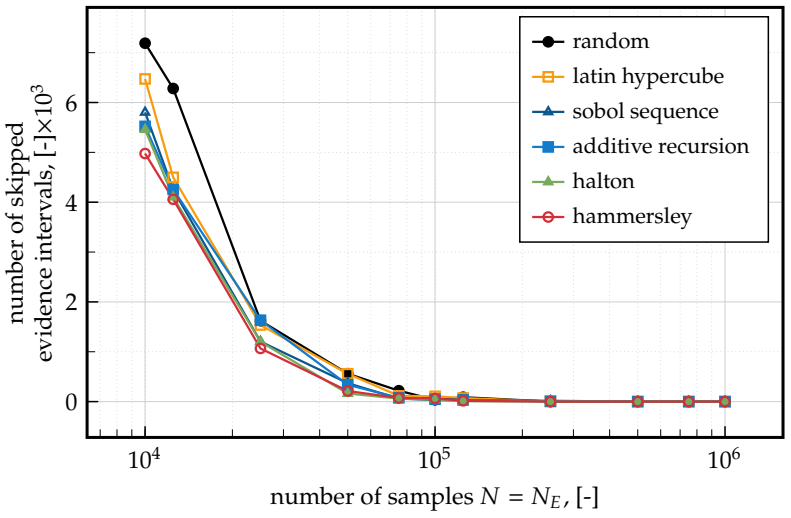
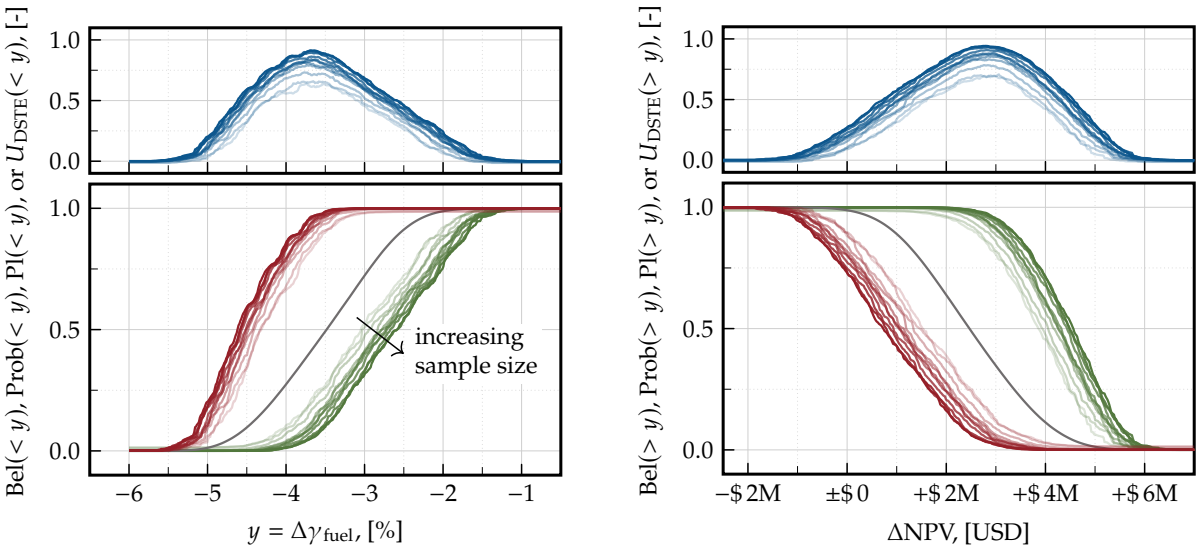


Figure 5.15 Coverage of combined evidence intervals with increasing sample size for different quasi-random sampling schemes.

Approaching a final array refers to what is usually meant by convergence. That is, increasing the sample size further does not lead to a significant change in the quantity of interest. Hereby, it is up to the user to define an acceptable level of accuracy. With the DSTE metrics being arrays (as opposed to scalar values), the accuracy threshold definition proves to be more difficult. For this study, the visual development of the (C)CBF and (C)CPF with increasing sample size has been used to investigate the convergence behavior. This is shown in Fig. 5.16, with the distinction of $\Delta\gamma_{\text{fuel}}$ in subplot (a) and ΔNPV in subplot (b). The increasing opacity of the

[13] This pertains to the input space without a hypothetical reduction in input uncertainty. For the cases with reduced uncertainty mentioned in the previous Section, the number of combined evidence intervals is 376,320 (case 1) and 62,720 (cases 2 and 3).
[14] Supplemental descriptions of the used sampling schemes can be found in Appendix E.3.

curves indicates an increasing sample size (the same as those in Fig. 5.15, i.e., from 10^4 to 10^6). For both, the (C)CDF arrays seem to have converged quite early. The (C)CBF and (C)CPF, however, vary significantly with increasing sample size. Even when N exceeds 10^5 (which was the aforementioned threshold for having nearly no skipped evidence intervals), the (C)CBF and (C)CPF seem to drift further apart with each increasing sample size, which leads to an increasing uncertainty curve u_{DSTE}. This is a flaw of the DSTE-based MCS and can be traced back to the numerical implementation of the definitions of the CCBF and CCPF in Eqns. (5.6) and (5.7), respectively. Because the minimum and maximum of the subset of the output y are numerically obtained, increasing the sample size increases the change of sampling inputs that are closer to the true minimum or maximum. While this behavior is asymptotic by nature, the sample sizes investigated here were not big enough to yield a (visually) stable set of CCXFs, i.e., one that did not change compared to the previous, smaller sample size.



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 5.16 Progression of (C)CXFs for increasing sample size N.

5.5 Conclusions

This Chapter has explored the potential of DSTE as a complementary or alternative approach to probabilistic UQ. Theoretical and practical aspects of the DSTE approach were discussed, followed by the implementation of expert interviews and the subsequent analysis of the obtained results. A modified MCS was employed to calculate uncertainty metrics using samples drawn from the interviews, and a comparison was made between the evidence-theoretic approach and the probabilistic counterpart. Furthermore, the verification and convergence of the evidence-theoretic approach were investigated.

The comparison of output distributions showed that the HLFC aircraft had a better performance when DSTE was used than when the probabilistic UQ approach was used. However, it is worth noting that the uncertainty, when measured by the standard deviation, increased. This indicates that while DSTE can enhance the UQ quality (by removing personal subjectivity and utilizing expert knowledge), it may introduce a higher level of uncertainty in the results. Despite this limitation, the evidence-theoretic approach demonstrated its value in constructing output statements using CCXFs. Although interpreting these statements can be challenging, they prove to be useful when assessing the probability of the output exceeding or falling below a particular threshold value.

The verification attempt revealed certain difficulties associated with the implementation of uncertainty reduction strategies. It is recommended to maintain a fixed number of statements to ensure comparability between different approaches. Additionally, the convergence analysis highlighted that the evidence-based approach necessitates a large number of samples to obtain a stable set of CCXFs. This finding emphasizes the computational requirements and potential limitations of DSTE in practical applications. It was also the reason for using a PCE-based surrogate of the lifecycle simulation framework.

One notable limitation of this approach is the difficulty of interpreting CCXFs, particularly for inexperienced users. While the concept of lower and upper bounds of probability provides some assistance, understanding uncertainty from the perspective of DSTE remains challenging, especially considering the modified MCS and its impact on U_{DSTE}. Consequently, it is essential for users to have a strong theoretical understanding of the method to avoid potential confusion and misinterpretation of results.

Although a Python package `dste` has been developed to facilitate the use of DSTE, further research and development are still needed. Specifically, attention should be given to addressing issues related to convergence. Runtime enhancements to the package, as well as additional studies and validation efforts, are necessary to enhance its usability and effectiveness.

In conclusion, this Chapter has demonstrated the potential of DSTE as an alternative or complementary approach to probabilistic uncertainty quantification, partially verifying the third research hypothesis. The comparison of output distributions revealed improved performance when employing DSTE, albeit with an increase in uncertainty. The use of CCXFs provided valuable insights into the probability of output exceeding or falling below certain thresholds. However, challenges in interpretation and the requirement for a large number of samples for convergence were noteworthy limitations. Nonetheless, with further research, development, and user education, DSTE holds promise as a valuable tool in UQ and decision-making processes.

6 Probabilistic and Non-Probabilistic Combination

This Chapter addresses the third and final research question, focusing on the integration of epistemic and aleatory uncertainties in TEA. The guiding hypothesis posits that by systematically differentiating between these uncertainties using evidence-theoretic methods and nested MCS, enhanced interpretability and actionable insights can be achieved. This approach is anticipated to facilitate more informed resource allocation and risk mitigation strategies, identifying dominant uncertainty types to guide strategic decisions on research investments or robust product designs.

To reiterate, when high epistemic uncertainties prevail in a TEA, enhancing internal knowledge through increased research and development can effectively reduce output uncertainty. Conversely, if aleatory uncertainties dominate, representing inherent and product-specific risks, further investment in knowledge acquisition is unlikely to be effective. Instead, efforts should focus on developing robust product designs, especially when economic stakes are high.

The Chapter is structured as follows: Section 6.1 outlines the approach and methodology, presenting the comparative approach of the study. Section 6.2 introduces the first of two combination methods, namely a DSTE-based approach where evidence-based uncertainty metrics (as discussed in the previous Chapter) are complemented by aleatory variation. In Section 6.3, a refined modification is presented that incorporates selected techniques from DSTE but utilizes UQ metrics originating from the probabilistic domain. Finally, Section 6.4 concludes the chapter by highlighting the identified advantages and limitations of each approach and summarizing the insights gained, particularly regarding the interpretation of different uncertainties. Note that, similar to the previous study, all analyses in this Chapter have been performed with the same PCE surrogate as in Chapter 5.

6.1 Methodology and Preparation

This Section begins with a description of the overall study design and continues with a brief description of the selected input uncertainties.

Study Design

To address the hypothesis, this study applies an evidence-based approach and compares it to a refined alternative in terms of uncertainty interpretation, implementation difficulty, execution speed, and other relevant factors. Identical boundary conditions are maintained to ensure a fair comparison. The methodology is depicted in Fig. 6.1, with the evidence-based approach highlighted in the top half, the refined approach at the bottom, and common steps in the center.

Initially, uncertain parameters from both epistemic and aleatory domains are selected and prepared for sampling. For aleatory uncertainties, the PDFs derived in Chapter 4 are utilized. Epistemic uncertainties are sampled using the Belief space obtained from expert elicitation and literature review discussed in Chapter 5. In the second step, these samples are fed into the simulation model, distinguishing between epistemic and aleatory uncertainties. For the evidence-based approach (top part of Fig.6.1), the calculation of DSTE-based uncertainty metrics is identical to the approach of Chapter 5, except for one modification: Instead of one final CCBF, CCPF, and CCDF, a collection of these functions is calculated using a nested MCS. Each set of CCXFs thus has its own (aleatory) uncertainty, detailed further in Section6.2. The

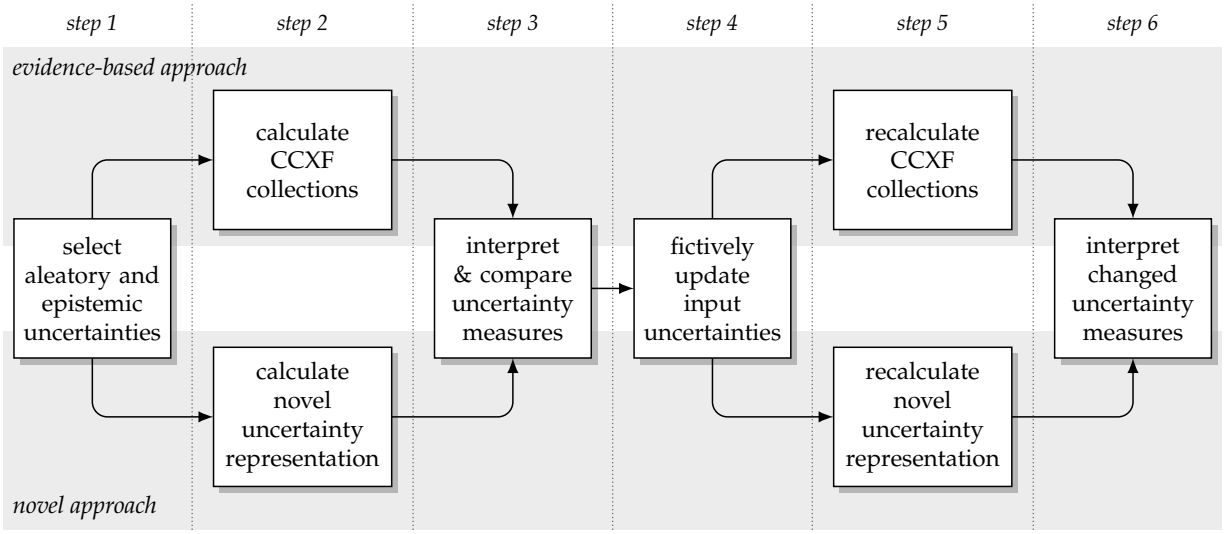


Figure 6.1 Methodology for the comparative uncertainty combination study.

refined approach (bottom part of Fig.6.1) similarly employs a nested MCS setup. The inner loop handles aleatory (or epistemic) samples, while the outer loop uses epistemic (or aleatory) samples. A more detailed description is provided in Section6.3. In the third step, results are converted to represent the uncertainty in an easily interpretable manner and compared to the uncertainty interpretation from DSTE. This step considers qualitative aspects, such as the ease of deriving investment-relevant uncertainty statements, as well as quantitative measures, including convergence behavior with increasing sample size. Steps four to six repeat the analyses with a hypothetically updated input uncertainty to verify and stress-test each approach and to analyze changes in the different output uncertainty measures.

Input Uncertainty Selection

The input uncertainties for the first study are summarized in Table 6.1. In total, eight of the eleven uncertainties have been selected to be varied. The parameters $\mathcal{O}_2$, $\mathcal{O}_3$, and $\mathcal{E}_1$ have been fixed to their average value since they had the lowest contribution in the GSA calculated in Chapter 4. One might argue that $\mathcal{M}_2$ and $\mathcal{M}_3$ also had a negligible contribution to the output uncertainty (see Fig. 4.19 on page 96). However, it should be noted that the GSA results

Table 6.1 Summary of uncertainties dealt with the combinatory UQ.

domain			**general information**			**theory**	
parameter	unit	type*	distribution	avg.	std.	DSTE	ProT
design							
$\mathcal{D}_1$: drag reduction	[%]	E	non-parametric	8.1	1.2	●	○
$\mathcal{D}_2$: mass increase	[kg]	E	non-parametric	750	194	●	○
$\mathcal{D}_3$: SFC penalty	[%]	E	non-parametric	1.13	0.22	●	○
operation							
$\mathcal{O}_1$: cruise mach	[-]	A	gen. normal	0.82	0.013	○	●
$\mathcal{O}_2$: extra cont. fuel†	[-]	E	*— fixed to 0.50 —*			○	○
$\mathcal{O}_3$: load factor	[-]	E	*— fixed to 0.80 —*			○	○
environment							
$\mathcal{E}_1$: insect contamin.‡	[-]	E	*— fixed to 0.50 —*			○	○
$\mathcal{E}_2$: cloud encounter	[-]	A	exponential	0.10	0.10	○	●
economic							
$\mathcal{M}_1$: fuel price§	[-]	A	skewnormal	0.46	0.26	○	●
$\mathcal{M}_2$: maint. increase	[USD/yr]	E	non-parametric	$79k	$21k	●	○
$\mathcal{M}_3$: price increase	[USD]	E	non-parametric	$1.3M	$405k	●	○

*E: epistemic, A: aleatory; †see Eqn. (4.22); ‡see Eqn. (4.23); §see Eqn. (4.27).

for these parameters were derived using ProT, whereas in this Chapter, their uncertainty is modeled with DSTE as discussed in Section 5.3.

Similarly to the previous analyses, the inputs are assumed to be independent from one another, i.e., there is no need for copula-based or any other correlation-considering approaches.

6.2 DSTE-Based Method for Combined UQ

To reiterate, the concept of capturing output uncertainty in Evidence theory revolves around the concepts of Belief and Plausibility. This allows for the calculation of CCBF and CCPF, which define a lower and upper probability for a threshold-based statement. These numerically obtained cumulative functions are designed to handle epistemic uncertainty but usually either neglect or substantially simplify the aleatory uncertainties. The goal of this study is to examine both types of uncertainties and their simultaneous impact on the output. To this end, a nested and DSTE-based MCS, as illustrated in Fig. 6.2, was developed.

First, a sample set for the aleatory uncertainties is created; that is, a value is determined for each of the three aleatory uncertain variables using their respective PDF. Subsequently, each of the five epistemic uncertainties is assigned a value sampled from their Belief space. With this complete sample set, the LYFE surrogate is executed once, yielding one output value for $\Delta\gamma_{\text{fuel}}$ and one for ΔNPV (i.e., y_{e_i,a_j}, where e_i and a_j represent current iteration count in the epistemic and aleatory loop, respectively). Next, a new set of epistemic samples is generated and fed to the LYFE surrogate, while the aleatory sample set remains as is, producing a new output value (e.g., y_{e_{i+1},a_j}). Once this inner loop (or epistemic loop) is completed, which is indicated by the respective count e_i exceeding the limit N_E, one column of the overall output matrix y is filled, see Eqn. (6.1). This resets the counter in the epistemic domain and initiates a new aleatory iteration until $a_j > N_A$.

$$y = \begin{bmatrix} y_{e_1,a_1} & y_{e_1,a_2} & \cdots & y_{e_1,N_A} \\ y_{e_2,a_1} & y_{e_2,a_2} & \cdots & y_{e_2,N_A} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N_E,a_1} & y_{N_E,a_2} & \cdots & y_{N_E,N_A} \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \tag{6.1}$$

$$\Rightarrow y_{\text{CCBF}} = \begin{bmatrix} \text{CCBF}_{a_1} & \text{CCBF}_{a_2} & \cdots & \text{CCBF}_{N_A} \end{bmatrix} \tag{6.2}$$

$$\Rightarrow y_{\text{CCDF}} = \begin{bmatrix} \text{CCDF}_{a_1} & \text{CCDF}_{a_2} & \cdots & \text{CCDF}_{N_A} \end{bmatrix} \tag{6.3}$$

$$\Rightarrow y_{\text{CCPF}} = \begin{bmatrix} \text{CCPF}_{a_1} & \text{CCPF}_{a_2} & \cdots & \text{CCPF}_{N_A} \end{bmatrix} \tag{6.4}$$

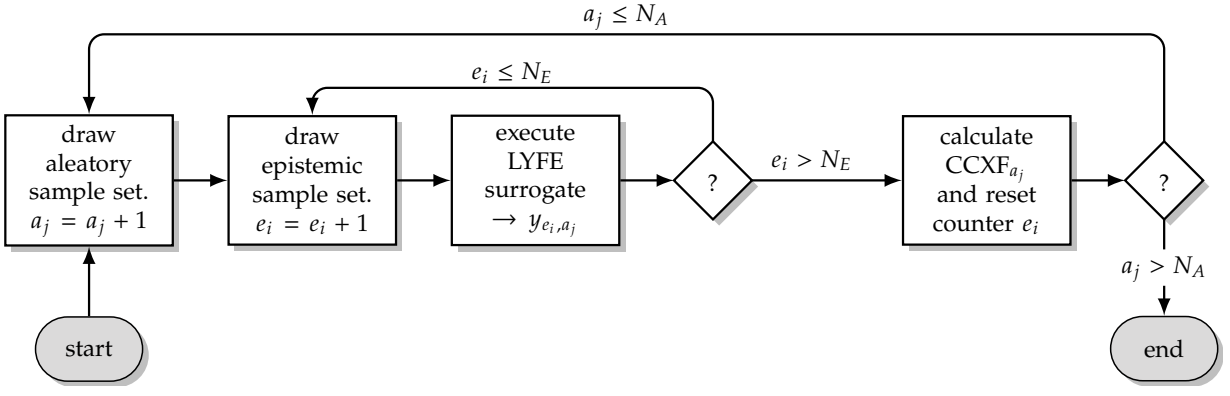
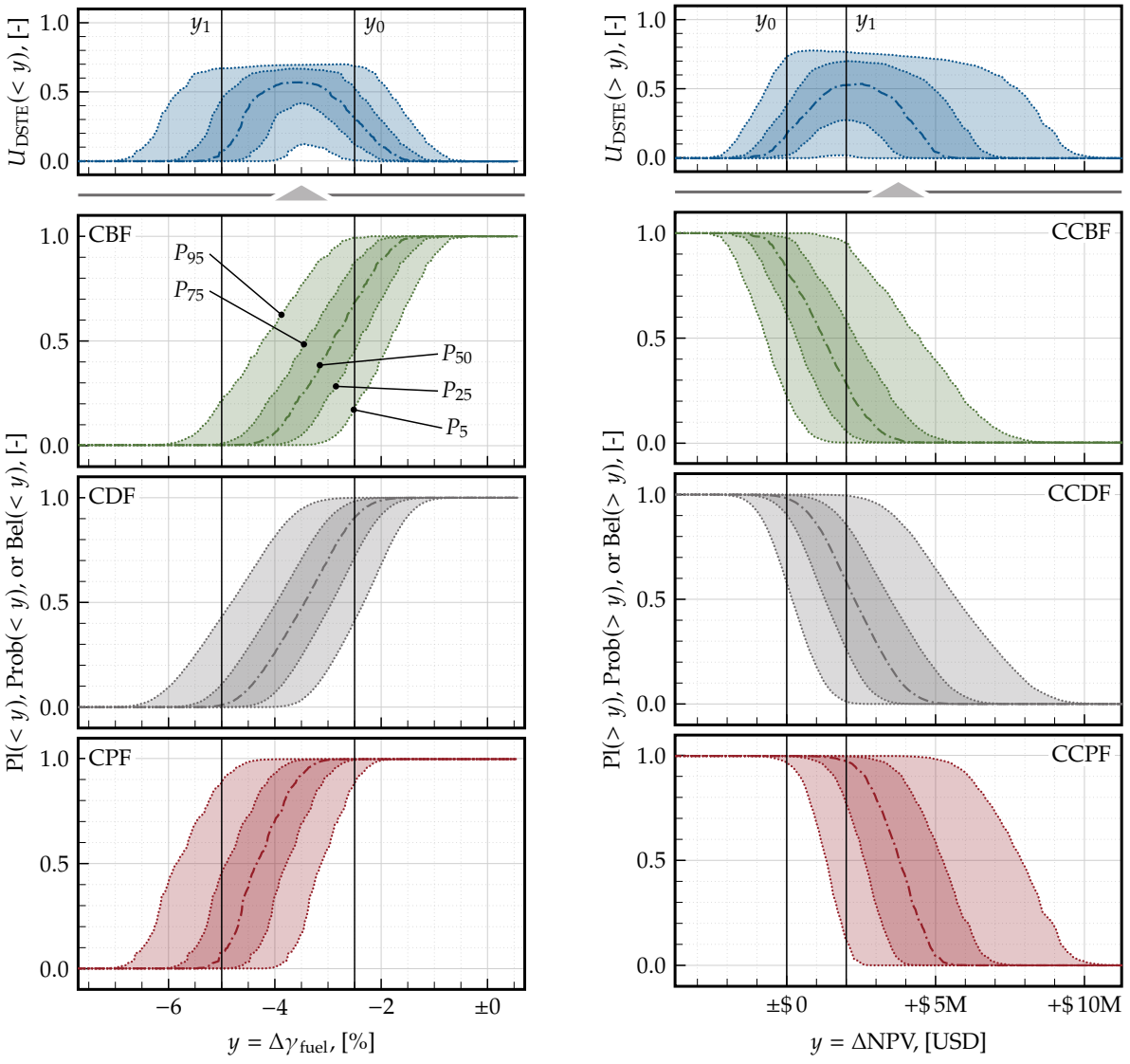


Figure 6.2 Flowchart for evidence-based uncertainty combination.

In other words, the overall output y holds the model's response to aleatory variations when viewed from column to column while the change from row to row represents the response to epistemic variations. Consequently, the first column of y contains the necessary data to calculate one CCBF, one CCDF, and one CCPF, all sharing an identical aleatory sample set. Each subsequent column of y is formulated with a new batch of aleatory samples, while retaining the exact same sequence of epistemic samples per row. Repeating the calculation of $\text{CCXF}_{a_j} \forall a_j$ results in the output space spanning matrices shown in Eqns. (6.2-6.4)[1].

6.2.1 Results

The first set of outputs was obtained using a total of 100M executions of the LYFE surrogate, resulting from $N_A = N_E = 10{,}000$ aleatory and epistemic samples. The overall results are shown in Fig. 6.3 (a) and Fig. 6.3 (b) for the $\Delta\gamma_{\text{fuel}}$ and ΔNPV, respectively. Here, the bottom three plots represent the possibility (red), density (gray), and belief curves (green). Since these are now function *collectives*, they are depicted using percentiles across the aleatory domain,



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 6.3 Overall result of the DSTE-based combinatory UQ showing the (C)CBF (green), (C)CDF (gray), (C)CPF (red), and the uncertainty (blue), each depicted with the percentiles P_i for $i \in [5, 25, 50, 75, 95]$.

[1]Supplemental information on this matrix operation can be found in Appendix F.1.

with the median depicted as a dashed line. The top plot shows the percentiles of the evidence-theoretic uncertainty U_{DSTE}. Each aleatory run leads to one uncertainty curve , which is calculated as

$$U_{\mathrm{DSTE}} = \begin{cases} \mathrm{CPF}\left(y_{a,\forall e}\right) - \mathrm{CBF}\left(y_{a,\forall e}\right) & \text{for } \Delta\gamma_{\mathrm{fuel}} \\ \mathrm{CCPF}\left(y_{a,\forall e}\right) - \mathrm{CCBF}\left(y_{a,\forall e}\right) & \text{for } \Delta\mathrm{NPV} \end{cases} \quad \forall a \in [1, 2, \ldots, N_A]\,. \tag{6.5}$$

To get started with the complex interpretation of this result depiction, it is helpful to focus on the median curves first and neglect the aleatory impact for now. For $\Delta\gamma_{\mathrm{fuel}}$ the median uncertainty U_{DSTE} is non-zero for output values between $-1\,\%$ and $-5\,\%$ and highest in the interval between $-3\,\%$ and $-4\,\%$, indicating that statements in these regions are particularly uncertain. For the economic KPI, the median U_{DSTE} is nonzero for values between $-\$\,1\mathrm{M}$ and $+\$\,5\mathrm{M}$ and highest between $+\$\,1\mathrm{M}$ and $+\$\,3.5\mathrm{M}$.

The aleatory impact can be observed by the relatively large spread of the (C)CBF, (C)CDF, and (C)CPF percentiles in both KPIs. This spread is in fact so large that it leads to significant overlapping of the (C)CBFs with the (C)CPFs and vice versa[2]. For $\Delta\gamma_{\mathrm{fuel}}$, the aleatory spread in the CDF seems to be relatively symmetrical to the median throughout the majority of the output range. In contrast, the aleatory spread in the CCDF of ΔNPV increases slightly with higher values of y. This is likely due to the effect of the fuel price uncertainty, which has a larger impact on the economic superiority when the epistemic parameter combination is more favorable for the HLFC aircraft. Consider, for example, an epistemic sample set that leads to high fuel savings (i.e., high drag reduction potential, low mass increase, and low SFC penalty). These fuel savings are incorporated in the ΔNPV through the sampled fuel price. As this parameter is modeled in a non-symmetrical manner (i.e., skewed normal distribution), higher ΔNPV values show a larger aleatory spread than lower ΔNPV values. Because the non-symmetrical fuel price does not affect $\Delta\gamma_{\mathrm{fuel}}$, the aleatory spread for the change in fuel consumption remains mostly symmetrical throughout the output space.

In order to derive more tangible statements out of these results, it is useful to select specific y values of interest and subsequently calculate the uncertainties. In line with the values of Chapter 5 (see Fig. 5.13), the selected thresholds are $y_1 = -2.5\,\%$ and $y_2 = -5\,\%$ and $y_1 = \pm\$\,0$ and $y_2 = +\$\,2\mathrm{M}$ for $\Delta\gamma_{\mathrm{fuel}}$ and ΔNPV, respectively. The results are shown in Fig. 6.4 (a) for the change in fuel consumption and Fig. 6.4 (b) for the change in economic superiority using boxplots. Starting with the change in fuel consumption and y_0, the statement could be as follows:

> *The probability of the fuel consumption improvement being greater than* $y^* = 2.5\,\%$ *is* $P_{med} = 0.90^{0.98}_{0.71}$ *and has a lower epistemic bound of* $P_{low} = 0.82^{0.95}_{0.47}$ *and an upper epistemic bound of* $P_{high} = 1$.

Here, the following notation for $P = x^b_a$ is used: x refers to the median value, a refers to the lower quartile and b to the upper quartile of the respective function. P_{med} represents the CDF, while P_{low} and P_{high} refer to the CBF and CPF, respectively. The range between a and b indicates the aleatory uncertainty, whereas the range between P_{low} and P_{high} could be interpreted as the epistemic uncertainty. It should be noted that the choice of a and b is not standardized and could equally be any other percentile. Alternatively, the mean value $\pm$ standard deviation could be used to construct the statements. This would, however, neglect potential asymmetries and is thus not recommended. The corresponding statement for y_1 is:

> *The probability of the HLFC aircraft to have a fuel consumption improvement of* $y^* = 5\,\%$ *or better is* $P_{med} = 0.01^{0.1}_{0}$. *This probability is bounded by* $P_{low} = 0$ *and* $P_{high} = 0.13^{0.45}_{0}$.

[2]Which is why each (C)CXF is shown in its own subplot in Fig. 6.3

A more detailed look into the boxplots in Fig. 6.4 (a) shows that the medians and quartiles of U_{DSTE} are somewhat similar for both thresholds y_i, which indicates that the output statements given above are comparable regarding both epistemic and aleatory uncertainty. For y_1, the aleatory uncertainty mainly affects the CBF, while the effect on the CPF is negligible. This is reversed for y_2. In other words, the aleatory parameters seem to introduce some uncertainty regarding the belief that the fuel consumption improvement *will be* better than 2.5 %, while the question of whether this threshold *can be* reached is unaffected. On the other hand, a fuel consumption of at least 5 % is unlikely to begin with, but the aleatory parameters add uncertainty to the plausibility of this statement, i.e., whether it *could* happen.

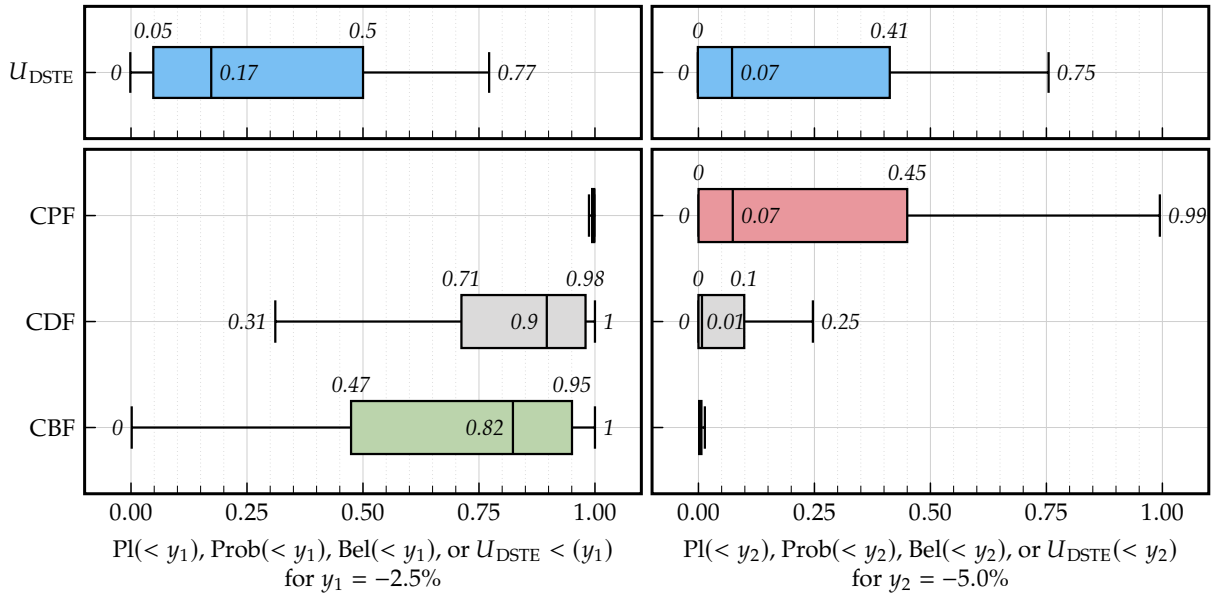
The statements regarding the economic superiority are derived similarly. Here, P_{low} refers to the CCBF and P_{high} represents the CCPF. The final statements for y_1 and y_2 are therefore:

> *The chance of the HLFC aircraft being economically superior to its turbulent counterpart is* $P_{med} = 0.98^{1}_{0.89}$. *This probability has a lower bound of* $P_{low} = 0.86^{0.96}_{0.55}$ *and an upper bound of* $P_{high} = 1$. *Considering a threshold of at least* +$2M, *the respective values are* $P_{med} = 0.59^{0.86}_{0.24}$, $P_{low} = 0.24^{0.55}_{0.03}$, *and* $P_{high} = 0.99^{1}_{0.82}$.

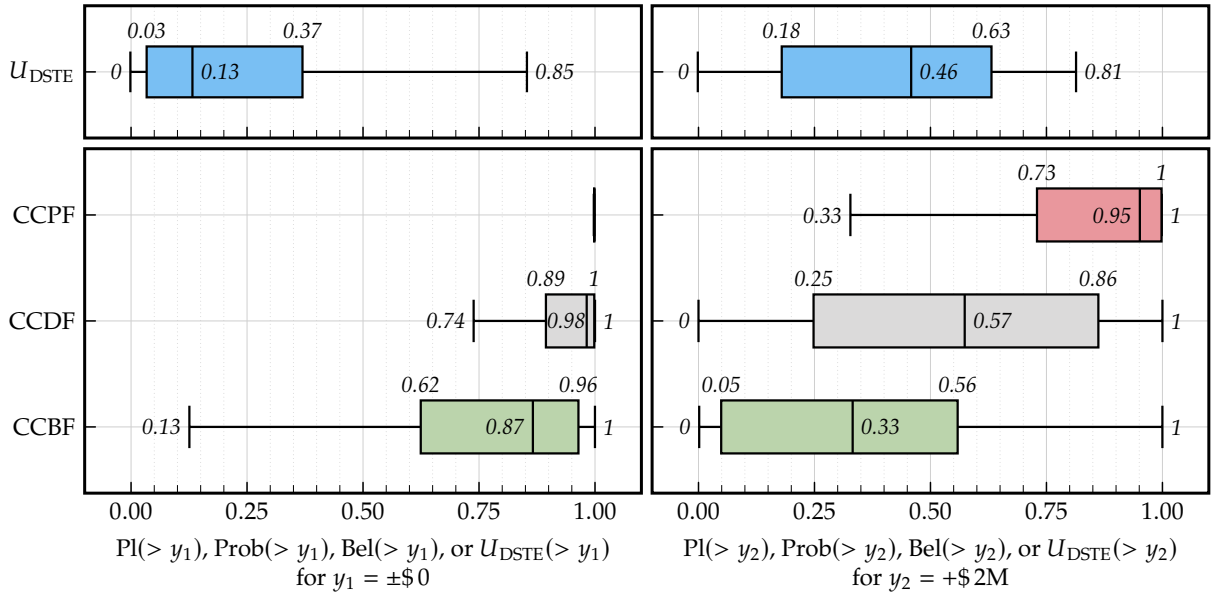
Here, it is observable that U_{DSTE} at y_1 is lower than at y_2, both in its median value as well as in the aleatory spread, indicated by the IQR. For the CCXFs of ΔNPV, the IQR is also larger at y_2 than at y_1, which likely results from the aforementioned fuel price uncertainty that has a larger impact on the economic superiority when ΔNPV values are higher.

Regarding the derivation of recommendations for future efforts with this uncertainty combination technique, it is clear that the threshold value plays a significant role, especially for ΔNPV. Considering a general superiority of the HLFC aircraft (i.e., $y^* = y_1 = \pm\$0$), where the median value for U_{DSTE} is quite low (i.e., 0.13), representing a low impact of knowledge-based uncertainty. The IQR at this threshold (i.e., $0.37 - 0.03 = 0.36$) indicates that the technology development could benefit from a more robust design. For the higher threshold in ΔNPV (i.e., $y^* = y_1 = +\$2\mathrm{M}$), the median value and IQR of U_{DSTE} are quite close (i.e., 0.46 and $0.63 - 0.18 = 0.45$, respectively). This represents a balanced impact of the different uncertainty types, i.e., investing in further research to improve the state of knowledge is equally advisable as improving the design to be more robust towards boundary conditions that are aleatory in nature.

Regarding the derivation of recommendations for future efforts with this uncertainty combination technique, it is clear that the threshold value plays a significant role, especially for ΔNPV. Considering a general superiority of the HLFC aircraft (i.e., $y \stackrel{=}{} y_1 = \pm\$0$), where the median value for U_{DSTE} is quite low (i.e., 0.13), representing a low impact of knowledge-based uncertainty. The IQR at this threshold (i.e., $0.37 - 0.03 = 0.36$) indicates that the technology development could benefit from a more robust design. For the higher threshold in ΔNPV (i.e., $y \stackrel{=}{} y_1 = +\$2\mathrm{M}$), the median value and IQR of U_{DSTE} are quite close (i.e., 0.46 and $0.63 - 0.18 = 0.45$, respectively). This balance indicates that both types of uncertainties have a similar impact, meaning that investing in further research to improve the state of knowledge is as advisable as improving the design to be more robust towards boundary conditions that are aleatory in nature. The recommendations for $\Delta\gamma_{\mathrm{fuel}}$ are derived in a similar manner but are less affected by the selected threshold value.



(a) change in fuel consumption.



(b) change in economic superiority.

Figure 6.4 Boxplot results of the DSTE-based combinatory UQ with two selected threshold values.

6.2.2 Results with Reduced Uncertainty

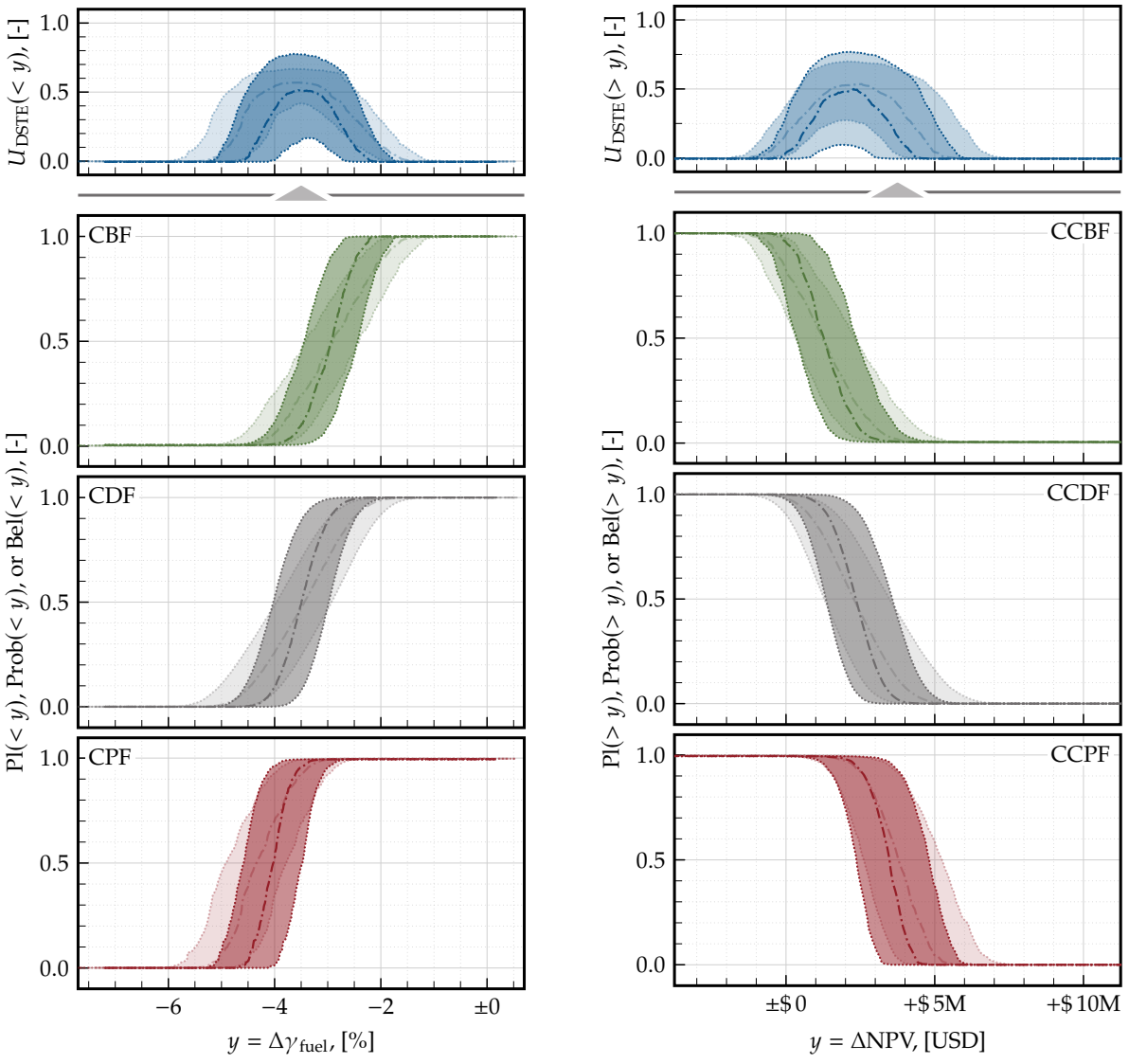
After discussing the results obtained using the original set of input uncertainties, an additional analysis is introduced. This analysis incorporates a hypothetical reduction in one selected input uncertainty, in line with the approach described in Section 5.4.3. That is, the drag reduction potential on the wing's upper side is fixed to its average value of 5.2 %[3].

With less epistemic input uncertainty, one would expect the medians of the (C)CBF and (C)CPF to move closer to each other, thereby reducing the median of the uncertainty U_{DSTE}. Furthermore, since no changes are made to uncertainties in the aleatory domain, the spread of the (C)CXFs is not expected to change. As Fig. 6.5 shows[4], this expectation holds true. The median

[3] The number of expert statements was kept constant, as the discussions on p. 117 showed that this approach allows for better comparisons between the results of the reduced and non-reduced analyses.

[4] For overview purposes, only the lower and upper quartile as well as the median are shown in Fig. 6.5.

U_{DSTE} of the modified campaign is slightly lower than that of the previous MCS across the output range of both KPIs. However, the distance between the quartiles appears to have increased for U_{DSTE}, suggesting a counter-intuitive increase in the aleatory domain. This behavior is likely associated with the issues discussed in Section 5.4.3, specifically the complexity behind the peak changes of U_{DSTE}.



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 6.5 Results of the DSTE-based combinatory uncertainty analysis with reduced epistemic uncertainty (foreground) and the initial results without reduced uncertainty (background), each depicted with the percentiles P_i for $i \in [25, 50, 75]$.

Interpreting the changes in the shape and/or progression of the CCXFs requires a deep understanding of both the theory and its numerical implementation. Upon detailed review and analysis of individual iterations from both the prior and updated MCSs, two possible (and interlinked) reasons can be identified for the lack of this DSTE-based combination approach to rigorously track uncertainty reduction efforts:

a) The lack of a proper mathematical foundation for interpreting U_{DSTE} in the presence of aleatory uncertainties poses a challenge. By definition, U_{DSTE} serves as a measure of knowledge-based uncertainty. Specifically, each U_{DSTE} curve in the collection of results measures the difference between the (C)CPF and the (C)CBF for a particular set of aleatory samples. Using percentiles to represent the effect of aleatory uncertainty makes sense for

one given TEA case but comparing percentiles over multiple TEAs (as done here) lacks a proper mathematical foundation with respect to the interpretation of U_{DSTE}. While results make sense within one TEA, the very nature of this DSTE-based uncertainty combination only allows for indicative and/or qualitative comparison when repeating a TEA with updated states of knowledge.

b) Even without aleatory uncertainties involved, the behavior of U_{DSTE} with reduced epistemic uncertainty can be counter-intuitive. As noted in Section 5.4.3, repeating an analysis with reduced epistemic uncertainty effectively reduces the width of U_{DSTE} but can lead to an increase in its peak due to the remaining BBAs being more concentrated within the remaining output space. This complicates the interpretation of such verification attempts in general. Adding the aleatory dimension further impedes the clarity of results, as it introduces an additional layer of variability that must be accounted for, potentially masking the effects of epistemic uncertainty reduction.

To conclude this verification attempt, it can be stated that while the reduction of epistemic uncertainties generally aligns with theoretical expectations, the observed increase in quartile spread for U_{DSTE} underscores the complexity of interpreting combined uncertainties. Future work should focus on developing a more robust mathematical framework for U_{DSTE} in the presence of aleatory uncertainties and refining percentile calculation methods to better capture the interplay between epistemic and aleatory domains. This will ultimately enhance the interpretability and actionable insights derived from uncertainty combination techniques.

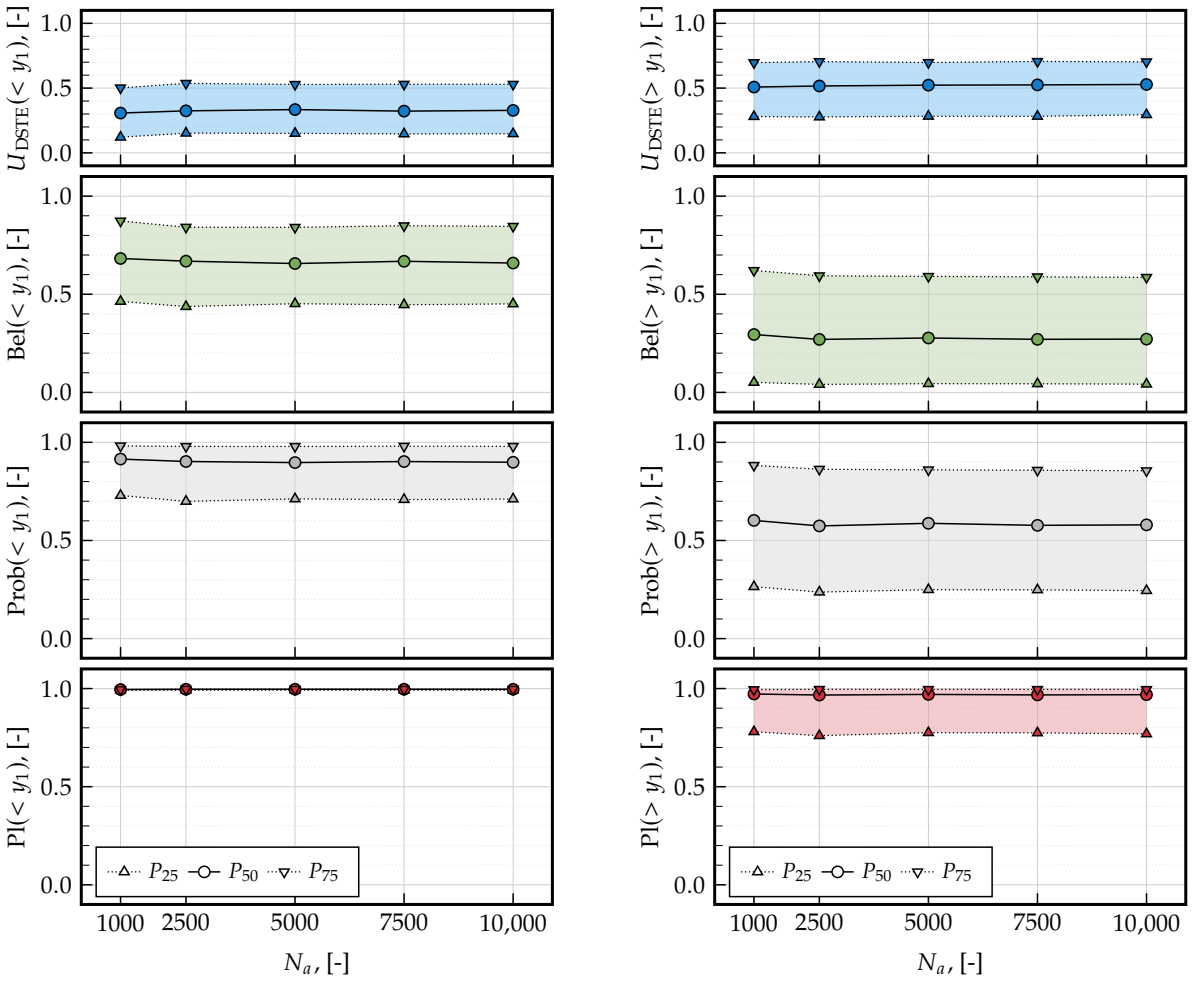
6.2.3 Convergence Behavior and Runtime Performance

Concluding the discussion on evidence-based uncertainty combination, this final Section focuses on two aspects: convergence and computational runtimes. To avoid redundancy, the convergence analysis was conducted solely in the aleatory domain. This is because the exploration of epistemic convergence has already been covered in Section 5.4.4. The results are illustrated in Fig. 6.6, with subplots (a) and (b) representing $\Delta\gamma_{\text{fuel}}$ and ΔNPV, respectively. These plots display the 25th, 50th, and 75th percentiles of the (C)CXFs, evaluated at $y_1 = -2.5\,\%$ for $\Delta\gamma_{\text{fuel}}$ and $y_1 = \pm\$\,0$ for ΔNPV. The sample size N_A was incrementally increased from 1000 to 10,000 in five steps. It should be noted that the total number of model executions is $N_{\text{tot}} = N_A \cdot N_E$, with the number of epistemic samples fixed at $N_E = 10{,}000$.

As the results show, the percentiles are almost constant throughout the range of varied sample sizes, indicating that $N_A = 2500$ is an already sufficient size for the aleatory uncertainties. This applies to both, $\Delta\gamma_{\text{fuel}}$ and ΔNPV. Only a minor change is observable from $N_A = 1000$ to $N_A = 2500$. Given that the DSTE-based uncertainty combination allows for changing N_A and N_E individually, users are recommended to utilize this capability to cut computational expenses, which can be prohibitively high as the next analysis shows.

The computational expense results, presented in Fig. 6.7, were initially obtained on a 60-core workstation using parallel computing. To make these results more universally applicable, they were adjusted to a single-core basis by dividing the runtimes by 60. This adjustment aims to present the computational efficiency in a manner that is representative of environments with varying computational resources. Additionally, the reported computational runtimes have been adjusted to exclude the model execution time for further comparability, since both of the uncertainty combination techniques can be used with any model utilizing discrete inputs and outputs.

As the convergence discussion of Section 5.4.4 showed, a very large epistemic sample size N_E is needed in the present case study to avoid a distortion of the CCXFs due to skipped intervals. Combining this, even with a small aleatory sample size, results in significant computational runtimes. Considering, for example, $N_A = 5000$ and $N_E = 10{,}000$ resulting in $\approx 234\,\text{h}$ when



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 6.6 Progression of percentiles P_{25}, P_{50}, and P_{75} of the DSTE metrics with increasing sample size and one fixed threshold value.

executed sequentially, the need for parallel execution becomes evident. The results from the previous sections, which were obtained with $N_A = N_E = 10{,}000$ samples took eight hours on the workstation to compute in parallel, whereas an estimated sequential runtime would take ≈20 days.

6.3 Novel Method for Combined Uncertainty Analysis

Similar to the evidence-based combinatory method, the novel technique also employs an inner epistemic loop and an outer aleatory loop, as depicted in Fig. 6.8. The method is inspired by the second-order probability design suggested by some authors, e.g., Dewey et al. [63] from Northrop Grumman. Initially, an aleatory sample set is drawn using the PDFs of the aleatory uncertain input parameters. The first of two novelties of this approach involves drawing epistemic samples according to their Belief space, thereby combining ProT and DSTE. This complete set of input parameters is then used as input to the LYFE surrogate model, yielding individual output values for $\Delta\gamma_{\mathrm{fuel}}$ and ΔNPV. Subsequently, a new set of epistemic samples is drawn while keeping the aleatory samples constant. The inner loop continues until a predetermined convergence criterion, represented by a fixed number of executions N in this analysis, is met. If $e_i > N$, the algorithm proceeds to the next iteration of the aleatory loop, drawing a new aleatory sample set and resetting e_i to zero. The outer loop also terminates when the number of executions n_a exceeds N. Unlike the purely evidence-based method, this approach mandates that the convergence thresholds N for both the epistemic and aleatory

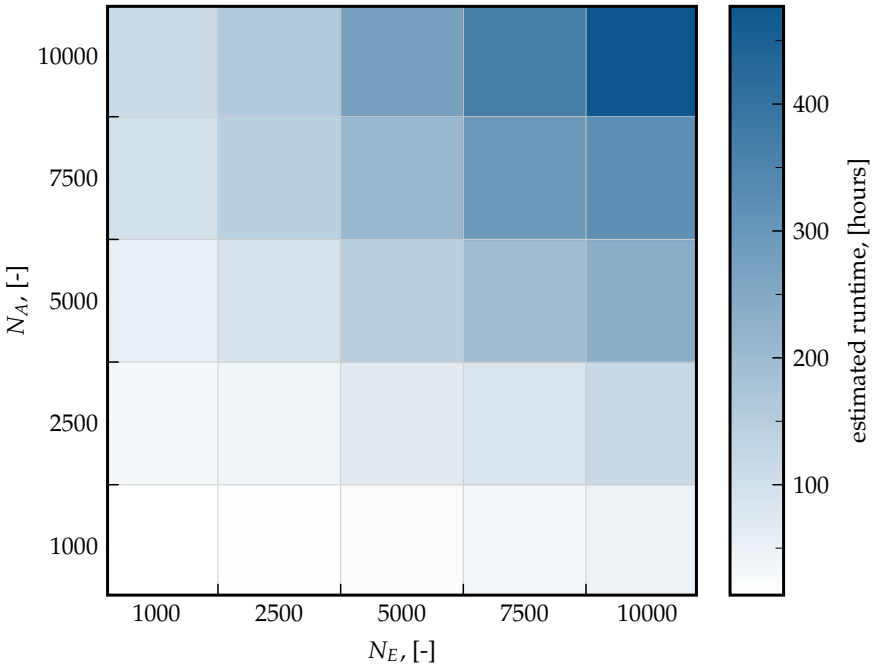


Figure 6.7 Estimated computational runtimes of the DSTE-based uncertainty combination for a sequential run.

loops be identical. This requirement stems from the subsequent post-processing steps, where mean values in each uncertain dimension (i.e., epistemic and aleatory) are calculated and visualized using a scatter plot and uncertainty contours. This marks the second novelty of this method. This scatter plot necessitates equal-length results, thereby defining N as the common convergence threshold for both loops.

Consequently, the output y of this approach has the same number of rows and columns, i.e.:

$$y = \begin{bmatrix} y_{e_1,a_1} & y_{e_1,a_2} & \cdots & y_{e_1,N} \\ y_{e_2,a_1} & y_{e_2,a_2} & \cdots & y_{e_2,N} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N,a_1} & y_{N,a_2} & \cdots & y_{N,N} \end{bmatrix} \tag{6.6}$$

To recall, the aleatory uncertainties vary from column to column, whereas the epistemic uncertainties vary from row to row. For the novel uncertainty representation, two vectors y_e and y_a are calculated. These represent the variation in the epistemic and aleatory dimensions, respectively[5]. They are calculated by computing the mean along the respective axis. For y_E,

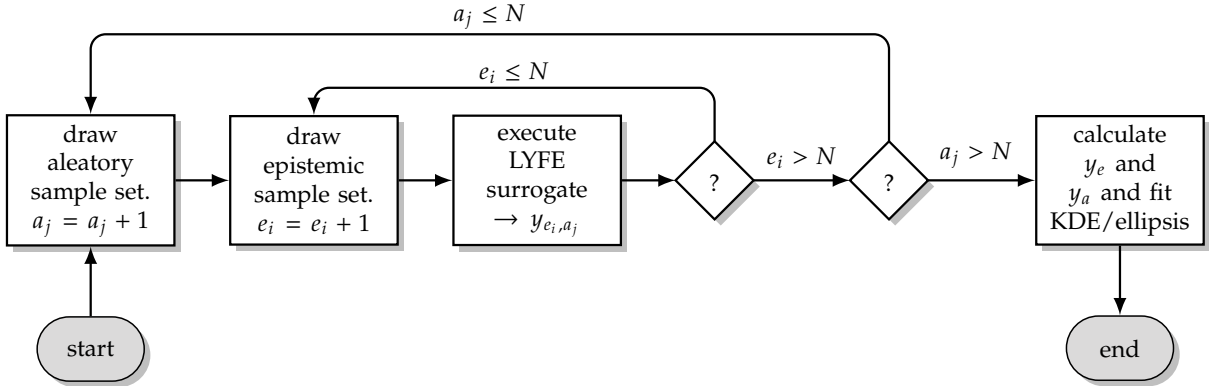


Figure 6.8 Flowchart for novel uncertainty combination approach.

[5]Supplemental information on this matrix operation can be found in Appendix F.2.

this is represented by the following equation:

$$y_E = \frac{1}{N}\left[\sum_{j=1}^{N} y_{e_1,a_j}, \sum_{j=1}^{N} y_{e_2,a_j}, \ldots, \sum_{j=1}^{N} y_{e_N,a_j}\right] \tag{6.7}$$

Similarly, y_A is obtained by averaging over all epistemic effects (rows) for each aleatory state (column), resulting in:

$$y_A = \frac{1}{N}\left[\sum_{i=1}^{N} y_{e_i,a_1}, \sum_{i=1}^{N} y_{e_i,a_2}, \ldots, \sum_{i=1}^{N} y_{e_i,a_N}\right]^T \tag{6.8}$$

The next step is to plot y_A against y_E. This process transforms the data into a bivariate dataset, represented as a scatter plot, which facilitates a direct visual assessment of the variations attributable to epistemic and aleatory factors. Each point in the plot corresponds to a pair (y_{E_i}, y_{A_j}), illustrating the mean effects across aleatory states for a given epistemic state and vice versa. Such plots simplify the interpretation of the complex interplay between different types of uncertainties, aiding in the identification of the balance between epistemic and aleatory variations as well as the overall spread of results.

To further elucidate the relationship between epistemic and aleatory variations depicted in the scatter plot, an uncertainty ellipse is superimposed. This ellipse is centered at the point (μ_{y_E}, μ_{y_A}), where μ_{y_E} and μ_{y_A} denote the mean values of y_E and y_A, respectively. Note that $\mu_{y_E} = \mu_{y_A} = \mu$ due to the methodical approach wherein, for each aleatory iteration within the nested MCS, an identical set of epistemic uncertainty samples is applied, and similarly, for each epistemic iteration, a consistent set of aleatory uncertainty samples is used. This ensures that across both dimensions of the matrix, the sum of all matrix elements remains invariant, whether aggregated row-wise or column-wise. Using the standard deviations $\sigma_{y_E} = \sigma(y_E)$ and $\sigma_{y_A} = \sigma(y_A)$ for the radii of the ellipse, its mathematical formulation is:

$$\frac{(\mathcal{X} - \mu_y)^2}{\sigma_{y_E}^2} + \frac{(\mathcal{Y} - \mu_y)^2}{\sigma_{y_A}^2} = 1, \tag{6.9}$$

where $\mathcal{X}$ and $\mathcal{Y}$ correspond to the coordinates on the abscissa and ordinate of the plot, respectively.

The inclusion of the uncertainty ellipse in the scatter plot not only highlights the dispersion of the data but also visually encapsulates the collective behavior of the system's uncertainties. By observing the placement and shape of the ellipse, we can compare the variance inherent in the epistemic and aleatory states.

A final aid in the interpretation of the output uncertainty is provided by an additional overlay of Kernel Density Estimate (KDE) plots. The KDE plots are particularly valuable for underlining the shape and density of the scattered data, providing insights into the distribution that extend beyond the symmetrical assumptions typically associated with standard-deviation-based ellipses.

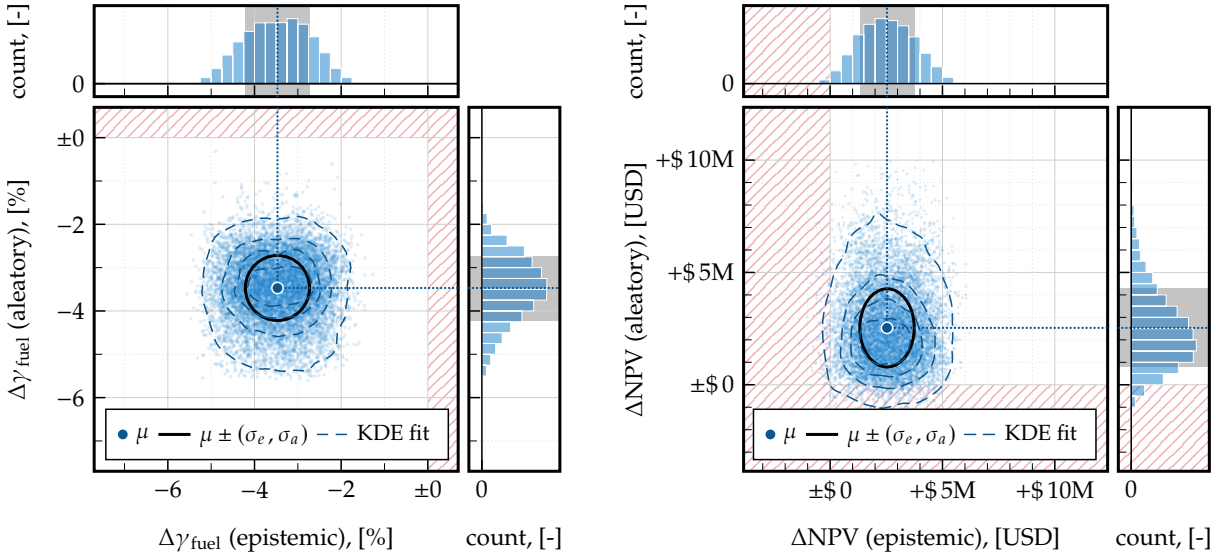
The mathematical formulation of the KDE, employing a Gaussian kernel, is as follows:

$$\hat{f}(\mathcal{X}, \mathcal{Y}) = \frac{1}{N}\sum_{i=1}^{N} K(\mathcal{X} - y_{E_i}, \mathcal{Y} - y_{A_i}), \tag{6.10}$$

where $\hat{f}(\mathcal{X}, \mathcal{Y})$ denotes the estimated density at the point $(\mathcal{X}, \mathcal{Y})$, and K represents the Gaussian kernel function applied to each data point (y_{E_i}, y_{A_i}), directly reflecting the pairs of epistemic and aleatory variations in our dataset. The selection of the Gaussian kernel is motivated by its efficacy in smoothly capturing the continuous probability density across the bivariate dataset of y_E and y_A.

6.3.1 Results

The results of the novel uncertainty combination method are displayed in Fig. 6.9, where the two outputs $\Delta\gamma_{\text{fuel}}$ and ΔNPV are distinguished as usual. In each bottom left subplot, the ordinate represents variation in the aleatory domain, while the abscissa shows the epistemic variation. A circle mark ("∘") at the center of the plot indicates the overall average μ. A black and solid ellipsis around μ illustrates the standard deviation in both the epistemic domain (with σ_e as the horizontal radius) and the aleatory domain (with σ_a as the vertical radius). Additionally, dashed lines represent uncertainty contours for each KPI. These contours indicate levels of similar density, based on a Gaussian KDE fitted to the scatter data. To further aid in interpreting these results, histograms of the marginal distributions for y_e and y_a are also shown. The gray area in these histograms indicates the range of the respective average ± the standard deviation.



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 6.9 Overall result of the novel combinatory UQ.

The $\Delta\gamma_{\text{fuel}}$ results, as depicted in Fig. 6.9 (a), show an overall average of −3.5 %. The standard deviations in both the epistemic and aleatory domains are very similar, resulting in a nearly circular ellipse and uncertainty contours. This suggests that the fuel consumption KPI is not dominated by either type of uncertainty. A closer examination of the histograms reveals that the epistemic variation is relatively symmetrical, while the aleatory variation has a slight tail toward lower (i.e., more negative and hence better for the HLFC aircraft) $\Delta\gamma_{\text{fuel}}$ values.

For the economic KPI shown in Fig. 6.9 (b), the overall average is $2.53M. The heights of the uncertainty contours and the standard deviation ellipse exceeds their width, indicating that this KPI is more sensitive to aleatory uncertainties than to epistemic ones. The histograms confirm this observation, showing that y_e is relatively symmetrical, while y_a is skewed and has a longer tail toward more positive ΔNPV values.

While these visual results provide a good overview, the summary in Table 6.2 allows for more quantifiable statements by showing the calculated statistics. In a direct comparison of the standard deviations in the epistemic and aleatory domains for $\Delta\gamma_{\text{fuel}}$, both are nearly identical with a value of approximately 0.75 %. On the other hand, the ΔNPV KPI shows a 48 % higher standard deviation in the aleatory domain compared to the epistemic domain. The percentiles for $\Delta\gamma_{\text{fuel}}$ are relatively close to one another.

Table 6.2 Summary of results of the novel combinatory UQ.

	μ	σ	$\mathbf{P_{25}}$	$\mathbf{P_{50}}$	$\mathbf{P_{75}}$
$\Delta\gamma_{\text{fuel}}$ [%]					
epistemic	−3.47	0.74	−4.03	−3.45	−2.91
aleatory	−3.47	0.75	−3.96	−3.45	−2.94
ΔNPV [USD]					
epistemic	$2.53M	$1.21M	$1.63M	$2.51M	$3.42M
aleatory	$2.53M	$1.77M	$1.26M	$2.27M	$3.50M

In terms of percentiles, for $\Delta\gamma_{\text{fuel}}$, the 25th, 50th, and 75th percentiles are almost identical between both uncertainty domains. For ΔNPV, the percentiles differ slightly, resulting in an IQR in the epistemic domain that is 20 % smaller than its aleatory counterpart.

With these results, it is suggested to summarize the overall results into statements as follows.

> *On average, the HLFC aircraft is expected to burn 3.5 % less fuel than its turbulent counterpart. The uncertainties are balanced with* $\pm\sigma_e = 0.74\,\%$ *and* $\sigma_a = 0.75\,\%$*. Regarding the economic KPI, the HLFC aircraft is estimated to yield $ 2.53M more, with uncertainties of* $\sigma_e = \$\,1.21\text{M}$ *and* $\sigma_a = \$\,1.77\text{M}$*.*

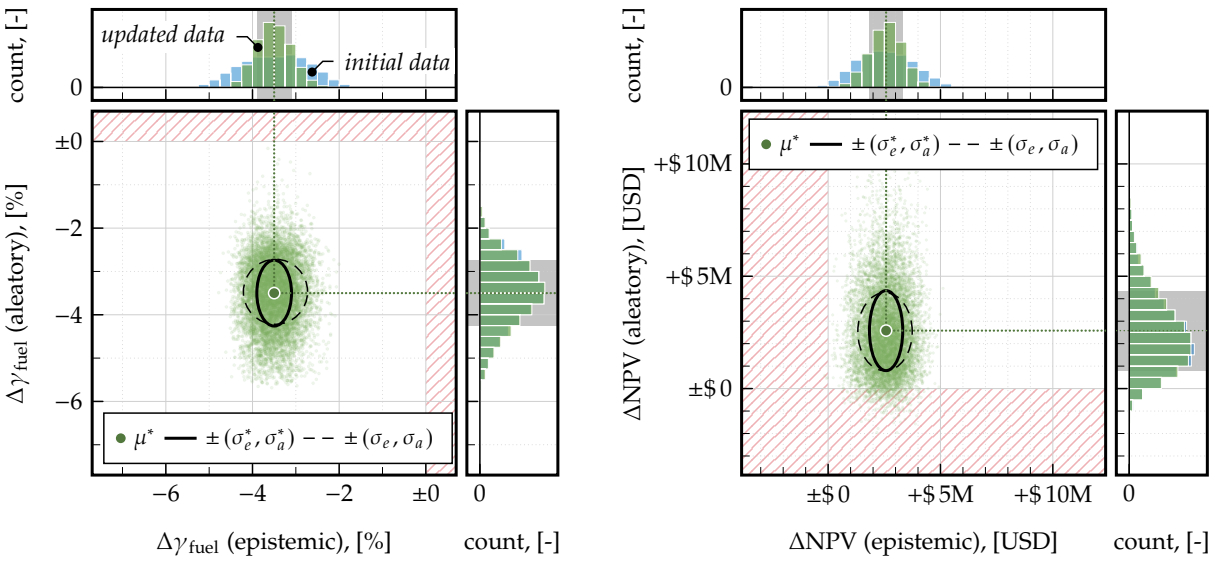
Thus, the key takeaway is that the fuel performance of the HLFC aircraft is expected to outperform that of its conventional turbulent counterpart. The overall uncertainty is balanced and relatively low, as evidenced by both σ_e/μ and σ_a/μ being well below 1. In terms of economic performance, although it is positive, it exhibits greater sensitivity to non-reducible (aleatory) uncertainties compared to reducible (epistemic) ones. With an epistemic uncertainty ratio of $\sigma_e/\mu \approx 0.47$, there is significant potential for reducing uncertainty through further research and knowledge acquisition. Conversely, the aleatory uncertainty ratio of $\sigma_a/\mu \approx 0.69$ suggests that design improvements could make the aircraft more robust to natural variations, such as those related to cruise Mach number.

6.3.2 Results with Reduced Uncertainty

To verify the effectiveness of the novel method for combined UQ, a similar approach to the one described in Section 6.2.2 is employed. Specifically, the epistemic input uncertainty $\mathcal{D}_1$ - which represents the overall drag reduction - is hypothetically reduced. This is achieved by fixing the drag reduction potential of the HLFC application on the upper wing side to its average value of 5.2 %. This modification is expected to decrease the standard deviation of both KPI in the epistemic domain, while the aleatory domain should remain unchanged.

The results of this verification attempt are depicted in Fig. 6.10, both as scattered data points and marginal histograms in green. Histograms from the initial campaign are also included for the sake of comparison in blue. To maintain clarity in the overview, uncertainty contours based on the fitted KDE are omitted. Instead, only the standard deviation ellipses for both the initial (dashed) and updated (solid) campaigns are shown.

The $\Delta\gamma_{\text{fuel}}$ results, as presented in subplot (a), show a significant reduction in epistemic uncertainty, as indicated by a narrower ellipse around μ^* and the histogram of y_e. The summary in Table 6.3 confirms that σ_e has been halved, while the aleatory uncertainty remains stable. Similar trends are observed for the economic results: the epistemic uncertainty has decreased from an initial $\$\,1.21\text{M}$ to $\$\,0.56\text{M}$, and the aleatory uncertainty has increased only marginally. The mean values for both KPI remain largely unchanged, affirming that the reduction in $\mathcal{D}_1$ did not alter its overall average.



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 6.10 Overall result of the novel combinatory UQ with reduced epistemic uncertainty (green), compared to the original results (blue).

Regarding percentiles, the IQR for ΔNPV in the epistemic domain is now 57 % smaller than in the aleatory domain. For $\Delta\gamma_{\text{fuel}}$, the epistemic IQR is also now smaller by 47 % compared to its aleatory counterpart. Both of these observations align with the expectations.

Table 6.3 Summary of results of the novel combinatory UQ with reduced epistemic uncertainty and σ values compared to the original results.

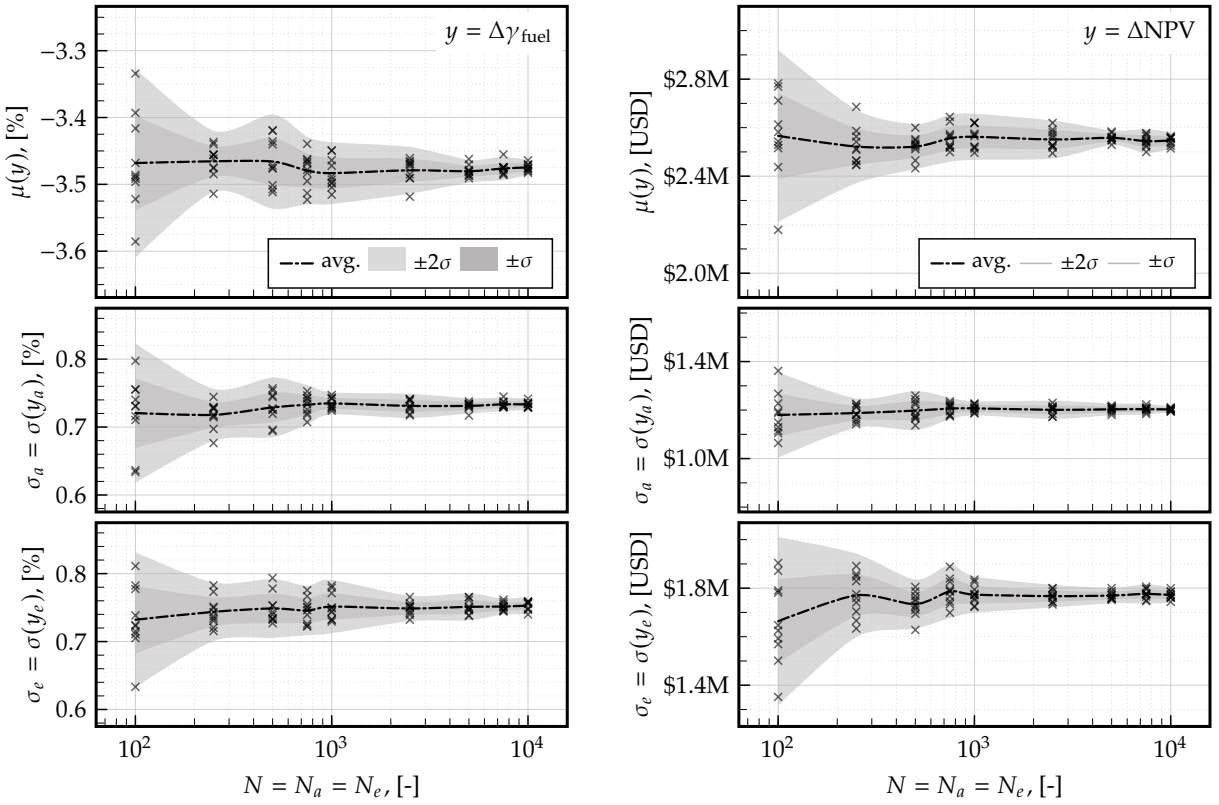
	μ	σ	P_{25}	P_{50}	P_{75}
$\Delta\eta_{\text{fuel}}$ [%]					
epistemic	−3.50	0.40 (−46 %)	−3.78	−3.51	−3.24
aleatory	−3.50	0.76 (± 0 %)	−3.99	−3.47	−2.98
ΔNPV [USD]					
epistemic	$2.58M	$0.74M (−38 %)	$2.08M	$2.60M	$3.09M
aleatory	$2.58M	$1.78M (+ 1 %)	$1.27M	$2.33M	$3.61M

These results confirm that the novel combination method successfully captures reductions in epistemic uncertainty, as anticipated. The scatter plot presentation, enhanced with uncertainty contours and a standard deviation ellipse, is easy to implement and interpret. Therefore, it is well-suited for recipients and decision-makers who may not be experts in the field of UQ.

6.3.3 Convergence Behavior and Runtime Performance

Given the computational challenges inherent in comprehensive UQ efforts, this Section investigates the convergence behavior of the novel method. To achieve this, the three key metrics μ, σ_e, and σ_a are repeatedly calculated with an increasing sample size N, where $N = N_A = N_E$. This leads to a total number of executions $N_{\text{tot}} = N^2$. To further refine the analysis, each calculation is repeated 10 times, allowing a quantification of the variation in μ, σ_e, and σ_a at each sample size, thereby enabling a more nuanced understanding of convergence behavior. The results are presented in Fig. 6.11, with changes in fuel consumption illustrated in subplot (a) and economic performance in subplot (b). The "×" marks indicate the result of a single repetition at a specific sample size N, the dash-dotted line represents the average value throughout each repetition of N, and the dark gray and light gray areas signify the average $\pm\sigma(N)$ and $\pm 2\sigma(N)$, respectively.

The convergence analysis for $\Delta\gamma_{\text{fuel}}$ reveals that all three result components μ, σ_a, and σ_e exhibit a relatively large spread at smaller sample sizes. For instance, at a sample size of $N = 100$ (or $N_{\text{tot}} = 10\text{k}$), the mean values $\mu(y)$ have an absolute spread of 0.25 %-points, and the standard



(a) change in fuel consumption.

(b) change in economic superiority.

Figure 6.11 Progression of μ, σ_e and σ_a with increasing sample size N.

deviations vary within a range of 0.18 %-points. As the sample size increases, for example at $N = 1000$ leading to $N_{\text{tot}} = 1\text{M}$, the spread in $\mu(y)$ narrows to 0.06 %-points, while the spreads in σ_a and σ_e reduce to 0.02 %-points and 0.05 %-points, respectively. Further increases in N continue to narrow these spreads.

It's important to note that the definition of convergence is contingent on the desired level of accuracy. For this analysis, an accuracy threshold of ±0.1 %-points suggests a required sample size of $N \approx 500$. The convergence behavior for ΔNPV is similar. A potential accuracy threshold of $ 100k would necessitate sample sizes of $N \geq 5000$, while $N \geq 500$ would suffice for an accuracy of $ 200k.

One of the major advantages of this approach is its low computational expense. Fig. 6.12 illustrates the overall execution runtime for various sample sizes, distinguishing between the time required for sample generation (depicted by red bars), model execution (gray bars), and the calculation of the overall mean and standard deviations σ_a and σ_e (blue bars). While the time for model execution increases exponentially, the time required for output UQ (i.e., calculating the uncertainty contours and uncertainty-domain-specific standard deviations) grows linearly. Importantly, the time spent on sample generation and output UQ—steps unique to this novel method—are negligible when compared to the model execution time. This point is further underscored by the use of the PCE surrogate model in this study, which is considerably less computationally demanding than the original LYFE model due to its analytical nature.

6.4 Comparison and Conclusion

To conclude this Chapter, this subsection provides a detailed comparison of the two methodologies and discusses their inherent strengths and weaknesses.

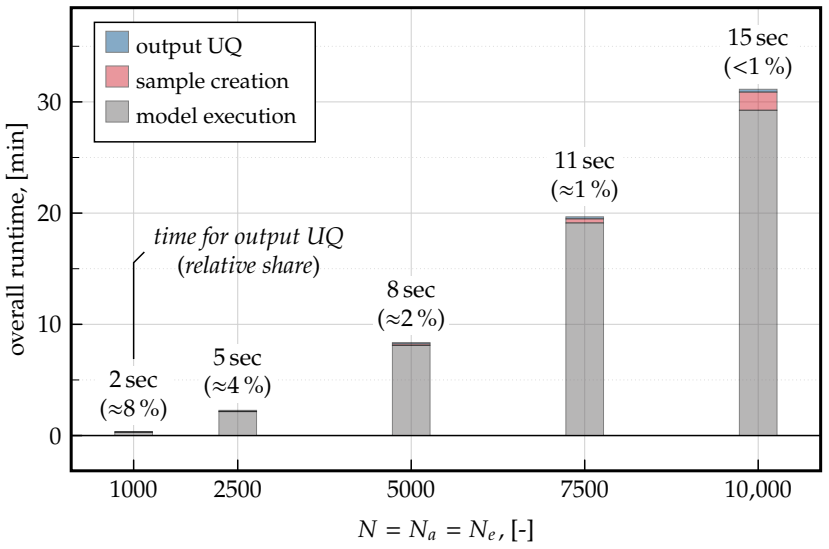


Figure 6.12 Runtimes of the novel uncertainty combination for different sample sizes with focus on the output UQ.

To recall, the first combination method utilizes the Belief spaces derived from the expert elicitation and essentially executes a DSTE-based UQ process as performed in Chapter 5 for each set of aleatory sampled uncertainties. This results in a set of (C)CXFs, which allow constructing statements that are useful for threshold-driven questions such as "*How likely is it that the HLFC aircraft burns at least 2.5 % less fuel than the conventional counterpart?*". The statements contain lower and upper bounds of probability (which represent the epistemic impact) as well as lower and upper quartiles for each bound (which represent the aleatory impact).

The second combination method also uses the Belief spaces for the epistemic uncertainties and PDFs for the aleatory uncertainties. The key difference to the first method lies in the post-processing, where, instead of the time-consuming calculation of (C)CXFs, the mean values, standard deviations, and percentiles are calculated for each of the two uncertainty domains (i.e., epistemic and aleatory). The results are then visualized with a scatter plot, where the two axes depict the variation in each uncertainty domain. As an additional interpretation aid, the scatter plot is complemented by a KDE fit and an ellipsis whose shape indicates which uncertainty domain is dominant.

Regarding the difficulty of interpreting the results, the evidence-based combination method has several drawbacks. The overall results, which span the entire output space (see Fig. 6.3), are comprehensive as they provide insights regarding various threshold values. This visual expression could, however, be overwhelming due to the amount of information it contains (i.e., (C)CBFs, (C)CDFs, (C)CPFs, U_{DSTE}, the aleatory variation of each, shown as percentiles). In order to construct more tangible statements, boxplots can be used, which essentially reduce the information to be specific to a particular threshold y. A noticeable level of understanding of DSTE and the combination methodology is required to fully understand these statements. The uncertainty metrics of the novel modification of the second order probability method, in contrast, are more straightforward as recipients are likely to be familiar with the concepts and interpretation of mean values and standard deviations. The visual expression is more easily explained (see Fig. 6.9), and the ellipses provide an immediate indication of the assessment's uncertainty, which is complemented by the tabular summary (see Table 6.2). Generally speaking, the interpretation superiority of the novel method largely originates from the fact that, while utilizing DSTE's input uncertainty definition, its outputs use probabilistic concepts, which are generally well understood by a larger community.

The traceability, which describes the capability to repeat the uncertainty combination at a later stage where one or more epistemic uncertainties are reduced, also differs between both methods. The evidence-based combination approach showed a behavior that is partially non-intuitive at first glance, as the aleatory spread of U_{DSTE} seemed to have increased for the updated analysis, although the aleatory inputs were identical to those of the original analysis (see Fig. 6.5). Again, a significant level of understanding of the theory as well as the implemen-

tation of the `dste` package are required to make some sense of the results. In contrast, the novel combination method showed consistent and meaningful results when repeated with less epistemic input uncertainty (see Fig. 6.10), rendering this approach superior to the evidence-based one.

The last aspect to compare is computational efficiency. In this regard, the evidence-based approach has one particular advantage over the novel method: the number of samples for the aleatory loop and epistemic loop do not have to be equal. For the analyses performed here, a relatively small aleatory sample size was needed to obtain a converged set of percentiles. However, the epistemic sample size needs to be extremely large for the calculation of (C)CXFs themselves to be converged, as the discussion in Section 5.4.4 showed. More specifically, a sufficient sample size for this combination method and the present use case is $N_E > 10^5$ (epistemic) and $N_A > 2500$ (aleatory), resulting in at least 250 M required samples. Considering the fact that the calculation of the (C)CXFs adds significantly to the overall execution time, this approach is likely to be infeasible for the majority of endeavors unless a quick-to-execute surrogate model and a high-performance cluster are used. The novel method, which necessitates N_E and N_A to be equal, seemed to have a converged result at $N_E = N_A > 5000$, which results in 25 M samples. While this is still a large number, it is ten times smaller than that of the previous approach. Furthermore, the post-processing of the novel method is negligible in terms of the execution time, as opposed to the evidence-theoretic technique.

In summary, the comparative study documented in this Chapter lead to the conclusion that the novel modification of the second order probability technique for UQ is superior to the evidence-based one in all investigated aspects, i.e., interpretation, traceability, and computational efficiency. Thereby, the fourth research hypothesis is verified. While some factors, such as the number of required samples, may change for different applications, the issues in interpretation and traceability are fundamental to the underlying theory, leading to the recommendation to use the novel method over the evidence-based one.

7 Conclusions, Limitations, and Outlook

This final chapter offers a thorough review of the research conducted throughout this thesis. It aims to synthesize the key findings, acknowledge the limitations, and suggest future directions in this field. To start, the overarching research objective is revisited, setting the stage for the discussion. Each of the three research questions and their associated hypotheses are then examined. This examination begins with a brief summary of the work performed, highlighting the key methodologies and results. The main conclusions drawn from the research follow, providing insights into the effectiveness and impact of the proposed approaches. Next, the limitations of the performed analyses are discussed, addressing the constraints and challenges encountered during the research process. Finally, the chapter explores the implications of these findings for uncertainty-driven techno-economic assessments in aeronautics, emphasizing the contributions and potential future developments enabled by this research.

The primary aim of this research was to enhance the transparency and reproducibility of prospective and lifecycle-based techno-economic assessments in aeronautics. To achieve this, an uncertainty management methodology was introduced in Chapter 1, which encompasses uncertainty identification, classification, combination, propagation, reduction, and representation for decision-makers. This methodology was designed to systematically address the complexities and nuances of uncertainty in techno-economic assessments, providing a structured approach for practitioners and enhancing the credibility of the assessments.

Several barriers hinder the effective implementation of this methodology. These barriers include the selection of appropriate methods for quantifying the influence of input parameter uncertainties, the integration of non-probabilistic theories, and the simultaneous consideration of both epistemic (knowledge-based) and aleatory (variability-based) uncertainties. To substantiate the need for overcoming these issues, a thorough review of the fundamentals of economic assessments in aeronautics, uncertainty quantification theories and techniques, and the prevalent literature in these fields was conducted and documented in Chapter 2.

This review highlighted that probability theory, particularly with its frequentist view, remains the predominant approach for dealing with uncertainty, effectively utilizing Monte-Carlo simulations when ample data is available. However, when data is limited, probability theory faces significant limitations, necessitating alternative approaches like non-probabilistic theories. Among these, evidence theory emerged as a complementary or potentially superior approach for uncertainty quantification in data-scarce situations. Despite its potential, its application in aeronautic techno-economic assessments remains limited, highlighting the need for demonstrative studies. The review also covered global sensitivity analysis, revealing a variety of fundamentally different methods with often unexplained rationale for their selection, indicating a need for systematic review and clearer guidelines. Additionally, most studies focus on either epistemic or aleatory uncertainties but rarely both, underscoring the necessity for a comprehensive method to simultaneously address and combine these types of uncertainty.

These barriers were addressed with one research question and hypothesis each, guiding the research efforts to enhance the transparency and reproducibility of techno-economic assessments in aeronautics. To demonstrate the application of the methodology, a case study was conducted evaluating the HLFC technology on a long-haul aircraft, utilizing the prospective and lifecycle-based techno-economic assessment framework LYFE at the preliminary aircraft design stage.

7.1 Efficient Uncertainty Reduction through Global Sensitivity Analysis

The first research question and hypothesis focused on identifying influential uncertainties and distinguishing them from negligible ones using global sensitivity analysis. For clarity and reference, they are restated here:

> **RQ1** *How can the adoption of global sensitivity analysis in uncertainty-addressing TEAs be facilitated or promoted, given the variety of available techniques and the lack of discussion among researchers about their specific choice?*
>
> **HYP1** *Through a comprehensive evaluation of the capabilities, interpretability of sensitivity, and computational efficiency of various GSA techniques on an example problem, followed by the demonstration of the selected method's effectiveness in a case study, it is anticipated that the insights gained will contribute to facilitating the adoption of GSA in uncertainty-addressing TEAs, informing future research directions and methodological choices in the field.*

These matters were addressed in Chapter 4, where a first study began with a comparative analysis of various GSA techniques, followed by an applied GSA of the HLFC case-study. The comparative analysis looked at the mathematical foundations of both variance-based and non-variance-based GSA techniques, introducing their definition of sensitivity. This analysis comprised a discussion and comparison of the provided sensitivity measures, convergence behaviors, computational demands, and challenges in interpretation, revealing significant differences. In the subsequent applied analysis, the most effective GSA method was used to apportion the uncertainty in the HLFC assessment results using LYFE. This part of the study demonstrated in detail how to quantify input uncertainties using PDFs from existing literature and databases. After analyzing the global sensitivities, the most influential input uncertainties were identified and separated from those with minimal impact on the outcome. As the uncertainty regarding the cruise Mach number was shown cause the majority of the output uncertainty, a hypothetical and more robust HLFC design modification was introduced, followed by a repetition of the global sensitivity analysis, demonstrating how sensitivity measures may change throughout the product development process.

The key takeaways of the first research question and hypothesis are:

- In general, GSA techniques are highly suitable for guiding uncertainty reduction efforts, as they systematically quantify the contribution of each input uncertainty to the output uncertainty. This facilitates the ranking of input parameters based on their criticality, helping to objectively identify and prioritize the most influential ones while also highlighting those with negligible impact.

- GSA techniques differ fundamentally in their mathematical basis, definition of sensitivity, provided sensitivity measures, computational expense, and practical aspects such as the capability to deal with correlated inputs or given data. Therefore, the choice of the optimal GSA technique depends on the assessment context and parameter interdependence.

- For the present work, the Structural and Correlative Sensitivity Analysis (SCSA) method emerged as the most suitable GSA technique, given its comprehensive sensitivity measures and efficient convergence. Due to its definition of sensitivity being a parameter-specific share of the output variance, its interpretation is more intuitive than that of other techniques.

The limitations of the work revolving around the first pillar of this research are:

- The GSA techniques assessed in this study represent only a subset of available methods. Given the dynamic nature of the GSA research field, there are numerous techniques continually being developed. For the sake of clarity and feasibility, this study focused solely on methods available within the Python package SALib.
- While the study delved into various characteristics of different GSA techniques, it did not explore every intricate detail. For instance, correlated sensitivity in the HLFC use case was not analyzed. As such, the insights and conclusions drawn from this work should be understood as being illustrative rather than exhaustive.
- Another shortcoming of the study deals with the number of uncertain input parameters, which was not varied within the comparative analysis nor in the applied case study. Therefore, the recommendations for choosing the SCSA method are not universal. However, the insights of the systematic comparison allow for a more informed selection of the global sensitivity technique.

Equipped with the insights gained from this part of the research, users are now better equipped to select a problem-appropriate global sensitivity analysis technique tailored to their specific needs. This selection process optimizes efforts in uncertainty reduction, a recurring task in the product development process. Consequently, the first hypothesis is verified, underscoring the key message: the selection of a global sensitivity analysis technique should be deliberate and informed, rather than arbitrary.

7.2 Epistemic Uncertainty Quantification using Evidence Theory

The second research question and hypothesis revolved around evidence theory as a potential candidate to complement or replace conventional probability theory in data-scarce situations, i.e.:

> **RQ2** *How can non-probabilistic theories, such as evidence theory, be more widely adopted in TEAs, given the challenges in interpreting their concepts and metrics and the lack of comparative studies with prevalent probabilistic methods?*
>
> **HYP2** *The development of publicly accessible and user-friendly tools for non-probabilistic UQ, such as for evidence theory, is expected to aid in the adoption of these methods in TEAs. Through demonstration, comparison with prevalent probabilistic methods, and analysis of key aspects such as interpretability, traceability, and computational efficiency, such tools have the potential to enhance decision-making processes in practical applications.*

This was addressed in Chapter 5, which begins with a discussion on the theoretical and practical aspects of using evidence theory to quantify epistemic uncertainties. The Chapter explains how evidence theory's uncertainty quantification metrics - such as CCBFs, CCDFs, and CCPFs (collectively known as CCXFs) - are derived from basic belief assignments obtained through systematic expert interviews. These metrics are particularly useful for assessing the likelihood of exceeding or falling below specific output thresholds of interest. To address the need for user-friendly tools, the Python package `dste` was developed. This package simplifies evidence-theoretic uncertainty quantification by treating the simulation model as a black box, enabling the quantification of inputs and outputs through basic belief assignments and CCXFs. The package was utilized in a case study evaluating the HLFC technology, where five experts were interviewed to obtain basic belief assignments, supplemented by literature values. The epistemic parameters were propagated through a LYFE surrogate model (due to computational constraints), and the resulting output sample space was compared to an approach using probability theory with uniform distributions where data is scarce. The chapter further discusses the interpretability and convergence behavior of the additional uncertainty quantification metrics provided by evidence theory. It concludes with a verification attempt

that hypothetically reduces one of the epistemic uncertainties, revealing limitations in using this approach repeatedly to track uncertainty throughout the product development process.

Key findings of this work are:

- The design of the systematic interviews is a crucial aspect, both in terms of avoiding expert bias and potential computational repercussions resulting from basic belief assignments. In this work, predefined intervals were provided for the selected uncertainties. Experts were then asked to assign a number representing their Belief that the true value of the uncertainty lies in a certain interval. However, the large number of basic belief assignments lead to an exponential increase of evidence intervals, which in turn heightens the computational expense of this approach.
- The computational expense is, in fact, one of the biggest difficulties of the evidence-theoretic approach. In the applied example, more than 10^6 model executions were needed for the result to be considered converged. This is likely to be prohibitive for most applications unless a surrogate model is combined with a powerful computational setup.
- Another challenge of the evidence-theoretic technique stems from the difficulty of interpreting the uncertainty metrics. A proper understanding of the theory as well as the coding implementation is needed, especially when repeating the assessment when new information arises. One example of this challenge is the novel definition of evidence-theoretic uncertainty U_{DSTE}, which is the difference between the Plausibility and Belief at a selected threshold. While it helped to visualize the epistemic uncertainty, it can change in non-intuitive ways when repeating the analysis with updated information.
- Generally speaking, using expert-based uncertainty quantification had a significant impact on the assessment results. While the probabilistic approach typically relies on simplified estimation techniques, often leading to uniform distributions, the belief spaces derived from evidence theory produced samples that resembled non-parametric probability density functions which would otherwise require substantial amounts of data to obtain. However, this does not imply that the true value is more accurately represented; rather, it indicates that this approach more effectively captures the true (lack of) knowledge about the uncertainty.

As with all research, the work revolving the second research pillar has some limitations:

- The provided guidelines for the elicitation of expert knowledge can be improved upon. Due to the large number of basic belief assignments, the computational expenses in this study were too high to reach a fully converged set of CCXFs. However, it should be noted that this convergence behavior is influenced by the very nature of the CCXFs calculation, which foresees the calculation of the output's minima and maxima for given evidence intervals. With a larger sample size, each evidence interval is covered by more samples, naturally pushing the resulting minima and maxima of the output further apart. Potential avenues to overcome this issue include the combination of interval analysis techniques (where extreme points are identified more efficiently) with advanced sampling methods to optimize the balance between computational cost and the accuracy of the uncertainty bounds.
- In the verification analysis, representing an assessment with updated knowledge, a specific epistemic input uncertainty was reduced. Since evidence theory does not prescribe the technique to deal with such an update, three alternative techniques were explored. However, these methods are not comprehensive, as the influence of different evidence combination techniques was not investigated. As a result, the previous remarks about the challenges of interpreting uncertainty in recurring assessments is not comprehensive.

To conclude, the investigation into evidence theory provided valuable insights into a non-probabilistic alternative for UQ. With the developed Python package, researchers are equipped

with a hands-on tool to explore this method in practice. The in-depth discussions and analyses around the approach highlighted both the potential benefits and inherent challenges of using evidence theory. While it presents a promising alternative, especially in contexts with limited data but available expert knowledge, it also comes with its own set of difficulties. As a result, the hypothesis can only be partially validated, highlighting the ongoing need for research and development in this area.

7.3 Combining Epistemic and Aleatory Uncertainties with Evidence and Probability Theory

As epistemic and aleatory uncertainties have different implications for decision-makers, their combination and simultaneous representation were addressed in the last research question and hypothesis, which were:

> **RQ3** *How can the systematic differentiation between epistemic and aleatory uncertainties be effectively addressed in TEAs, considering the challenges associated with combining and propagating these different types of uncertainties in a manageable and interpretable manner?*
>
> **HYP3** *By systematically differentiating between epistemic and aleatory uncertainties in TEA using evidence-theoretic methods and nested Monte Carlo simulations, it is anticipated that enhanced interpretability and actionable insights will enable more informed resource allocation and risk mitigation strategies. Specifically, this approach is expected to identify dominant uncertainty types and guide strategic decisions on research investments or robust product designs.*

To address these, Chapter 6 investigated and compared two alternative techniques for combining epistemic and aleatory uncertainties. Both use a nested Monte-Carlo design, where the inner and outer loop deal with the sampling of the epistemic and aleatory uncertainties, respectively. The first of the two combination techniques remains in the evidence-theoretic domain and can be seen as an extension of the work in Chapter 5. As such, the selected epistemic uncertainties were sampled using the same Belief spaces obtained from expert interviews. Additionally, aleatory uncertainties were sampled using the PDFs derived in Chapter 4. After executing the nested MCS, the outputs were structured in a two-dimensional matrix, out of which complementary cumulative function collectives were calculated. This resulted in a set of CCXFs for each aleatory sample set. After visualizing these evidence-based uncertainty metric collectives over the output space using percentiles, statements regarding the assessment of the HLFC aircraft were constructed, demonstrating how to interpret CCXFs when aleatory variation is present. A verification attempt, in line with the one performed in Chapter 5, introduced a hypothetically reduced epistemic uncertainty, followed by an analysis of the changes in CCXFs. The discussion of this combination method was concluded with a convergence and computational runtime analysis, where the respective sample sizes were varied. The second approach to uncertainty combination represented a novel modification of the nested Monte-Carlo simulation. It incorporates elements from evidence theory for quantifying epistemic inputs but remains in the probabilistic domain for representing the output uncertainty. The results, which were structured in a 2D matrix as well, were post-processed by calculating the averages alongside the epistemic and aleatory axis. These were then visualized using scatter plots, enriched by uncertainty contours and ellipses as well as marginal histograms. This novel technique was verified using the same approach, i.e., reducing one selected uncertainty and discussing the expected change in the provided uncertainty measures. A final convergence and computational expense analysis concluded this work.

The conclusions of this last set of studies are as follows:

- The aforementioned difficulties of interpreting Belief and Plausibility are heightened when aleatory uncertainties are included. In this work, Belief and Plausibility were interpreted as lower and upper bounds for probability, representing the epistemic spread in the results. This lead to the notation of x_a^b notation, where the difference between a and b represent the epistemic uncertainty. Extending the evidence-theoretic uncertainty quantification by aleatory uncertainties introduces an additional upper and lower limit for x, a, and b. Therefore, a complete statement regarding the likelihood of exceeding (or falling below) a certain threshold value requires the simultaneous consideration of nine different values to address both the epistemic and aleatory impact. This intricate notation underscores the complexity of accurately representing uncertainty and highlights the importance of careful interpretation.

- The visualization of the results of the evidence-theoretic combination method over the entire output space is comprehensive, as it allows users to construct statements for all possible threshold values. For cases where only specific thresholds are of interest, boxplots can effectively illustrate the relevant distribution characteristics. However, to ensure that decision-makers fully grasp the implications, either form of depiction may necessitate a thorough explanation of the methodology. This need for detailed understanding becomes more pronounced with recurring assessments, as the behavior of uncertainty measures can be counterintuitive due to the interconnections among evidence intervals. Consequently, a deep understanding of the process is essential.

- The computational expense of the evidence-theoretic combination is likely to be a challenging, if not prohibitive factor for most applications. The calculation of CCXFs, which foresees matching each output sample to its causing evidence intervals, is computationally demanding. Introducing aleatory samples in a nested Monte-Carlo design essentially means repeating this demanding process many times. Without a quick-to-evaluate surrogate model and a powerful computational setup, this uncertainty combination approach is likely not feasible, especially when the number of input uncertainties is higher than investigated here.

- The novel method, which used probabilistic measures such as mean values and standard deviations to represent the output uncertainty, was significantly easier to understand and interpret. The scatter-plot visualization, which covers the entire output space as well, is swiftly explained to decision-makers. Its uncertainty contours and ellipses allow for a quick evaluation of the assessment result. Constructing statements about the results was also straightforward, as only three values are of interest: The overall average, the standard deviation in the epistemic domain and the standard deviation in the aleatory domain.

- The verification of the novel combination method was successful and showed intuitive behavior. However, this method also suffers from computational issues, albeit to a lesser degree. The sample size required for convergence is roughly 10 times smaller compared to the evidence-theoretic approach, but the computational runtime of the method-specific post-processing is negligibly small. Total runtimes for a converged result did not significantly exceed 30 min, rendering this technique the more appropriate one for most applications.

- In comparison to its evidence-theoretic counterpart, the novel method has a notable limitation: it offers less flexibility regarding sample sizes. The evidence-theoretic approach permits different sample sizes for epistemic and aleatory uncertainties, allowing for a more customized Monte-Carlo design. Conversely, the novel combination method requires equal sample sizes for both types of uncertainty. This constraint may result in a higher number of executions than necessary for convergence, particularly if the distribution of epistemic and aleatory uncertainties is imbalanced.

The limitations of the work revolving the last research pillar are:

- The CCXF collectives were visualized using medians and selected percentiles for representing the aleatory output uncertainty. Alternative techniques, such as averages ± multiples of the standard deviation or a closer look into the variance, were not investigated, limiting the takeaways regarding interpretation to the approach specifically employed in this study.
- Both approaches assume independent parameters. In situation where correlation exists, the output uncertainty may be misrepresented. This underlines the necessity of further research in the domain of uncertainty combination when inputs are not independent, both in the theoretical domain as well as with respect to the development of publicly available toolboxes.

Wrapping up the final part of the research, the introduction of a modified nested MCS has paved the way for a comprehensive combination of epistemic and aleatory uncertainties, seamlessly integrating both probabilistic and non-probabilistic elements. When compared with an alternative method, which solely relies on evidence theory for both input and output UQ, the novel approach stands out in terms of its interpretability, traceability, and computational efficiency. Yet, it is imperative to acknowledge the inherent computational demands of simultaneously addressing epistemic and aleatory uncertainties, especially when discerning their individual impacts. This underlines the advisability of employing a surrogate model that ensures rapid evaluations during uncertainty analyses. In summation, the final hypothesis can be verified. Researchers can easily adopt the suggested method to quantify and visualize the influence of either uncertainty nature. The incorporation of uncertainty contours and ellipses equips decision-makers with a direct means to gauge the uncertainty inherent in assessments. The characteristics and positioning of these graphical representations can guide tailored recommendations, enhancing the decision-making process.

7.4 Recommendations for Future Work

As this thesis concludes, it is essential to look forward and identify potential areas for future exploration in this domain. The following suggestions aim to inspire and guide further research rather than provide an exhaustive list.

1. While probability theory and its techniques are well established, there is a clear need for more research on the non-probabilistic side. This includes a deeper understanding of evidence theory's mathematical foundations, especially when considering the measure for epistemic uncertainty when aleatory variations exist.
2. The `dste` package, a central tool in this research, could benefit from further development. Beyond improving the computational efficiency, integrating it with the SCSA method for GSA could be valuable, given the method's ability to work with existing data.
3. Correlation and its impact on UQ methods remain active research areas. This is crucial, as correlated inputs are common in integrative aeronautic cost-benefit assessments. Thus, another potential avenue for future research is the modification of evidence theory's techniques for correlated inputs.

In conclusion, the research presented in this thesis has addressed many of the technical challenges that previously hindered the effective application of the uncertainty management methodology. This work contributes significantly to promoting transparency and trust in an era where computer model simulations are becoming increasingly prevalent and uncertainties are inevitable.

Appendices

A Supplemental Information for Chapter 1

This part of the appendix provides supplemental information on topics that were covered or touched upon in Chapter 1, including the delineation of TEA with ROA and TPS and a description of the foundational Clean Sky 2 projects.

A.1 Real Option Analysis and Techno-Economic Assessments

Uncertainty-addressing ROA and TEA are related evaluation approaches but differ in their foci and objectives. This section provides an overview of their commonalities and differences.

TEA, as used in this thesis, models the impact of a technology from the operator's perspective. This involves creating a detailed model of the operational environment and considering repercussions from various domains such as design, operation, environmental, and economic factors. This results in KPIs such as γ_{fuel} for fuel consumption and NPV for economic value. By comparing these to a reference, the Δ values quantify the relative inferiority or superiority of a technology. Introducing parameter uncertainties aims to shed some light on the impact of knowledge gaps and natural variability on the techno-economic outcomes. This can be achieved using MCS, which produce output PDFs for $\Delta\gamma_{\text{fuel}}$ and ΔNPV.

ROA also models the operator's expected cash flow but introduces a different perspective by emphasizing the value of managerial flexibility and strategic decision-making under uncertainty [74]. Originating from financial options theory, it applies these principles to real assets, such as aircraft, to evaluate the value of making specific decisions at various points throughout the asset's operational life. This approach is particularly useful given the long operational phase of aircraft, where such flexibility can significantly impact overall value. MCSs can be utilized here as well, but included uncertainties tend to focus on parameters affecting flexible decision-making rather than solely techno-economic performance.

While conventional TEAs and ROA address different aspects of uncertainty and value, they are complementary tools. TEA provides a robust foundation for understanding the baseline performance and economic feasibility of a project under uncertainty. It offers insights into the distribution of potential outcomes, helping identify critical risk factors and areas of uncertainty. ROA builds on this baseline understanding to evaluate the additional value created by managerial flexibility and adaptive decision-making. It helps decision-makers appreciate the strategic value of maintaining flexibility, especially in long-term projects like aircraft operations, where future conditions can change significantly.

A.2 Technology Selection and Economic Evaluation

Both TPS, which is often found in integrated aircraft design frameworks, and TEA, when applied within the IPDP, quantify the performance and cost-effectiveness of technologies, but they do so from different perspectives and with different foci. This section delineates the application of TEA within the IPDP and TPS in aircraft design. It discusses their multi- and interdisciplinary nature and examines relevant aspects of parameter UQ.

Technology Portfolio Selection and Aircraft Design

TPSs involves modeling, evaluating, optimizing, and choosing from a broad set of potential technologies that can be integrated into aircraft design. This process often considers a dozen technologies or more, each with its own set of parameters and performance metrics. Thus, the nature of TPS is inherently multidisciplinary. Each technology considered can influence the overall system performance, and changes in one technology can affect the parameters and uncertainties of others. For example, the integration of a new propulsion system can impact the aircraft's weight, fuel efficiency, and aerodynamics, leading to cascading effects on other subsystems.

To select a portfolio of technologies that yield the best and most reliable outcome, the impacts of different technologies T_i on the output parameters k_j are quantified first, resulting in a TIM as shown in Tab. A.1 (a). Subsequently, the compatibility of the technologies with each other is tabulated in a TCM, as illustrated in Tab. A.1 (b). Finally, any the expected non-linear effects are quantified and summarized in a TSM, see Tab. A.1 (c).

Table A.1 Matrices used in TPS, based on [11].

		technologies			
		T_1	T_2	...	T_n
KPIs	k_1	2%	-	...	6%
	k_2	-	1%	...	3%
	⋮	⋮	⋮	⋱	⋮
	k_m	1%	1%	...	-

(a) Technology Impact Matrix (TIM)

		technologies			
		T_1	T_2	...	T_n
technologies	T_1	-	●	...	○
	T_2	●	-	...	●
	⋮	⋮	⋮	⋱	⋮
	T_n	○	●	...	-

(b) Technology Compatibility Matrix (TCM)

		technologies			
		T_1	T_2	...	T_n
technologies	T_1	-	.95	...	0
	T_2	.95	-	...	1.05
	⋮	⋮	⋮	⋱	⋮
	T_n	0	1.05	...	-

(c) Technology Synergy Matrix (TSM)

Due to this multi-technological and multi-disciplinary approach, UQ in TPS is a challenging task. Each technology's performance and interactions need to be modeled accurately, considering parameter uncertainties and their repercussions on the aircraft design. However, the complexity of these interactions and the high number of involved technologies make a detailed lifecycle modeling difficult, often limiting the uncertainty inclusion to those from the design domain. Detailed inclusions of lifecycle-based uncertainties and their repercussions are rarely seen in uncertainty-enabled TPS.

TEA in IPDP

As mentioned previously, TEA focuses on evaluating the economic viability and performance of a single technology within its operational environment. This approach integrates technical performance data with economic factors to provide a comprehensive evaluation of the technology's potential.

TEAs are interdisciplinary, involving inputs from various fields such as engineering, economics, and environmental science. Unlike the multidisciplinary approach of TPS, this process focuses on one technology at a time and its interactions with the environment. The integration of information from various disciplines provides insights into the lifecycle performance and cost of the technology without necessarily feeding back into the design loop. Instead, decision-makers from industrial or academic research projects, who are faced with questions about how to proceed with the IPDP, are the primary recipients of this information.

UQ in TEA is crucial for providing reliable evaluations. Focusing on a single technology allows for a more detailed examination of specific uncertainties, often involving experts from different fields to quantify these uncertainties accurately. These uncertainties can include fuel price forecasts, which impact the economic viability of the technology, weather changes that affect

operational performance, and variations in technological performance due to manufacturing tolerances or operational wear and tear.

In summary, TPS in aircraft design and TEA in IPDP are distinct yet complementary processes. While the former deals with multiple technologies in a multidisciplinary context, the latter focuses on a single technology in an interdisciplinary manner. Both approaches require careful consideration of uncertainties to ensure robust and reliable outcomes. This thesis resides in the domain of TEA, emphasizing the integration of UQ to provide comprehensive evaluations of aeronautical technologies.

A.3 Foundational Clean Sky 2 Projects

The Clean Sky 2 program was a large-scale European aeronautical research initiative under the Horizon 2020 framework, aimed at developing innovative, cutting-edge technologies to significantly reduce aviation's environmental footprint. This initiative brings together major stakeholders from the aerospace industry, including manufacturers, research institutions, and academia, to collaborate on projects that address critical challenges in aviation, such as reducing CO_2 and NO_x emissions, noise pollution, and improving fuel efficiency.

Two projects, both originating from the Clean Sky 2 program, have been foundational for the use-case and TEA described in this thesis: ECHO (from WP 1.4.1.) and HLFC-Win from (WP 1.4.4.). Both projects focus on the development and integration of the HLFC technology, each on different application areas of a large passenger aircraft. The consortium comprised Industry including Airbus, Aernnova, and SONACA as well as research institutions such as DLR, ONERA, and Fraunhofer [47].

ECHO

The project ECHO (Evaluation of a Certified HLFC Elevator Operation), conducted from July 2015 to December 2019, aimed to develop and evaluate a HLFC system for the HTP of an Airbus A350. The main objective was to design, manufacture, and test an HLFC-equipped elevator in a real-flight environment to achieve technology readiness level 5, supporting the Clean Sky 2 goals of reducing CO_2, NO_x, and noise emissions in aviation.

ECHO focused on developing a manufacturing process for HLFC purposes. This involved advanced technologies such as micro-perforation for titanium sheets, end-to-end assembly processes, and bonding and tolerance management for carbon fibre reinforced polymers and titanium. These technologies were validated through the creation of a ground-based demonstrator, a HLFC leading edge segment based on A350 geometry.

This project provided valuable parameter and modeling knowledge including, but not limited to:

- Overall aircraft lifecycle modeling
- Customer airline definition and information
- Modeling of insect contamination and cloud encounter degradation
- Incorporating cost for acquisition and maintenance

HLFC-Win

The project HLFC-Win, conducted from July 2017 to December 2023, aimed to integrate all necessary major sub modules and components such as the HLFC leading edge, the joint between the front spar and leading edge, the high-lift system design, the HLFC system itself, and the wing ice-protection system. In this project, HLFC was applied on the outer and upper wing portion of an Airbus A330 similar aircraft, complementing the research on the HTP of an A350 in ECHO. Verification and validation was conducted with a large-scale ground-based demonstrator and large-scale wind tunnel tests as well as various sub-scale demonstrators, resulting in a technology readiness level 4.

HLFC-Win aimed to address three main challenges: fulfilling HLFC-wing related performance requirements through a simplified suction system, realizing this system with an appropriate mix of materials, and developing manufacturing approaches that are cost-effective and suitable for serial production. The project employs a interdisciplinary approach, integrating aerodynamic, structural, and systems design to achieve a wing configuration that meets stringent lightweight, cost, and maintainability criteria.

In addition to ECHO, the involvement in HLFC-Win provided further insights for the TEA in this thesis, including, but not limited to:

- Off-design behavior when flying at different cruise speeds
- Additional contingency fuel carriage for HLFC
- Improved fuel projection UQ
- Use of evidence theory in combination with advisory board experts

B Supplemental Information for Chapter 2

This part of the Appendix offers supplemental insights on topics introduced or discussed in Chapter 2, such as the limitations of DOC methods, fundamental statistical terms, and interpretation aids for belief and plausibility intervals.

B.1 Limitations of DOC methods

As discussed in Section 2.1.1, DOC methods are known for their simplicity and quick computational evaluation. However, it is important to understand the limitations of DOC methods to accurately interpret their results and applicability.

The following table outlines the scope and applicability of DOC methods, highlighting the aspects that are within the scope of these methods and those that fall outside their scope. This table is based on the work of Schnieder [241] with additional insights.

Table B.1 Scope and applicability of DOC methods, based on Schnieder [241] with own additions.

In Scope	Out of Scope
Period of Consideration	
The settled state of an aircraft, i.e., between the 5^{th} and 15^{th} year of operation.	Initial and late phases as well as age-related effects, e.g., increase of maintenance cost or decrease of fuel efficiency.
Operational Content	
One representative flight connection, characterized by a fixed range and average utilization.	Different routes, regions, and flight schedules. Network and fleet compilation as well as commonality effects.
Object of Interest	
Technologies or aircraft alternatives that affect the independent variables of the CERs, e.g., maximum takeoff weight or fuel burn.	OoIs that do not solely affect these independent variables, e.g., operational procedures and maintenance strategies.
Level of (temporal) Detail	
Relationship between the average flight time and average annual utilization.	Particular and discrete temporal effects, e.g., from flight delays or unscheduled maintenance.

B.2 Fundamental Statistical Terms

This section provides definitions of key concepts in statistics and complements the foundations given in Section 2.2.2.

Expected Value

The mean value, often referred to as the expected value, is a measure of the central tendency of a random variable. For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$ and corresponding probabilities $p_1, p_2, \ldots, p_n$, the expected value $E(X)$ is defined as:

$$E(X) = \sum_{i=1}^{n} x_i p_i \tag{B.1}$$

For a continuous random variable X with probability density function $f(x)$, the expected value is given by:

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx \tag{B.2}$$

In the context of MCS, the expected value is typically approximated by the mean:

$$\hat{E}(X) = \frac{1}{N} \sum_{i=1}^{N} X_i \tag{B.3}$$

where N is the number of simulations and X_i represents the outcome of the i-th simulation.

The expected value is crucial in UQ as it provides a single summary measure of the central location of a distribution of possible outcomes. It helps in understanding the long-term average behavior of a system under uncertainty. However, it does not provide information about the variability or dispersion of the outcomes.

Variance and Standard Deviation

Variance is a measure of the spread or dispersion of a random variable's values around the mean. It quantifies the degree to which each value in the distribution differs from the mean value. For a random variable X, the variance is defined as:

$$\text{Var}(X) = E[(X - E(X))^2] \tag{B.4}$$

For a discrete random variable:

$$\text{Var}(X) = \sum_{i=1}^{n} (x_i - E(X))^2 p_i \tag{B.5}$$

For a continuous random variable:

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)\, dx \tag{B.6}$$

The standard deviation is the square root of the variance:

$$\sigma = \sqrt{\text{Var}(X)} \tag{B.7}$$

Variance and standard deviation are key metrics in UQ as they provide insights into the variability and spread of possible outcomes. A higher variance or standard deviation indicates greater uncertainty and dispersion around the mean value.

Covariance

Covariance is a measure of the joint variability of two random variables X and Y. It indicates the extent to which the two variables change together. The covariance is defined as:

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] \tag{B.8}$$

For discrete random variables:

$$\text{Cov}(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i - E(X))(y_j - E(Y)) p_{ij} \tag{B.9}$$

where p_{ij} is the joint probability of $X = x_i$ and $Y = y_j$.

Covariance is useful in UQ to understand the degree to which two variables influence each other. Positive covariance indicates that the variables tend to increase or decrease together, while negative covariance indicates an inverse relationship.

Correlation

Correlation is a normalized measure of the covariance and indicates the strength and direction of a linear relationship between two variables. The correlation coefficient ρ is defined as:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \tag{B.10}$$

where σ_X and σ_Y are the standard deviations of X and Y, respectively.

The correlation coefficient ranges from -1 to 1. A value of 1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship. Correlation is dimensionless, making it easier to interpret compared to covariance. In UQ, correlation helps identify and quantify the strength of dependencies between variables, which is crucial for multivariate risk analysis.

Law of Large Numbers

The LLN states that as the number of trials or observations increases, the sample average converges to the expected value (mean) of the population. There are two main forms of the LLN.

The weak LLN asserts that for a large number of trials, the sample average will be close to the expected value with high probability. This is known as convergence in probability. Formally, for a sequence of independent and identically distributed random variables $X_1, X_2, \ldots, X_n$ with expected value μ,

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) = 0 \tag{B.11}$$

for any $\epsilon > 0$. This means that as the number of observations n increases, the probability that the sample average deviates from the population mean by at least ϵ becomes smaller and smaller, approaching zero.

The strong LLN provides a stronger form of convergence, known as almost sure convergence. It states that the sample average almost surely (with probability 1) converges to the expected value as the number of observations goes to infinity. Formally,

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu\right) = 1 \tag{B.12}$$

This means that the sequence of sample averages $\frac{1}{n}\sum_{i=1}^{n} X_i$ will converge to the population mean μ for almost every possible sequence of outcomes (with probability 1).

The LLN has profound implications for statistical practices. It ensures that as sample sizes grow, the sample mean becomes a more accurate estimate of the population mean, thereby reducing variability and underpinning the reliability of statistical inference. This principle is foundational in constructing confidence intervals and conducting hypothesis tests, providing a theoretical basis for using sample data to make inferences about a population.

When data is scarce, the assumptions and limitations of the LLN become particularly significant. Its reliability hinges on having a large number of observations. With small samples, the sample mean may not approximate the population mean well, leading to unreliable estimates and inferences. This increased variability can cause the sample mean to be heavily influenced by outliers or random fluctuations, making the results less stable.

To mitigate these issues, efforts should be made to collect as much data as possible. When data collection is challenging or expensive, acknowledging the limitations in any analysis or conclusions drawn is crucial. Results derived from small samples should be interpreted with caution, considering the potential for increased variability and bias, and avoiding overgeneralizing findings.

Central Limit Theorem

The CLT describes the behavior of the sum (or average) of a large number of independent, identically distributed random variables. It states that, regardless of the original distribution of the variables, the distribution of the sample mean will approach a normal distribution as the sample size becomes large.

Formally, let $X_1, X_2, \ldots, X_n$ be a sequence of independent, identically distributed random variables with mean μ and finite variance σ^2. The sample mean $\overline{X}_n$ is given by

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{B.13}$$

Formally, the CLT states:

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1) \tag{B.14}$$

This means that for a sufficiently large sample size, the distribution of $\overline{X}_n$ is approximately normal with mean μ and variance σ^2/n.

The implications of the CLT are highly relevant for statistical practices. Because the sample mean $\overline{X}_n$ is approximately normally distributed for large n, even if the underlying data are not, the CLT allows for applying methods that assume normality.

Similarly to LLN, the assumptions and limitations of the CLT become significant. In situations with limited data, the approximation to normality may be poor, especially if the original data are heavily skewed or have heavy tails. Consequently, statistical tests and confidence intervals based on the normal approximation may be unreliable, leading to incorrect conclusions. Small samples may not capture the variability of the population adequately, resulting in biased estimates.

B.3 Interpretation of Belief and Plausibility Intervals

Belief and plausibility intervals are fundamental concepts in evidence theory. To reiterate the discussion from section 2.2.2, belief represents the minimum probability that an event will occur based on available evidence, while plausibility represents the maximum probability that an event is not contradicted by the evidence. The difference between these two measures indicates the level of epistemic uncertainty.

The table below presents various example intervals of belief and plausibility, along with their interpretations. These interpretations help in understanding how different combinations of

Bel(A) and Pl(A) reflect different levels of certainty and uncertainty about the occurrence of event A. The examples are adapted from Worden [287] and illustrate scenarios ranging from complete certainty to total ignorance.

Table B.2 Example uncertainty intervals in DSTE and their interpretation, adapted from Worden [287].

[Bel(A), Pl(A)]	**Interpretation**
[0.0, 0.0]	It is believed with certainty that A will not occur. There is no epistemic uncertainty since Bel(A) = Pl(A).
[1.0, 1.0]	It is believed with certainty that A will occur. There is no epistemic uncertainty since Bel(A) = Pl(A).
[0.8, 0.8]	There is a strong belief and high plausibility that A will occur. The epistemic uncertainty is zero but there is a small (probabilistic) chance that A will not occur.
[0.2, 0.2]	There is a slight belief and slight plausibility that A will occur. The epistemic uncertainty is zero but there is a high (probabilistic) chance that A will not occur.
[0.0, 1.0]	There is total ignorance or uncertainty regarding the occurrence of A. There is high epistemic uncertainty due to the large difference between Bel(A) and Pl(A).
[0.3, 1.0]	Available evidence does not contradict A from occurring. There is some epistemic uncertainty due to the difference between Bel(A) and Pl(A).
[0.0, 0.6]	Available evidence does not support the occurrence of A. There is also some level of doubt due to Pl(A) < 1. There exists some level of epistemic uncertainty.
[0.4, 0.6]	Some of the available evidence supports the occurrence of A. In addition, other pieces of evidence contradict A from occurring. The epistemic uncertainty is moderate.

C Supplemental Information for Chapter 3

This part of the Appendix offers supplemental information on topic from Chapter 3, including an error estimate of underlying the mission simulation tool, the determination of the reference aircraft, as well as a list of assumptions for the deterministic study.

C.1 Accuracy of the Mission Simulation Tool

As discussed in Section 3.1.2, LYFE utilizes response surfaces with data calculated using a commercial mission simulation tool, namely Piano-X [247]. To estimate the accuracy of Piano-X, particularly for the A330-343 aircraft, a comparison with literature values was performed. These values, sourced from Aircraft Commerce [5], involve two city-pairs: Los Angeles International (LAX) to La Guardia, New York (LGA), with a track distance of 2188 NM. The literature accounted for a tailwind of 20 knots on the eastbound route (LAX-LGA), resulting in an equivalent air distance of 2100 NM. Conversely, the westbound route assumed a headwind of 50 knots, yielding an equivalent air distance of 2420 NM. The second city-pair distance involved LAX to Stockholm Arlanda airport (ARN), with a track distance of 4900 NM and corresponding equivalent air distances of 4850 NM (LAX-ARN) and 5160 NM (ARN-LAX). Piano-X was then executed for these missions, matching these distances and documented payloads. The results of this comparison are summarized in Table C.1. The maximum discrepancy observed across the four distances, under identical conditions, was 62 kg or 0.21 %.

Table C.1 Comparison of block fuel calculation of A330-343 between Piano-X and literature [5, p. 18].

route	**distance**	**block fuel**		**difference**	
		literature	Piano-X	absolute	relative
	[NM]	[kg]	[kg]	[kg]	[%]
LGA-LAX	2420	28840	28902	+62	+0.21
LAX-LGA	2100	25066	25105	+39	+0.15
LAX-ARN	4850	60557	60573	+16	+0.03
ARN-LAX	5160	64976	64990	+14	+0.02

C.2 Reference Aircraft Definition

This Section complements the information about the reference aircraft provided in Section 3.3.2 including the flight profile and maintenance schedule.

In order for the flight schedule to be considered *representative*, a profile has to be chosen, so that the aircraft's utilization is close to the average of all A330-300. Because it is a highly versatile aircraft, different operators use the A330 on very different ranges. On the shorter end, airlines such as Asia Pacific, China Airlines, and Philippine Airlines operate their aircraft with average FH/FC times of 2 to 3 h. Long-Range operators include Lufthansa, SAS, Air Canada, and Qantas, with average FH/FC of 5 to 8 h. After a selection of Lufthansa as the operating airline, their A330 fleet was analyzed for its average flight times, yielding six potential aircraft. Subsequently, the route networks of these six aircraft from the year 2019 were extracted[1] and analyzed for compatibility with LYFE's flight schedule input. Although all aircraft were

[1] Using FlightRadar24.com

operated on very similar routes, the tail sign D-AIKI has shown the least amount of errors in the data (which occur, for instance, due to diversions or ADSB signal errors in certain regions). Therefore, the route network of this aircraft, which is plotted and summarized in Fig. 3.8, was chosen for the simulation. Due to the nature of the LYFE framework, the simulated flights are defined by their origin, destination, and great circle distance, i.e., detailed routing procedures are neglected. Furthermore, the flight performance is not calculated on an one by one basis, but instead determined using the available response surfaces and the multi-linear interpolation within it, as described in Section 3.1.2.

Furthermore, the application on an A330-300 type of aircraft defines the underlying maintenance schedule, consisting of line, base, heavy component, and unscheduled maintenance, which is also explained in Section 3.3.2. Both the flight and maintenance schedule are kept identical for both aircraft and are needed for a realistic simulation of the operators' and aircraft-dependent cash flow.

Table C.2 Complete flight schedule for both aircraft fed to LYFE.

entry	route information				durations				load factors	
	origin	destin.	dist.	freq.	flight	taxi in	turnar.	taxi out	pax	cargo
	[−]	[−]	[NM]	[%]	[hrs]	[hrs]	[hrs]	[hrs]	[%]	[%]
1	AUS	FRA	8537	110	10.42	0.5	2.0	0.5	80	0
2	FRA	AUS	8537	110	10.42	0.5	2.0	0.5	80	0
3	FRA	IAH	8412	388	10.27	0.5	2.0	0.5	80	0
4	IAH	FRA	8412	384	10.27	0.5	2.0	0.5	80	0
5	DFW	FRA	8267	158	10.10	0.5	2.0	0.5	80	0
6	FRA	DFW	8267	154	10.10	0.5	2.0	0.5	80	0
7	FRA	SEA	8207	357	10.02	0.5	2.0	0.5	80	0
8	SEA	FRA	8207	360	10.02	0.5	2.0	0.5	80	0
9	FRA	WDH	8120	1	9.92	0.5	2.0	0.5	80	0
10	WDH	FRA	8120	1	9.92	0.5	2.0	0.5	80	0
11	DEN	FRA	8098	33	9.90	0.5	2.0	0.5	80	0
12	FRA	DEN	8098	33	9.90	0.5	2.0	0.5	80	0
13	FRA	YVR	8072	424	9.87	0.5	2.0	0.5	80	0
14	YVR	FRA	8072	424	9.87	0.5	2.0	0.5	80	0
15	FRA	MLE	7893	42	9.65	0.5	2.0	0.5	80	0
16	MLE	FRA	7893	42	9.65	0.5	2.0	0.5	80	0
17	FRA	MIA	7772	241	9.52	0.5	2.0	0.5	80	0
18	MIA	FRA	7772	244	9.52	0.5	2.0	0.5	80	0
19	FRA	MAA	7602	5	9.32	0.5	2.0	0.5	80	0
20	MAA	FRA	7602	5	9.32	0.5	2.0	0.5	80	0
21	ATL	FRA	7418	223	9.10	0.5	2.0	0.5	80	0
22	FRA	ATL	7418	221	9.10	0.5	2.0	0.5	80	0
23	BLR	FRA	7410	331	9.08	0.5	2.0	0.5	80	0
24	FRA	BLR	7410	280	9.08	0.5	2.0	0.5	80	0
25	FRA	ORD	6979	2	8.58	0.5	2.0	0.5	80	0
26	ORD	FRA	6979	2	8.58	0.5	2.0	0.5	80	0
27	DTW	FRA	6687	133	8.25	0.5	2.0	0.5	80	0
28	FRA	DTW	6687	135	8.25	0.5	2.0	0.5	80	0
29	FRA	LAD	6571	205	8.12	0.5	2.0	0.5	80	0
30	LAD	FRA	6571	198	8.12	0.5	2.0	0.5	80	0
31	BOM	FRA	6576	312	8.12	0.5	2.0	0.5	80	0
32	FRA	BOM	6576	234	8.12	0.5	2.0	0.5	80	0
33	FRA	IAD	6558	340	8.10	0.5	2.0	0.5	80	0
34	IAD	FRA	6558	340	8.10	0.5	2.0	0.5	80	0
35	FRA	NBO	6324	672	7.82	0.5	2.0	0.5	80	0
36	NBO	FRA	6324	685	7.82	0.5	2.0	0.5	80	0
37	FRA	YYZ	6351	188	7.85	0.5	2.0	0.5	80	0

(continued on next page)

Table C.2 Complete flight schedule for both aircraft fed to LYFE (continued).

entry	route information				durations				load factors	
	origin	destin.	dist.	freq.	flight	taxi in	turnar.	taxi out	pax	cargo
	[−]	[−]	[NM]	[−]	[hrs]	[hrs]	[hrs]	[hrs]	[%]	[%]
38	YYZ	FRA	6351	188	7.85	0.5	2.0	0.5	80	0
39	EWR	FRA	6218	674	7.70	0.5	2.0	0.5	80	0
40	FRA	EWR	6218	583	7.70	0.5	2.0	0.5	80	0
41	FRA	JFK	6196	447	7.67	0.5	2.0	0.5	80	0
42	JFK	FRA	6196	358	7.67	0.5	2.0	0.5	80	0
43	DEL	FRA	6127	24	7.58	0.5	2.0	0.5	80	0
44	FRA	DEL	6127	24	7.58	0.5	2.0	0.5	80	0
45	BOS	FRA	5896	428	7.32	0.5	2.0	0.5	80	0
46	FRA	BOS	5896	448	7.32	0.5	2.0	0.5	80	0
47	FRA	YUL	5861	20	7.28	0.5	2.0	0.5	80	0
48	YUL	FRA	5861	20	7.28	0.5	2.0	0.5	80	0
49	ALA	FRA	5095	117	6.38	0.5	2.0	0.5	80	0
50	FRA	ALA	5095	131	6.38	0.5	2.0	0.5	80	0
51	ALA	TSE	952	83	1.53	0.5	2.0	0.5	80	0
52	FRA	LOS	4862	370	6.12	0.5	2.0	0.5	80	0
53	LOS	FRA	4862	374	6.12	0.5	2.0	0.5	80	0
54	DXB	FRA	4849	565	6.10	0.5	2.0	0.5	80	0
55	FRA	DXB	4849	599	6.10	0.5	2.0	0.5	80	0
56	TSE	FRA	4314	82	5.47	0.5	2.0	0.5	80	0
57	ABV	FRA	4569	149	5.77	0.5	2.0	0.5	80	0
58	FRA	ABV	4569	150	5.77	0.5	2.0	0.5	80	0
59	BAH	FRA	4445	7	5.62	0.5	2.0	0.5	80	0
60	FRA	BAH	4445	4	5.62	0.5	2.0	0.5	80	0
61	DMM	FRA	4372	56	5.53	0.5	2.0	0.5	80	0
62	FRA	DMM	4372	74	5.53	0.5	2.0	0.5	80	0
63	FRA	TSE	4314	98	5.47	0.5	2.0	0.5	80	0
64	FRA	RUH	4297	388	5.45	0.5	2.0	0.5	80	0
65	RUH	FRA	4297	361	5.45	0.5	2.0	0.5	80	0
66	FRA	JED	4136	30	5.27	0.5	2.0	0.5	80	0
67	JED	FRA	4136	30	5.27	0.5	2.0	0.5	80	0
68	FRA	KWI	4025	346	5.13	0.5	2.0	0.5	80	0
69	KWI	FRA	4025	315	5.13	0.5	2.0	0.5	80	0
70	FRA	IKA	3778	18	4.85	0.5	2.0	0.5	80	0
71	IKA	FRA	3778	18	4.85	0.5	2.0	0.5	80	0
72	BEY	FRA	2841	2	3.75	0.5	2.0	0.5	80	0
73	TSE	ALA	952	98	1.53	0.5	2.0	0.5	80	0
74	LOS	SSG	675	89	1.22	0.5	2.0	0.5	80	0
75	SSG	LOS	675	88	1.22	0.5	2.0	0.5	80	0
76	ABV	SSG	606	1	1.13	0.5	2.0	0.5	80	0
77	SSG	ABV	606	2	1.13	0.5	2.0	0.5	80	0
78	ABV	PHC	446	122	0.95	0.5	2.0	0.5	80	0
79	PHC	ABV	446	122	0.95	0.5	2.0	0.5	80	0
80	BAH	KWI	421	40	0.92	0.5	2.0	0.5	80	0
81	KWI	BAH	421	43	0.92	0.5	2.0	0.5	80	0
82	BAH	RUH	421	352	0.92	0.5	2.0	0.5	80	0
83	RUH	BAH	421	350	0.92	0.5	2.0	0.5	80	0
84	DMM	KWI	356	125	0.83	0.5	2.0	0.5	80	0
85	KWI	DMM	356	142	0.83	0.5	2.0	0.5	80	0
86	FRA	MUC	299	93	0.77	0.5	2.0	0.5	80	0
87	MUC	FRA	299	130	0.77	0.5	2.0	0.5	80	0
88	BAH	DMM	86	78	0.52	0.5	2.0	0.5	80	0
89	DMM	BAH	86	73	0.52	0.5	2.0	0.5	80	0

Table C.3 Complete maintenance schedule for both aircraft (excluding HLFC maintenance), based on Aircraft Commerce [5, 6].

entry	name	general type	downtime	intervals FH	intervals FC	intervals time	fiscal year	total cost
	[−]	[−]	[hrs]	[hrs]	[−]	[days]	[−]	[US$]
1	Transit	L	0.5	0	1	0	2008	120
2	Daily	L	1.0	0	0	1	2008	330
3	Weekly	L	2.0	0	0	7	2008	920
4	A1	L	7.5	600	0	0	2008	36,000
5	A2	L	8.1	1200	0	0	2008	39,500
6	A3	L	9.4	1800	0	0	2008	44,750
7	A4	L	10	2400	0	0	2008	48,250
8	A5	L	7.5	3000	0	0	2008	36,000
9	A6	L	8.1	3600	0	0	2008	39,500
10	A7	L	8.8	4200	0	0	2008	43,000
11	A8	L	10	4800	0	0	2008	48,250
12	C1	B	72	0	0	540	2008	215,000
13	C2	B	89	0	0	1080	2008	420,000
14	C3	B	103	0	0	1620	2008	235,000
15	C4	B	523	0	0	2160	2008	1,450,000
16	C5	B	72	0	0	2700	2008	235,000
17	C6	B	88	0	0	3240	2008	500,000
18	C7	B	72	0	0	3780	2008	255,000
19	C8	B	720	0	0	4320	2008	2,205,000
20	Tire Inspect.	H	2.0	0	230	0	2008	1000
21	Tire Retreat	H	2.1	0	300	0	2008	5700
22	Tire Replace	H	2.1	0	1200	0	2008	11,600
23	Brakes	H	4.8	0	1100	0	2008	320,000
24	Landing Gear	H	10	0	7000	0	2008	900,000
25	ThrustRevs	H	5.8	0	6000	0	2008	430,000
26	APU	H	4.4	0	4000	0	2008	275,000
27	ESV 1	H	6.0	30,100	4300	0	2012	4,750,000
28	ESV 2	H	6.0	53,200	7600	0	2012	5,600,000
29	ESV 3	H	6.0	75,600	10,800	0	2012	4,700,000
30	ESV 4	H	6.0	98,000	14,000	0	2012	5,600,000

Abbreviations L: line maintenance, B: base maintenance, H: heavy component maintenance.

C.3 List of Assumptions for the Deterministic Study

For transparency and reproducibility purposes, this Section lists all assumptions for the deterministic study documented in Section 3.3. The assumptions are categorized into general, design, operational, environmental, and economic.

General Assumptions

- The TEA is performed with LYFE on aircraft level in a prospective and comparative manner, using the difference in fuel consumption $\Delta\gamma_{\text{fuel}}$ as a technological KPI (see Eqn. 3.5) and the difference in NPV as an economic KPIs (see Eqn. 3.6).
- The simulation lifetime is set to 20 years, starting from the year 2030, as described in Section 3.3.2.
- The reference aircraft is an Airbus A330-343 with Rolls-Royce Trent 772 engines as described in Table 3.1. A novel baseline aircraft design with assumed technological improvements that are expected for the considered entry into service was not used, as described on p. 57.

- The HLFC aircraft is a derivative of the reference aircraft without a comprehensive redesign, as described on p. 58.
- HLFC is assumed to be incorporated on the upper side of the wings (outboard of the engines) and both sides of the HTPs and VTP, as shown in Fig. 3.7.
- The customer airline is Lufthansa, which defined the operational flight profile shown in Fig. 3.8 and tabulated in Table C.2. The maintenance schedule for both aircraft (excluding HLFC-specific maintenance efforts) is shown in Table C.3.
- The fuel burn calculation uses a multi-linear interpolation from response surfaces as described in Section 3.1.2.

Design Assumptions

- The drag reduction potential of the HLFC aircraft on overall aircraft level is 7.4 % and includes the loss of laminarity due to turbulent wedges between the leading edge panels as described on p. 60 and shown in Fig. 3.9.
- The mass difference between the HLFC aircraft and the turbulent reference is 744 kg and was obtained with the help of literature values and scaled using the suction area as a driving parameter, as shown in Fig. 3.10.
- The SFC penalty due to the operation of HLFC is 1.18 % and includes electrical efficiencies for the compressors, inverters, generators, and power lines, as described on p. 61 and Eqn. 3.7.
- The drag reduction and SFC penalty affect only the cruise portion (including the step climbs) only, where the HLFC system is turned on. The mass difference affects the fuel burn throughout the entire mission.

Operational Assumptions

- The cruise speed is assumed to be constant at Ma = 0.82, rendering potential losses of laminarity due to varying cruise speeds irrelevant for this study, as described in Section 3.3.2.
- When operating the HLFC aircraft, it is assumed that the operators carry an HLFC-specific contingency fuel of 2.5 % of the trip fuel, as derived in Section 3.3.2.
- For both aircraft, it is assumed that the payload is comprised of passengers only (i.e., no additional cargo other than the personal belongings of the passengers), with a passenger load factor of 80 %, as described in Section 3.3.2 and listed in Table C.2.

Environmental Assumptions

- The HLFC aircraft is assumed to lose 35 % of its laminar efficacy on the tails due to insect contamination, while the efficacy on the wing application remains unaffected, as described in Section 3.3.2.
- Cloud encounter is assumed to degrade the HLFC efficacy by 7 % for all HLFC application areas. This results in an overall laminar efficacy of 82 %, as described in Eqn. (3.11).
- The turbulent reference aircraft not affected by insect contamination or cloud encounter.

Economic Assumptions

- The future fuel price is considered using time-series projections, ranging from $0.79/kg in 2030 to $0.94/kg in 2050, as shown in Fig. 3.11.
- The maintenance cost due to the HLFC system is derived using literature values and the mass increase of 744 kg as a driving parameter, resulting in $16.8k per year. Unscheduled maintenance cost are considered using LYFE's default mechanism, as described in Section 3.3.2.
- The increase in aircraft price for the HLFC aircraft (compared to the reference aircraft) is derived similarly and results to $1.3M, as described in Section 3.3.2. The capital cost mechanism assumes that the aircraft are bought and hence comprise ordering payments, delivery payments, and distributed payments in between. Insurance and interest cost are also modeled.

D Supplemental Information for Chapter 4

This part of the Appendix offers supplemental information on topic from Chapter 4, covering aspects regarding the sampling sequence of the non-variance-based GSA method from Morris, a comparative analysis of the investigated GSA methods when inputs are (artificially) correlated, the required modification of the Python package SALib, and a list of assumptions and uncertainties regarding the applied GSA.

D.1 Sampling Sequence of Morris

As described in Section 4.1.2, the Morris method uses its own specific sampling approach, which uses trajectories to explore the input space. This section explains the sampling methodology and the impact of the key parameters: the number of trajectories n_r, and the number of levels p. To do so, a parameter variation for $u = 2$ input variables X_1 and X_2 with n_r and p being increased in the columns and rows, respectively. The output is shown in Fig. D.1.

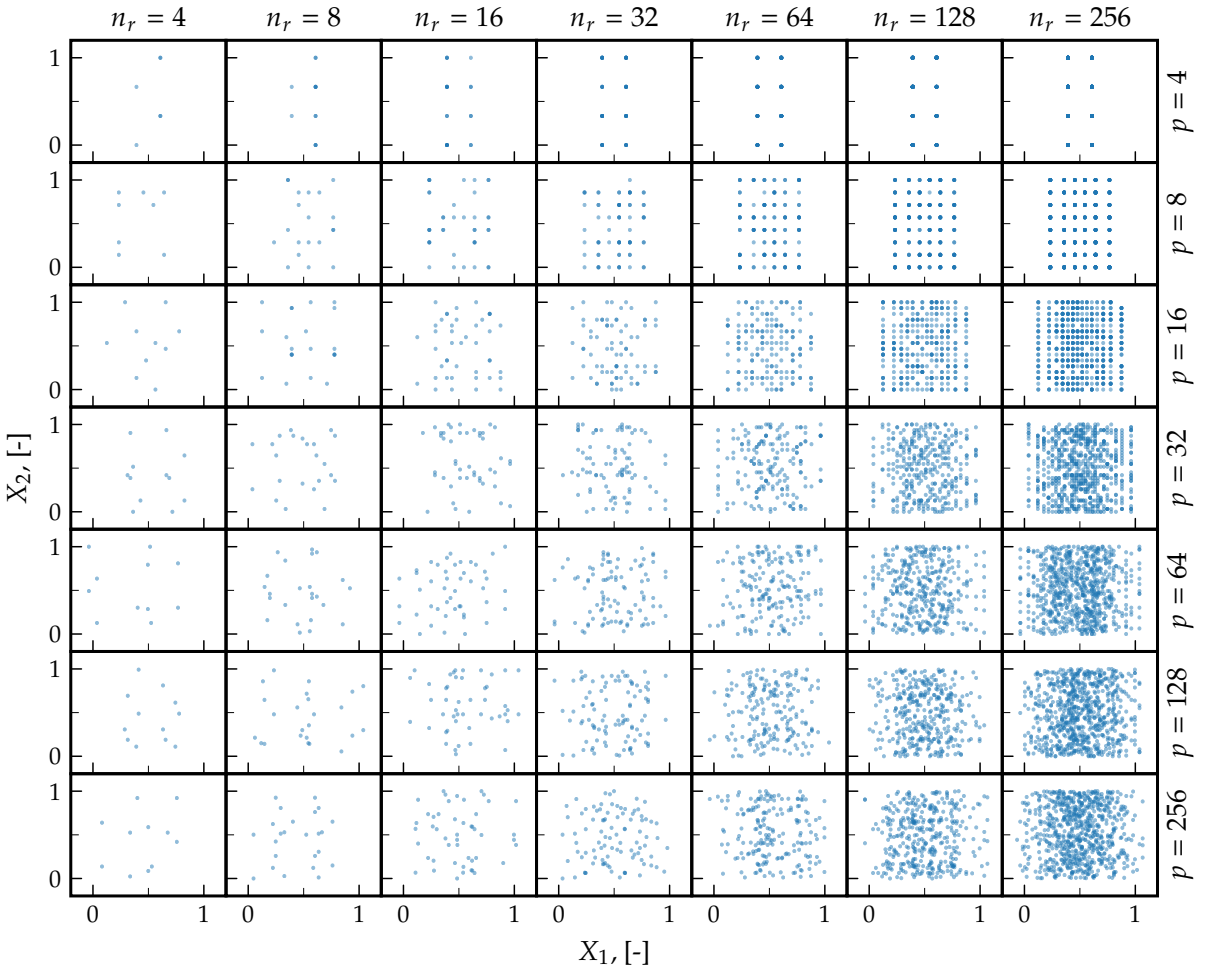
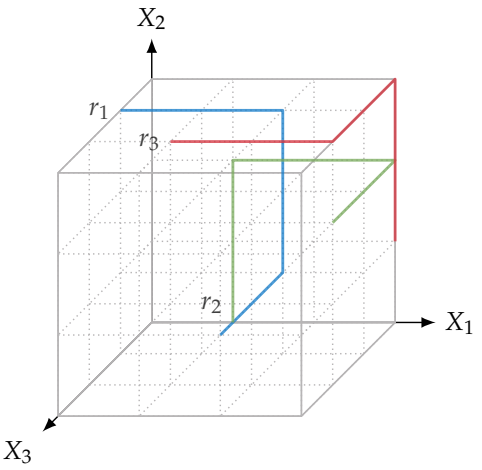


Figure D.1 Morris' sample space of a $u = 2$ parameter model where $X_1 \sim \mathcal{N}(0.5, 0.25)$ and $X_2 \sim \mathcal{U}$ and a variation of the number of trajectories (columns) and number of levels (rows) by 2^x for $x = 2, 3, \ldots, 8$.

The influence of the two varied parameters can be summarized as follows. As the number of trajectories n_r increases from left to right, the density of sampling points within the input space also increases. With only a few trajectories (e.g., $n_r = 4$), the sample points are sparse, potentially missing significant patterns, such as those stemming from the marginal distributions. Increasing n_r to higher values (e.g., $n_r = 256$) ensures a more thorough exploration, revealing the underlying normal and uniform marginal distributions from X_1 and X_2, respectively. Similarly, as the number of levels p increases from top to bottom, the granularity of the input space discretization improves. Lower values of p (e.g., $p = 4$) result in a coarser grid, which might overlook subtle effects. Higher values of p (e.g., $p = 256$) create a finer grid, enabling the method to detect more nuanced influences of input variables.

The patterns in Figure D.1 show how the combination of n_r and p affects the sampling distribution. Sparse trajectories with low p levels leave significant gaps, while higher values for both parameters provide comprehensive coverage, essential for robust sensitivity analysis.

Figure D.2 complements the Morris method's supplemental material by visualizing sampling trajectories in a three-dimensional unit cube for a model with three input variables ($u = 3$). Each trajectory starts at a random point and varies one input variable at a time, ensuring balanced exploration of the input space. The table in Figure D.2 details the sampling points for each trajectory, illustrating how this method isolates the effect of each input variable.



(a) Visualization in the unit cube.

Trajectory	Samples		
	X_1	X_2	X_3
r_1	0	1	1/3
	2/3	1	1/3
	2/3	1/3	1/3
	2/3	1/3	1
r_2	1/3	0	0
	1/3	2/3	0
	1	2/3	0
	1	2/3	2/3
r_3	1/3	1	2/3
	1	1	2/3
	1	1	0
	1	1/3	0

(b) Sample space of each trajectory.

Figure D.2 Examples of trajectories used in the Morris sampling method with $u = 3$, $p = 4$ and $\Delta = 2/3$.

Figures D.1 and D.2 together highlight the Morris method's rigorous approach in GSA. Adjusting n_r and p allows researchers to balance sampling density, computational cost, and analysis precision. This flexibility makes the Morris method a powerful tool for identifying key input variables and their interactions in complex models. However, determining the optimal combination of parameters requires a thorough understanding of the sampling approach, necessitating an investigative preprocessing step.

D.2 GSA with Correlated Inputs

This Section complements the comparative GSA study described in Section 4.2, investigating the effects of input parameter correlation, an aspect that is often ignored but can make a big difference when present [64, p. 2,4]. As Saltelli et al. [234] explained, their consideration is not trivial as many conventional methods make use of assumptions and simplifications which are only valid when input parameters are independent. The purpose of this study is to determine if the techniques that promise to capture correlation can, in fact, do so, and what happens if an independence assuming method is fed with correlated samples.

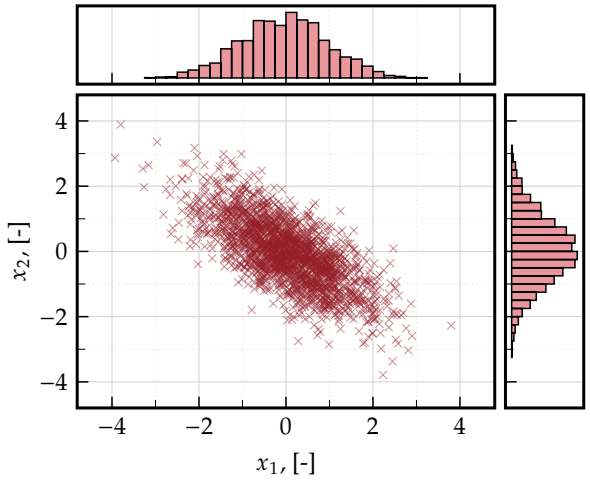
Parameter Dependencies and Sampling

While the bounds and individual distributions of the uncertain parameters are kept identical to those of the previous analyses (see Table 4.2), two of them are selected and made dependent: namely, the SFC factor t_s and the price of the technology p_{tech}. This dependency is modeled as a negative correlation; that is, high values of p_{tech} are more likely to occur with low values of t_s. A possible interpretation could be as follows: The higher the price of a technology, the more research and development effort was made, resulting in the technology likely being more performant.

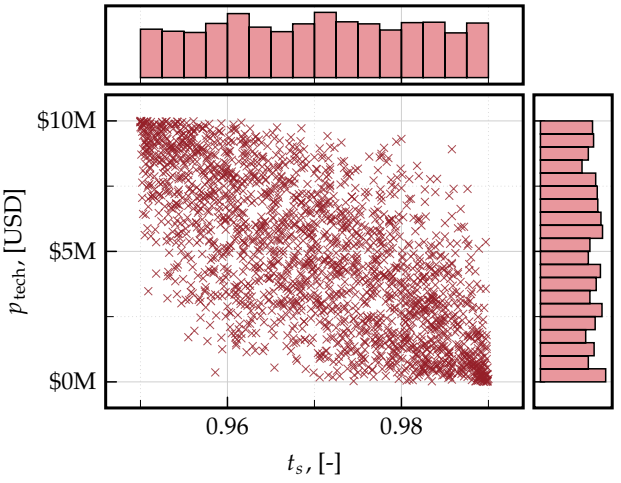
To maintain comparability with the previous analyses, the marginal distributions of t_s and p_{tech} are kept uniform. To create the samples, a bivariate standard normal distribution was first generated, whose probability density function is defined as:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[x_1^2 - 2\rho x_1 x_2 + x_2^2\right]\right) . \tag{D.1}$$

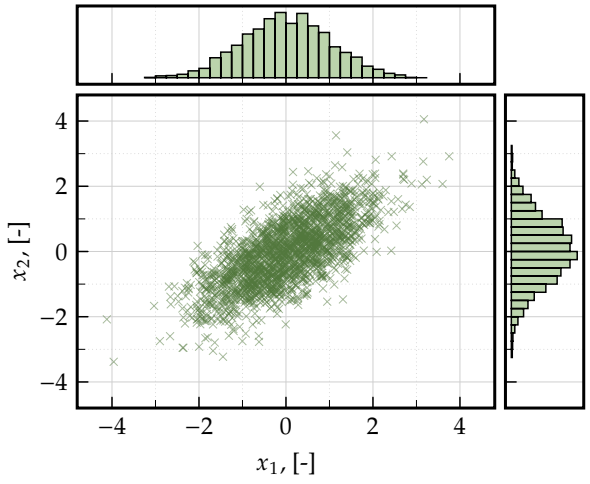
The marginal distributions of Eqn. (D.1) are, by definition, standard normal distributions themselves. This is illustrated with a correlation coefficient of $\rho = -0.7$ in Fig. D.3 (a). To transform them to be uniform, the inverse probability integral transformation method, described later



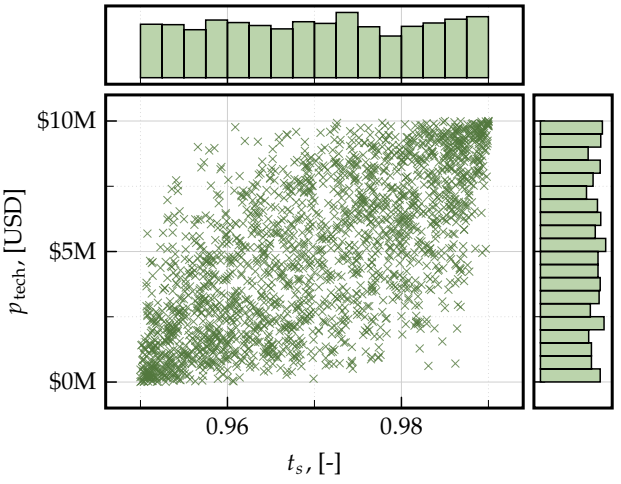
(a) bivariate standard normal distribution with $\rho = -0.7$.



(b) resulting desired sample space for $\rho = -0.7$.



(c) bivariate standard normal distribution with $\rho = +0.7$.



(d) resulting desired sample space for $\rho = +0.7$.

Figure D.3 Samples used to create a bivariate distribution with uniform marginals with positive (red) and negative (green) correlation.

(see Appendix 4.2.1), was adapted for use in the two-dimensional space. The resulting modified bivariate distribution, shown in Fig. D.3 (b), displays uniform marginal distributions as desired. For the purpose of demonstrating the GSA capabilities with both positive and negative correlations, the process was repeated with $\rho = +0.7$, and the respective sample spaces are shown in Fig. D.3 (c) and (d).

Because of this sampling strategy, only GSA methods which support given data can be utilized. These are SCSA, DMIM, and FAST-RBD. The former two are expected to capture the correlation correctly, whereas the latter requires independent samples. Therefore, this study will demonstrate the impact of violating this assumption, knowingly or not. The sample size was set to $N = 2^{15}$ to ensure that all methods are converged[1].

Results

The results of the SCSA method are depicted in Fig. D.4 (a) with negative correlation, no correlation, and positive correlation in red (bottom plot), blue (center plot), and green (top plot), respectively. For each uncertain input parameter, three bars are shown. The left one (solid) represents S_a being the uncorrelated measurement of sensitivity, the middle one (hatched), S_b, being the correlation portion, and the right one (colored and hatched) showing $S_c = S_a + S_b$.

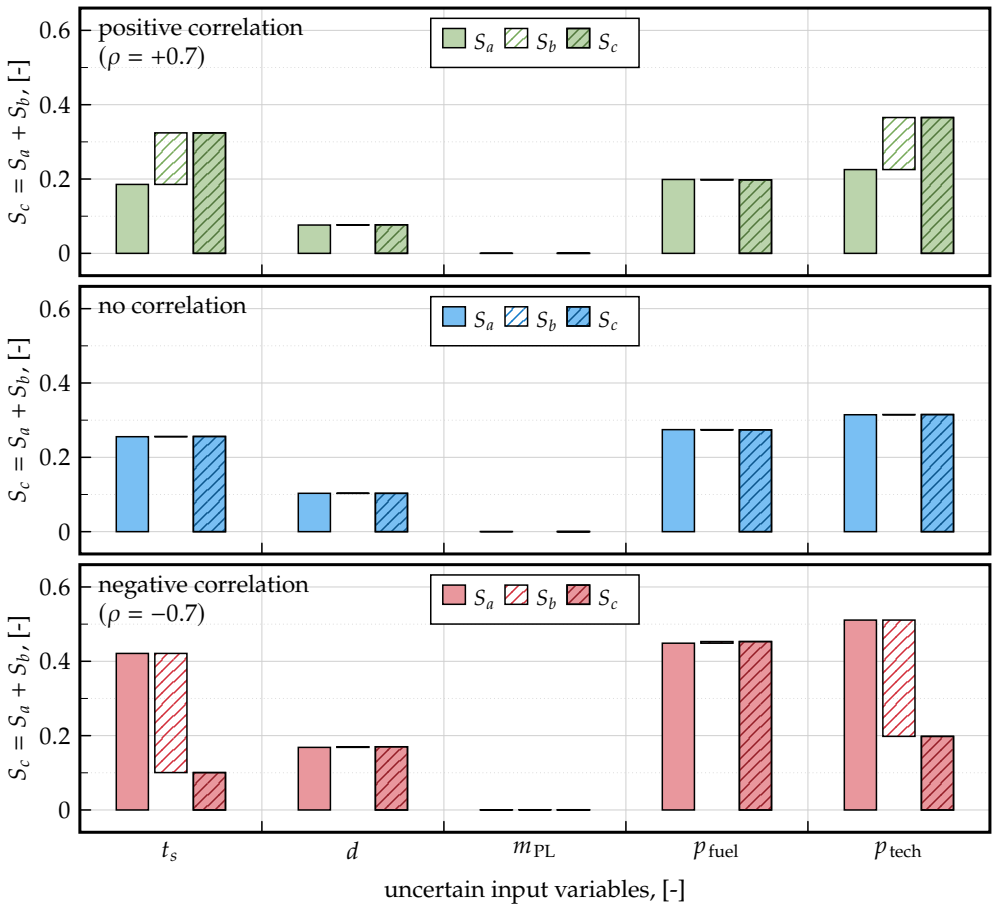
The results in the no correlation case (center) are identical to those shown in the independent analysis (see Fig. 4.2). Apart from minor numerical noise, the method correctly quantifies S_b as zero, i.e., no correlation has been identified. For the case with positive correlation (top), S_b is nonzero for the two correlated parameters t_s and p_{tech} and remains zero for the remaining ones, verifying the positive correlation capture capability. In this case, correlation increases S_c for the correlated parameters. This is reversed in the case of negative correlation (bottom), where the value of S_b correctly shows negative values for t_s and p_{tech}, verifying this case as well. It is important to note that the ranking, when it is done using S_a, remains identical for each correlation case. For $\rho = +0.7$, the bar representing S_b seems to be half as big as for $\rho = -0.7$ which is an indicator of a non-linear model response. However, S_c is the only sensitivity measure which can be interpreted as the ratio of parameter caused variance over total variance, as it is the only one for which

$$\sum S_c = 1 \tag{D.2}$$

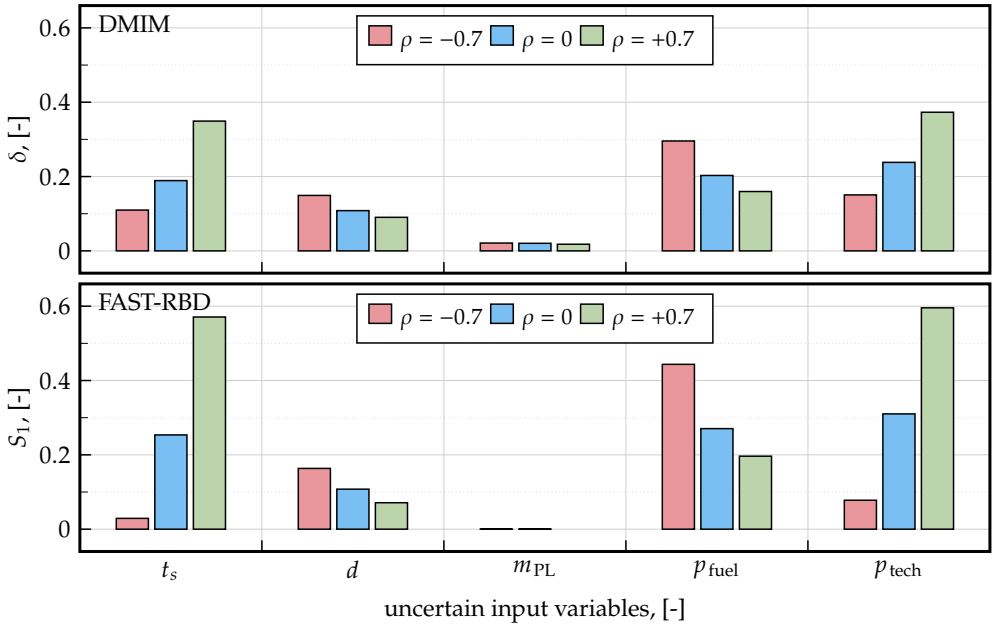
holds. Therefore, it is advised to always take into account the full extent of SCSA results.

The response of the DMIM and FAST-RBD techniques to the same inputs are shown in Fig. D.4 (b). Here, each bar represents one correlation case. There seems to be, apart from the absolute value of the sensitivity measures, virtually no difference between both methods, which is surprising as DMIM is said to capture correlation, whereas FAST-RBD necessitates independence for the sensitivity measures to be meaningful. For both methods, the sensitivity measures of the correlated parameters grow for higher values of ρ, while those of the other parameters decrease, although their samples did not change with correlation. For FAST-RBD, the positive correlation increased the sensitivity measure for t_s and p_{tech} so much that the sum of all S_1 exceeds 1. Therefore, it is concluded that these methods not only cannot capture correlation, but their use when it is present leads to erroneous and misleading results.

[1] In fact, the previous analysis has shown that the DMIM and FAST-RBD techniques may not be converged for values below $N = 2^{14}$, see Table 4.4.



(a) Response of the SCSA method for different cases of correlation.



(b) Response of the DMIM and FAST-RBD methods for different cases of correlation.

Figure D.4 Comparative GSA analysis for different cases of correlation.

D.3 Implementation (and Modification) of SALib

As mentioned before, the open source Python package SALib [105] is used for the GSA in Chapter 4. SALib itself is designed to be non-intrusive and its implementation hence does not interfere with the mathematical or computational model (i.e., FastLYFE in this case). More explicitly, SALib takes care of the sample generation, which may or may not be GSA method specific, as well as breaking down the total output uncertainty into sensitivity measures. For the former task, the package provides uniform, triangular, normal, and lognormal distributions per default. As actual data can be present in all forms of distributions, the sampling approach may need to be modified. This modification is based on the inverse probability integral transformation method and makes use of the following fact: Assume a random variable X with a CDF termed F_X. Regardless of the distribution of X, the output of $Y = F_X(X)$ has a uniform distribution on $[0, 1]$. This process can be inversed. That is, given a random variable Y which has a uniform distribution on $[0, 1]$, the execution of the inverse CDF F_X^{-1} on Y results in an output that has the same distribution as X. This is illustrated in Fig. D.5 with X having a standard normal distribution.

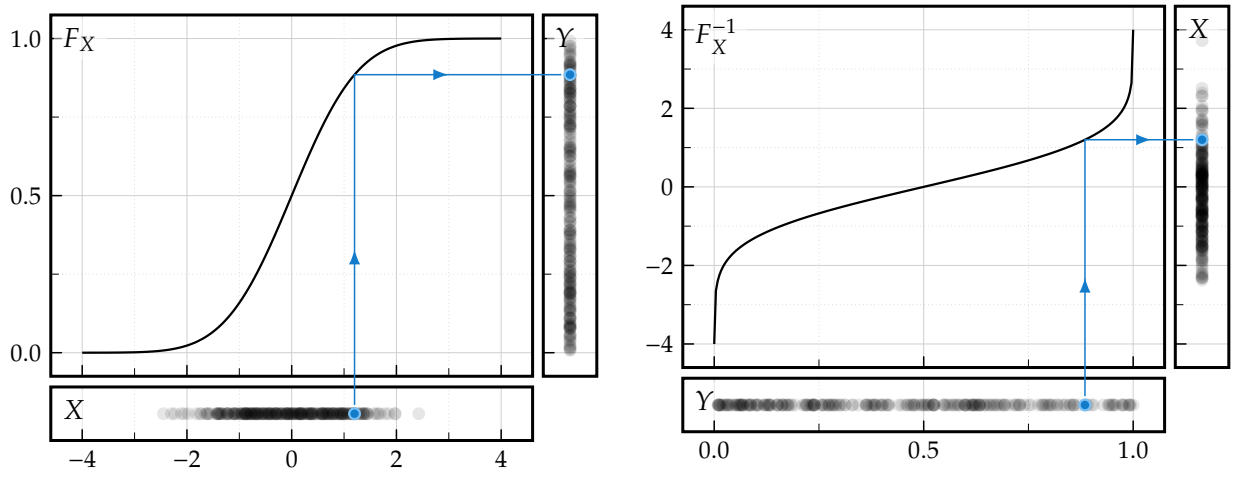
For the present GSA investigation, the first step is to select a distribution F_X that is appropriate for each uncertain variable. Assuming one or more of these is not supported by SALib, the next step is to define the bounds $X_{\min}$ and $X_{\max}$ of the desired distribution. These need to be transformed to the respective bounds in the $[0, 1]$ domain, i.e.,

$$Y_{\min} = F_X(X_{\min}) \quad \text{and} \quad Y_{\max} = F_X(X_{\max}). \tag{D.3}$$

Subsequently, these bounds are used to create the uniformly distributed samples

$$Y \sim \mathcal{U}(Y_{\min}, Y_{\max}) \tag{D.4}$$

together with all other uncertain inputs using the GSA method specific sampler. This ensures the correct sequencing and input space preparation which some of the methods require. Next, Y is transformed with the inverse CDF of the initially selected distribution. The new samples, now having the desired distribution F_X, are then fed to the model in the same manner as the other uncertain parameters.



(a) forward: from X to Y using $F_X(X)$.

(b) inverse: from Y to X using $F_X^{-1}(Y)$.

Figure D.5 Illustration of the inverse probability integral transformation method with the random variables $X \sim \mathcal{N}(0, 1)$ and $Y \sim \mathcal{U}(0, 1)$ and the (inverse) CDFs F_X and F_X^{-1}.

D.4 Assumptions and Uncertainties for the GSA study

Table D.1 shows the overview of the uncertainties considered in the GSA in Chapter 4. It should be noted that all are treated with ProT and are assumed to be independent.

Table D.1 Summary of uncertain inputs for GSA with parameters for reproducibility.

domain		**general Information**			**scipy parameters**		
parameter	Unit	Type*	distribution	Bounds	`loc`	`scale`	`shape`
design							
$\mathcal{D}_1$: drag reduction	[%]	E	uniform	$[6.0, 8.6]$	–	–	–
$\mathcal{D}_2$: mass increase	[kg]	E	uniform	$[384, 1104]$	–	–	–
$\mathcal{D}_3$: SFC penalty	[%]	E	gen. normal	$[0.73, 1.73]$	1.213	0.424	10.137
operation							
$\mathcal{O}_1$: cruise mach	[-]	A	gen. normal	$[0.79, 0.85]$	0.819	0.0140	1.4023
$\mathcal{O}_2$: extra cont. fuel†	[-]	E	uniform	$[0, 1]$	–	–	–
$\mathcal{O}_3$: load factor	[-]	E	gen. logistics	$[0.66, 0.90]$	0.835	0.0158	0.3968
environment							
$\mathcal{E}_1$: insect contamin.‡	[-]	E	normal	$[0, 1]$	0.5	0.1667	–
$\mathcal{E}_2$: cloud encounter	[-]	A	exponential	$[0, 0.6]$	25e-6	0.093	–
economic							
$\mathcal{M}_1$: fuel price§	[-]	A	skewnormal	$[0, 1]$	0.19	0.390	2.6
$\mathcal{M}_2$: maint. increase	[\$/yr]	E	nagasaki	[14k, 82k]	14,060	29,626	0.8955
$\mathcal{M}_3$: price increase	[\$]	E	nagasaki	[300k, 2.1M]	311,033	800,129	1.0467

*E: epistemic, A: aleatory; †see Eqn. (4.22); ‡see Eqn. (4.23); §see Eqn. (4.27).

E Supplemental Information for Chapter 5

This part of the Appendix offers supplemental information on topic from Chapter 5, covering aspects such as the PCE-based surrogate model of LYFE, biographies of the interviewed experts, and a brief description of the investigated quasi-random sampling schemes.

E.1 Polynomial Chaos Expansion for LYFE

As it was discussed in Section 5.4.4, quantifying the uncertainty with evidence theory requires a very large number of executions to be conclusive. With more than 250,000 runs for a converged result, neither the original lifecycle model LYFE, nor its mechanistic surrogate FastLYFE allow an uncertainty-enabled simulation with reasonable execution times. To mitigate this, an analytical PCE-based surrogate model was developed, whose execution time is a fraction of that of FastLYFE.

General Considerations

Creating a PCE-based surrogate for a computationally expensive model can be done in multiple ways. For the creation of LYFE's PCE surrogate model, the regression-based non-intrusive technique is used due to its simplicity. Code E.1 shows a minimal example with the Ishigami function[1] (lines 7 and 8) introduced earlier as the original model. Note that the majority of the underlying PCE algorithms is provided by the imported `chaospy` package (line 1).

The overall process is as follows. After defining the number of samples (n = 10,000 in this example, line 11) and the polynomial order ($p = 5$, line 12), the input distributions are defined (lines 16-18). Here, x_1 and x_2 are selected to be uncertain, following a normal and uniform distribution, respectively.The third parameter, x_3, is fixed at $\pi/4$. Afterwards, the joint distribution J of x_1 and x_2 is created (line 19) and sampled from (line 23). These samples are then fed to the original model while recording the outputs y (lines 24-26). Using the polynomial order p and the joint distribution J, the polynomial basis for the PCE is then generated (line 30), for which the coefficients subsequently are determined using least squares regression (line 31), resulting in the surrogate model `sm`, which can be evaluated using the same syntax as the original function.

PCE Surrogate

For the LYFE surrogates, the original model comprises two runs of LYFE, one with the reference case, and one with the laminar aircraft, resulting in one scalar output for each KPI, i.e., $\Delta\eta_{\text{fuel}}$ and ΔNPV. Hence, two surrogate models are created, one for each KPI. The uncertain input parameters are the same as those in the GSA performed in Chapter 4 (see Table D.1 on page 173). The MCS was fed with 32,768 samples, which are used, in combination with the MCS outputs, to generate the surrogate model. The polynomial order p was varied from 2 to 6 and finally selected to be $p = 3$ as this order had the lowest Root Mean Square Error (RMSE) in the test data[2].

A qualitative perspective on the accuracy of the surrogate model is shown in Fig. E.1, where the predicted data (i.e., generated by the surrogate model) is plotted over the observed data (i.e., generated by LYFE). Fig. E.1 (a) shows that the surrogate model seems to both overestimate and

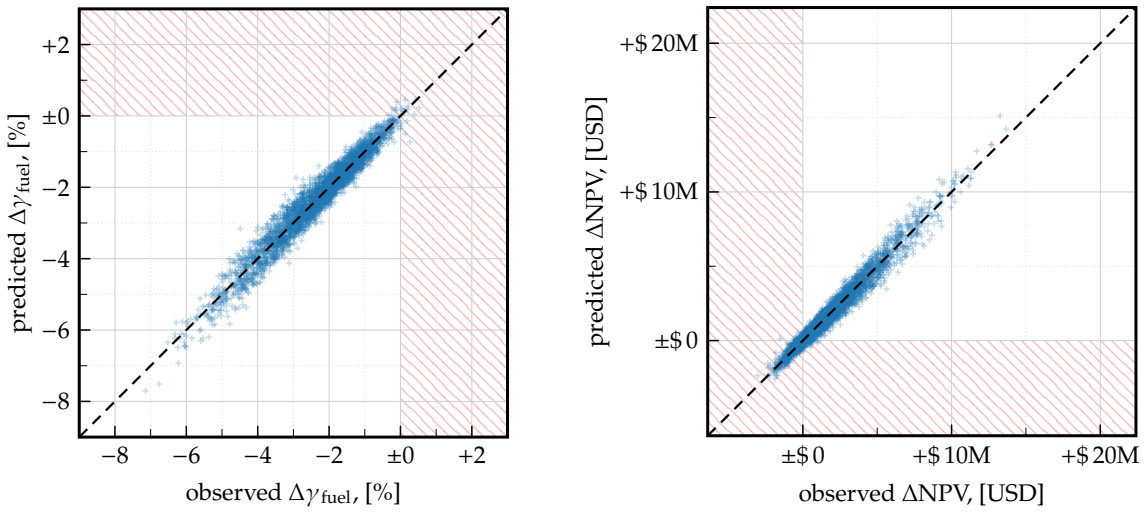
[1] The function was proposed by Ishigami et al. in 1990 and was designed to have a highly non-linear, non-monotonic response, making it useful for testing the performance of uncertainty quantification and sensitivity analysis methods.

[2] 80 % of the data were used to train the model and the remaining 20 % are used for the error estimation.

Code E.1 Minimal example for creating a PCE-based surrogate

```python
import chaospy as cp
import numpy as np

# Step 0: Preparatory Work
# ------------------------
# Defining the model (which is the Ishigami function in this example)
def blackboxmodel(x1, x2, x3):
    return np.sin(x1) + 7*(np.sin(x2))**2 + (0.3*x3**4)*np.sin(x1)

# Setting up further parameters
n = 10000                                       # Number of samples to create
p = 5                                           # polynomial order

# Step 1: Defining the input distributions
# ----------------------------------------
X1 = cp.Normal(mu=0, sigma=np.pi/3)             # first uncertain variable
X2 = cp.Uniform(-np.pi, np.pi/3)                # second uncertain variable
X3 = np.pi/4                                    # this is fixed
J = cp.J(X1, X2)                                # joint distribution

# Step 2: Running the Monte-Carlo Simulation
# ------------------------------------------
samples = J.sample(rule='sobol', size=n)        # Generate samples
y = np.zeros(n)
for idx in range(n):
    y[idx] = blackboxmodel(samples[0, idx], samples[1, idx], X3)

# Step 3: Create the polynomial basis and determine the coefficients
# ------------------------------------------------------------------
expansion = cp.generate_expansion(order=p, dist=J)  # Generate expansion
sm = cp.fit_regression(expansion, samples, y)

# Evaluate output exemplarily
print(sm(np.pi/8, -np.pi/2, X3))                # prints 7.36
print(blackboxmodel(np.pi/8, -np.pi/2, X3))     # prints 7.43
```

underestimate the $\Delta\eta_{\text{fuel}}$ values, leading to a RMSE of 0.128 %. The ΔNPV surrogate model shown in Fig. E.1 (b) has a similar, albeit slightly better performance with a RMSE of $157k. In conclusion, neither surrogate model is perfect, but considering the significant reduction in computational runtime, their errors are accepted for the studies in this Chapter. Note that this prediction error is not included in the overall UQ effort, as commonly done in most PCE applications [178].



(a) operational KPI. **(b)** economic KPI.

Figure E.1 Results of the PCE-based surrogate model (*predicted* on the ordinate) vs. the original model (*observed* on the abscissa).

E.2 Interviewed Experts

Roger Taplin Roger Taplin began his career as a British Aircraft Corporation apprentice in 1964. He has extensive experience in system development, having worked on the Concorde Engine Air Intake Control System, VAAC Harrier digital flight control system, A320 Gust Loads Alleviation System, and A340/A330 airframe systems concepts and development. In 1987, he became the UK Chief System Engineer and Head of the Aircraft Systems department. By 1993, he was the UK Chief Engineer for FLA, later becoming the A400M Wing Chief Engineer. From 2006 to 2014, he served as the Senior Aircraft Architect – Wings, where he was responsible for all aspects of wings, including leading future aircraft concept studies such as A3xx, which led to Wing of Tomorrow research, ensuring the satisfactory integration of new projects like A320 Sharklet wing tip device, A350, and BLADE – laminar flow test aircraft using A340 MSN 1, conducting architectural assessments of all wing research studies including Laminar Flow drag reduction, and leading task forces investigating wing and landing gear in-service problems.

Heinz Hansen Heinz Hansen received his scientific education at the University of Bochum and the RWTH Aachen, where he graduated as a Diplom-Ingenieur in Aeronautics. He worked for nearly 35 years at Airbus in the field of Flight Physics and Research and Technology, contributing to various aircraft projects with a focus on aerodynamic design, wind tunnel, and flight testing. For nearly 20 years, he specialized in laminar flow technology, participating in numerous national and European research projects. During his last eight years at Airbus, he led all laminar flow and riblet research activities within the internal project TOP-Low Drag Aircraft. Hansen is the inventor of several patents and the author of many publications, recognized as an authority on laminar flow. Since his retirement in 2018, he remains very active in various technology topics.

Thomas Kilian Thomas Kilian has been a scientific employee at the DLR Institute of Aerodynamics and Flow Technology since 2010, focusing on aerodynamic research with a significant emphasis on maintaining laminar flow and exploring HLFC structures and concepts. Since 2016, he has concentrated on novel HLFC designs, enhancing the aerodynamic efficiency of aircraft. His contributions span various national and European projects, such as those under the Clean Sky 2 initiative, aimed at developing environmentally friendly and fuel-efficient aircraft technologies. Kilian's work includes the development of tubeless suction systems and inductive heating for HLFC systems, crucial for reducing drag and improving fuel efficiency in long-range passenger aircraft. His expertise in both experimental and numerical methods has been showcased in numerous national and international conferences, making him a well-regarded figure in the field of advanced aerodynamics.

Alexander Bismark Alexander Bismark is a System Engineer with a research focus on the assessment and design of various innovative aircraft systems at an early development stage. In 2018, he became involved in the Clean Sky 2 projects dealing with HLFC and was mainly responsible for the suction system development, from evaluating solution candidates to the specific component design and testing. Moreover, he was the focal point for technical discussions at the overall project level in HLFC-Win.

Tim Effing Since early 2018, Tim Effing has been employed as a scientific assistant at the Institute of Aerospace Systems at RWTH Aachen, a position he held until the end of April 2024. He has been researching HLFC for a total of six years, with the last three years focusing on the combination of HLFC & variable camber, which is also the subject of his dissertation to be submitted in 2024. He is a project partner in the LuFo V-3 project AVACON (Advanced Aircraft Concepts) and project leader in the German LuFo VI-1 project CATeW (Coupled Aerodynamic Technologies for Aircraft Wings). His research primarily evaluates the potential of individual and combined technologies at the aircraft pre-design level.

E.3 Investigated Sampling Schemes

Exploration and analysis of complex datasets in the fields of data science and statistics require careful consideration of sampling techniques. The selection of suitable sampling schemes plays a crucial role in obtaining representative samples at smaller sample sizes, thereby potentially improving convergence behavior. This Section provides a concise overview of a variety of sampling strategies commonly used in data science and engineering. Random sampling, Latin-Hypercube sampling, Sobol sequence sampling, additive recursion sampling, Halton sampling, and Hammersley sampling are among the sampling methods discussed. Each of these techniques provides distinct strategies for systematically selecting samples from the data space, thereby ensuring exhaustive coverage and minimizing biases. Understanding the advantages and disadvantages of these sampling schemes is essential for practitioners to make informed decisions about their sampling methodologies, thereby facilitating robust analysis and inference in a variety of applications.

Random Sampling Scheme

Random sampling is a widely used approach in which samples are selected randomly from the entire parameter space. This technique is simple to implement and provides a representative sampling of the design space. However, since the samples are chosen randomly, there is a possibility of clustering or sparse distribution of samples, which may not efficiently explore the entire design space.

Latin-Hypercube Sampling Scheme

Latin-hypercube sampling is a stratified sampling technique that ensures a more even distribution of samples across the parameter space. It partitions the design space into equally sized intervals and randomly selects one sample from each interval. This method reduces the clustering effect and enhances the coverage of the design space, making it suitable for exploring multi-dimensional problems.

Sobol Sequence Sampling Scheme

The Sobol sequence sampling scheme employs low-discrepancy quasi-random numbers to generate samples. It is designed to provide a more even distribution of points throughout the parameter space compared to random sampling. The Sobol sequence is a deterministic sequence that exhibits low discrepancy properties, ensuring a better coverage of the design space with fewer samples.

Additive Recursion Sampling Scheme

The additive recursion sampling scheme is another method based on quasi-random numbers. It involves using a series of arithmetic operations to generate a sequence of points that fills the parameter space in a uniform manner. By iteratively adding different offsets to the initial seed point, the resulting sequence exhibits a more regular distribution, minimizing clustering and enhancing coverage.

Halton Sampling Scheme

The Halton sampling scheme is a deterministic quasi-random sampling technique that generates sequences of points based on prime numbers. It provides a more even distribution of samples across the parameter space compared to random sampling. The Halton sequence uses different prime numbers as bases for each dimension, resulting in an improved exploration of the design space.

Hammersley Sampling Scheme

The Hammersley sampling scheme is a stratified low-discrepancy technique that generates points using a combination of a regular grid and the Van der Corput sequence. It is particularly effective for sampling in higher dimensions. The Hammersley sequence ensures a more uniform coverage of the design space compared to random sampling, reducing the clustering effect and enhancing the accuracy of the sampling process.

F Supplemental Information for Chapter 6

This last part of the Appendix offers supplemental information regarding the matrices for uncertainty combination as discussed in Chapter 6.

F.1 Matrices for Belief-Based Uncertainty Combination

In Section 6.2, the DSTE-based method for uncertainty combination was introduced. The nested MCS, illustrated in Fig. 6.2, samples the epistemic uncertainties in the inner loop and the aleatory samples in the outer loop. Each iteration then executes the LYFE surrogate with one set of epistemic and aleatory parameters, resulting in one y_{e_i,a_j} value for ΔNPV and $\Delta\gamma_{\text{fuel}}$. Once both loops are finished, the outputs are structured in a matrix as shown in Eqn.(6.1), which is then used to calculate the CCXF collectives.

In order to shed more light on the matrix creation and CCXF derivation, this Section illustrates these steps using a simplified example. Consider the following model:

$$y = f(\alpha_1, \alpha_2, \epsilon_1, \epsilon_2, \epsilon_3) = \alpha_1 + 2\alpha_2 - \epsilon_1 + \epsilon_2 \cdot \epsilon_3 \quad . \tag{F.1}$$

Here, it is assumed that α_1 and α_2 are aleatory uncertainties, while ϵ_1, ϵ_2, and ϵ_3 are of epistemic nature. For simplicity reasons, all uncertainties in this example are modeled using a uniform distribution. The respective bounds are as follows:

- α_1 (aleatory): $\mathcal{U}(1, 5)$
- α_2 (aleatory): $\mathcal{U}(0.5, 2.5)$
- ϵ_1 (epistemic): $\mathcal{U}(0, 3)$
- ϵ_2 (epistemic): $\mathcal{U}(0.1, 1)$
- ϵ_3 (epistemic): $\mathcal{U}(1, 4)$

In the first step of the nested MCS, one aleatory sample set is created, resulting in one sample vector for α_1 and one for α_2, each with N_A entries:

$$\vec{\alpha_1} = \begin{bmatrix} 2.498 \\ 4.803 \\ 3.928 \\ \vdots \\ 1.432 \end{bmatrix} \quad \text{and} \quad \vec{\alpha_2} = \begin{bmatrix} 0.563 \\ 1.773 \\ 1.129 \\ \vdots \\ 2.060 \end{bmatrix} \tag{F.2}$$

Next, the epistemic uncertainties ϵ_1, ϵ_2, and ϵ_3 are sampled, resulting in one sample vector with N_E entries for each parameter:

$$\vec{\epsilon_1} = \begin{bmatrix} 1.926 \\ 0.252 \\ 0.485 \\ \vdots \\ 0.256 \end{bmatrix} \quad , \quad \vec{\epsilon_2} = \begin{bmatrix} 0.147 \\ 0.578 \\ 0.587 \\ \vdots \\ 0.779 \end{bmatrix} \quad \text{and} \quad \vec{\epsilon_3} = \begin{bmatrix} 1.309 \\ 3.708 \\ 2.516 \\ \vdots \\ 3.959 \end{bmatrix} \quad . \tag{F.3}$$

Before executing the model, the aleatory and epistemic samples have to be combined. For the first iteration (where $a_j = a_1$ and $e_i = e_1$), the sample set is thus as follows:

$$S_{e_1,a_1} = [\alpha_1, \alpha_2, \epsilon_1, \epsilon_2, \epsilon_3] = [2.498, 0.563, 1.926, 0.147, 1.309] \quad . \tag{F.4}$$

The resulting output for this sample set is

$$y_{e_1,a_1} = 1.89 \quad . \tag{F.5}$$

The next execution of the model only foresees a change in the epistemic parameters, i.e., $e_i = e_2$ while $a_j = a_1$, resulting in the second sample set and output value:

$$S_{e_2,a_1} = [\alpha_1, \alpha_2, \epsilon_1, \epsilon_2, \epsilon_3] = [2.498, 0.563, 0.252, 0.578, 3.708] \tag{F.6}$$

$$y_{e_2,a_1} = 5.515 \tag{F.7}$$

This is repeated until the predefined epistemic sample size N_E is reached, which results in the first column of the overall output matrix of Eqn. (6.1):

$$S_{e_i,a_1} = \begin{bmatrix} 2.498 & 0.563 & 1.926 & 0.147 & 1.309 \\ 2.498 & 0.563 & 0.252 & 0.578 & 3.708 \\ 2.498 & 0.563 & 0.485 & 0.587 & 2.516 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2.498 & 0.563 & 0.256 & 0.779 & 3.959 \end{bmatrix} \Rightarrow y_{e_i,a_1} = \begin{bmatrix} 1.890 \\ 5.515 \\ 4.616 \\ \vdots \\ 6.452 \end{bmatrix} \tag{F.8}$$

The second column of the output matrix is then obtained with a new aleatory sample set, which is represented by the next iteration in the outer (aleatory) loop of the nested MCS. In matrix notation, this is:

$$S_{e_i,a_1} = \begin{bmatrix} 4.803 & 1.773 & 1.926 & 0.147 & 1.309 \\ 4.803 & 1.773 & 0.252 & 0.578 & 3.708 \\ 4.803 & 1.773 & 0.485 & 0.587 & 2.516 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 4.803 & 1.773 & 0.256 & 0.779 & 3.959 \end{bmatrix} \Rightarrow y_{e_i,a_2} = \begin{bmatrix} 6.615 \\ 10.240 \\ 9.341 \\ \vdots \\ 11.177 \end{bmatrix} \tag{F.9}$$

Note how all values of α_1 and α_2 are row-wise identical in Eqns. (F.8) and (F.9). This ensures that the variation in each output column y_{e_i,a_1} and y_{e_i,a_2} represents the epistemic uncertainty. This is necessary for the subsequent CCXF calculation, which calculates one CCBF, CCDF, and CCPF for each column of the output matrix

$$y = \begin{bmatrix} 1.890 & 6.615 & 4.452 & \cdots & 3.818 \\ 5.515 & 10.24 & 8.077 & \cdots & 7.443 \\ 4.616 & 9.341 & 7.178 & \cdots & 6.544 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 6.452 & 11.177 & 9.014 & \cdots & 8.380 \end{bmatrix} \tag{F.10}$$

The collection of CCBFs, denoted with y_{CCBF}, can be expressed as a matrix as well. It holds

one CCBF in each column:

$$y_{\mathrm{CCBF}} = \begin{bmatrix} 1.000 & 1.000 & 1.000 & \cdots & 1.000 \\ 1.000 & 1.000 & 0.972 & \cdots & 1.000 \\ 1.000 & 0.948 & 0.972 & \cdots & 0.898 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0.000 & 0.000 & 0.000 & \cdots & 0.000 \end{bmatrix} \tag{F.11}$$

The collectives y_{CCDF} and y_{CCPF} are calculated in the same column-wise manner. The calculation of the percentiles, which are used to visualize the result of the DSTE-based combinatory UQ, is performed by looping through y_{CCXF} from row to row and finding the respective percentile over all columns. This ensures that the distribution of percentiles reflects the aleatory uncertainty, whereas the CCXFs themselves represent the epistemic uncertainty.

F.2 Matrices for Hybrid Uncertainty Combination

The hybrid-theoretic uncertainty combination, introduced in Section 6.3, also foresees a nested MCS, as illustrated in Fig. 6.8. The creation of aleatory and epistemic samples and their combination is identical to the previously described DSTE-based approach. Picking up the example from Section F.1 and the resulting output matrix from Eqn. (F.10), the hybrid-theoretic uncertainty combination proceeds to calculate the marginal average vectors as defined in Eqns. (6.7) and (6.8) as follows:

$$y = \begin{bmatrix} 1.890 & 6.615 & 4.452 & \cdots & 3.818 \\ 5.515 & 10.240 & 8.077 & \cdots & 7.443 \\ 4.616 & 9.341 & 7.178 & \cdots & 6.544 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 6.452 & 11.177 & 9.014 & \cdots & 8.380 \end{bmatrix} \Rightarrow y_E = \begin{bmatrix} 4.138 \\ 7.763 \\ 6.684 \\ \vdots \\ 8.700 \end{bmatrix} \tag{F.12}$$

$$\Rightarrow y_A = \begin{bmatrix} 3.499 & 8.224 & 6.061 & \cdots & 5.427 \end{bmatrix}^T \tag{F.13}$$

This representation shows how each entry in y_A is calculated by averaging one column of y. As y holds the aleatory variation from column to column, the vector y_A represents (epistemically averaged) aleatory uncertainty. In a similar manner, the calculation of y_E represents the (aleatory averaged) epistemic uncertainty. By plotting y_A over y_E, the scatter plots from Figs. 6.9 and 6.10 can be obtained.

Bibliography

[1] R. K. Abdelhalim Boussabaine. *Whole Life-Cycle Costing*. John Wiley & Sons, May 8, 2008. 264 pp. ISBN: 9780470759158.

[2] R. Adams. *If you build it will they come? three steps to test and validate any market opportunity. three steps to test and validate any market opportunity*. Wiley, 2010, p. 215. ISBN: 9780470563632.

[3] H. Agarwal et al. "Uncertainty Quantification Using Evidence Theory in Multidisciplinary Design Optimization." In: *Reliability Engineering & System Safety* 85.1-3 (July 2004), pp. 281–294. DOI: 10.1016/j.ress.2004.03.017.

[4] Airbus S.A.S. *A330 Aircraft Characteristics Airport And Maintenance Planning*. Ed. by Customer Services, Technical Data Support and Services. June 1, 2020.

[5] Aircraft Commerce. "Owners and Operators Guide: A330-200/-300." In: *The Journal for Commercial Aircraft Business* (57 Apr. 2008): *Wave off? The dilemma of replaing 200- to 300 seat widebodies*.

[6] Aircraft Commerce. "Owners and Operators Guide: Rolls-Royce Trent Family." In: *The Journal for Commercial Aircraft Business* (83 Aug. 2012): *The graceful queen: Ageing 747-400 base maintenance requirements*.

[7] F. Akram and D. N. Mavris. "Surrogate Modeling Based Uncertainty Quantification In Subject Matter Expert Systems." In: *55th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, Jan. 2014. DOI: 10.2514/6.2014-0218.

[8] F. Akram and D. N. Mavris. "Uncertainty Propagation in Technology Valuation Process." In: *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, Jan. 2017. DOI: 10.2514/6.2017-0975.

[9] F. Akram, M. Prior, and D. Mavris. "An Improved Methodology for Gas Turbine Technology Portfolio Planning, Including Technology Synergy Matrices and Real Options Analysis." In: *49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*. American Institute of Aeronautics and Astronautics, Jan. 2011. DOI: 10.2514/6.2011-578.

[10] F. Akram, M. A. Prior, and D. N. Mavris. "An Improved Approach to Technology Portfolio Prioritization Process Under Uncertainty." In: *Volume 1: Aircraft Engine; Ceramics; Coal, Biomass and Alternative Fuels; Wind Turbine Technology*. GT2011. ASMEDC, Jan. 2011. DOI: 10.1115/gt2011-46156.

[11] M. F. Akram. "A methodology for uncertainty quantification in quantitative technology valuation based on expert elicitation." PhD thesis. Georgia Institute of Technology, Mar. 28, 2012.

[12] R. Ali and O. Al-Shamma. "A Comparative Study of Cost Estimation Models Used for Preliminary Aircraft Design." In: *Global Journal of Research In Engineering* (2014).

[13] D. Allaire and K. Willcox. "Surrogate Modeling for Uncertainty Assessment with Application to Aviation Environmental System Models." In: *AIAA Journal* 48.8 (Aug. 2010), pp. 1791–1803. DOI: 10.2514/1.j050247.

[14] D. Allaire et al. "Uncertainty Quantification of an Aviation Environmental Toolsuite." In: *Reliability Engineering & System Safety* 126 (June 2014), pp. 14–24. DOI: 10.1016/j.ress.2014.01.002.

[15] S. Altavilla, F. Montagna, and M. Cantamessa. "A Multilayer Taxonomy of Cost Estimation Techniques, Looking at the Whole Product Lifecycle." In: *Journal of Manufacturing Science and Engineering* 140.3 (Dec. 2017). DOI: 10.1115/1.4037763.

[16] S. Amaral et al. "A Decomposition-based Uncertainty Quantification Approach for Environmental Impacts of Aviation Technology and Operation." In: *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 31.3 (Aug. 2017), pp. 251–264. DOI: 10.1017/s0890060417000154.

[17] M. Antivachis et al. "Novel High-speed Turbo Compressor with Integrated Inverter for Fuel Cell Air Supply." In: *Frontiers in Mechanical Engineering* 6 (Feb. 2021). DOI: 10.3389/fmech.2020.612301.

[18] P. C. Arcara, D. W. Bartlett, and L. A. McCullers. "Analysis for the Application of Hybrid Laminar Flow Control to a Long-range Subsonic Transport Aircraft." In: *SAE Technical Paper Series*. SAE International, Sept. 1991. DOI: 10.4271/912113.

[19] Y. Asiedu and P. Gu. "Product Life Cycle Cost Analysis: State of the Art Review." In: *International Journal of Production Research* 36.4 (Apr. 1998), pp. 883–908. DOI: 10.1080/002075498193444.

[20] Association of European Airlines. *Long Range Aircraft - AEA Requirements*. 1989.

[21] A. Babucke. *Kuenstliches Schlierenbild des Umschlags laminar-turbulent in einer Plattengrenzschicht, berechnet mittels DNS*. (CC BY 3.0 DE), Adjustments made. Sept. 19, 2008. URL: https://upload.wikimedia.org/wikipedia/commons/a/a1/Dns_schlierenimage.png (visited on 01/10/2021).

[22] H.-R. Bae, R. V. Grandhi, and R. A. Canfield. "Epistemic Uncertainty Quantification Techniques Including Evidence Theory for Large-scale Structures." In: *Computers & Structures* 82.13-14 (May 2004), pp. 1101–1112. DOI: 10.1016/j.compstruc.2004.03.014.

[23] H.-R. Bae, R. V. Grandhi, and R. A. Canfield. "Uncertainty Quantification of Structural Response Using Evidence Theory." In: *AIAA Journal* 41.10 (Oct. 2003), pp. 2062–2068. DOI: 10.2514/2.1898.

[24] S. H. Barr et al. "Bridging the Valley of Death: Lessons Learned from 14 Years of Commercialization of Technology Education." In: *Academy of Management Learning & Education* 8.3 (Mar. 2009), pp. 370–388. DOI: 10.1109/emr.2014.6758259.

[25] T. Bayes and R. Price. "An Essay Towards Solving a Problem in the Doctrine of Chances." In: *Philosophical Transactions of the Royal Society of London* 53 (Dec. 1763), pp. 370–418. DOI: 10.1098/rstl.1763.0053.

[26] C. Beierle and G. Kern-Isberner. *Methoden Wissensbasierter Systeme*. Springer Fachmedien Wiesbaden, 2019. DOI: 10.1007/978-3-658-27084-1.

[27] L. Bertsch, B. Schäffer, and S. Guérin. "Uncertainty Analysis for Parametric Aircraft System Noise Prediction." In: *Journal of Aircraft* 56.2 (Mar. 2019), pp. 529–544. DOI: 10.2514/1.c034809.

[28] N. Bilal. "Implementation of Sobol's Method of Global Sensitivity Analysis to a Compressor Simulation Model." In: *International Compressor Engineering Conference*. West Lafayette, Indiana, 2014.

[29] Boeing. *Hybrid Laminar Flow Control Study - Final Technical Report*. Tech. rep. NASA-CR-165930. Boeing Commercial Airplane Company, 1982.

[30] J. M. Booker and T. J. Ross. "An Evolution of Uncertainty Assessment and Quantification." In: *Scientia Iranica* 18.3 (June 2011), pp. 669–676. DOI: 10.1016/j.scient.2011.04.017.

[31] E. Borgonovo. "A New Uncertainty Importance Measure." In: *Reliability Engineering & System Safety* 92.6 (June 2007), pp. 771–784. DOI: 10.1016/j.ress.2006.04.015.

[32] R. Brealey and A. Marcus. *Ise Fundamentals of Corporate Finance*. McGraw-Hill Education Ltd, May 5, 2019. ISBN: 1260566099.

[33] R. Bui et al. "Uncertainty and Sensitivity Analysis Applied to a Rammed Earth Wall: Evaluation of the Discrepancies between Experimental and Numerical Data." In: *E3S Web of Conferences* 172 (2020). Ed. by J. Kurnitski and T. Kalamees, p. 17004. DOI: 10.1051/e3sconf/202017217004.

[34] F. Campolongo, J. Cariboni, and A. Saltelli. "An Effective Screening Design for Sensitivity Analysis of Large Models." In: *Environmental Modelling & Software* 22.10 (Oct. 2007), pp. 1509–1518. DOI: 10.1016/j.envsoft.2006.10.004.

[35] M. Cartagena, J. Rosario, and D. Mavris. "A Method for Technology Identification, Evaluation, and Selection of Aircraft Propulsion Systems." In: *36th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*. American Institute of Aeronautics and Astronautics, July 2000. DOI: 10.2514/6.2000-3712.

[36] S. Castagne et al. "A Generic Tool for Cost Estimating in Aircraft Design." In: *Research in Engineering Design* 18.4 (Jan. 2008), pp. 149–162. DOI: 10.1007/s00163-007-0042-x.

[37] I. Chakraborty and D. N. Mavris. "Assessing Impact of Epistemic and Technological Uncertainty on Aircraft Subsystem Architectures." In: *Journal of Aircraft* 54.4 (July 2017), pp. 1388–1406. ISSN: 1533-3868. DOI: 10.2514/1.c034007.

[38] I. Chakraborty and D. N. Mavris. "Integrated Assessment of Aircraft and Novel Subsystem Architectures in Early Design." In: *Journal of Aircraft* 54.4 (July 2017), pp. 1268–1282. ISSN: 1533-3868. DOI: 10.2514/1.c033976.

[39] G. Chastaing, F. Gamboa, and C. Prieur. "Generalized Hoeffding-sobol Decomposition for Dependent Variables - Application to Sensitivity Analysis." In: *Electronic Journal of Statistics* 6.none (Jan. 2012). DOI: 10.1214/12-ejs749.

[40] P. Cheema, G. Mathews, and G. A. Vio. "Bayesian Inferencing on an Aircraft T-tail Using Probabilistic Surrogates and Uncertainty Quantification." In: *AIAA Journal* 56.3 (Mar. 2018), pp. 1185–1197. DOI: 10.2514/1.j056026.

[41] J. Chen et al. "Multi-objective Fuzzy Rule-based Prediction and Uncertainty Quantification of Aircraft Taxi Time." In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE. IEEE, Oct. 2017, pp. 1–5. DOI: 10.1109/itsc.2017.8317826.

[42] S. Chen et al. "Comparison of Probabilistic and Fuzzy Set Methods for Designing under Uncertainty." In: *40th Structures, Structural Dynamics, and Materials Conference and Exhibit*. American Institute of Aeronautics and Astronautics, Apr. 1999. DOI: 10.2514/6.1999-1579.

[43] W. Chen, R. Jin, and A. Sudjianto. "Analytical Variance-based Global Sensitivity Analysis in Simulation-based Design under Uncertainty." In: *Journal of Mechanical Design* 127.5 (2005), p. 875. DOI: 10.1115/1.1904642.

[44] Y. H. Chiang, E. W. L. Cheng, and P. T. I. Lam. "Employing the Net Present Value-consistent IRR Methods for PFI Contracts." In: *Journal of Construction Engineering and Management* 136.7 (July 2010), pp. 811–814. DOI: 10.1061/(asce)co.1943-7862.0000179.

[45] R. Chowdhury and S. Adhikari. "Fuzzy Parametric Uncertainty Analysis of Linear Dynamical Systems: A Surrogate Modeling Approach." In: *Mechanical Systems and Signal Processing* 32 (Oct. 2012), pp. 5–17. DOI: 10.1016/j.ymssp.2012.05.002.

[46] P. Clark. *Buying the Big Jets : Fleet Planning for Airlines. Fleet Planning for Airlines*. 3rd ed. Abingdon, Oxon New York, NY: Routledge, an imprint of the Taylor & Francis Group, 2017. 372 pp. ISBN: 9781315570662.

[47] Clean Aviation. *HLFC on Tails and Wing*. Accessed: 2024-05-26. 2024. URL: https://www.clean-aviation.eu/clean-sky-2/demonstrators/key-demonstrators/hlfc-on-tails-wing.

[48] F. Colmenares et al. "Future Aero-engines' Optimisation for Minimal Fuel Burn." In: *Proceedings of ASME Turbo Expo 2008*. Jan. 2008. DOI: 10.1115/gt2008-50126.

[49] R. Cooke. "The Anatomy of the Squizzel: The role of operational definitions in representing uncertainty." In: *Reliability Engineering & System Safety* 85.1-3 (July 2004), pp. 313–319. DOI: 10.1016/j.ress.2004.03.019.

[50] C. C. Croom and B. J. Holmes. *Insect Contamination Protection for Laminar Flow Surfaces*. Tech. rep. N88-14954. Hampton, Virginia: NASA Langley Research Center, 1988.

[51] C. Cuerno-Rejado, L. Alonso-Albir, and P. Gehse. "Conceptual Design of a Medium-sized Joined-wing Aircraft." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 224.6 (Dec. 2009), pp. 681–696. DOI: 10.1243/09544100jaero648.

[52] R. I. Cukier, H. B. Levine, and K. E. Shuler. "Nonlinear Sensitivity Analysis of Multiparameter Model Systems." In: *Journal of Computational Physics* 26.1 (Jan. 1978), pp. 1–42. DOI: 10.1016/0021-9991(78)90097-9.

[53] R. I. Cukier, J. H. Schaibly, and K. E. Shuler. "Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. III. Analysis of the Approximations." In: *The Journal of Chemical Physics* 63.3 (Aug. 1975), pp. 1140–1149. DOI: 10.1063/1.431440.

[54] R. I. Cukier et al. "Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. I Theory." In: *The Journal of Chemical Physics* 59.8 (Oct. 1973), pp. 3873–3878. DOI: 10.1063/1.1680571.

[55] Q. C. Curran, D. Allaire, and K. E. Willcox. "Sensitivity Analysis Methods for Mitigating Uncertainty in Engineering System Design." In: *Systems Engineering* 21.3 (Jan. 2018), pp. 191–209. DOI: 10.1002/sys.21422.

[56] R. Curran et al. "Integrating Aircraft Cost Modeling into Conceptual Design." In: *Concurrent Engineering* 13.4 (Dec. 2005), pp. 321–330. DOI: 10.1177/1063293x05060698.

[57] R. Curran et al. "Costing Tools for Decision Making within Integrated Aerospace Design." In: *Concurrent Engineering* 9.4 (Dec. 2001), pp. 327–338. DOI: 10.1177/1063293x0100900408.

[58] R. E. Davis, D. V. Maddalon, and R. D. Wagner. *Performance of laminar-flow leading-edge test articles in cloud encounters*. Tech. rep. No.: N90-12511, ID: 19900003195. NASA Langley Research Center, Dec. 1, 1987. 32 pp.

[59] A. M. DeGennaro, C. W. Rowley, and L. Martinelli. "Uncertainty Quantification for Airfoil Icing Using Polynomial Chaos Expansions." In: *Journal of Aircraft* 52.5 (Sept. 2015), pp. 1404–1411. DOI: 10.2514/1.c032698.

[60] A. P. Dempster. "A Generalization of Bayesian Inference." In: *Journal of the Royal Statistical Society: Series B (methodological)* 30.2 (1968), pp. 205–232.

[61] A. P. Dempster. "Upper and Lower Probabilities Induced by a Multivalued Mapping." In: *The Annals of Mathematical Statistics* 38.2 (1967), pp. 325–339. DOI: 10.1214/aoms/1177698950.

[62] Q. Deng, B. F. Santos, and W. J. C. Verhagen. "A Novel Decision Support System for Optimizing Aircraft Maintenance Check Schedule and Task Allocation." In: *Decision Support Systems* 146 (July 2021), p. 113545. DOI: 10.1016/j.dss.2021.113545.

[63] H. H. Dewey, D. R. DeVries, and S. R. Hyde. "Uncertainty Quantification in Prognostic Health Management Systems." In: *2019 IEEE Aerospace Conference*. IEEE. IEEE, Mar. 2019, pp. 1–13. DOI: 10.1109/aero.2019.8741821.

[64] N. C. Do and S. Razavi. "Correlation Effects? A Major but Often Neglected Component in Sensitivity and Uncertainty Analysis." In: *Water Resources Research* 56.3 (Mar. 2020). DOI: 10.1029/2019wr025436.

[65] R. Doganis. *Flying off Course : Airline Economics and Marketing*. 5th ed. Taylor & Francis Ltd., Jan. 10, 2019. 358 pp. ISBN: 9781315402970.

[66] D. Dubois and H. Prade. "Operations on Fuzzy Numbers." In: *International Journal of Systems Science* 9.6 (June 1978), pp. 613–626. DOI: 10.1080/00207727808941724.

[67] D. Dubois and H. Prade. "Possibility Theory and Its Applications: A Retrospective and Prospective View." In: *Decision Theory and Multi-Agent Planning*. Springer Vienna, 2006, pp. 89–109. DOI: 10.1007/3-211-38167-8_6.

[68] D. Dubois and H. Prade. "Possibility Theory in Information Fusion." In: *Data Fusion and Perception*. Springer Vienna, 2001, pp. 53–76. DOI: 10.1007/978-3-7091-2580-9_3.

[69] D. Dubois and H. Prade. "Unfair coins and necessity measures: Towards a possibilistic interpretation of histograms." In: *Fuzzy Sets and Systems* 10.1-3 (Jan. 1983), pp. 15–20. DOI: 10.1016/s0165-0114(83)80099-2.

[70] T. Effing, F. Schültke, and E. Stumpf. "HLFC-optimized Retrofit Aircraft Design of a Medium-range Reference Configuration within the AVACON Project." In: *CEAS Aeronautical Journal* 12.2 (Apr. 2021), pp. 441–456. DOI: 10.1007/s13272-021-00510-0.

[71] M. Eldred and J. Burkardt. "Comparison of Non-intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Quantification." In: *47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition*. American Institute of Aeronautics and Astronautics, Jan. 2009. DOI: 10.2514/6.2009-976.

[72] A. Elham and M. J. L. van Tooren. "Winglet Multi-objective Shape Optimization." In: *Aerospace Science and Technology* 37 (Aug. 2014), pp. 93–109. DOI: 10.1016/j.ast.2014.05.011.

[73] A. Elsenaar and H. N. Haasnoot. "Survey on Schiphol Airport of the Contamination of Wing Leading Edges of Three Different Aircraft Types under Operating Conditions." In: *First European Forum on Laminar Flow Technology*. 1992.

[74] S. Esber and D. Baier. "Real Options in the Assessment of New Products." In: *Studies in Classification, Data Analysis, and Knowledge Organization*. Springer Berlin Heidelberg, 2009, pp. 537–544. ISBN: 9783642010446. DOI: 10.1007/978-3-642-01044-6_49.

[75] European Commission. Directorate General for Research and Innovation. *Fly the Green Deal: Europe's Vision for Sustainable Aviation.* Publications Office, 2022. DOI: 10.2777/732726.

[76] M. E. Ewing, B. C. Liechty, and D. L. Black. "A General Methodology for Uncertainty Quantification in Engineering Analyses Using a Credible Probability Box." In: *Journal of Verification, Validation and Uncertainty Quantification* 3.2 (June 2018). DOI: 10.1115/1.4041490.

[77] W. J. Fabrycky. *Life-cycle Cost and Economic Analysis*. Englewood Cliffs, N.J: Prentice Hall, 1991. ISBN: 9780135383230.

[78] J. L. Fleiss. "Design and Analysis of Reliability Studies: The Statistical Evaluation of Measurement Errors." In: *Statistics in Medicine* 10.1 (Jan. 1991), pp. 155–156. DOI: 10.1002/sim.4780100120.

[79] Flightradar24. *Flight History for Aircraft D-AIKI. Live Air Traffic*. Dec. 31, 2019. URL: https://www.flightradar24.com/data/aircraft/d-aiki (visited on 01/11/2021).

[80] T. L. Galloway and D. N. Mavris. *Aircraft Life Cycle Cost Analysis (ALCCA) Program*. 1993.

[81] B. Gao et al. "A Direct Random Sampling Method for the Fourier Amplitude Sensitivity Test of Nonuniformly Distributed Uncertainty Inputs and Its Application in C/c Nozzles." In: *Aerospace Science and Technology* 100 (May 2020), p. 105830. DOI: 10.1016/j.ast.2020.105830.

[82] R. Gao et al. "The Economic Competitiveness of Promising Nuclear Energy System: A Closer Look at the Input Uncertainties in LCOE Analysis." In: *International Journal of Energy Research* 43.9 (Mar. 2019), pp. 3928–3958. DOI: 10.1002/er.4393.

[83] A. Gaspar et al. "Methodology for a Probabilistic Analysis of an RCC Gravity Dam Construction. Modelling of Temperature, Hydration Degree and Ageing Degree Fields." In: *Engineering Structures* 65 (Apr. 2014), pp. 99–110. DOI: 10.1016/j.engstruct.2014.02.002.

[84] W. Gibson and P. Morrell. "Theory and Practice in Aircraft Financial Evaluation." In: *Journal of Air Transport Management* 10.6 (Nov. 2004), pp. 427–433. DOI: 10.1016/j.jairtraman.2004.07.002.

[85] G. Gigerenzer. *Gut Feelings: The Intelligence of the Unconscious*. New York: Viking, 2007. ISBN: 1429535660.

[86] G. Gigerenzer and J. N. Marewski. "Surrogate Science." In: *Journal of Management* 41.2 (Sept. 2014), pp. 421–440. DOI: 10.1177/0149206314547522.

[87] C. Goldberg et al. "Assessment of an Energy-efficient Aircraft Concept from a Techno-economic Perspective." In: *Applied Energy* 221 (July 2018), pp. 229–238. DOI: 10.1016/j.apenergy.2018.03.163.

[88] C. Gong and B.-F. Ma. "Shape Optimization and Sensitivity Analysis of a Morphing-wing Aircraft." In: *International Journal of Aeronautical and Space Sciences* 20.1 (Jan. 2019), pp. 57–69. DOI: 10.1007/s42405-018-0110-7.

[89] J. R. Graham and C. R. Harvey. "The Theory and Practice of Corporate Finance: Evidence from the Field." In: *Journal of Financial Economics* 60.2-3 (May 2001), pp. 187–243. DOI: 10.1016/s0304-405x(01)00044-7.

[90] A. Grenyer et al. "A systematic review of multivariate uncertainty quantification for engineering systems." In: *CIRP Journal of Manufacturing Science and Technology* 33 (May 2021), pp. 188–208. DOI: 10.1016/j.cirpj.2021.03.004.

[91] K. M. Groth, C. L. Smith, and L. P. Swiler. "A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods." In: *Reliability Engineering & System Safety* 128 (Aug. 2014), pp. 32–40. ISSN: 0951-8320. DOI: 10.1016/j.ress.2014.03.010.

[92] J. H. A. Guillaume et al. "Introductory Overview of Identifiability Analysis: A Guide to Evaluating Whether You Have the Right Type of Data for Your Modeling Purpose." In: *Environmental Modelling & Software* 119 (Sept. 2019), pp. 418–432. DOI: 10.1016/j.envsoft.2019.07.007.

[93] H. Gupta and S. Razavi. "Challenges and Future Outlook of Sensitivity Analysis." In: *Sensitivity Analysis in Earth Observation Modelling*. Elsevier, 2017, pp. 397–415. DOI: 10.1016/b978-0-12-803011-0.00020-3.

[94] Y. P. Gupta. "Life Cycle Cost Models and Associated Uncertainties." In: *Electronic Systems Effectiveness and Life Cycle Costing*. Springer Berlin Heidelberg, 1983, pp. 535–549. DOI: 10.1007/978-3-642-82014-4_29.

[95] P. K. J. Han et al. "Communication of Uncertainty Regarding Individualized Cancer Risk Estimates." In: *Medical Decision Making* 31.2 (July 2010), pp. 354–366. DOI: 10.1177/0272989x10371830.

[96] L. U. Hansen, W. Heinze, and P. Horst. "Blended Wing Body Structures in Multidisciplinary Pre-design." In: *Structural and Multidisciplinary Optimization* 36.1 (Nov. 2007), pp. 93–106. DOI: 10.1007/s00158-007-0161-z.

[97] F. D. Harris. *An Economic Model of US Airline Operating Expenses*. Tech. rep. NASA/CR-2005-213476. NASA Ames Research Center, Moffet Field, California: National Aeronautics and Space Administration, 2005.

[98] J. Hartmann et al. "Collaborative Conceptual Design of a Mid-Range Aircraft Under Consideration of Advanced Methods for Technology Assessment." In: *31st Congress of the international Council of the Aeronautical Sciences*. Sept. 2018.

[99] M. Hassan, H. Pfaender, and D. Mavris. "Probabilistic assessment of aviation CO2 emission targets." In: *Transportation Research Part D: Transport and Environment* 63 (Aug. 2018), pp. 362–376. DOI: 10.1016/j.trd.2018.06.006.

[100] S. Hawer, A. Schönmann, and G. Reinhart. "Guideline for the Classification and Modelling of Uncertainty and Fuzziness." In: vol. 67. Elsevier BV, 2018, pp. 52–57. DOI: 10.1016/j.procir.2017.12.175.

[101] J. C. Helton, J. D. Johnson, and W. L. Oberkampf. "An Exploration of Alternative Approaches to the Representation of Uncertainty in Model Predictions." In: *Reliability Engineering & System Safety* 85.1-3 (July 2004), pp. 39–71. DOI: 10.1016/j.ress.2004.03.025.

[102] J. C. Helton and W. L. Oberkampf. "Alternative Representations of Epistemic Uncertainty." In: *Reliability Engineering & System Safety* 85.1-3 (July 2004), pp. 1–10. DOI: 10.1016/j.ress.2004.03.001.

[103] J. C. Helton et al. "A Sampling-based Computational Strategy for the Representation of Epistemic Uncertainty in Model Predictions with Evidence Theory." In: *Computer Methods in Applied Mechanics and Engineering* 196.37-40 (Aug. 2007), pp. 3980–3998. DOI: 10.1016/j.cma.2006.10.049.

[104] J. C. Helton, W. L. Oberkampf, and J. D. Johnson. "Competing Failure Risk Analysis Using Evidence Theory." In: *Risk Analysis* 25.4 (Aug. 2005), pp. 973–995. DOI: 10.1111/j.1539-6924.2005.00644.x.

[105] J. Herman and W. Usher. "SALib: An open-source Python library for Sensitivity Analysis." In: *The Journal of Open Source Software* 2.9 (Jan. 2017), p. 97. DOI: 10.21105/joss.00097.

[106] H. Hersbach et al. *ERA5 Hourly Data on Pressure Levels from 1979 to Present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed on 03/01/2021. 2018.

[107] P. Hester. "Epistemic Uncertainty Analysis: An Approach Using Expert Judgment and Evidential Credibility." In: *International Journal of Quality, Statistics, and Reliability* 2012 (2012), pp. 1–8. DOI: 10.1155/2012/617481.

[108] P. Himmelfarb. *Survival of the Fittest : New Product Development during the 90's*. Englewood Cliffs, N.J: Prentice Hall, 1992. ISBN: 9780138793135.

[109] P. Hollingsworth and D. N. Mavris. "A Method for Concept Exploration of Hypersonic Vehicles in the Presence of Open and Evolving Requirements." In: *SAE Transactions* 109 (2000), pp. 1148–1160. ISSN: 0096736X, 25771531.

[110] D.-K. Hong et al. "Ultra High Speed Motor Supported by Air Foil Bearings for Air Blower Cooling Fuel Cells." In: *IEEE Transactions on Magnetics* 48.2 (Feb. 2012), pp. 871–874. DOI: 10.1109/tmag.2011.2174209.

[111] S. Hong et al. "Wide-body Aircraft Trip Direct Operating Cost Model Based on Translog Function." In: *Proceedings of the International Conference on Logistics, Engineering, Management and Computer Science (IEMS 2015)*. Ed. by S. Yingying, C. Guiran, and L. Zhen. Vol. 117. Advances in Intelligent Systems Research. 2015, pp. 726–729. ISBN: 978-94-6252-102-5. DOI: 10.2991/lemcs-15.2015.143.

[112] L. Hong-Qi. "Applications of Random Balance Design Fourier Amplitude Sensitivity Test and Extended Fourier Amplitude Sensitivity Test in the Parameter Sensitivity Analysis of Land Surface Process Model." In: *Acta Physica Sinica* 64.6 (2015), p. 069201. DOI: 10.7498/aps.64.069201.

[113] IATA. *Economic Performance of the Airline Industry. 2021 End-year report*. Semi-Annual Report. International Air Transportation Association (IATA), Oct. 4, 2021.

[114] B. Iooss and P. Lemaitre. "A Review on Global Sensitivity Analysis Methods." In: *Uncertainty Management in Simulation-Optimization of Complex Systems*. Springer US, 2015, pp. 101–122. DOI: 10.1007/978-1-4899-7547-8_5.

[115] T. Ishigami and T. Homma. "An Importance Quantification Technique in Uncertainty Analysis for Computer Models." In: *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*. IEEE Comput. Soc. Press, 1990. DOI: 10.1109/isuma.1990.151285.

[116] A. T. Isikveren. "Identifying Economically Optimal Flight Techniques of Transport Aircraft." In: *Journal of Aircraft* 39.4 (July 2002), pp. 528–544. DOI: 10.2514/2.2982.

[117] A. Iwaniuk and W. Wiśniowski. "Optimization of Small Aircraft Parameters in the Initial Phase of the Project." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 231.12 (June 2017), pp. 2248–2258. DOI: 10.1177/0954410017716481.

[118] J. Jacques, C. Lavergne, and N. Devictor. "Sensitivity Analysis in Presence of Model Uncertainty and Correlated Inputs." In: *Reliability Engineering & System Safety* 91.10-11 (Oct. 2006), pp. 1126–1134. DOI: 10.1016/j.ress.2005.11.047.

[119] P. H. M. Janssen et al. "Towards Guidance in Assessing and Communicating Uncertainties." In: *4th International Conference on Sensitivity Analysis of Model Output (SAMO)* (Mar. 8–11, 2004). Santa Fe, 2004.

[120] W. H. Jasperson et al. "GASP cloud encounter statistics - Implications for laminar flow control flight." In: *Journal of Aircraft* 21.11 (Nov. 1984), pp. 851–857. DOI: 10.2514/3.45054.

[121] W. H. Jasperson et al. "Variability of Cloudiness at Airline Cruise Altitudes from GASP Measurements." In: *Journal of Climate and Applied Meteorology* 24.1 (1985), pp. 74–82.

[122] V. S. Johnson. "Minimizing Life Cycle Cost for Subsonic Commercial Aircraft." In: *Journal of Aircraft* 27.2 (Feb. 1990), pp. 139–145. DOI: 10.2514/3.45909.

[123] R. D. Joslin. "Aircraft Laminar Flow Control." In: *Annual Review of Fluid Mechanics* 30.1 (Jan. 1998), pp. 1–29. DOI: 10.1146/annurev.fluid.30.1.1.

[124] R. D. Joslin. *Overview of Laminar Flow Control*. Tech. rep. NASA/TP-1998-208705. Hampton, VA: NASA Langley Research Center, Jan. 10, 1998.

[125] J. C. June, R. H. Thomas, and Y. Guo. "System Noise Prediction Uncertainty Quantification for a Hybrid Wing–body Transport Concept." In: *AIAA Journal* 58.3 (Mar. 2020), pp. 1157–1170. DOI: 10.2514/1.j058226.

[126] C. Justin, E. Garcia, and D. Mavris. "Aircraft Valuation: A Network Approach to the Evaluation of Aircraft for Fleet Planning and Strategic Decision Making." In: *10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*. American Institute of Aeronautics and Astronautics, Sept. 2010. DOI: 10.2514/6.2010-9061.

[127] C. Y. Justin. "A Quantitative Real Options Method for Aviation Technology Decision-Making in the Presence of Uncertainty." PhD thesis. Atlanta, Georgia: Georgia Institute of Technology, Dec. 2015.

[128] C. Y. Justin and D. N. Mavris. "Aircraft and Engine Economic Evaluation for Fleet Renewal Decision-making and Maintenance Contract Valuation." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 229.11 (Jan. 2015), pp. 2051–2065. DOI: 10.1177/0954410014564403.

[129] A. S. Kangas and J. Kangas. "Probability, Possibility and Evidence: Approaches to Consider Risk and Uncertainty in Forestry Decision Analysis." In: *Forest Policy and Economics* 6.2 (Mar. 2004), pp. 169–188. DOI: 10.1016/s1389-9341(02)00083-7.

[130] R. L. Keeney and D. von Winterfeldt. "On the Uses of Expert Judgment on Complex Technical Problems." In: *IEEE Transactions on Engineering Management* 36.2 (May 1989), pp. 83–86. DOI: 10.1109/17.18821.

[131] T. Kilian, T. Streit, and G. Schrauf. *Deliverable D1.4.4.1-29: Maximizing HLFC Performance - Report on Inner Wing Laminarization*. Tech. rep. Approved by Airbus. Clean Sky 2, LPA Platform 1, Nov. 2, 2022.

[132] M. Kirby. "A Methodology for Technology Identification, Evaluation, and Selection in Conceptual and Preliminary Aircraft Design." PhD thesis. Georgia Institute of Technology, Mar. 1, 2001.

[133] M. R. Kirby and D. N. Mavris. "A Technique for Selecting Emerging Technologies for a Fleet of Commercial Aircraft to Maximize R&D Investment." In: *SAE Transactions* 110 (2001), pp. 674–687. ISSN: 0096736X, 25771531.

[134] M. R. Kirby and D. N. Mavris. "Forecasting the Impact of Technology Infusion on Subsonic Transport Affordability." In: *SAE Transactions* 107 (1998), pp. 1733–1743. ISSN: 0096736X, 25771531.

[135] G. J. Klir and D. Harmanec. "Types and Measures of Uncertainty." In: *International Series in Intelligent Technologies*. Ed. by J. Kacprzyk, H. Nurmi, and M. Fedrizzi. Boston, MA: Springer US, 1997, pp. 29–51. ISBN: 978-1-4615-6333-4. DOI: 10.1007/978-1-4615-6333-4_3.

[136] M. Kok et al. "Critical Considerations in the Mitigation of Insect Residue Contamination on Aircraft Surfaces – a Review." In: *Progress in Aerospace Sciences* 75 (May 2015), pp. 1–14. DOI: 10.1016/j.paerosci.2015.02.001.

[137] K. S. G. Krishnan and O. Bertram. "Preliminary Design and System Considerations for an Active Hybrid Laminar Flow Control System." In: *Aerospace* 6.10 (Oct. 2019), p. 109. DOI: 10.3390/aerospace6100109.

[138] M. Krosche and W. Heinze. "Robustness Analysis of an Aircraft Design for Short Takeoff and Landing." In: *Journal of Aircraft* 52.4 (July 2015), pp. 1235–1246. DOI: 10.2514/1.c032876.

[139] R. Kruse. "Fuzzy-systeme - Positive Aspekte Der Unvollkommenheit." In: *Informatik-Spektrum* 19.1 (Feb. 1996), pp. 4–11. DOI: 10.1007/s002870050011.

[140] S. Kucherenko, S. Tarantola, and P. Annoni. "Estimation of Global Sensitivity Indices for Models with Dependent Variables." In: *Computer Physics Communications* 183.4 (Apr. 2012), pp. 937–946. DOI: 10.1016/j.cpc.2011.12.020.

[141] S. Kucherenko and B. Iooss. "Derivative Based Global Sensitivity Measures." In: (Dec. 8, 2014).

[142] J. H. Kwakkel, W. E. Walker, and V. A. W. J. Marchau. "Classifying and Communicating Uncertainties in Model-based Policy Analysis." In: *International Journal of Technology, Policy and Management* 10.4 (2010), p. 299. DOI: 10.1504/ijtpm.2010.036918.

[143] K. G. Kyprianidis et al. "Assessment of Future Aero-engine Designs with Intercooled and Intercooled Recuperated Cores." In: *Journal of Engineering for Gas Turbines and Power* 133.1 (Sept. 2010). DOI: 10.1115/1.4001982.

[144] H.-C. Lahne et al. "Design of a 50000 Rpm High-speed High-power Six-phase PMSM for Use in Aircraft Applications." In: *2016 Eleventh International Conference on Ecological Vehicles and Renewable Energies (EVER)*. IEEE, Apr. 2016. DOI: 10.1109/ever.2016.7476345.

[145] S. Langhans. "A Systems Engineering Approach for Economic Assessment of Air Transportation Concepts." PhD thesis. German Aerospace Center (DLR), 2012. 300 pp.

[146] S. Langhans et al. "System Analysis for an Intermediate Stop Operations Concept on Long Range Routes." In: *Journal of Aircraft* 50.1 (Jan. 2013), pp. 29–37. DOI: 10.2514/1.c031446.

[147] B. Lee et al. "Economic Evaluation with Uncertainty Analysis Using a Monte-carlo Simulation Method for Hydrogen Production from High Pressure PEM Water Electrolysis in Korea." In: *International Journal of Hydrogen Energy* 42.39 (Sept. 2017), pp. 24612–24619. DOI: 10.1016/j.ijhydene.2017.08.033.

[148] M. Lee, L. K. B. Li, and W. Song. "Analysis of Direct Operating Cost of Wide-body Passenger Aircraft: A Parametric Study Based on Hong Kong." In: *Chinese Journal of Aeronautics* 32.5 (May 2019), pp. 1222–1243. DOI: 10.1016/j.cja.2019.03.011.

[149] G. Li et al. "Global Sensitivity Analysis for Systems with Independent And/or Correlated Inputs." In: *The Journal of Physical Chemistry A* 114.19 (Apr. 2010), pp. 6022–6032. DOI: 10.1021/jp9096919.

[150] R. H. Liebeck et al. *Advanced Subsonic Airplane Design and Economic Studies*. Contractor Report 195443. NASA Lewis Research Center, Cleveland, Ohio: National Aeronautics and Space Administration (NASA), Apr. 1995.

[151] V. Lohweg, K. Voth, and S. Glock. "A Possibilistic Framework for Sensor Fusion with Monitoring of Sensor Reliability." In: *Sensor Fusion - Foundation and Applications*. InTech, June 2011. DOI: 10.5772/17384.

[152] I. Lopez and N. Sarigul-Klijn. "A Review of Uncertainty in Flight Vehicle Structural Damage Monitoring, Diagnosis and Control: Challenges and Opportunities." In: *Progress in Aerospace Sciences* 46.7 (Oct. 2010), pp. 247–273. DOI: 10.1016/j.paerosci.2010.03.003.

[153] R. Mahmud et al. "Integration of techno-economic analysis and life cycle assessment for sustainable process design – A review." In: *Journal of Cleaner Production* 317 (Oct. 2021), p. 128247. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2021.128247.

[154] K.-G. Mäler and A. Fisher. "Environment, Uncertainty, and Option Values." In: *Handbook of Environmental Economics*. Ed. by K.-G. Mäler and J. R. Vincent. Vol. 2. Elsevier, 2005. Chap. 13, pp. 571–620. DOI: 10.1016/s1574-0099(05)02013-9.

[155] E. Mamdani and S. Assilian. "An experiment in linguistic synthesis with a fuzzy logic controller." In: *International Journal of Man-Machine Studies* 7.1 (Jan. 1975), pp. 1–13. DOI: 10.1016/s0020-7373(75)80002-2.

[156] R. W. Marczewski. "Bridging the Virtual Valley of Death for Technology R&d." In: *The Scientist* 2.11 (1997).

[157] J. Markish. "Valuation Techniques for Commercial Aircraft Program Design." MA thesis. Massachusetts Institute of Technology (MIT), 2000. 153 pp.

[158] R. Martinez-Val et al. "Optimization of Planform and Cruise Conditions of a Transport Flying Wing." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 224.12 (June 2010), pp. 1243–1251. DOI: 10.1243/09544100jaero812.

[159] W. Marx, D. Mavris, and D. Schrage. "A Hierarchical Aircraft Life Cycle Cost Analysis Model." In: *Aircraft Engineering, Technology, and Operations Congress*. American Institute of Aeronautics and Astronautics, Sept. 1995. DOI: 10.2514/6.1995-3861.

[160] W. J. Marx, D. N. Mavris, and D. P. Schrage. "A Knowledge-based System Integrated with Numerical Analysis Tools for Aircraft Life-cycle Design." In: *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 12.3 (June 1998), pp. 211–229. DOI: 10.1017/s0890060498123016.

[161] D. Mavris and D. DeLaurentis. "Methodology for Examining the Simultaneous Impact of Requirements, Vehicle Characteristics, and Technologies on Military Aircraft Design." In: *22nd Congress of the International Council on the Aeronautical Sciences (ICAS)*. ICAS Paper 2000-1.4.4. Harrogate, England, Aug. 27, 2000.

[162] D. Mavris, A. Baker, and D. Schrage. "Technology Infusion and Resource Allocation for a Civil Tiltrotor." In: *Vertical Lift Aircraft Design Specialist's Meeting*. San Francisco, CA, Jan. 19, 2000.

[163] D. Mavris and M. Kirby. "Preliminary Assessment of the Economic Viability of a Family of Very Large Transport Configurations." In: *1st world aviation congress*. Oct. 22, 1996.

[164] D. Mavris and M. Kirby. "Technology Identification, Evaluation, and Selection for Commercial Transport Aircraft." In: *58th Annual Conference of Society of Allied Weight Engineers*. SAWE Paper 99-2456. San Jose, CA, May 25, 1999.

[165] D. N. Mavris and D. DeLaurentis. "A Stochastic Design Approach for Aircraft Affordability." In: 1998.

[166] D. N. Mavris, M. R. Kirby, and S. Qiu. "Technology Impact Forecasting for a High Speed Civil Transport." In: *SAE Transactions* 107 (1998), pp. 1645–1658. ISSN: 0096736X, 25771531.

[167] S. McCrary. *Mastering Corporate Finance Essentials: The Critical Quantitative Methods and Tools in Finance*. Hoboken, N.J: Wiley, 2010. ISBN: 9780470393338.

[168] K. Meifarth and S. Heinrich. "The Environment for Aircraft with Laminar Flow Technology within Airline Service." In: *First European Forum on Laminar Flow Technology*. Hamburg, 1992.

[169] G. Moore. *Dealing with Darwin : How Great Companies Innovate at Every Phase of Their Evolution*. New York: Portfolio, 2008. ISBN: 9781591842149.

[170] M. D. Morris. "Factorial Sampling Plans for Preliminary Computational Experiments." In: *Technometrics* 33.2 (May 1991), pp. 161–174. DOI: 10.1080/00401706.1991.10484804.

[171] Z. P. Mourelatos and J. Zhou. "A Design Optimization Method Using Evidence Theory." In: *Journal of Mechanical Design* 128.4 (Dec. 2005), pp. 901–908. DOI: 10.1115/1.2204970.

[172] Z. P. Mourelatos and J. Zhou. "Reliability Estimation and Design with Insufficient Data Based on Possibility Theory." In: *AIAA Journal* 43.8 (Aug. 2005), pp. 1696–1705. DOI: 10.2514/1.12044.

[173] J.-K. Muller and A. Mertens. "Power Electronics Design for a Direct-driven Turbo Compressor Used As Advanced High-lift System in Future Aircraft." In: *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, Oct. 2017. DOI: 10.1109/iecon.2017.8216756.

[174] NASA Langley Research Center. *Flight Optimization System (FLOPS)*. May 1, 2024.

[175] J. Newcamp et al. "Retirement Optimization through Aircraft Transfers and Employment." In: *Journal of Air Transport Management* 79 (Aug. 2019), p. 101680. DOI: 10.1016/j.jairtraman.2019.101680.

[176] A.-T. Nguyen and S. Reiter. "A Performance Comparison of Sensitivity Analysis Methods for Building Energy Models." In: *Building Simulation* 8.6 (July 2015), pp. 651–664. DOI: 10.1007/s12273-015-0245-4.

[177] A. Niazi et al. "Product Cost Estimation: Technique Classification and Methodology Review." In: *Journal of Manufacturing Science and Engineering* 128.2 (Sept. 2005), pp. 563–575. DOI: 10.1115/1.2137750.

[178] A. O'Hagan. "Polynomial Chaos: A Tutorial and Critique from a Statistician's Perspective." In: *SIAM/ASA Journal of Uncertainty Quantification* 20 (2013), pp. 1–20.

[179] W. Oberkampf and J. Helton. "Investigation of Evidence Theory for Engineering Applications." In: *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, Apr. 2002. DOI: 10.2514/6.2002-1569.

[180] W. Oberkampf, J. Helton, and K. Sentz. "Mathematical Representation of Uncertainty." In: *19th AIAA Applied Aerodynamics Conference*. American Institute of Aeronautics and Astronautics, June 2001. DOI: 10.2514/6.2001-1645.

[181] T. O. Odedele, K. B. Ajoku, and H. D. Ibrahin. "Applying Soft Computing Approach to Uncertainty Analysis in Oil and Gas Economic Evaluations." In: *Proceedings of the World Congress on Engineering and Computer Science 2017* (Oct. 25–27, 2017). Vol. 1. San Francisco, CA, 2017.

[182] J. M. V. Oliveira. "Development of Operating Cost Models for the Preliminarydesign Optimization of an Aircraft." MA thesis. Lisbon, Portugal: Tècnico Lisboa, 2015. 82 pp.

[183] M. M. J. Opgenoord, D. L. Allaire, and K. E. Willcox. "Variance-based Sensitivity Analysis to Support Simulation-based Design under Uncertainty." In: *Journal of Mechanical Design* 138.11 (Sept. 2016). DOI: 10.1115/1.4034224.

[184] M. M. J. Opgenoord and K. E. Willcox. "Sensitivity Analysis Methods for Uncertainty Budgeting in System Design." In: *AIAA Journal* 54.10 (Oct. 2016), pp. 3134–3148. DOI: 10.2514/1.j054743.

[185] N. R. Ortiz et al. "Use of Expert Judgment in NUREG-1150." In: *Nuclear Engineering and Design* 126.3 (May 1991), pp. 313–331. DOI: 10.1016/0029-5493(91)90023-b.

[186] H.-U. Park et al. "Multidisciplinary Wing Design Optimization Considering Global Sensitivity and Uncertainty of Approximation Models." In: *Journal of Mechanical Science and Technology* 28.6 (June 2014), pp. 2231–2242. DOI: 10.1007/s12206-014-0127-1.

[187] B. Pearce. *Profitability an dthe Air Transport Value Chain. IATA Economics Briefing No. 10*. Tech. rep. Version 1.1. International Air Transportation Association (IATA), June 2013.

[188] P. Perkins and U. Gustafsson. *An automated atmospheric sampling system operating on 747 airliners*. NASA Technical Memorandum. Cleveland, OH, United States: NASA Lewis Research Center, Jan. 1, 1975.

[189] M. Petr et al. "Inconsistent recognition of uncertainty in studies of climate change impacts on forests." In: *Environmental Research Letters* 14.11 (Nov. 2019), p. 113003. DOI: 10.1088/1748-9326/ab4670.

[190] T. Pfeiffer, B. Nagel, and V. Gollnick. "Konfigurationsanalyse Im Flugzeugvorentwurf Unter Der Berücksichtigung Von Unsicherheiten." In: *Deutscher Luft- und Raumfahrtkongress*. 281388. Deutsche Gesellschaft für Luft- und Raumfahrt (DGLR). Sept. 10, 2012.

[191] T. Pfeiffer et al. "Aircraft Configuration Analysis Using a Low-fidelity, Physics Based Aerospace Framework under Uncertainty Considerations." In: *29th Congress of the International Council on the Aeronautical Sciences*. international Council of the Aeronautical Sciences (ICAS). St. Petersburg, Russia, Sept. 7, 2014.

[192] T. Pfeiffer et al. "Analysis of Aircraft Configurations Including Propagated Uncertainties." In: *5th CEAS Air & Space Conference*. Delft, Netherlands, Sept. 2015.

[193] F. Pianosi et al. "Sensitivity Analysis of Environmental Models: A Systematic Review with Practical Workflow." In: *Environmental Modelling & Software* 79 (May 2016), pp. 214–232. DOI: 10.1016/j.envsoft.2016.02.008.

[194] E. Plischke. "An Effective Algorithm for Computing Global Sensitivity Indices (EASI)." In: *Reliability Engineering & System Safety* 95.4 (Apr. 2010), pp. 354–360. DOI: 10.1016/j.ress.2009.11.005.

[195] E. Plischke, E. Borgonovo, and C. L. Smith. "Global Sensitivity Measures from Given Data." In: *European Journal of Operational Research* 226.3 (May 2013), pp. 536–550. DOI: 10.1016/j.ejor.2012.11.047.

[196] A. A. Pohya. "Selected Current Challenges in the Development of Hybrid Laminar Flow Control on Transport Aircraft." en. In: *Deutscher Luft- und Raumfahrtkongress 2019*. Deutsche Gesellschaft für Luft- und Raumfahrt - Lilienthal-Oberth e.V., Oct. 1, 2019. DOI: 10.25967/490089.

[197] A. A. Pohya and K. Wicke. "An Impact Assessment of Degrading Elements on the Overall Benefit of Aircraft with Hybrid Laminar Flow Control." In: *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics, Jan. 2019. DOI: 10.2514/6.2019-1589.

[198] A. A. Pohya, K. Wicke, and V. Gollnick. "Bewertung Der Effizienz Von Flugzeugen Mit Hybridlaminarhaltung Unter Berücksichtigung Von Wolkeneinflüssen." In: *Deutscher Luft- und Raumfahrtkongress 2017*. Munich, Germany, Sept. 2017.

[199] A. A. Pohya, K. Wicke, and J. Hartmann. "Comparison of Direct Operating Cost and Life Cycle Cost-benefit Methods in Aircraft Technology Assessment." In: *2018 AIAA Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics, Jan. 2018. DOI: 10.2514/6.2018-0282.

[200] A. A. Pohya, K. Wicke, and T. Kilian. "Introducing Variance-based Global Sensitivity Analysis for Uncertainty Enabled Operational and Economic Aircraft Technology Assessment." In: *Aerospace Science and Technology* 122 (Feb. 2022), p. 107441. ISSN: 1270-9638. DOI: 10.1016/j.ast.2022.107441.

[201] A. A. Pohya et al. "A Modular Framework for the Life Cycle Based Evaluation of Aircraft Technologies, Maintenance Strategies, and Operational Decision Making Using Discrete Event Simulation." In: *Aerospace* 8.7 (July 2021), p. 187. DOI: 10.3390/aerospace8070187.

[202] A. A. Pohya et al. "Cloud Encounter Impact on Operational and Economical Effectiveness of Hybrid-laminar-flow-control Aircraft." In: *Journal of Aircraft* 56.4 (July 2019), pp. 1513–1523. DOI: 10.2514/1.c035205.

[203] A. A. Pohya et al. *LYFE Documentation*. Ed. by German Aerospace Center. https://lyfe.pages.gitlab.dlr.de/airlyfe/. German Aerospace Center (DLR). Aug. 12, 2021.

[204] M. C. Politi, P. K. J. Han, and N. F. Col. "Communicating the Uncertainty of Harms and Benefits of Medical Interventions." In: *Medical Decision Making* 27.5 (Sept. 2007), pp. 681–695. DOI: 10.1177/0272989x07307270.

[205] S. W. Popper, R. J. Lempert, and S. C. Bankes. "Shaping the Future." In: *Scientific American* 292.4 (Apr. 2005), pp. 66–71. DOI: 10.1038/scientificamerican0405-66.

[206] L. I. U. Qiao and T. HOMMA. "A New Importance Measure for Sensitivity Analysis." In: *Journal of Nuclear Science and Technology* 47.1 (Jan. 2010), pp. 53–61. DOI: 10.1080/18811248.2010.9711927.

[207] S. G. Rabinovich. *Measurement Errors and Uncertainties*. Springer-Verlag GmbH, Dec. 26, 2006. ISBN: 9780387291437.

[208] S. Raghunathan et al. "Key Aerodynamic Technologies for Aircraft Engine Nacelles." In: *The Aeronautical Journal* 110.1107 (May 2006), pp. 265–288. DOI: 10.1017/s0001924000013154.

[209] L. P. Raj, K. Yee, and R. S. Myong. "Sensitivity of Ice Accretion and Aerodynamic Performance Degradation to Critical Physical and Modeling Parameters Affecting Airfoil Icing." In: *Aerospace Science and Technology* 98 (Mar. 2020), p. 105659. DOI: 10.1016/j.ast.2019.105659.

[210] J. Ramm et al. "Uncertainty quantification in hydrogen tank exchange: Estimating maintenance costs for new aircraft concepts." In: *International Journal of Hydrogen Energy* 68 (May 2024), pp. 159–169. ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2024.04.157.

[211] D. P. Raymer. *Aircraft Design: A Conceptual Approach*. American Institute of Aeronautics and Astronautics, Inc., 1992.

[212] S. Razavi and H. V. Gupta. "What Do We Mean by Sensitivity Analysis? The Need for Comprehensive Characterization of Global Sensitivity in Earth and Environmental Systems Models." In: *Water Resources Research* 51.5 (May 2015), pp. 3070–3092. DOI: 10.1002/2014wr016527.

[213] S. Razavi et al. "The Future of Sensitivity Analysis: An Essential Discipline for Systems Modeling and Policy Support." In: *Environmental Modelling & Software* 137 (Mar. 2021), p. 104954. DOI: 10.1016/j.envsoft.2020.104954.

[214] G. Reinhart, P. Krebs, and M. F. Zaeh. "Fuzzy-logic-based Integration of Qualitative Uncertainties into Monetary Factory-evaluations." In: *2009 IEEE International Conference on Control and Automation*. IEEE, Dec. 2009. DOI: 10.1109/icca.2009.5410328.

[215] M. G. J. Repko and B. F. Santos. "Scenario Tree Airline Fleet Planning for Demand Uncertainty." In: *Journal of Air Transport Management* 65 (Oct. 2017), pp. 198–208. DOI: 10.1016/j.jairtraman.2017.06.010.

[216] M. Rezaei et al. "Fuzzy Uncertainty Analysis in the Flutter Boundary of an Aircraft Wing Subjected to a Thrust Force." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 233.6 (May 2019), pp. 2185–2197. DOI: 10.1177/0954410018773898.

[217] S. Richardson. *Markov Chain Monte Carlo in Practice (Interdisciplinary Statistics)*. Chapman & Hall/CRC, 1995, p. 512. ISBN: 9780412055515.

[218] K. Risse. *Preliminary Overall Aircraft Design with Hybrid Laminar Flow Control*. en. RWTH Aachen University, 2016. DOI: 10.18154/RWTH-2017-00974.

[219] K. Risse et al. "Central Reference Aircraft Data System (CeRAS) for Research Community." In: *CEAS Aeronautical Journal* 7.1 (Nov. 2015), pp. 121–133. DOI: 10.1007/s13272-015-0177-9.

[220] M. Roelofs and R. Vos. "Uncertainty-based Design Optimization and Technology Evaluation: A Review." In: *2018 AIAA Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics, Jan. 2018. DOI: 10.2514/6.2018-2029.

[221] J. Roskam. *Airplane Design Part Vii: Airplane Cost Estimation: Design, Development, Manufacturing and Operating*. Lawrence, Kansas: DARcorporation, 2002. ISBN: 978-1884885556.

[222] S. Ross. *Introductory Statistics*. LondonSan Diego, CA: Academic Press, 2017. ISBN: 0128043172.

[223] T. Ross. *Fuzzy Logic and Probability Applications : Bridging the Gap*. Philadelphia, Pa. Alexandria, Va: Society for Industrial and Applied Mathematics American Statistical Assoxciation, 2002. ISBN: 0898715253.

[224] B. Roth et al. "Adaptive Selection of Engine Technology Solution Sets from a Large Combinatorial Space." In: *37th Joint Propulsion Conference and Exhibit*. American Institute of Aeronautics and Astronautics, July 2001. DOI: 10.2514/6.2001-3208.

[225] C. J. Roy and W. L. Oberkampf. "A Comprehensive Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing." In: *Computer Methods in Applied Mechanics and Engineering* 200.25-28 (June 2011), pp. 2131–2144. DOI: 10.1016/j.cma.2011.03.016.

[226] C. A. A. Sa, B. F. Santos, and J.-P. B. Clarke. "Portfolio-based Airline Fleet Planning under Stochastic Demand." In: *Omega* 97 (Dec. 2020), p. 102101. DOI: 10.1016/j.omega.2019.08.008.

[227] E. N. Saatlou et al. "On the Trade-off between Minimum Fuel Burn and Maximum Time between Overhaul for an Intercooled Aeroengine." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 228.13 (Jan. 2014), pp. 2424–2438. DOI: 10.1177/0954410013518509.

[228] Sabre Airline Solutions. *Sabre Data & Analytics Market Intelligence*. 2021.

[229] A. Saltelli, S. Tarantola, and K. P.-S. Chan. "A Quantitative Model-independent Method for Global Sensitivity Analysis of Model Output." In: *Technometrics* 41.1 (Feb. 1999), pp. 39–56. DOI: 10.1080/00401706.1999.10485594.

[230] A. Saltelli et al. *Global Sensitivity Analysis*. John Wiley & Sons, Feb. 4, 2008. 306 pp. ISBN: 0470059974.

[231] A. Saltelli. "Making Best Use of Model Evaluations to Compute Sensitivity Indices." In: *Computer Physics Communications* 145.2 (May 2002), pp. 280–297. DOI: 10.1016/s0010-4655(02)00280-1.

[232] A. Saltelli and P. Annoni. "How to Avoid a Perfunctory Sensitivity Analysis." In: *Environmental Modelling & Software* 25.12 (Dec. 2010), pp. 1508–1517. DOI: 10.1016/j.envsoft.2010.04.012.

[233] A. Saltelli and R. Bolado. "An Alternative Way to Compute Fourier Amplitude Sensitivity Test (FAST)." In: *Computational Statistics & Data Analysis* 26.4 (Feb. 1998), pp. 445–460. DOI: 10.1016/s0167-9473(97)00043-1.

[234] A. Saltelli and S. Tarantola. "On the Relative Importance of Input Factors in Mathematical Models." In: *Journal of the American Statistical Association* 97.459 (Sept. 2002), pp. 702–709. DOI: 10.1198/016214502388618447.

[235] A. Saltelli et al. "Five Ways to Ensure That Models Serve Society: A Manifesto." In: *Nature* 582.7813 (June 2020), pp. 482–484. DOI: 10.1038/d41586-020-01812-9.

[236] A. Saltelli et al. "Sensitivity Analysis Practices: Strategies for Model-based Inference." In: *Reliability Engineering & System Safety* 91.10-11 (Oct. 2006), pp. 1109–1125. DOI: 10.1016/j.ress.2005.11.014.

[237] A. Saltelli et al. "Variance Based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index." In: *Computer Physics Communications* 181.2 (2010), pp. 259–270. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2009.09.018.

[238] A. Saltelli et al. "Why so Many Published Sensitivity Analyses Are False: A Systematic Review of Sensitivity Analysis Practices." In: *Environmental Modelling & Software* 114 (Apr. 2019), pp. 29–39. DOI: 10.1016/j.envsoft.2019.01.012.

[239] R. Saltelli Campolongo. *Sensitivity Analysis in Practice*. John Wiley & Sons, Mar. 24, 2004. 232 pp. ISBN: 0470870931.

[240] B. F. Santos et al. "Airline Delay Management Problem with Airport Capacity Constraints and Priority Decisions." In: *Journal of Air Transport Management* 63 (Aug. 2017), pp. 34–44. DOI: 10.1016/j.jairtraman.2017.05.003.

[241] H. Schnieder. "Methode Zur Bewertung Von Projekten Und Technologien Im Zivilen Flugzeugbau." In: *DGLR Jahrestagung*. Daimler-Benz Aerospace Airbus GmbH. 1999, pp. 653–660.

[242] G. Schrauf. "Status and Perspectives of Laminar Flow." In: *The Aeronautical Journal* 109.1102 (Dec. 2005), pp. 639–644. DOI: 10.1017/s000192400000097x.

[243] O. Schwabe, E. Shehab, and J. Erkoyuncu. "Uncertainty Quantification Metrics for Whole Product Life Cycle Cost Estimates in Aerospace Innovation." In: *Progress in Aerospace Sciences* 77 (Aug. 2015), pp. 1–24. DOI: 10.1016/j.paerosci.2015.06.002.

[244] K. Sentz and S. Ferson. *Combination of Evidence in Dempster-Shafer Theory*. Tech. rep. Sandia National Laboratories, Apr. 2002. DOI: 10.2172/800792.

[245] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Apr. 21, 1976. 314 pp. ISBN: 069110042X.

[246] S. Sibdari, I. Mohammadian, and D. F. Pyke. "On the Impact of Jet Fuel Cost on Airlines' Capacity Choice: Evidence from the U.s. Domestic Markets." In: *Transportation Research Part E: Logistics and Transportation Review* 111 (Mar. 2018), pp. 1–17. DOI: 10.1016/j.tre.2017.12.009.

[247] D. Simos. *Piano-X: Aircraft Emissions and Performance*. last accessed on 2021/25/04. 2021.

[248] P. Slovic et al. "Affect, Risk, and Decision Making." In: *Health Psychology* 24.4, Suppl (2005), S35–S40. DOI: 10.1037/0278-6133.24.4.S35.

[249] R. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*. CAMBRIDGE, Mar. 6, 2014. 400 pp. ISBN: 161197321X.

[250] D. S. Soban and D. N. Mavris. "Assessing the Impact of Technology on Aircraft Systems Using Technology Impact Forecasting." In: *Journal of Aircraft* 50.5 (Sept. 2013), pp. 1380–1393. DOI: 10.2514/1.c031871.

[251] I. M. Sobol. "Sensitivity Analysis for Non-linear Mathematical Models." In: *Mathematical modelling and computational experiment* 1 (1993). (translated), pp. 407–414. DOI: 10.1002/wilm.42820050114.

[252] I. M. Sobol' et al. "Estimating the Approximation Error When Fixing Unessential Factors in Global Sensitivity Analysis." In: *Reliability Engineering & System Safety* 92.7 (July 2007), pp. 957–960. DOI: 10.1016/j.ress.2006.07.001.

[253] S. Song and L. Wang. "A Novel Global Sensitivity Measure Based on Probability Weighted Moments." In: *Symmetry* 13.1 (Jan. 2021), p. 90. DOI: 10.3390/sym13010090.

[254] P. Soundappan et al. "Comparison of Evidence Theory and Bayesian Theory for Uncertainty Modeling." In: *Reliability Engineering & System Safety* 85.1-3 (July 2004), pp. 295–311. DOI: 10.1016/j.ress.2004.03.018.

[255] D. Spiegelhalter, M. Pearson, and I. Short. "Visualizing Uncertainty about the Future." In: *Science* 333.6048 (Sept. 2011), pp. 1393–1400. DOI: 10.1126/science.1191181.

[256] K. Staehr. *Risk and Uncertainty in Cost Benefit Analysis*. Tech. rep. Environmental Assessment Institute, 2006.

[257] G. A. Stevens and J. Burley. "3,000 Raw Ideas = 1 Commercial Success!" In: *Research-Technology Management* 40.3 (May 1997), pp. 16–27. ISSN: 1930-0166. DOI: 10.1080/08956308.1997.11671126.

[258] E. Suwondo. "LCC-OPS : Life Cycle Cost Application in Aircraft Operations." PhD thesis. Bandung: TU Delft, 2007. ISBN: 9793507926.

[259] L. Swiler et al. "Epistemic Uncertainty in the Calculation of Margins." In: *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, May 2009. DOI: 10.2514/6.2009-2249.

[260] M. Szycher. *The Guide to Entrepreneurship*. Taylor & Francis Ltd., Feb. 7, 2018. 427 pp. ISBN: 148220908X.

[261] Z.-C. Tang et al. "A Non-probabilistic Solution for Uncertainty and Sensitivity Analysis on Techno-economic Assessments of Biodiesel Production with Interval Uncertainties." In: *Energies* 11.3 (Mar. 2018), p. 588. DOI: 10.3390/en11030588.

[262] S. Tarantola, D. Gatelli, and T. Mara. "Random balance designs for the estimation of first order global sensitivity indices." In: *Reliability Engineering & System Safety* 91.6 (June 2006), pp. 717–727. DOI: 10.1016/j.ress.2005.06.003.

[263] S. Tarantola et al. "Can Global Sensitivity Analysis Steer the Implementation of Models for Environmental Assessments and Decision-making?" In: *Stochastic Environmental Research and Risk Assessment (SERRA)* 16.1 (Feb. 2002), pp. 63–76. DOI: 10.1007/s00477-001-0085-x.

[264] M. M. Taskin and A. Cansiz. "Design and Optimization of Generator for Narrow Body Commercial Aircraft." In: *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*. IEEE, Nov. 2019. DOI: 10.23919/eleco47770.2019.8990621.

[265] I. The Mathworks. *Fuzzy Interference Process*. Fuzzy Logic Toolbox. 2020. URL: https://de.mathworks.com/help/fuzzy/fuzzy-inference-process.html#responsive_offcanvas (visited on 04/04/2020).

[266] P. Thokala, J. Scanlan, and A. Chipperfield. "Life Cycle Cost Modelling As an Aircraft Design Support Tool." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 224.4 (Nov. 2009), pp. 477–488. DOI: 10.1243/09544100jaero574.

[267] P. Thokala, J. Scanlan, and A. Chipperfield. "Framework for Aircraft Cost Optimization Using Multidisciplinary Analysis." In: *Journal of Aircraft* 49.2 (Mar. 2012), pp. 367–374. DOI: 10.2514/1.c000187.

[268] J. Thorbeck and D. Scholz. "Doc Assessment Method." In: *3rd Symposium on Collaboration in Aircraft Design*. TU Berlin. Linkoeping, Sweden, 2013.

[269] J.-Y. Tissot and C. Prieur. "Bias Correction for the Estimation of Sensitivity Indices Based on Random Balance Designs." In: *Reliability Engineering & System Safety* 107 (Nov. 2012), pp. 205–213. DOI: 10.1016/j.ress.2012.06.010.

[270] U.S. Department of Defence. *Parametric Cost Estimating Handbook*. Ed. by B. Brudnick. 1995. 272 pp.

[271] U.S. Department of Transportation. *Air Carrier Financial Reports (Form 41 Financial Data)*. Ed. by Bureau of Transport Statistics. Washington, DC, 2020.

[272] U.S. Department of Transportation. *Annual Energy Outlook 2022. Table 12: Petroleum & Other Liquid Prices*. Ed. by Energy Information Administration (EIA). 2022.

[273] U.S. Department of Transportation. *T-100 Segment Data - All Airlines - All Airports*. Ed. by Bureau of Transport Statistics. last accessed on 2021/16/04. 2021.

[274] US Department of Defense. *Operating and Support Cost-Estimating Guide - Cost Assessment and Program Evaluation*. Guide. Office of the Secretary of Defense, Sept. 2020. 67 pp.

[275] G. Van Bodegraven. "Commercial Aircraft DOC Methods." In: *Aircraft Design, Systems and Operations Conference*. American Institute of Aeronautics and Astronautics, Sept. 1990. DOI: 10.2514/6.1990-3224.

[276] J. P. Van der Sluijs et al. *RIVM/MNP Guidance for Uncertainty Assessment and Communication*. Tech. rep. NUSAP, 2003.

[277] Verein Deutscher Ingenieure. *Entwicklung technischer Produkte und Systeme - Modell der Produktentwicklung*. 2221. Blatt 1. VDI. Düsseldorf, Germany, Mar. 1, 2018.

[278] C. Viavattene and H. Faulkner. "An Uncertainty Index to Measure the Feasibility of Whole-life Cycle Costing Approach in Flood Risk Management." In: *Journal of Flood Risk Management* 5.3 (Apr. 2012), pp. 215–225. DOI: 10.1111/j.1753-318x.2012.01140.x.

[279] J. Vink et al. "Dynamic Aircraft Recovery Problem - an Operational Decision Support Framework." In: *Computers & Operations Research* 117 (May 2020), p. 104892. DOI: 10.1016/j.cor.2020.104892.

[280] H.-W. M. Vos, B. F. Santos, and T. Omondi. "Aircraft Schedule Recovery Problem – a Dynamic Modeling Framework for Daily Operations." In: *Transportation Research Procedia* 10 (2015), pp. 931–940. DOI: 10.1016/j.trpro.2015.09.047.

[281] W. E. Walker et al. "Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-based Decision Support." In: *Integrated Assessment* 4.1 (Mar. 2003), pp. 5–17. DOI: 10.1076/iaij.4.1.5.16466.

[282] E. Walling and C. Vaneeckhaute. "Developing successful environmental decision support systems: Challenges and best practices." In: *Journal of Environmental Management* 264 (June 2020), p. 110513. DOI: 10.1016/j.jenvman.2020.110513.

[283] Z. H. Wen, X. Yin, and Y. P. Liu. "The Uncertainty Analysis of Health Management Strategies for Aero-engines." In: *Applied Mechanics and Materials* 201-202 (Oct. 2012), pp. 34–38. DOI: 10.4028/www.scientific.net/amm.201-202.34.

[284] T. K. West et al. "Uncertainty Quantification and Certification Prediction of Low-boom Supersonic Aircraft Configurations." In: *Journal of Aircraft* 54.1 (Jan. 2017), pp. 40–53. DOI: 10.2514/1.c033907.

[285] K. Wicke. "Bewertung der natürlichen Laminarhaltung an Verkehrsflugzeugen im operationellen Umfeld." German. PhD thesis. Hamburg, Germany: German Aerospace Center, May 30, 2016. 216 pp.

[286] K. Wicke et al. "Insect Contamination Impact on Operational and Economic Effectiveness of Natural-laminar-flow Aircraft." In: *Journal of Aircraft* 53.1 (Jan. 2016), pp. 158–167. DOI: 10.2514/1.c033237.

[287] K. Worden. "Uncertainty Analysis in Structural Dynamics." In: *Key Engineering Materials* 588 (Oct. 2014), pp. 318–332. DOI: 10.4028/www.scientific.net/kem.588.318.

[288] B. Wüsthoff. "Definition Wirtschaftlicher Kenngrößen Und Ihre Quantifizierung in Der Risikobewertung Von Technologieprojekten." Berlin University of Technology, 2003.

[289] J. Xie, E. Harrison, and D. N. Mavris. "Quantifying Impacts of Uncertainties on Certification-Driven Design." In: *AIAA AVIATION 2023 Forum*. American Institute of Aeronautics and Astronautics, June 2023. DOI: 10.2514/6.2023-4194.

[290] C. Xu and G. Gertner. "Extending a Global Sensitivity Analysis Technique to Models with Correlated Parameters." In: *Computational Statistics & Data Analysis* 51.12 (Aug. 2007), pp. 5579–5590. DOI: 10.1016/j.csda.2007.04.003.

[291] C. Xu and G. Gertner. "Understanding and Comparisons of Different Sampling Approaches for the Fourier Amplitudes Sensitivity Test (FAST)." In: *Computational Statistics & Data Analysis* 55.1 (Jan. 2011), pp. 184–198. DOI: 10.1016/j.csda.2010.06.028.

[292] J. Xu et al. "Aircraft Route Optimization for Formation Flight." In: *Journal of Aircraft* 51.2 (Mar. 2014), pp. 490–501. DOI: 10.2514/1.c032154.

[293] W. Yao et al. "Review of Uncertainty-based Multidisciplinary Design Optimization Methods for Aerospace Vehicles." In: *Progress in Aerospace Sciences* 47.6 (Aug. 2011), pp. 450–479. DOI: 10.1016/j.paerosci.2011.05.001.

[294] T. M. Young and J. P. Fielding. "Potential Fuel Savings Due to Hybrid Laminar Flow Control under Operational Conditions." In: *The Aeronautical Journal* 105.1052 (Oct. 2001), pp. 581–588. DOI: 10.1017/s0001924000012525.

[295] T. M. Young and B. Humphreys. "Liquid Anti-contamination Systems for Hybrid Laminar Flow Control Aircraft—a Review of the Critical Issues and Important Experimental Results." In: *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 218.4 (Apr. 2004), pp. 267–277. DOI: 10.1243/0954410041872825.

[296] T. Young. "Investigations into the Operational Effectiveness of Hybrid Laminar Flow Control Aircraft." en. PhD thesis. Cranfield, UK: Cranfield University, 2002. DOI: 10.13140/RG.2.2.16036.68484.

[297] L. Zadeh. "Fuzzy Sets As the Basis for a Theory of Possibility." In: *Fuzzy Sets and Systems* 1 (1978), pp. 3–28.

[298] L. A. Zadeh. "Fuzzy Sets." In: *Information and Control* 8.3 (1965), pp. 338–353. DOI: 10.1016/s0019-9958(65)90241-x.

[299] T. A. Zaidi. "A methodology for probabilistic aircraft technology assessment and selection under uncertainty." PhD thesis. Georgia Institute of Technology, Aug. 1, 2016.

[300] W. B. Zehner and G. Pletcher. "Successful Technology Commercialization – Yes or No? Improving the Odds. The Quick Look Methodology and Process." In: *Marketing of Scientific and Research Organizations* (2017). DOI: 10.14611/minib.25.09.2017.13.

[301] A. W. Zimmermann et al. "Techno-Economic Assessment Guidelines for CO2 Utilization." In: *Frontiers in Energy Research* 8 (Jan. 2020). ISSN: 2296-598X. DOI: 10.3389/fenrg.2020.00005.

[302] E. Zio and N. Pedroni. *Literature Review of Methods for Representing Uncertainty*. Tech. rep. Foundation for an Industrial Safety Culture, 2013.