POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Comparison of Uncertainty Quantification and Validation Methodologies on a Civil Tiltrotor Flight Simulation Model

Tesi di Laurea Magistrale in
Aeronautical Engineering - Ingegneria Aeronautica

Author: Matteo Mamino

Student ID: 994517
Advisor: Prof. Giuseppe Quaranta
Co-advisors: Ing. Andrea Ragazzi, Ing. Federico Porcacchia
Academic Year: 2022-23

Abstract

The recent advancements of simulation tools made them a mainstay of aircraft development and pilot training in the industry. It is predicted that demonstration of compliance to certification requirements through flight simulations may take advantage of the reduction of cost, risk and required time offered by modelling tools. To capitalize on these advantages, effort shall be devoted to the development and the validation of simulation tools of sufficient fidelity. In the present work, the most widely acknowledged methodologies in the field of verification and validation of computational models are revised and applied to a state-of-the-art flight simulation model of a civil tiltrotor developed by Leonardo Helicopters Division. The verification and validation procedures are carried out through an application of the certification by simulation process presented in the guidelines developed by RoCS project. The different validation methodologies are eventually compared in terms of fundamental assumptions, computational expense, validation metrics values and suitability to RoCS guidelines.

Keywords: certification by simulation, uncertainty quantification, model validation

Abstract in lingua italiana

Il crescente avanzamento degli strumenti di simulazione nell'industria aeronautica ha reso questi ultimi un elemento imprescindibile nello sviluppo delle macchine volanti e nell'addestramento dei piloti. Si prevede che la dimostrazione di conformità ai requisiti certificativi attraverso simulazioni di volo possa beneficiare della riduzione di costi, rischi e tempi richiesti offerta dagli strumenti di modellazione. Tuttavia, per godere di questi vantaggi, è necessario dedicare energie allo sviluppo e alla validazione di strumenti di simulazione adeguati. Nel presente lavoro, le metodologie più diffuse nel campo della verifica e della validazione dei modelli computazionali vengono presentate ed applicate ad un modello di meccanica del volo allo stato dell'arte di un velivolo tiltrotor. Le procedure di verifica e validazione sono dimostrate attraverso l'applicazione del processo di certificazione attraverso simulazione presentato nelle linee guida del progetto RoCS. Le diverse metodologie di validazione vengono infine confrontate in termini di assunzioni fondamentali, costo computazionale, valori delle metriche di validazione e applicabilità alle linee guida di RoCS.

Parole chiave: certificazione attraverso simulazione, quantificazione delle incertezze, validazione

Contents

Abstract **i**

Abstract in lingua italiana **iii**

Contents **v**

Nomenclature **1**

1 Introduction **5**
1.1 Research Objective 5
1.2 Problem Statement 6
1.3 Computational Tools 6
1.4 Sensitive Information Disclosure 6

2 Elements of Probability Theory **9**
2.1 Probability Density Function 9
2.2 Cumulative Distribution Function 9

3 RoCS Process Overview **11**
3.1 Phase 1: Requirements-Capture and Build 11
3.2 Phase 2a: Flight Simulation Model Development 15
3.3 Phase 3: Credibility Assessment 19

4 Verification and Validation Methods Review **21**
4.1 ASME Standard 21
4.1.1 Definitions 22
4.1.2 Validation Methodology 23
4.1.3 Model Adequacy Assessment 25
4.1.4 Observations 26
4.2 Roy-Oberkampf Approach 29

4.2.1 Definitions . . . 29
4.2.2 Validation Methodology . . . 30
4.2.3 Model Adequacy Assessment . . . 34
4.2.4 Observations . . . 36

5 RoCS Guidelines Application - Phase 1 39
5.1 Certification Requirements and Fidelity Metrics . . . 39
5.2 Flight Data Selection . . . 41
5.3 Domains Definition and Influence, Predictability and Confidence Levels . . . 43
5.4 Model Requirements . . . 44

6 RoCS Guidelines Application - Phase 2a 47
6.1 Flight Simulation Model . . . 48
6.2 Solution Verification . . . 51
6.2.1 Parameters of Influence . . . 53
6.2.2 Numerical Error Estimation Procedure . . . 57
6.2.3 Results . . . 61
6.3 Model Validation . . . 66
6.3.1 Input Uncertainty Identification and Characterization . . . 66
6.3.2 Sensitivity Analysis . . . 69
6.3.3 ASME Solution Validation . . . 75
6.3.4 Roy-Oberkampf Solution Validation . . . 82
6.3.5 Interval Analysis Solution Validation . . . 97
6.3.6 Validation Metrics Comparison . . . 101

7 Conclusions 107

Bibliography 109

A Flight Data Time Histories 113

B ASME VV Standard 20 - Validation Results 117

C Roy-Oberkampf Approach - Validation Results 121

D Interval Analysis - Validation Results 125

E Validation Metrics Comparison **129**

List of Figures **133**

List of Tables **139**

Nomenclature

Acronyms and Abbreviations

δ	Error
μ	Mean
σ	Standard Deviation
D	Best Estimate of the Measurand
E	Comparison Error
S	Nominal Simulation Prediction
U	Expanded Uncertainty
u	Standard Uncertainty
ACR	Applicable Certification Requirement
ASME	American Society of Mechanical Engineers
BEM	Blade Element Method
CbS	Certification by Simulation
CDF	Cumulative Distribution Function
CFD	Computational Fluid Dynamics
CG	Center of Gravity
CHT	Computational Heat Transfer
CR	Confidence Ratio
CS	Certification Specifications
DoE	Domain of Extrapolation
DoP	Domain of Prediction

DoR	Domain of Physical Reality
DoV	Domain of Validation
FCS	Flight Control System
FS	Flight Simulator
FSM	Flight Simulation Model
FTMS	Flight Test Measurement System
IPC	Influence, Predictability and Confidence
LHD	Leonardo Helicopters Division
MOAT	Morris One-At-a-Time
MoC	Means of Compliance
OGE	Out of Ground Effect
PDF	Probability Density Function
RoCS	Rotorcraft Certification by Simulation
SA	Sensitivity Analysis
UQ	Uncertainty Quantification
VV	Verification and Validation

Physical Quantities

α	Angle of Attack [°]
β	Angle of Sideslip or Rotor Flapping Angle [°]
Ω	Rotor Speed of Rotation $[\frac{rad}{s}]$
ϕ	Roll Angle [°]
ψ	Azimuth or Yaw Angle [°]
θ	Pitch Angle [°]
$\theta_{3/4}$	Rotor Blade Collective Pitch [°]
$\underline{F}$	Force Vector $[N]$
$\underline{f}$	Force per Unit Length Vector $[\frac{N}{m}]$

$A1$	Rotor Longitudinal Flapping Angle [°]
$B1$	Rotor Lateral Flapping Angle [°]
g	Acceleration of Gravity $[\frac{m}{s^2}]$
I_{xx}	Product of Inertia $\int x^2\, dm$ $[kg\ m^2]$
I_{xy}	Product of Inertia $\int xy\, dm$ $[kg\ m^2]$
I_{yy}	Product of Inertia $\int y^2\, dm$ $[kg\ m^2]$
I_{yz}	Product of Inertia $\int yz\, dm$ $[kg\ m^2]$
I_{zx}	Product of Inertia $\int zx\, dm$ $[kg\ m^2]$
I_{zz}	Product of Inertia $\int z^2\, dm$ $[kg\ m^2]$
L, M, N	Components of Resultant Moments in Aircraft Body Frame $[N\ m]$
n_{az}	Number of Integration Steps in One Rotor Revolution [-]
n_{har}	Number of Inflow Harmonics [-]
n_{rad}	Highest Power of Inflow Radial Variation [-]
n_{seg}	Number of Aerodynamic Sections [-]
n_{simps}	Number of Intervals in the Composite Simpson Integration rule [-]
p, q, r	Angular Velocity Components in Aircraft Body Frame $[\frac{rad}{s}]$
Q	Rotor Torque $[N\ m]$
t	Time $[s]$
v_i	Induced Velocity $[\frac{m}{s}]$
v_x, v_y, v_z	Velocity Components in Aircraft Body Frame $[\frac{m}{s}]$
X, Y, Z	Components of Resultant Forces in Aircraft Body Frame $[N]$
xt_{col}	Total Collective Control [°]
xt_{lat}	Total Lateral Control $[in]$
xt_{lon}	Total Longitudinal Control $[in]$
xt_{ped}	Total Directional Control $[in]$
y	Generic System Response Quantity$[-]$

ACW, m	Aircraft Mass $[kg]$
BL	CG Butt Line Station $[m]$
STA	CG Fuselage Station $[m]$
WL	CG Waterline Station $[m]$

Subscripts and Superscripts

fus	Fuselage
ht	Horizontal Tail
L	Left Rotor
nac	Nacelle
R	Right Rotor
rot	Rotors
tol	Tolerance
vt	Vertical Tail
w	Main Wing

1 | Introduction

The demonstration of compliance to certification requirements is a fundamental milestone in the development of any rotorcraft and is required to testify that the vehicle meets the safety requirements set by the certification authority. Nevertheless, the compliance demonstration process is generally the most expensive and time demanding part of the certification activity, due to the amount of necessary ground and flight testing. At the same time, the recent advancements of simulation tools made them a mainstay of aircraft development and pilot training in the industry. Hence, it is predicted that demonstration of compliance through flight simulations may take advantage of the reduction of cost, risk and required time offered by modelling tools. However, in order to deliver these benefits, some effort shall be devoted to the development and the validation of simulation tools of sufficient fidelity. Building on this view, the Rotorcraft Certification by Simulation (RoCS) project aims to explore the challenges and opportunities associated to the use of flight modelling during certification and to provide guidelines [22] for the application of flight simulation to support, either directly or indirectly, the compliance demonstration activity for helicopters and tiltrotors.

1.1. Research Objective

The objective of the present work is to revise the most widely acknowledged procedures in the field of model Verification and Validation (VV) and to frame them in the Certification by Simulation (CbS) guidelines proposed by RoCS [22], in order to establish what methodology may be overall best suited for an application in the industry and, possibly, implementation into future revisions of the CbS process.

Indeed, as also stated in [18], a distinctive aspect of the VV research field is that different, apparently irreconcilable approaches exist. In this dissertation, two approaches have been elected to primary reference work: the ASME standard for VV in Computational Fluid Dynamics (CFD) and Computational Heat Transfer (CHT) [8] and the VV approach for scientific computing proposed by Roy and Oberkampf [29]. Once framed in the CbS process, the two standards are applied to a real case scenario and compared in terms of

simplicity, fundamental assumptions, computational expense, validation metrics values and suitability to the different uses of model VV tools within the guidelines proposed by RoCS. Moreover, their applicability to the current industry operating standards is assessed, making use of a state-of-the-art Flight Simulation Model (FSM) and a set of experimental flight data both supplied by Leonardo Helicopters Division (LHD).

1.2. Problem Statement

The investigation of the research objective is pursued by emulating the implementation of the starting phases (namely, phase 1 and 2a) of the CbS process on a civil tiltrotor. EASA CS.29.143 (d) [7] low speed controllability and maneuverability certification requirements were chosen for the application presented in this work. As stated in [33], the typically expensive and time-consuming relocation to high-altitude sites, the involved risks and the modelling challenges associated to this Certification Specification (CS) makes it an interesting test bench for RoCS guidelines application.

1.3. Computational Tools

All modelling activities and simulation results presented in this dissertation are obtained with FLIGHTLAB [1]. FLIGHTLAB is a state-of-the-art, component-based, selective fidelity modelling and analysis software package, specifically tailored for the rotorcraft industry.

In addition, all the parametric, optimization, sensitivity and uncertainty quantification analyses presented in this dissertation were performed coupling FLIGHTLAB with Dakota. Dakota is an open-source software suited for a wide variety of engineering problems, including, among others, design optimization, model calibration and uncertainty quantification. It also comes with a flexible interface, based on the exchange of simple text files, which allows the user to couple it with external codes and softwares, effectively managed as "*black boxes*". More information about Dakota is available in the Dakota theory [11] and users manuals [3].

1.4. Sensitive Information Disclosure

The flight simulation model and the flight data exploited in the validation phase were both supplied by LHD. As a consequence, several quantitative results of the present work are omitted on purpose, to avoid disclosure of sensitive information. Similarly, quantitative

conclusions about the model adequacy and the validation metrics values are not presented.

2 | Elements of Probability Theory

In this section, the concepts of probability theory exploited in the course of the present dissertation are briefly presented.

2.1. Probability Density Function

In probability theory, the Probability Density Function (PDF) of a continuous variable X conveys information about the probability of X falling within a particular range. In essence, given two values a and b, the probability that X lies within the interval $[a;b]$ can be computed according to the following equation.

$$P(a \leq X \leq b) = \int_a^b \mathrm{PDF}(x)\, dx \tag{2.1}$$

The PDF satisfies two key properties, the first one being non-negativity. Namely, the value of the PDF is non-negative for all values of X, as formalized in eq. (2.2). The second property, conversely, is normalization, stating that the total area under the probability density function is equal to 1, as expressed in eq. (2.3).

$$\mathrm{PDF}(X) > 0 \quad \text{for all } X \tag{2.2}$$

$$\int_{-\infty}^{\infty} \mathrm{PDF}(x)\, dx = 1 \tag{2.3}$$

2.2. Cumulative Distribution Function

Strictly related to the PDF is the Cumulative Distribution Function (CDF). The latter is defined as the integral of the PDF, as formalized in the following equation.

$$\mathrm{CDF}(X) = \int_{-\infty}^{X} \mathrm{PDF}(x)\, dx \tag{2.4}$$

As a consequence, building on eq. (2.1), the value of CDF directly quantifies the probability of X being smaller than a given value b, as reported below.

$$P(X \leq b) = \text{CDF}(b) \tag{2.5}$$

The CDF also exhibits two important properties. First of all, it is a monotonically increasing function of X. Secondly, as X approaches negative infinity, the CDF approaches 0, while as X tends to positive infinity, the CDF approaches 1. The latter property is delivered mathematically in terms of limits, as reported in the equations below.

$$\lim_{X \to -\infty} \text{CDF}(X) = 0 \tag{2.6}$$

$$\lim_{X \to \infty} \text{CDF}(X) = 1 \tag{2.7}$$

3 | RoCS Process Overview

In the course of this dissertation, the structured CbS process proposed by RoCS committee, as it is presented in [22], is followed. The three main phases in which the guide is organized are reported in fig. 3.1. Nevertheless, only phase 1, phase 2a and part of phase 3 are briefly summarized in this section, being the ones more involved in the development of the present work.

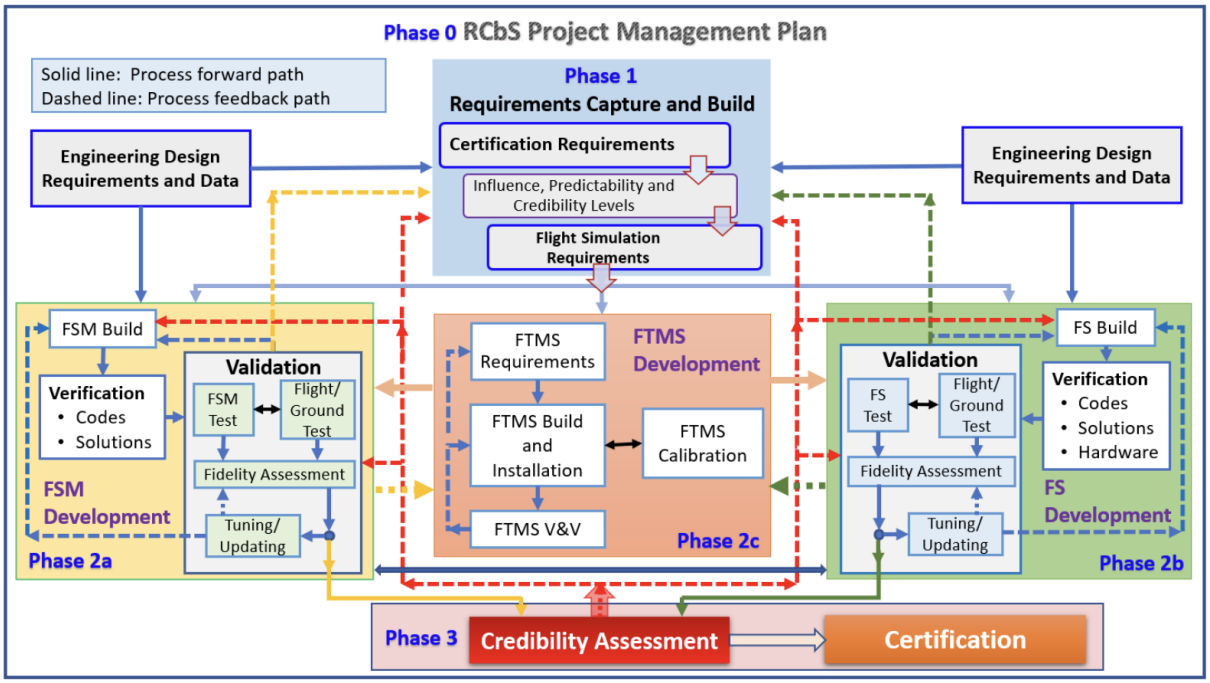


Figure 3.1: CbS process summary (from [22])

3.1. Phase 1: Requirements-Capture and Build

Before embarking in the CbS process, the applicant shall understand the problem under consideration and determine the objectives and the desired outcomes and accuracy of the simulation activity. In particular, these understandings and determinations are to be framed both in a specific perspective, strictly related to an Applicable Certification Requirement (ACR), and in a general perspective, which relates to the aircraft behaviour

throughout the flight envelope. Eventually, these understandings and determination come together in the definition of a set of requirements which the FSM, FS and FTMS must measure up to.

Thus, the requirements-capture and build phase starts with the identification of the ACRs, directly drawn from the CS, for which simulation is foreseen to play a role in compliance demonstration. Then, the identification of flight simulation Influence, Predictability and Confidence (IPC) levels for each ACR follows. The definitions of these levels are used to convey meaning to the underlying consequences of the application of the CbS process. Indeed, their choice affects the simulation-aided certification campaign in terms of safety, efficiency and effort partition between simulation and flight testing. Within RoCS framework, influence and predictability levels are effectively recognized exploiting the concept of four domains. This domains structure not only may guide the applicant in the identification of simulation influence and predictability levels, but comes in handy in the understanding of influence and predictability concepts themselves. The definitions of RoCS application domains is reported below and a sketch of the domains, together with their mutual relationship, is displayed in fig. 3.2.

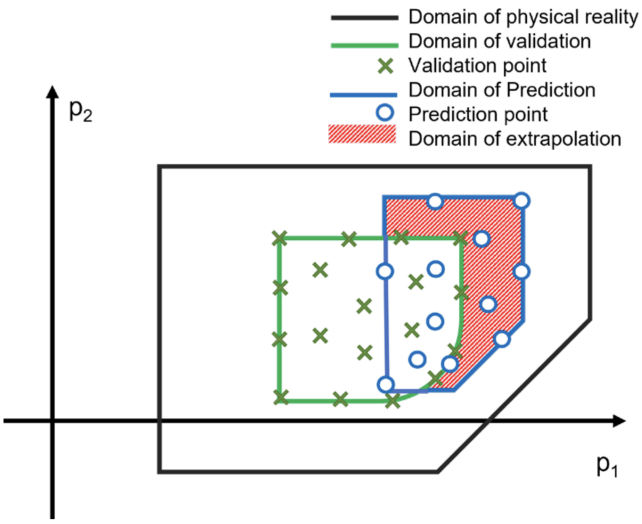


Figure 3.2: Sketch of RoCS application domains and their mutual relationship (from [22])

1. The domain of physical reality (DoR) is the domain within which the laws of physics being used are adequately represented in the flight model and flight simulator. Since all models and simulations used in the CbS process will include approximations to physical reality, this domain is strictly the region where the approximations are valid, reflecting the description 'adequately represented'. All the other domains mentioned below must be included within DoR.

2. The domain of prediction (DoP) represents the domain within which the applicant intends to predict the behaviour of the aircraft[1] and to use these predictions to achieve certification at the defined influence level for an ACR.

3. The domain of validation (DoV) is the domain within which test data that are involved in flight model validation lie.

4. The domain of extrapolation (DoE) represents the domain within which extrapolation of predictions are made to achieve certification at the defined influence level for an ACR.

Thus, the operational envelope of the rotorcraft, together with engineering data and ACRs requirements, lay the basis for defining the conditions under which the components and the features of the flight simulation model may operate (and so the complexity of the physics that is to be modelled). Then, building on the flight points in which exhibition of compliance to an ACR is requested by the certification authority, the applicant identifies the simulation influence level by deciding to what extent simulation is to be exploited[2]. In the present guidance, the levels of influence are organized into four options, defined in table 3.1. In effect, this decision explicitly marks simulation prediction points and traces the DoP boundaries within the DoR.

Influence Levels		**Description**
I1	De-risking	The simulation is used to develop/familiarise with flight test procedures and to obtain an understanding of possible problems, hazards, or the need for additional data gathering etc. No certification credit is obtained.
I2	Critical point analysis	The simulation is used to explore the flight envelope to be tested for a specific ACR and to perform a down-selection of critical points to be tested in flight, yielding improvements in test efficiency and safety.
I3	Partial credit	The simulation is used to receive certification credit for a portion of the flight-envelope/aircraft-configuration matrix, or an aspect of an ACR. Supplementary flight tests will need to be performed to obtain full credit.
I4	Full credit	Full credit This category is for cases where certification flight tests for a specific ACR are replaced by simulation.

Table 3.1: Influence levels within RoCS framework

[1] or a component.

[2] Namely, which compliance demonstrations are to be assessed with flight simulation predictions.

Within phase one, the flight testing activity is also scheduled building on the necessary flight data (either required for model validation and/or certification compliance demonstration, depending on the influence level). As an outcome of these choices, the DoV boundaries are traced as well and, eventually, the DoE is identified as the portion of DoP outside the DoV. As soon as the DoE is defined, the simulation predictability level can be identified. In the framework of RoCS guidance, the predictability level is picked out among the options presented in table 3.2.

Predictability Levels		**Description**
P1	Full Interpolation	Predictions performed within the DoV, the (interpolation) errors for the quantities of interest can be estimated with high confidence
P2	Extensive interpolation and limited extrapolation	All cases of acceptable extrapolation as per the current CS-29 and CS-27 Means of Compliance (MoC) are of predictability level P2.
P3	Limited interpolation and extensive extrapolation	A possible interpretation of extensive is - when an extrapolation model can be built from validation data that do not fall in the P2 level.
P4	Full extrapolation	All points used in simulated tests are outside the DoV and so no direct comparison of the complete FSM with flight test data is available, e.g. failure testing.

Table 3.2: Predictability levels within RoCS framework

It is important to stress that, despite the order in which the concepts are exposed above, the identification of simulation influence and predictability levels made by the applicant do not take place independently. Conversely, their choice, together with the scheduling of the flight test campaign, takes place in the same moment of the CbS process and shall reflect, among others:

1. the current modelling and simulation capability of the applicant, including expertise in modelling, fidelity assessment and model-updating techniques;
2. the logistical, economical and safety related advantages the applicant benefits from when simulation is used to support certification;
3. the will of the applicant to invest resources in flight simulation modelling, methodologies and hardware development;

4. the opinion of the certification authority (concerning the selected CSs);
5. the credibility and confidence level on simulation predictions that can be achieved by the applicant (as expanded in section 3.3).

Possibly, the selection of the influence and predictability levels conditions the minimum level of credibility expected from the results of the CbS process. The concept of FSM credibility and how minimum credibility levels are related to influence and predictability levels is tackled in section 3.3.

In the end, building on the selected ACRs, flight test campaign program and simulation IPC levels, the applicant can draft FSM, FS and FTMS requirements, fidelity metrics[3] for model validation and credibility assessment and collect all relevant aircraft design data which are necessary for the modelling activity.

3.2. Phase 2a: Flight Simulation Model Development

The FSM development process is effectively summarized in fig. 3.3. The FSM employed for ACRs compliance demonstration should capture the physics necessary to achieve sufficient fidelity (for the cases and conditions of interest). For a high level of confidence in the simulation predictions, the FSM is applied within the DoV subset of the DoP. Beyond this, in the DoE, physics should guide the model content, and the levels of confidence in the results will depend on the credibility analysis. Therefore, the FSM should be *physics-based*, i.e. "*expressed in terms of, or derived from, the physical laws applied in the creation of the mathematical model and in the operation of the numerical simulation*", as stated in [22]. Ultimately, the limits of validity of the FSM shall reflect the DoR boundaries. These limits are expressed in terms of both global variables, such as those that define the flight envelope (e.g. airspeed), and local variables (e.g. angle of attack of a specific aerodynamic component). Multiple models, with different levels of complexities and components, may also be used according to [22], with the complication driven by, and adapted to, the specific application. Thus, accounting for these considerations and relying on the FSM requisites drafted in phase one, the applicant defines and develops the flight simulation model. This particular step of the process is referred to in fig. 3.3 as "*Definition*".

[3]Namely, the system quantities whose simulation-predicted and experimental values are somehow compared to assess the model fidelity.

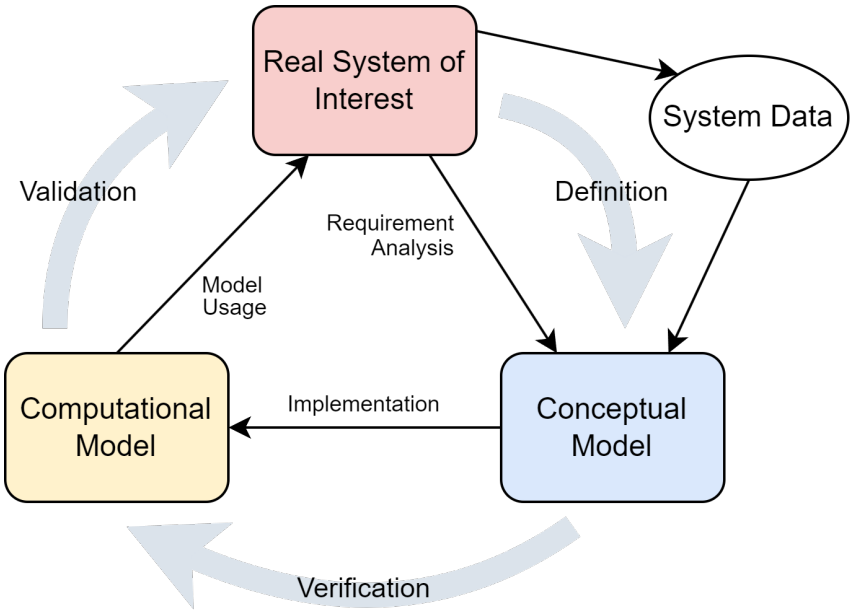


Figure 3.3: Process to create a simulation model (from [22])

Upon completion of a FSM, the Verification and Validation (VV) process is undergone, as in fig. 3.3. RoCS guidelines, in agreement with the VV framework in the literature ([8], [28]), divide the verification phase into two parts, namely *code verification* and *solution verification*.

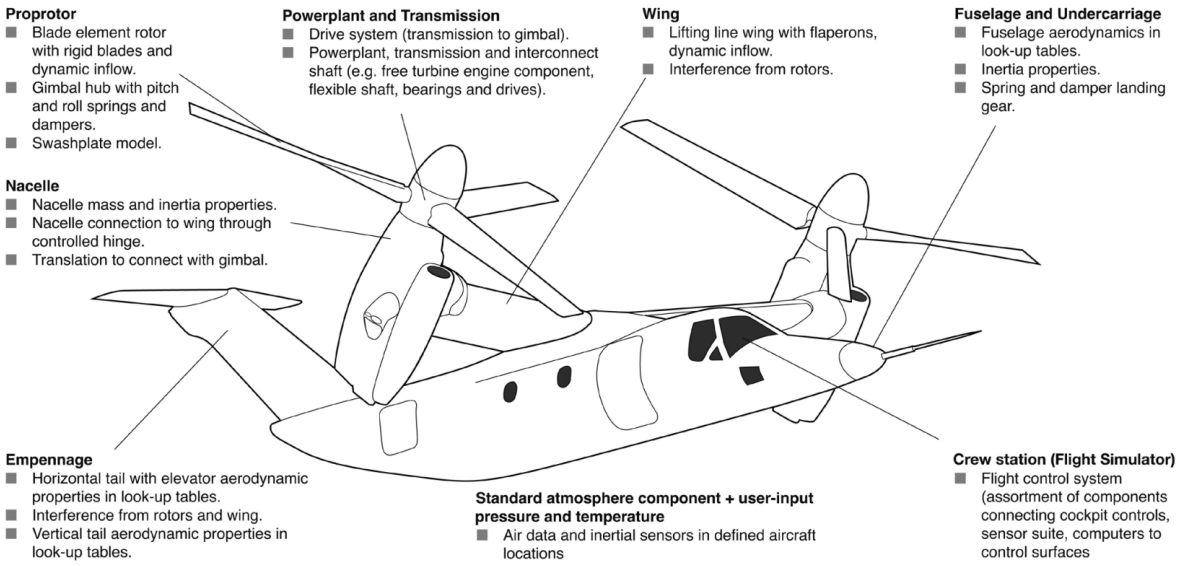


Figure 3.4: Outline of a component-based tilt-rotor flight simulation model (from [22])

Code verification, according to the definition in [28], represents "*the process of determining that the numerical algorithms are correctly implemented in the computer code and of*

identifying errors in the software". Typically, its assessment is carried out in two stages. At first, software quality assurance operations are performed to collect evidence of the correct implementation of the code and to verify that the code produces repeatable results in a specified hardware and software environment. Then, in the course of the second stage, model results are compared against so-called *benchmark solution* or *verification benchmarks*. These benchmarks can either be analytical solutions or numerical solutions which have been generated appropriately. Whenever commercial softwares are exploited, the better part of code verification is often performed by the code provider. Nevertheless, the applicant needs to collect evidence that the verification has been undertaken.

On the other hand, always according to [28], solution verification represents "*the process of determining the correctness of input data, the numerical accuracy of the solution obtained, and the correctness of the output data for a particular simulation*". Within solution verification, particular attention is generally devoted to the characterization of the *numerical accuracy*. Indeed, due to finite tolerances for iterative methods, temporal and spatial discretizations and finite machine precision, every computational model prediction is affected by some degree of numerical uncertainty and its assessment is crucial for model validation.

Upon completion of the verification process, as reported in fig. 3.3, solution validation takes place. Within the CbS process, this stage is identified with the process referred to as *model validation* in the *restricted view of validation* promoted in [28]. It consists in the quantification of the accuracy of the computational model and, from a practical standpoint, it involves the comparison of fidelity metrics[4] via an appropriate operator, which accounts for simulation and test measurements uncertainties and is referred to as *validation metric*. Despite the aforementioned definition being generally accepted in literature (fig. 3.5 and 3.6), no unanimous agreement on the methodologies that are to be exploited for *model validation* has been found to these days. Indeed, as stated in [18], different approaches exist which are irreconcilable at a fundamental level: from basic definitions to the validation metric. A review of the most widely recognized approaches and their differences is reported in chapter 4.

[4]selected during phase 1.

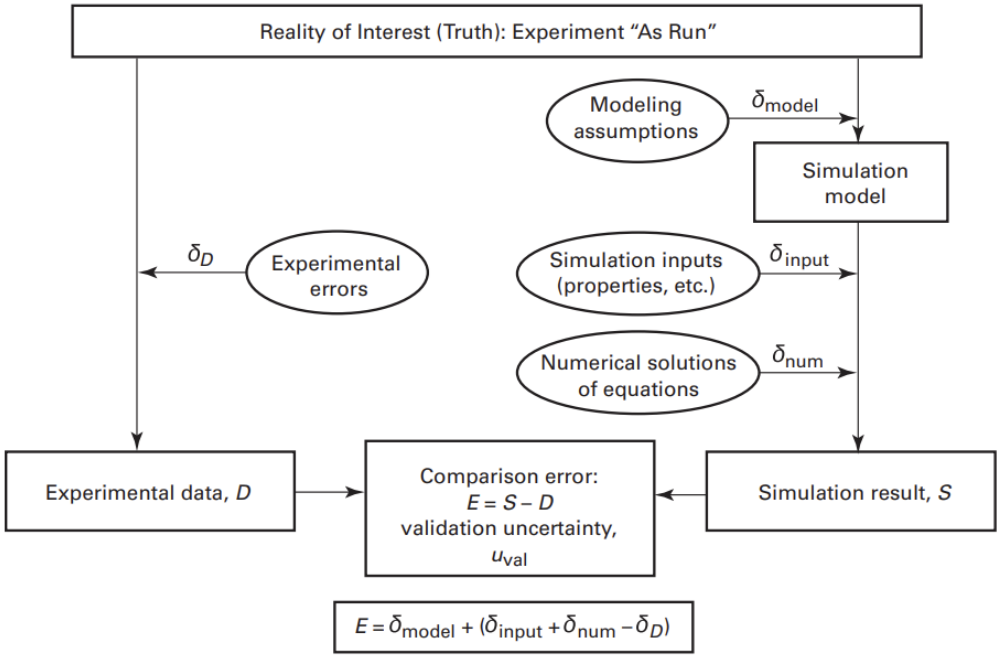


Figure 3.5: Overview of the sources of uncertainties and model validation process proposed in the ASME VV Standard 20 [8]

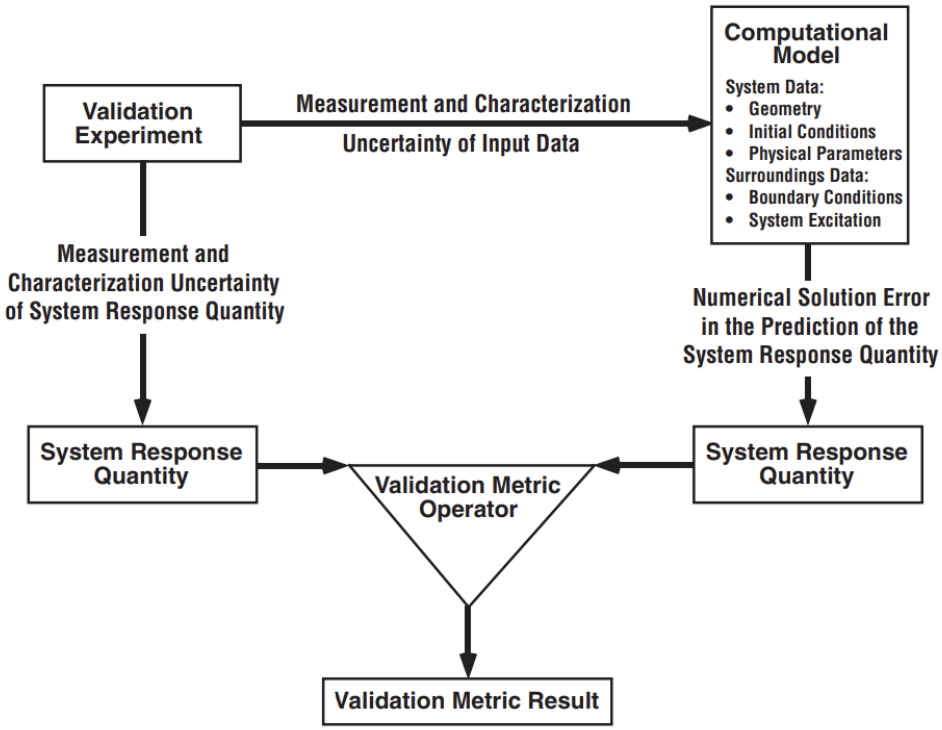


Figure 3.6: Overview of the sources of uncertainty and model validation process proposed by Roy and Oberkampf in [28]

Nevertheless, regardless the methodology chosen by the applicant for model validation, an iterative model tuning and updating phase is expected in the CbS process. Indeed, during pre-certification FSM development, validation results may convey evidence that the FSM fidelity is insufficient. In this event, the applicant should investigate the causes of the discrepancy and carry out physics-based updates to the FSM. Improvements in the model fidelity may be achieved via relaxation of modelling assumptions, addition of previously neglected dynamics, calibration of model parameters.

At completion of phase 2a, the applicant proceeds to perform the credibility assessment and to exploit the validated FSM to support (according to the intended levels of influence and predictability) the certification activity.

3.3. Phase 3: Credibility Assessment

Within the CbS process, phase 3 encompasses both to the model credibility assessment and certification. However, in the present work, the part related to credibility assessment only is briefly summarized. Moreover, in order to deliver a technical and unambiguous description of what the credibility assessment is, the author has to rely on some definitions and concepts (e.g. model-input, numerical and model-form uncertainty) that are foundational to any validation framework and are yet to be formally introduced in this report. For all these elements, an in-depth description is provided in chapter 4.

The bulk of credibility assessment is associated to what is referred to as *model adequacy assessment* in Roy and Oberkampf *restricted view of validation* [28]. In this phase, the applicant focuses on quantifying the model predictive capability accuracy level (affected by model-input, numerical and model-form uncertainties) and on establishing whether or not the accuracy level is satisfactory for the purpose of the simulation activity. Hence, the credibility assessment shall be judged directly in the DoP and requires usage of Uncertainty Quantification (UQ) techniques to account for model-input uncertainties in FSM predictions. Moreover, it is important to stress that also model-form error interpolation/extrapolation becomes relevant at this stage, in order to include model-form uncertainties, evaluated within the DoV in the course of model validation, within the overall model prediction uncertainty. In this regard, it is critical for the applicant to maximise the confidence related to model extrapolation, by developing a strong background in the understanding of the physics behind model predictions variations and by exploiting a sufficient number of point in the DoV to develop the extrapolation.

From a practical perspective, model adequacy or credibility is quantitatively judged through a comparison with prescribed margins of acceptability. According to RoCS guide-

lines, model adequacy requirement may be defined in terms of conventional acceptability margins, alike the ones typically exploited to judge model fidelity, or directly phrased in terms of ACRs, as exemplified by the concept of Confidence Ratio (CR). The CR, as reported in eq. (3.1), relates the relative size of a prediction uncertainty U and the prediction margin M to ACR compliance, so that a minimum acceptable value of CR may be appropriate to define a minimum acceptable confidence level of the simulation prediction.

$$\mathrm{CR} = \frac{\mathrm{M}}{\mathrm{U}} \tag{3.1}$$

Building on this concept, [22] proposes a system for minimum model prediction confidence definition which is based on selected influence and predictability levels. An example of a possible application of this concept is reported in table 3.3.

Influence Level	**Predictability Level**			
	P1	P2	P3	P4
I1	Low	Low	Low	Low
I2	Low	Low	Medium	Medium
I3	Low	Medium	High	High
I4	Medium	Medium	High	Very High

Table 3.3: Influence-predictability levels matrix with minimum confidence levels (from [22])

Finally, similarly to model validation, several iterations may be necessary to the applicant to demonstrate model credibility as well. As a matter of fact, as soon as the adequacy assessment conveys evidence that the FSM prediction uncertainty is too large, the applicant is required to either update the FSM or reduce the uncertainty associated to model-inputs and numerical approximations (depending on what the most dominant contribution is).

4 | Verification and Validation Methods Review

Verification and validation in scientific computing is still an ongoing domain of research. However, as anticipated in section 3.2, a distinctive aspect of VV research field is that different, apparently irreconcilable approaches exist. In this dissertation, two approaches have been elected to primary reference work: ASME standard for verification and validation in CFD and CHT [8] and Roy and Oberkampf verification and validation for scientific computing book [28] and conference paper [29]. The reasons behind this selection can be traced back to the evidence that both approaches are somehow referred to in RoCS guide [22]. Therefore, in this section, these two methodologies are presented and discussed, devoting particular attention to the proposed validation process (and metrics) and stressing the irreconcilable differences in uncertainty definitions and management.

4.1. ASME Standard

As highlighted in [18], the foundational idea of ASME VV 20 is to extend the standard approaches adopted for experimental uncertainties to include numerical and model-input uncertainties in a comprehensive validation framework. Thus, within this structure, numerical, input-related and experimental uncertainties do not belong to *independent uncertainty domains* but, conversely, they are interpreted as the very same class of objects. This idea was first proposed and rigorously structured by Coleman and Stern [31] in 1997. Specifically for CFD applications, building on the ground-breaking work of Richardson [24] and Roache [27], [25], [26] on the quantification of discretization uncertainty, Coleman and Stern devised a VV methodology and put forward two validation metrics: the comparison error E and the validation standard uncertainty u_{val}. Then, these methodology and validation metrics evolved into the current ASME VV standard for CFD presented in [8].

4.1.1. Definitions

Error and uncertainty definitions adopted by this standard are directly inherited by [17] and their concepts are extended in order to be applied to the solution variable from a simulation as well. In particular:

- the *error (of a measurement)* δ is defined as the "*result of the measurement minus a true value of the measurand*";
- the *uncertainty (of a measurement)* u is identified as the "*parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand*".

Moreover, depending on the method exploited for its evaluation, the uncertainty u can be classified into two categories:

- *Type A*, if the method of evaluation of uncertainty is based on the statistical analysis of series of observations
- *Type B*, if the method of evaluation of uncertainty relies on means other than the statistical analysis of series of observations

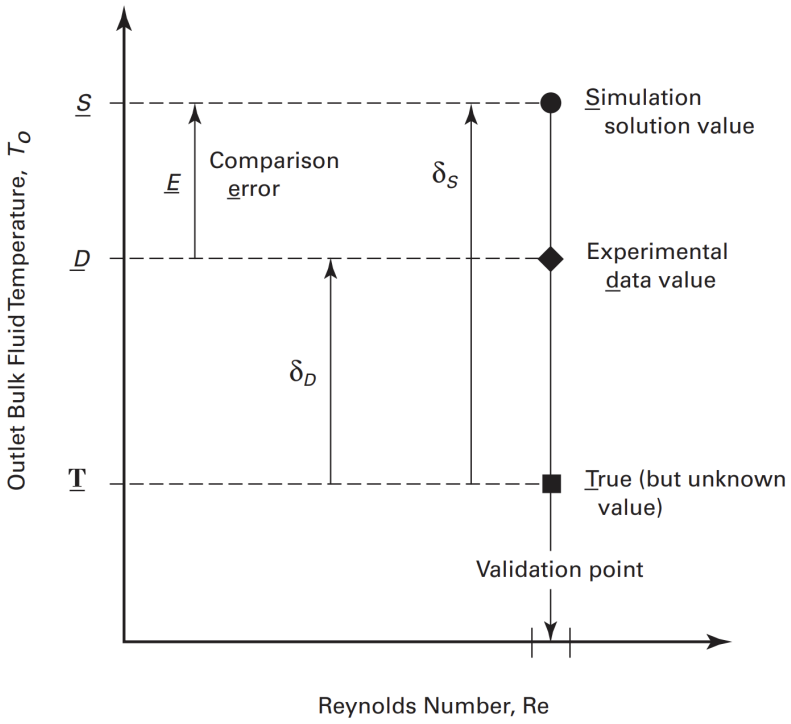


Figure 4.1: Partial nomenclature schematic for ASME VV 20 approach (from [8])

Thus, in this context, an error δ is a quantity characterized by a specific sign and magni-

tude (which, if known, allow the error to be removed by correction). On the other hand, in agreement with [16] and [17], the uncertainty u corresponds to the *standard uncertainty*, estimate of the standard deviation σ of the parent distribution (of unknown form) from which δ is a single realization. Building on these foundational concepts and definitions, the whole nomenclature system adopted by ASME standard, summarized in fig. 3.5 and fig. 4.1, can be devised.

Once the System Response Quantities (SRQ) upon which the applicant desires to apply the validation process (namely, the fidelity metrics) are identified, the following quantities can be derived.

- S represents the simulation prediction of the fidelity metric value in a given point of the validation domain.
- D represents the experimental datum (namely, the best estimate of the measurand, according to [16]) of the fidelity metric in the same validation point.
- T represents the true (but unknown) value of the fidelity metric.

As a consequence, the simulation error δ_S, the experimental error δ_D and the comparison error E, displayed in fig. 4.1, are devised through eq. (4.1), (4.2) and (4.3).

$$\delta_S = S - T \tag{4.1}$$

$$\delta_D = D - T \tag{4.2}$$

$$E = S - D = \delta_S - \delta_D \tag{4.3}$$

Then, as in [31], the simulation error is decomposed into three separate contributions, as reported in eq. (4.4).

- the error δ_{model}, due to modelling assumptions and approximations
- the error δ_{num}, due to the numerical solution of the model equations
- the error δ_{input}, due to error in the simulation input parameters (e.g. boundary conditions, geometrical and physical quantities related to the modelled system)

$$\delta_S = \delta_{model} + \delta_{num} + \delta_{input} \tag{4.4}$$

4.1.2. Validation Methodology

The goal of the validation process is to estimate the model contribution δ_{model} to the simulation error, which, rearranging eq. (4.3) and (4.4), can be expressed through eq. (4.5).

$$\delta_{model} = E - (\delta_{num} + \delta_{input} - \delta_D) \tag{4.5}$$

Focusing on the right-hand side of eq. (4.5), it is possible to observe that once S and D are determined (the former via model prediction and the latter according to the methodologies reported in [16]), the comparison error E is known from eq. (4.3). Nevertheless, the magnitude and signs of δ_{num}, δ_{input} and δ_D are unknown. Still, despite this evidence, standard uncertainties u_{num}, u_{input} and u_D associated to the corresponding errors parent populations can be defined, thus leading to the definition of a validation standard uncertainty u_{val}. This validation standard uncertainty effectively represents an estimate of the standard deviation of the parent population whose one single realization is obtained as $(\delta_{num} + \delta_{input} - \delta_D)$. Thus, its estimation, which is a pivotal point of the methodology hereby presented, provides an indication of the dispersion of δ_{model} parent population around the comparison error E. Equation (4.6) formalizes this idea, providing an explicit relation between δ_{model}, objective of the model validation, and E and u_{val}, the selected validation metrics for this validation methodology.

$$\delta_{model} = E \pm u_{val} \tag{4.6}$$

Thus, at this stage, the computation of u_{val} represents the only open question. If δ_{num}, δ_{input} and δ_D are independent, as it is typically assumed, the validation uncertainty can be expressed as a function of the other errors uncertainties through the Root Sum of Squares (RSS), eq. (4.7), hence effectively shifting the problem to the need of quantifying u_{num}, u_{input} and u_D.

$$u_{val} = \sqrt{u_{num}^2 + u_{input}^2 + u_D^2} \tag{4.7}$$

The computation of u_{num} takes place during solution verification and methodologies which strictly apply to CFD and CHT problems are presented in [8]. Nevertheless, it is important to stress that these methods, which rely on Richardson extrapolation and Roache's grid convergence index, lead to an expanded uncertainty estimate U_{num} (with a confidence interval of approximately 95%) of the numerical contribution. Thus, an assumption on the value of the coverage factor necessary to compute u_{num} starting from U_{num} is needed. In regard of u_{input}, both local and global uncertainty propagation methods are presented in [8]. Thus, as soon as model inputs uncertainties are characterized, u_{input} can be directly computed exploiting these techniques. Eventually, concerning u_D, the standard procedures for measurand uncertainty estimation reported in [16] are accepted.

Upon completion of the validation process, the applicant has direct access to the values of the validation metrics E and u_{val} and two corollaries follow:

- if $|E| >> u_{val}$, then it is safe to assume that δ_{model} has the same order of magnitude of E and that its estimation is reliable;
- if $|E| \leq u_{val}$, probably δ_{model} is of the same order of magnitude, or even smaller, than $(\delta_{num} + \delta_{input} - \delta_D)$.

Thus, from a practical standpoint, it is possible to conclude that reducing the validation uncertainty[1] is beneficial to obtain a reliable quantification of the modelling error and set up model improvements (i.e. updates to reduce the modelling error). Otherwise, whenever δ_{model} is within the "*noise level imposed by numerical, input and experimental uncertainty*", as stated in [8], formulating and measuring modelling improvements is difficult.

Eventually, since E and u_{val} characterize the model-form contribution to the simulation prediction error, any assessment aimed at proving the sufficiency of FSM fidelity (intended as the compliance of the model-form error to prescribed acceptable margins) can be directly carried out on these metrics. Indeed, whenever the acceptable margin for the FSM model-form error encompasses both E and its uncertainty (derived from u_{val}) for all fidelity metrics, it is possible to state that the prescribed fidelity requirements have been met.

4.1.3. Model Adequacy Assessment

Concerning model adequacy assessment and model-form error extrapolation, no explicit procedure is provided in [8]. However, building on the definition of model adequacy assessment delivered in section 3.3, E and u_{val} alone are not suited for the completion of this task, since at this stage the analyst aims at evaluating the total uncertainty on the FSM predictions (represented by δ_S, and not limited to δ_{model}). As a consequence, the information that E and u_{val} convey, appropriately extrapolated into the DoP, shall be combined with the quantification of δ_{input} and δ_{num} to obtain a reliable estimate of δ_S on certification-aimed model predictions. Then, building on this estimate, the total uncertainty on model prediction U can be quantified, and model adequacy assessment can be carried out relying, for example, on the CR (fig. 4.2).

[1] e.g. via rigorous model inputs uncertainty estimation and validation-oriented flight testing.

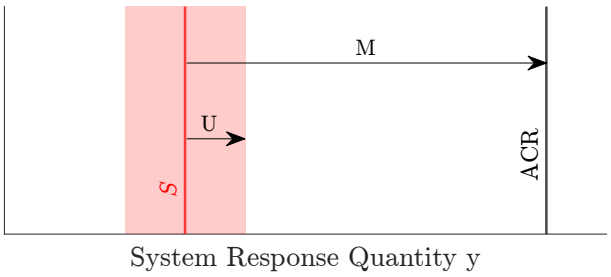


Figure 4.2: Example of model adequacy assessment within ASME validation framework

4.1.4. Observations

One fundamental assumption of the methodology exposed in section 4.1.1 concerns the definition of uncertainty and its quantitative characterization. Since the process is based on the extension to simulations of the experimental approach [16], it completely relies on the mathematical framework of the latter. Thus, regardless the type (random or systematic) and the nature (numerical, model-form, due to inputs or experimental) of the error, and independently from the available knowledge of its population[2], δ must be characterized by a probability density function. Namely, quantities characterized by a single or multiple intervals, with no assumption on their PDF, are not conceived in this framework. This foundational definition represents a crucial difference with the approach proposed by Roy and Oberkampf, as it is cleared in section 4.2, and, as a consequence, results in significant dissimilarities in terms of how uncertainties are handled and of validation metrics.

A second observation concerns the management of the numerical uncertainty. Indeed, being the VV standard tailored for CFD and CHT applications, no procedure for flight mechanics problems and multi-body models is reported. As a consequence, no direct instruction to estimate the numerical error standard uncertainty starting from a generic error magnitude estimate is proposed. In addition, also the methodologies based on Richardson extrapolation and Roache's grid convergence index eventually require the analyst to make assumptions on the PDF of the discretization error. Nevertheless, this event is well characterized within this framework, and it is typically referred to as *type B uncertainty evaluation*, as reported in [16]. On such occasion, the uncertainty is evaluated by means other than statistical analysis, thus representing an educated guess the analyst can make building on the information available.

Concerning the account of model-form uncertainty on simulation predictions in other

[2] which in turn, according to the classification proposed in [16], lead to either type A or type B uncertainty evaluation.

points within the validation domain (interpolation) and beyond its boundaries (extrapolation), no instructions are reported in the standard. However, putting aside the issue of characterizing the trend of the model-form error beyond validation points, it is reasonable to think that a practical methodology may be conceived within this framework, based on eq. (4.1). Indeed, if realizations of δ_{input} and δ_{num} can be obtained in the prediction point of interest[3] and samples of δ_{model} can be derived building on its statistics (E, u_{val}), it is in principle possible to characterize δ_S affecting the nominal simulation prediction S outside the DoV.

Concerning the validation corollaries, whenever $|E| >> u_{val}$ is not achieved, but the analyst wants to perform some model update and tuning operations, the validation framework allows to inspect the different contributions to u_{val}, reported in eq. (4.7). Then, building on this information, focused interventions may be carried out in order to reduce the validation uncertainty effectively, acting on the dominant contributions.
In addition, starting from the corollaries, it is also possible to formulate some guidelines concerning fidelity assessment, as stated in [18]. Indeed, if the uncertainties contributing to u_{val} are foreseen to be of the same order of magnitude of the required model fidelity level (e.g. maximum/minimum allowable δ_{model} for a set of fidelity metrics), at the end of model validation it might not be possible for the applicant to carry out any sensible conclusion. In fact, even in the best case scenario ($E = 0$), the uncertainty associated to δ_{model} would prevent the analyst to confidently state that the model fidelity target has been reached. Similar conclusions, also apply to the model adequacy assessment. If FSM adequacy requisites are formulated in absolute terms[4], and not referred to specific ACRs, it is important to ensure that the model-input and numerical uncertainties alone do not exceed the prescribed margins in the prediction point of interest. Otherwise, even neglecting model-form error, the applicant will not be able to carry out predictions that are useful to support the certification activity. Hence, from a practical perspective, whenever possible, it is worthy to devote some effort to a preliminary estimation of expected experimental and input uncertainties, in order to gain awareness of what the best achievable fidelity and adequacy levels are.

Eventually, it is worthy to mention one observation regarding the possible uses the analyst can make of E and u_{val} (and their relationship with the procedure corollaries) in RoCS. As far as the author is concerned, two major alternatives take shape.
The first one is to exploit the validation technique to derive a quantitative, accurate prediction of the model error (and of its uncertainty) for model fidelity and adequacy assessment

[3] through the same methodologies used for the quantification of u_{inp} and u_{num}.
[4] e.g. a given SRQ shall have an associated predictive uncertainty of 5%.

or certification-aimed model predictions. For these intended uses, the assumptions upon which the uncertainty quantification and combination process is founded[5] should be at least assessed in order to increase the credibility of the final result.
The second use, on the other hand, is provide the analyst an effective tool to measure improvements in the model update and tuning phases and to avoid drawing misleading conclusions due to interference of numerical, experimental and model input errors in the process. When exploited in this sense, the assumptions regarding the uncertainty propagation and combinations procedures move to the background, since the computation of u_{val} may serve only to provide evidence that $u_{val} << E$ and that E is a good proxy for the model-form error. At the same time, the computational cost of the validation procedure becomes of paramount importance, since many calibration data may be included in the model tuning phase and since model update might require several iterations to achieve the desired level of fidelity.
Building on these views, cheaper alternatives to the algorithm proposed by ASME may be conceived for the second use.

[5]To name a few: numerical error estimates combination and conversion to uncertainty, independence of the numerical uncertainty from model input uncertainty in their aggregation, extension of the validation standard uncertainty, etc.

4.2. Roy-Oberkampf Approach

The VV approach proposed by Roy and Oberkampf[6] is summarized in their 2010 conference paper [29] and is based on the work of Ferson et al. [15], [13] and [14]. In this section, a high-level overview of the process is provided. More details about their VV framework and related worked-out examples are tackled in [28].

At last, it is also worthy to mention that the Roy-Oberkampf framework definitions and procedures are explicitly referred to in AIAA recommended practices for certification supported by flight simulation [4].

4.2.1. Definitions

At the foundation of the Roy-Oberkampf VV framework stands the classification of uncertainties into two categories: aleatory (or irreducible) and epistemic (or reducible).

- Aleatory uncertainties represent "*the inherent variation in a quantity that, given sufficient samples of the stochastic process, can be characterized via a probability distribution*". In this framework, this kind of uncertainty is also referred to as irreducible, due to its inherent randomness, and it is mathematically characterized with a PDF.
- Epistemic uncertainties identify the instances "*where there is insufficient information concerning the quantity of interest to specify either a fixed value*[7] *or a precisely known probability distribution*". This kind of uncertainties are also referred to as reducible, since they could, in principle, be reduced if additional knowledge is supplied. Epistemic uncertainties can be represented mathematically as intervals with no associated probability distribution or with a PDF, which, as a consequence, does not reflect the probability of occurrence of a given value but, conversely, the degree of belief of the analyst. However, in this framework, all epistemic uncertainties are treated as intervals.

Then, in agreement with the ASME framework presented in section 4.1, Roy and Oberkampf decompose the uncertainty affecting the simulation prediction into three contributions.

- Uncertainty due to model inputs can be either aleatory or epistemic and, in principle, should involve all inputs whose associated uncertainty is non negligible. Its evaluation occurs in the course of model validation.

[6] From now on referred to as Roy-Oberkampf approach in this work.
[7] e.g. a model parameter, such as a turbulence model coefficient in a RANS CFD simulation.

- Uncertainty due to numerical approximation includes discretization errors, iterative convergence errors and round-off errors. In addition, whenever a sampling technique is exploited to propagate input uncertainties through the model, the numerical error associated to the convergence of the propagation technique shall be considered as well. Within this framework, numerical uncertainties are always treated as epistemic, thus with no associated PDF, and their estimation takes place during the solution verification.
- Uncertainty due to model form error is treated as epistemic and it is estimated during solution validation.

4.2.2. Validation Methodology

The verification and validation process begins with the identification and characterization of all model inputs uncertainties. In this stage, only the inputs for which there is convincing evidence[8] that associated uncertainties would result in minimal uncertainty in all fidelity metrics may be excluded. For all the other instances, a mathematical structure shall be assigned.

Then, solution verification is undertaken. Within this stage of the VV process, estimates for discretization, iterative convergence and round-off error shall be computed and explicitly converted into epistemic uncertainties. The reason why numerical errors are managed as epistemic uncertainties (within this framework) should be traced back to the inherent difficulties associated with obtaining accurate estimates of the error themselves. Indeed there are techniques for the estimation of discretization and iterative convergence error in CFD and CHT ([28], [24], [27], [25] and [26]). However, the reliability of these procedures ultimately depends on the specific case and SRQ of interest while, for some other models (e.g. multi-body models), rigorous methodologies for numerical error quantification may not even exist. Thus, starting from a numerical error estimate, the simplest method to quantify associated epistemic uncertainty is to use the magnitude of the error estimate to define an interval. Then, the different contributions from discretization, iterative convergence and round-off may be combined as reported in the following equation.

$$U_{NUM} = U_{DE} + U_{IT} + U_{RO} \tag{4.8}$$

The numerical error bound provided by eq. (4.8) shall be computed for each SRQ of interest and, in principle, as a function of simulation input quantities (since, in general, the numerical error may change within the model domain of physical reality). Nevertheless,

[8] e.g. as an outcome of a preliminary input sensitivity analysis on the model.

it is fairly common to explicitly compute this estimate in a limited number of conditions only, then extrapolating the resulting uncertainty over the entire physical domain.

Upon completion of the solution verification stage, input uncertainties (previously identified and characterized) are propagated through the model. For this stage, Roy and Oberkampf propose Monte Carlo sampling method in [29], due to its flexibility and capability of dealing with both aleatory and epistemic input uncertainties. In [28] and [29], methodologies for aleatory or epistemic input uncertainties only are proposed as well. However, since they are simplified cases of the more general instance[9], and considering that the general case is of interest for the rest of this work, the mixed uncertainty case only is hereby reported.

Whenever both aleatory and epistemic uncertainties occur, the sampling of each type of uncertainty shall be separated. Indeed, considering that there is no PDF associated to an epistemic model input, a sample represents only a single possible realization over the interval-valued range. On the other hand, inputs which are characterized by a PDF shall be sampled considering the different probability of occurrence of different values. Hence, a nested sampling method is proposed in [29]. For each realization of the epistemic interval-valued inputs, a single discrete CDF for each SRQ is realized by propagating the remaining inputs through the model. Eventually, after all epistemic and aleatory samples have been computed, an ensemble of discrete CDFs for each SRQ of interest is obtained. Then, the boundary envelope of the ensemble is used to form a so called *probability box* (sometimes also referred to as *p-box* or *imprecise probability function*). A summary of the sampling process, which involves two nested sampling loops, is reported in fig. 4.3.

The probability box is a peculiar type of CDF. It is an interval-valued probability structure which delivers information on both aleatory and epistemic uncertainties but without confounding the two. The typical appearance of a probability box is reported in fig. 4.4. The possible values of the SRQ for a given percentile of the CDF quantify the epistemic uncertainties contribution to the SRQ. On the other hand, if a single realization of the interval-valued inputs is considered, the corresponding single CDF represents the uncertainty due to aleatory inputs. The two contributions remain distinct and can be accessed by the analyst by the solely interrogation of the probability box[10].

[9] at least for Monte Carlo sampling.

[10] As anticipated before, the instances in which only PDF-prescribed or interval-valued uncertainties are present represent simplified cases of the general approach summarized in fig. 4.3. In these cases a single sampling loop is required and, eventually, the resulting probability structure is a degenerate probability box. If no interval-valued uncertainty is present, the CDFs ensemble of the general case is constituted of a single CDF only, and no interval-valued uncertainty on the SRQ is present. On the other hand, if only interval-valued uncertainties are present, the pbox degenerates into a rectangle, effectively representing the interval of possible values for the SRQ with no associated probability.

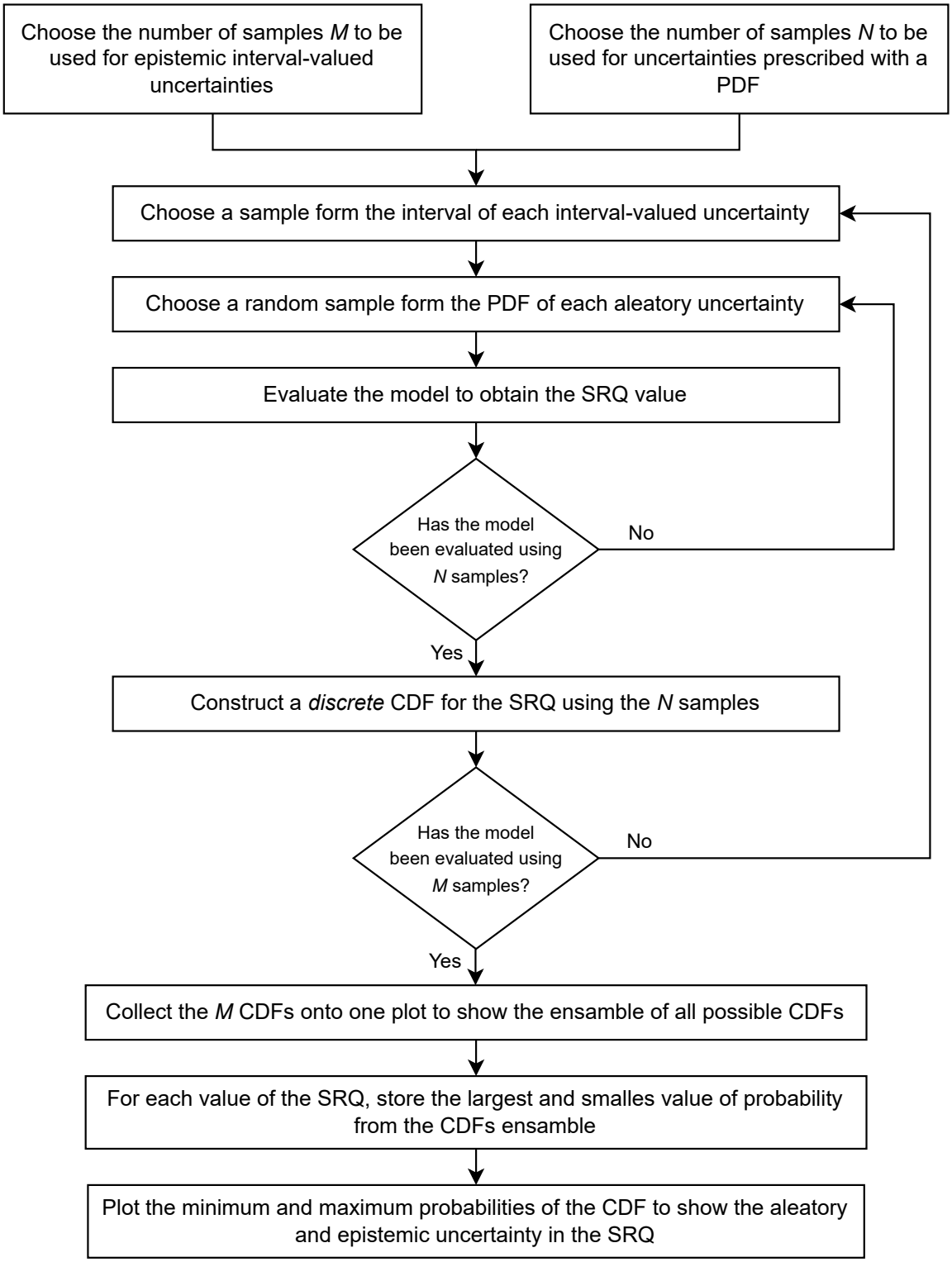


Figure 4.3: Flow chart of nested sampling loops for propagation of both aleatory and epistemic uncertainties (from [28])

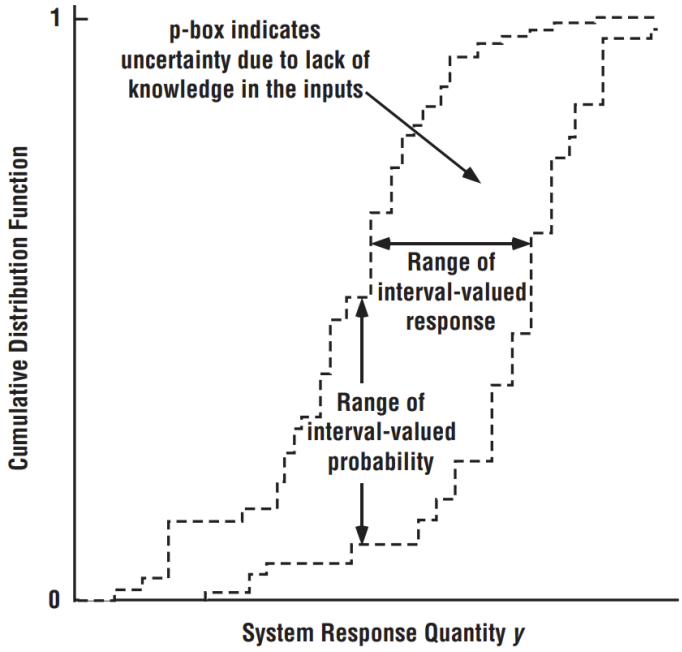


Figure 4.4: Example of the p-box obtained for a SRQ y (from [29])

As the probability box for each SRQ is available, the model form uncertainty can be estimated via model validation. At first, the numerical epistemic uncertainty U_{NUM}, estimated in the course of solution verification, is appended to the probability box. From a practical standpoint, this operation results in an enlargement of the epistemic uncertainty associated to the SRQ prediction, since both the left and right boundaries of the p-box are shifted outwards by U_{NUM}.
Then, the experimental measurement and its uncertainties are reported against the p-box graph. In this framework, all experimental realization of the SRQ are exploited to construct another discrete CDF.
Eventually, the minimum area of non-overlap included between the two probability structures (the p-box from the simulation and the discrete CDF from experimental data) is identified as the validation metric d (also referred to as *area metric*). A visual representation of the area metric concept is reported in fig. 4.5.

The area metric value, dimensional and characterized by the same measurement unit of the SRQ (alike a comparison error), effectively represents the measure of disagreement between the model and the experiment according to the knowledge available to the analyst. Any disagreement, as stated in [29], is attributed to model-form error. As a consequence, model fidelity requirements compliance can be checked, in principle, comparing the area metric value with appropriately delivered acceptability margins.

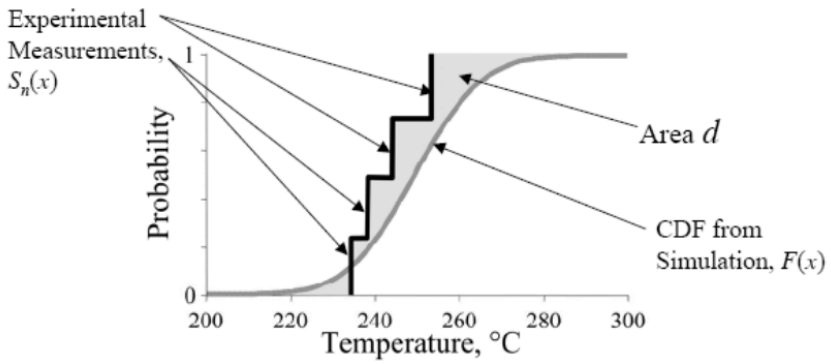


Figure 4.5: Area validation metric example with four experimental measurements available for the SRQ (from [29]). The uncertain simulation prediction is represented by a single CDF, meaning that there is no interval-valued uncertainty in model inputs and that numerical uncertainties have been neglected

4.2.3. Model Adequacy Assessment

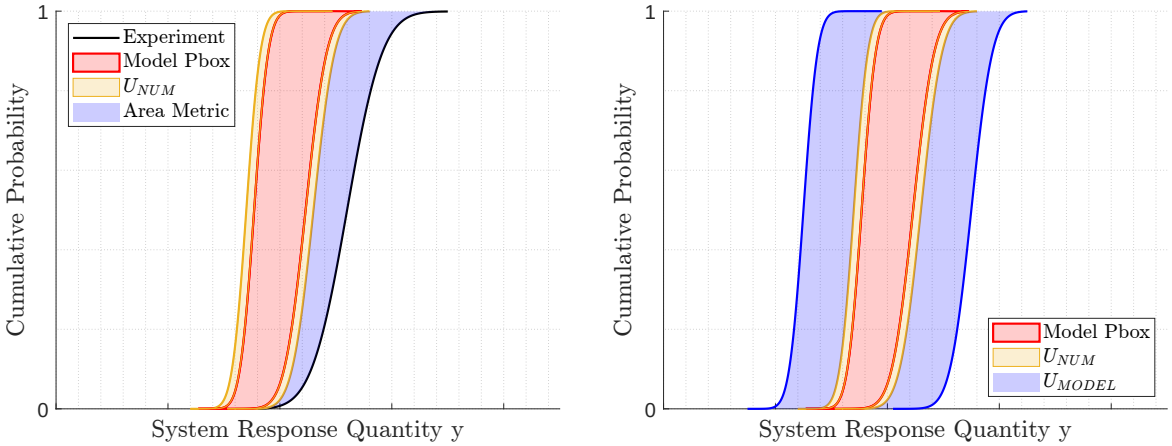
Unlike in [8], Roy and Oberkampf provide general guidelines to account for the model-form error also in the domain of application (thus in points other than validation ones). For these predictions, the model error is managed as an epistemic uncertainty with magnitude equal to the area metric value, as reported in eq. (4.9).

$$U_{MODEL} = d \tag{4.9}$$

By "correcting" the area metric value for the increased uncertainty due to the interpolation/extrapolation process, the epistemic model-form uncertainty can be derived and applied to model predictions beyond the validation points. Thus, the total uncertainty in the SRQ at the condition of interest for the analysis may be computed as follows.

1. At first, the SRQ probability box is generated propagating input uncertainties, according to the algorithm summarized in fig. 4.3.
2. Secondly, the numerical U_{NUM} and the extrapolated model-form U_{MODEL} uncertainties are appended to either sides of the pbox (according to the rules of interval analysis), thus providing an increase in the epistemic uncertainty associated to the SRQ.

The increase in predictive uncertainty originating from this process is exemplified in fig. 4.6.

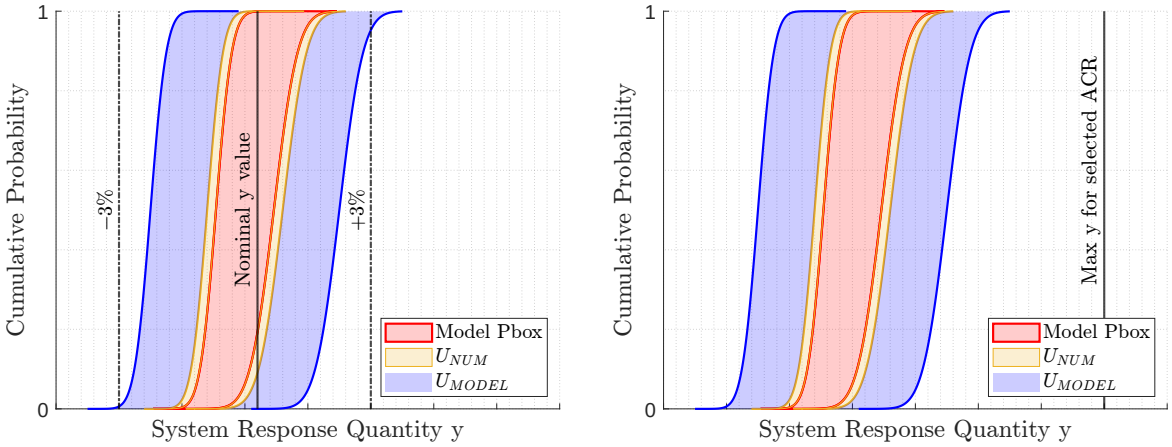


(a) Example of computation of the area validation metric. The area metric is represented by the region of non-overlap between the experimental CDF and the probability box associated to model prediction

(b) The account for the model-form epistemic uncertainty U_{MODEL} enlarges the width of the probability box associated to the model prediction

Figure 4.6: Increase in epistemic uncertainty associated to model prediction due the addition of numerical and model-form uncertainty

Then, the model adequacy assessment can be carried out directly comparing the uncertain model prediction with the adequacy margins and/or the ACR, as displayed in fig. 4.7. Similar examples of adequacy assessment within this framework can be consulted at [30].



(a) Example of model adequacy assessment, with adequacy margin delivered in terms of distance from the nominal simulation prediction

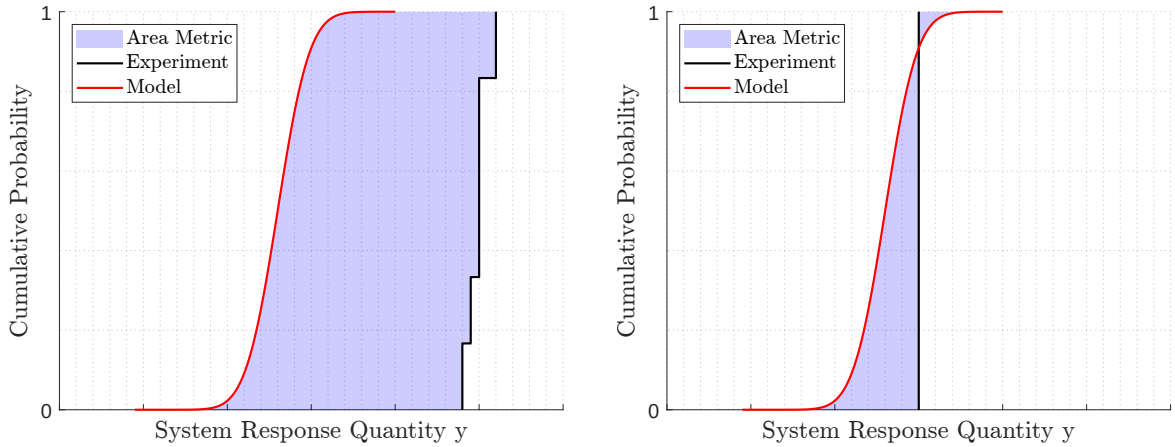
(b) Example of model adequacy assessment, with adequacy margin delivered as function of the selected ACR

Figure 4.7: Examples of model adequacy assessments within the Roy-Oberkampf framework

4.2.4. Observations

The epistemic uncertainty conceived in the Roy-Oberkampf approach effectively represent an extension of the concept of type B evaluation of uncertainties in [16] and ASME VV 20 standard. Indeed, they represent reducible uncertainties and their characterization reflects the belief of the analyst. However, their mathematical structure is limited to an interval with no assumption made by the analyst on the PDF (and thus relying only on the definition of the bounds). From a mathematical standpoint, this feature results in a different UQ methodology and representation of propagated uncertainties.

Concerning the area validation metric d, it represents the second fundamental difference with the approach proposed by ASME and it exhibits a number of interesting properties. First of all, it is defined relying on an integral definition, which makes it very robust with respect to extreme value of the SRQ at the tails of the discrete CDFs. Secondly, it can be evaluated regardless the structure of the simulation and experimental probability boxes. Then, unlike other metrics for CDFs comparison, such as the Kolmogorov-Smirnov test, it can be applied even to probability structures which does not have any value of the SRQ in common, as in fig. 4.8a, or identified by vertical spikes, as in fig. 4.8b. At last, the fact it retains the dimensional physical unit of the SRQ allows the analyst to have a more intuitive appeal with it when it comes to its application during model interpolation and/or extrapolation.



(a) Example of application of the area validation metric on non-overlayed simulation and experimental CDFs

(b) Example of application of the area validation metric with vertical spike (single realization) as experimental CDF

Figure 4.8: Example of application of the area metric to *extreme* conditions

Another advantage of the Roy-Oberkampf approach is that the probability box arising from model inputs propagation allows the analyst to quantitatively distinguish the impact

of aleatory and epistemic uncertainties on each model SRQ. Nevertheless, this edge over the ASME approach becomes less useful as soon as a large number of physically unrelated uncertain inputs is considered or given that sensitivity analysis are typically performed in tandem with uncertainty propagation.

Moreover, some observations regarding the interpretation of d, the definition of the model-form error and model extrapolation are worthy of mention.
From [29], an explicit and quantitative connection between the model-form error and the validation metric is solely addressed when $d \neq 0$. However, no clear guideline is proposed about what conclusions regarding the model-form error shall be drawn when $d = 0$. Nevertheless, it is straightaway clear that whenever $d = 0$, no model update can be conceived solely basing on the value of d.
Moreover, it is important to stress that no information about the size of the pbox and the uncertainty of the measurement is either explicitly or implicitly delivered with d (unlike u_{val} in ASME framework). As a consequence, the information about the uncertainties involved in the validation process does not make it to the model adequacy assessment. So, for example, if for any reason the FSM pbox is overestimated during model validation (e.g. due to overestimated reducible model-input uncertainties), the area metric undergoes a reduction, and, as a result, the analyst may inadvertently end up with an underestimation of the overall FSM prediction uncertainty in the DoP (due to the underestimation of the model-form contribution). In addition, it is not true either to state that the underestimation of U_{MODEL} that occurs on the aforementioned occasions is compensated, during model adequacy assessment, by the same overestimation of the pbox that caused it in the first place (hence, leading to a net null error on the evaluation of the total model prediction uncertainty). Sure enough, either due to FSM non-linearity or due to the fact that model-input uncertainties may be simply different during certification-aimed predictions[11] the size of the input-related pbox might even reduce in the DoP. This feature of the area metric is recalled and discussed more in detail in section 6.3.6, where d is compared with the other model validation metrics.
On the other hand, concerning the model extrapolation technique proposed by Roy and Oberkampf, the following evidence is noteworthy. If the area metric value $d \neq 0$ retrieved in a validation point (fig. 4.9a) is applied to the very same probability box it has been

[11] e.g. in the course of validation the applicant shall consider uncertainties associated to the simulation boundary conditions (density altitude, air speed, etc.) since the values are retrieved from the flight measurements upon which the validation is carried out. However, some of these inputs may turn deterministic in certification-aimed predictions, with value assigned directly from the ACR. As a consequence, it is possible that the pbox shrinks in the DoP, since the overall uncertainty on simulation inputs has reduced.

derived from[12] to account for the evaluated model-form uncertainty, the result reported in fig. 4.9b is obtained. Indeed, the different shapes of the simulation pbox and experimental CDFs lead to a residual disagreement (highlighted in green in fig. 4.9b) between the model and the experiment. However, apparently, this disagreement is not contemplated within this framework, given that all conceived uncertainties (inputs related, numerical, model-form and experimental) have been already taken into account.

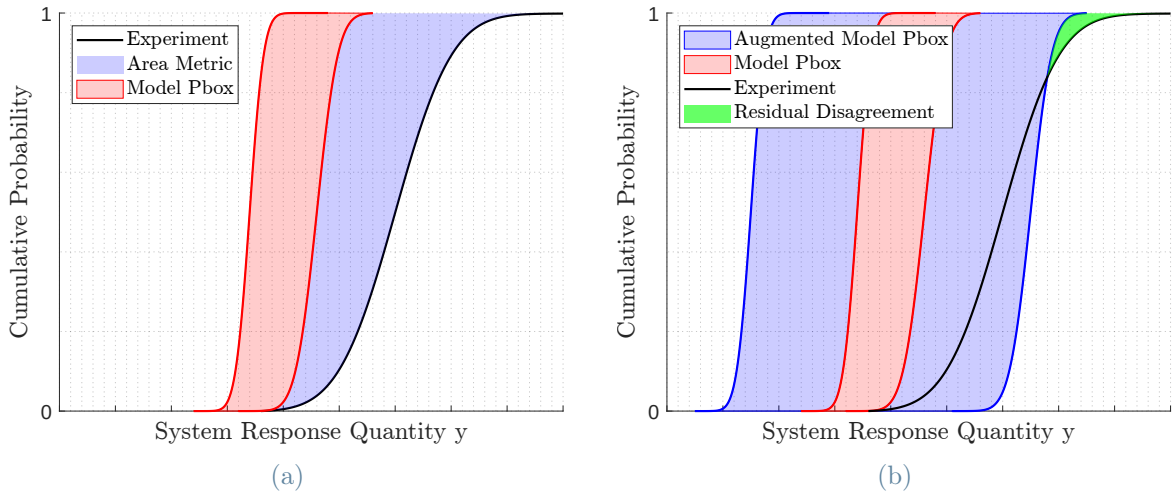


Figure 4.9: Example area validation metric computation and model-form uncertainty definition and application

Finally, as anticipated in section 4.1.4, considering that expensive sampling techniques are involved in this framework as well, it is be possible that simpler and cheaper alternatives may be conceived for model update purposes and preliminary quantification of the model prediction uncertainty.

[12]Thus, without the need of any correction to to interpolation/extrapolation of the model-form error.

5 | RoCS Guidelines Application - Phase 1

In this section, a simplified application example of the phase 1 of the CbS follows. Since the work presented in this dissertation was carried out in collaboration with LHD, the proposed application could not ignore, among others, the wide availability of flight test data and the state-of-the-art tiltrotor FSM provided by the company. As a consequence, less attention was devoted to the decision making process an actual applicant would follow[1]. Nevertheless, the pivotal concepts exposed in RoCS guidelines (namely, application domains, flight data for model validation, IPC levels and model requirements) are tackled.

5.1. Certification Requirements and Fidelity Metrics

Certification specifications for the hereby presented demonstration were extracted from [7]. In particular, CS.29.143 (d) for OGE low speed controllability and maneuverability, also reported below, was considered.

"Wind velocities from zero to at least 31 km/h (17 knots), from all azimuths, must be established in which the rotorcraft can be operated without loss of control out-of-ground effect, with:

1. *Weight selected by the applicant;*
2. *Critical center of gravity;*
3. *Rotor rpm selected by the applicant; and*
4. *Altitude, from standard sea-level conditions to the maximum take-off and landing altitude capability of the rotorcraft"*

The CS requires the applicant to interpret what *loss of control* means in order to be translated into ACRs. LHD takes into consideration several factors when it comes to

[1] Also considering, as specified in section 3.1, than the most critical decisions within this phase of the process depend on the applicant resources, modelling experience, etc. and the certification authority.

the interpretation of this specific CS. Static trim controls margins, minimum achievable rates and accelerations in single-axis maneuvers, aircraft trim attitudes and pilot's opinion and workload all contribute to the definition of the low speed controllability envelope for the tiltrotor configuration. Thus, to derive simple and directly applicable certification requirements for the aforementioned specifications alone is a great task indeed. So, for the sake of simplicity, the author decided to exclude any stability and aircraft dynamics related argument, hence rephrasing the controllability requirement as a sole function of the aircraft trim static control margins[2]. As a consequence, model validation is based on the comparison of fidelity metrics at trim and involves the time domain only.

Concerning the fidelity metrics, the preliminary set proposed by RoCS [32] was considered. This set of metrics is directly derived from CS-FSTD(H) [6], the certifications standard for training flight simulators, and was developed in collaboration with LHD. The set of fidelity metrics proposed in [32] for trim accuracy assessment is summarized in table 5.1, together with the associated margins proposed to assess model fidelity. In addition, wherever present, the acceptability margins provided in [6] are reported for comparison.

Parameter	**Acceptability Margins**	
	RoCS & LHD [32]	**CS-FSTD(H) [6][3]**
Rotor Mast Torque[4]	3%	3%
Pitch Angle	1.5°	1.5°
Roll Angle	1.5°	2°
Longitudinal (Total) Control Position	5%	5%
Lateral (Total) Control Position	5%	5%
Directional (Total) Control Position	5%	5%
Collective (Cockpit) Control Position	5%	5%
Rotor Longitudinal Flapping	1° or 10%	-
Rotor Lateral Flapping	1° or 10%	-
Angle of Attack	1°	-
Angle of Sideslip	1°	-

Table 5.1: Summary of time domain fidelity metrics (and associated acceptable margins) proposed by RoCS and CS-FSTD(H) for trim fidelity assessment

With respect to CS-FSTD(H), this set of metrics incorporates angles of attack and sideslip,

[2]Thus, relying only on a subset of the controllability criteria exploited by LHD.
[3]Margins selected from "Low Airspeed Handling Qualities - Trimmed Flight Control Positions"
[4]Acceptable margins in % of the rotor maximum continuous operating torque.

together with rotor flapping angles. The latter are considered of paramount importance in LHD, due to their impact in simulation-based rotor clearance assessments.

However, starting from the summary in table 5.1, some modifications were carried out. First of all, the angle of sideslip and angle of attack were excluded from the set, due to the unavailability of reliable airspeed data in low speed flight conditions. Secondly, the collective control position (xt_{col}) was replaced with the rotors blades collective pitch. Indeed, within the FSM, xt_{col} directly controls the rotors blades pitch. On the other hand, on the flying aircraft prototype, the collective control position is processed differently, and represents an input signal for the control system of the turboshaft engines. As a consequence, a direct comparison of the total collective control between FSM and flight data was not viable, and the actual blade collective pitch input at the swash-plate was used instead as fidelity metric. The summary of fidelity metrics exploited in this work for model validation is reported in table 5.2.

Parameter		**RoCS & LHD Margins [32]**
Rotor Mast Torque[5]	Q	3%
Pitch Angle	θ	1.5°
Roll Angle	ϕ	1.5°
Longitudinal (Total) Control Position	xt_{lon}	5%
Lateral (Total) Control Position	xt_{lat}	5%
Directional (Total) Control Position	xt_{ped}	5%
Rotor Blades Collective Pitch	$\theta_{3/4}$	-
Rotor Longitudinal Flapping	$A1$	1° or 10%
Rotor Lateral Flapping	$B1$	1° or 10%

Table 5.2: Summary of time domain fidelity metrics exploited in this work

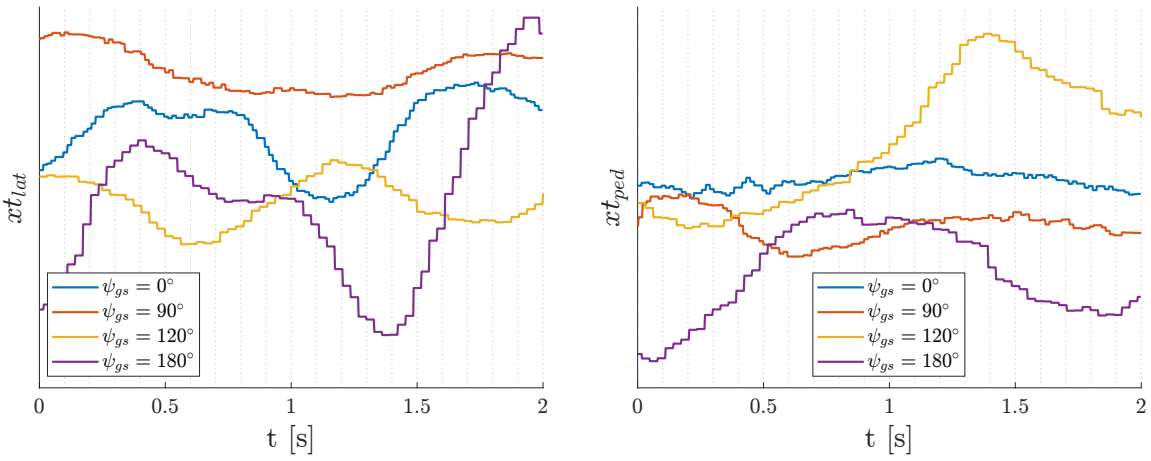
5.2. Flight Data Selection

The choice of influence and predictability levels within the CbS process, as mentioned in section 3.1, depends on several factors and it is carried out together with the flight test campaign planning. Thus, starting from the data-set of low speed flight data made available by LHD, the author selected a subset upon which to perform the validation. Then, the IPC levels were defined as a consequence.

The flight data selected for the guidelines application are represented by a set of mea-

[5]Acceptable margins in % of the rotor maximum continuous operating torque.

surements carried over four time windows. Each window has a duration of 2 seconds and corresponds to an attempt of horizontal, uniform, rectilinear flight at an assigned ground speed azimuth angle. The selection process took into account several factors, the most important, aircraft configuration (vertical tail size, control chain version, FCS release, etc.), the OGE conditions and the quality of the flight datum itself. In addition, the will to explore different flight conditions ultimately led the author to select a set of data spanning different ground speed azimuths. In fig. 5.1, the time histories of aircraft pitch attitude and longitudinal control position associated with the selected validation windows are displayed. As reported in the legend, different flights have been referred to using the corresponding nominal ground speed azimuth. For the sake of convenience, this way of referencing is exploited from here on to identify the different validation points. Flight data time histories of other quantities interest can be consulted in appendix A.



(a) Lateral control position

(b) Directional control position

Figure 5.1: Examples of flight data time histories for the selected validation flights

The measurand is defined as a steady value, corresponding to the system response quantity at trim. However, great variability can be spotted in the time histories of fig. 5.1 and appendix A[6]. Building on this evidence and considering the body axis acceleration data, it is straightaway clear that, mathematically speaking, the trim condition was not strictly achieved. Thus, among others, indeed some errors (reasonably of both random and systematic nature) associated to the "*imperfect realization of the measurand*", as it is referred to in [16], are present. The validation procedures presented in chapter 4 both account for experimental data uncertainty, but in different ways. Hence, in this work, the problem of experimental uncertainty quantification is carried out within model validation in section 6.3.

[6]The lateral control position in $\psi_{gs} = 0°$ and $\psi_{gs} = 180°$, the directional control position in $\psi_{gs} = 120°$ and $\psi_{gs} = 180°$ and the roll angle in $\psi_{gs} = 180°$ are the greatest examples.

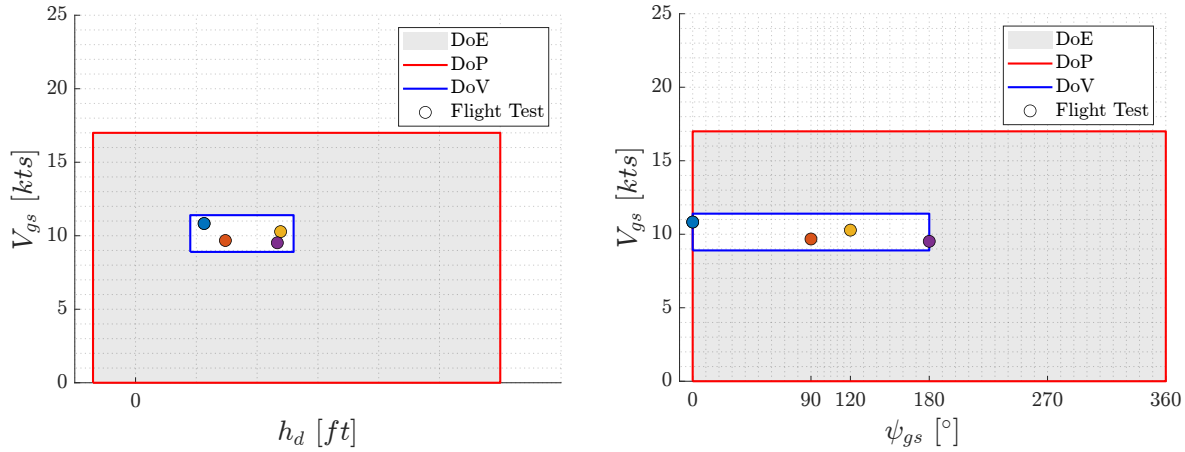
It is also interesting to observe that the flights selected for validation purposes in this work do not belong to a validation-oriented flight test program. On the other hand, they derive from testing campaigns that LHD carried out for vehicle development and certification compliance demonstration purposes. Nevertheless, the application of the validation procedures to this set of data is indeed of use to shed light on the challenges associated to uncertainty definition for these kind of time histories and to figure out whether this level of uncertainty may be adequate for model validation, or if the implementation of more strict validation-oriented testing procedures shall be considered for future, real case scenario applications of the guidelines. Indeed, building on the observations carried out in section 4.1.4, if an applicant was to find out that its standard procedures for flight testing result in experimental uncertainties which are unacceptable for the desired margin of model fidelity, greater effort shall be accounted for validation-oriented flight tests during the CbS planning. Moreover, it is also possible that the measurement error observed in these tests, including the contribution associated to the aforementioned "*imperfect realization of the measurand*", may not be significantly reducible, since largely due to piloting technique, limit cycle oscillations, etc.

5.3. Domains Definition and Influence, Predictability and Confidence Levels

Building on the validation flight tests data discussed in section 5.2, the application domains and the IPC levels have been derived. In fig. 5.2, RoCS application domains defined in terms of (nominal) density altitude and (nominal) ground speed magnitude and azimuth are reported.

In fig. 5.3, on the other hand, the (nominal) validation flights collocation in the weight-CG position envelope is displayed.

Considering the ACRs discussed in section 5.1, and relying upon these sole selected flights, the applicant may exploit the simulation model for a critical point analysis within the compliance demonstration, thus identifying the influence and predictability levels as I2-P3. Sure enough, considering fig. 5.2, it is possible to conclude that all validation data points are far from both the maximum speed of 17 knots and the maximum (and minimum) landing density altitude of the aircraft. In addition, critical center of gravity positions are not achieved either, as reported in fig. 5.3. As a consequence, according to RoCS proposal, upon acceptance by the certification authority, it may also be hypothesised that the minimum confidence level required on model predictions may be high (table 3.3).



(a) Prediction, validation and extrapolation domains as a function of ground speed magnitude and density altitude

(b) Prediction, validation and extrapolation domains as a function of ground speed magnitude and azimuth

Figure 5.2: Prediction, validation and extrapolation domains

5.4. Model Requirements

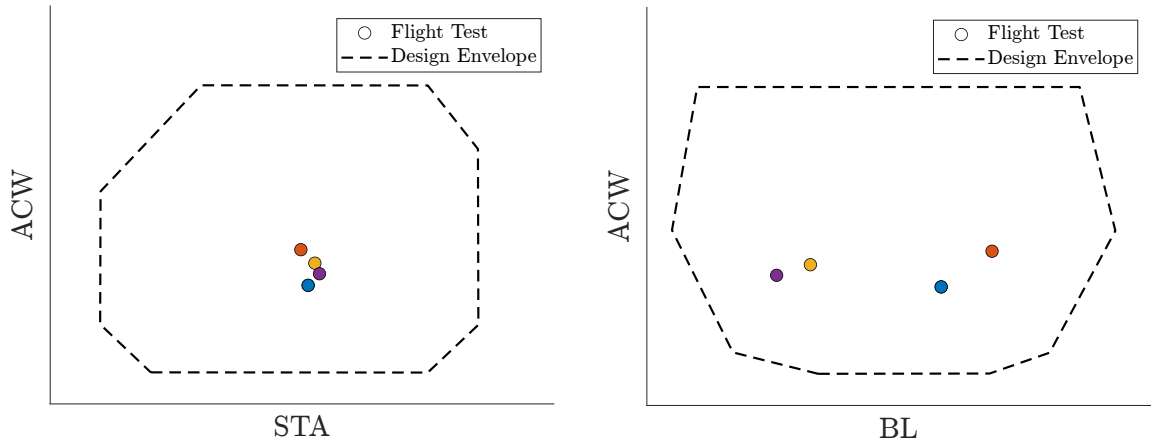
In the framework of this dissertation, the first phase of the CbS can be considered completed upon selection of the FSM requirements. Building on the elements exposed in section 5.1, 5.2 and 5.3, the following requisites for the FSM can be established.

Certainly, the model shall be physics-based, especially considering the amount of extrapolation (and interpolation) that is necessary for compliance demonstration. Moreover, all FSM components shall always operate within the DoR required for this CS.

Secondly, considering the fact that no pilot-in-the-loop simulation is forecasted (due to the simplification hypothesis discussed in section 5.1), no explicit model run time constraints are to be placed. Nevertheless, keeping in mind that Sensitivity Analyses (SA) and Uncertainty Quantification (UQ) are to be carried out in the VV phase, reducing, whenever possible, the FSM trim evaluation time is in the best interest of the users of this CbS process.

In addition, since the total control positions have been selected as fidelity metrics in section 5.1, a faithful reproduction of the aircraft flight control system (at least for the trim related part) shall be included within the FSM. Otherwise, no direct comparison between the total control position can be carried out.

Moreover, since all selected validation points and ACRs do not involve ground effect, its modelling may not be included in the FSM. Nevertheless, considering the low speed environment of the CS, interference aerodynamics are expected to play a crucial role in



(a) Position of the validation flight points in the CG fuselage station and aircraft weight design envelope

(b) Position of the validation flight points in the CG butt line station and aircraft weight design envelope

Figure 5.3: Validation points position within the weight-balance envelope of the rotorcraft

the aircraft behavior; hence, its physics-based modelling shall be a pivotal component of the FSM. It may be reasonable to conclude that at least the interactions between the rotor wake, wing, tail and nacelles are necessary to achieve sufficiently high confidence levels in the model predictions.

6 | RoCS Guidelines Application - Phase 2a

In the following sections, the phase 2a of RoCS guidelines is tackled. At first, the FSM supplied by LHD is introduced and assessed in terms of compliance to the requirements assembled in section 5.4. Then, the solution verification is carried out, devoting particular attention to the error estimation algorithm and the model parameters involved. Eventually, the solution validation is performed following different approaches and the obtained results are presented and compared in terms of validation metrics values, computational expense and suitability to the CbS process.

Both in this chapter and the following, the author refers to the set of parameters summarized in table 6.1 as SRQs of the FSM.

Parameter	**Symbol**
Rotor Mast Torque	Q
Pitch Angle	θ
Roll Angle	ϕ
Longitudinal (Total) Control Position	xt_{lon}
Lateral (Total) Control Position	xt_{lat}
Directional (Total) Control Position	xt_{ped}
Rotor Blades Collective Pitch	$\theta_{3/4}$
Rotor Longitudinal Flapping	$A1$
Rotor Lateral Flapping	$B1$

Table 6.1: Summary of the SRQs of the FSM

Rotor flapping angles are defined according to the following equation.

$$\beta = \beta_0 + A1 \cos\psi + B1 \sin\psi \tag{6.1}$$

where β_0 is the rotor pre-cone and ψ is the azimuthal position of the rotor blade in the fixed rotor frame.

6.1. Flight Simulation Model

The phase 2a of the CbS process begins with the FSM development. However, since, as already mentioned, LHD provided the baseline FSM exploited in this work, this section is mainly devoted to the description of the model and to the acknowledgment of its observance of the phase 1 requirements.

The FSM is developed in FLIGHTLAB. Within the model, each aircraft functional part is implemented and integrated with other systems in a component-based approach, as exemplified in fig. 6.1. The model is conceived as a time-marching simulator, capable of the time integration of the aircraft equations of motion, reported in eq. (6.2), together with the equations of motion of the rotors subsystem. All aerodynamic components are modelled as rigid, as typical in the flight mechanics framework, and a list of their main features, together with the associated modelling hypotheses, is reported in table 6.2.

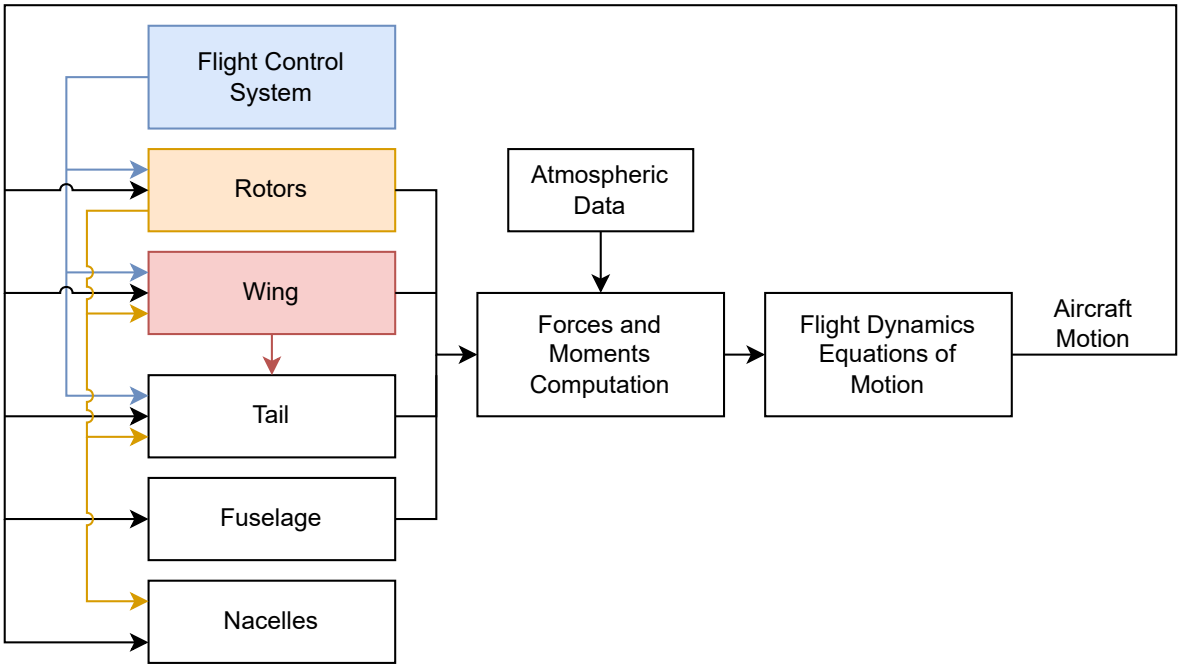


Figure 6.1: FSM components breakdown

$$
\begin{aligned}
m(\dot{v}_x + qv_z - rv_y) + mg\sin\theta &= X_{rot} + X_{wing} + X_{tail} + X_{nac} + X_{fus} \\
m(\dot{v}_y + rv_x - pv_z) + mg\cos\theta\sin\phi &= Y_{rot} + Y_{wing} + Y_{tail} + Y_{nac} + Y_{fus} \\
m(\dot{v}_z + pv_y - qv_x) + mg\cos\theta\cos\phi &= Z_{rot} + Z_{wing} + Z_{tail} + Z_{nac} + Z_{fus} \\
I_{xx}\dot{p} - I_{zx}\dot{r} + q(I_{zz}r - I_{zx}p) - qrI_{yy} &= L_{rot} + L_{wing} + L_{tail} + L_{nac} + L_{fus} \\
I_{yy}\dot{q} + r(I_{xx}p - I_{xz}r) - p(I_{zz}r - I_{xz}p) &= M_{rot} + M_{wing} + M_{tail} + M_{nac} + M_{fus} \\
I_{zz}\dot{r} - I_{xz}\dot{p} + q(-I_{xx}p + I_{xz}r) + pqI_{yy} &= N_{rot} + N_{wing} + N_{tail} + N_{nac} + N_{fus} \\
\dot{\phi} &= p + (q\sin\phi + r\cos\phi)\tan\theta \\
\dot{\theta} &= q\cos\phi - r\sin\phi \\
\dot{\psi} &= \frac{q\sin\phi + r\cos\phi}{\cos\theta}
\end{aligned}
\tag{6.2}
$$

The rotors are modelled with blade element theory. Each blade is subdivided in a finite number of sections with prescribed aerodynamic properties. Fixed lifting surfaces, namely the wing and the tail, are modelled as lifting lines. Also in this instance, the surfaces are subdivided in sections with prescribed aerodynamic properties. Finally, fuselage and nacelles aerodynamic loads are managed via specific look-up tables and making use of single control points for airspeed acquisition.

Component	**Feature**	**Model**
Rotor blade	Airloads	Unsteady BEM with non linear look-up tables
	Induced velocity	Peters-He finite state
	Active aerodynamic interference	Peters-He finite state
Wing & tail	Airloads	Lifting line
	Induced velocity	Peters-He finite state
	Active aerodynamic interference	Peters-He finite state
Fuselage	Airloads	Steady non linear look-up tables
Nacelle	Airloads	Steady non linear look-up tables

Table 6.2: FSM aerodynamic components features and models

The induced velocity field of the lifting components (namely, blades, wing and tail) is modelled with the Peters-He finite state model ([19], [23]), a physics-based time-domain dynamic wake model originally conceived as a tool for the aeroelasticity and aeromechanics of lifting rotors. Nevertheless, within FLIGHTLAB, the possibility to apply this model to fixed lifting surfaces is contemplated. Within the Peters-He model, the induced velocity at the rotor disk (or at a fixed lifting surface) is expressed as a combination of shape functions which account for both azimuthal and spanwise variations, as reported in the following equation.

$$v_i(\overline{r},\psi,t) = \sum_{k=0}^{\infty} \sum_{j=k+1,k+3,\ldots}^{\infty} \Psi_j^k(\overline{r})[a_j^k(t)\cos(k\psi) + b_j^k \sin(k\psi)] \tag{6.3}$$

where $\overline{r}$ is the non dimensional spanwise coordinate, ψ the azimuthal coordinate, Ψ_j^k are radial shape functions and a_j^k and b_j^k time dependent states of the induced flow model. A system of first order differential equation, directly derived from the incompressible potential flow equations, is solved to compute the time varying coefficients of the shape functions. The closed-form analytical expression of this model, together with its finite state state-space formulation as a system of non-linear differential equations, makes it suitable for wide variety of applications (including, e.g., time-marching simulations and rotor stability calculations).
In addition, as described in [20], this inflow model can be expanded in order to account for aerodynamic interference. Considering the relative position of the lifting device (either the rotor disk or a fixed wing) and the point of interest for interference, the integral relations summarized in eq. (6.4) can be derived[1].

$$\begin{cases} v_{i\ x_w}(x_w,y_w,z_w) = \dfrac{1}{V_\infty}\int_\xi^\infty \Phi_\xi^V\, d\xi \\ v_{i\ y_w}(x_w,y_w,z_w) = -\dfrac{1}{V_\infty}\int_\xi^\infty \Phi_{y_w}^V\, d\xi \\ v_{i\ z_w}(x_w,y_w,z_w) = -\dfrac{1}{V_\infty}\int_\xi^\infty \Phi_{z_w}^V\, d\xi \end{cases} \tag{6.4}$$

where x_w, y_w, z_w represent the position of the interference point with respect to the lifting device in the latter's wind axes frame, V_∞ represents the non dimensional freestream velocity, Φ the non dimensional pressure distribution (see [19]), ξ the non dimensional

[1] It is important to stress that this interference model establish a one-way connection between the *inducing* lifting body and the interference point of interest. In the latter, the experienced dynamic pressure depends on both an induced contribution and a freestream contribution, but no feed-back of the presence of a body in the interference point of interest is experienced by the inducing lifting device.

coordinate along the freestream line and v_i the induced velocity components at the interference point of interest. To compute these integrals, the pressure distribution Φ shall be known in advance.
Being aerodynamic interference modelling a fundamental FSM requisite, as pointed out in section 5.4, this wake interference formulation is integrated in the model to compute the active interference of lifting devices[2] on other components. As highlighted in fig. 6.2, active interference of the rotors on the nacelles, the wing and the tail has been considered. In addition, the active interference of the main wing on the tail was modelled, as well as the mutual interference between the vertical and horizontal tail surfaces. No nacelle or wing blockage effects on the induced velocity at the rotor disk were accounted for, as well as any active interference on the fuselage body. Similarly, no active interference of the tail on either the wing, the rotors or the nacelles is modelled[3]. No interference between the two rotors is considered either and no ground interaction with any component of the model is included, given that IGE flight conditions are beyond the scope of the CS exposed in section 5.1.

Concerning the aerodynamic loads look-up tables (namely for the rotor, wing and tail sections, for the fuselage and for the nacelles), the attitude and speed envelopes ensure that only interpolation is carried out during their interrogation in the DoP. Similar conclusions can be carried out for the other model components, designed to be suitable for flight conditions in the whole aircraft low-speed envelope.

For the FCS, a replica is included in the FSM within the FLIGHTLAB environment.

6.2. Solution Verification

Being FLIGHTLAB a commercial software package, code verification was not been explicitly undertaken by the author. Conversely, the evidence provided in [12] was considered sufficient for the scope of this work. Hence, the work presented in this section focuses on the assessment of the FSM numerical accuracy.

As acknowledged in [33], the widely recognized solution verification practices outlined in existing standards (e.g. [8]) are aimed at CFD and structural analysis problems and

[2]Indeed, in order to account for the active interference of bluff bodies with a physics based formulation, a fundamentally different approach would be required.

[3]Indeed, it may be argued that in backward flight, the tail wake may result in some influence on the wing and rotor behaviour. Nevertheless, being the tail an aerodynamic body designed for airplane mode flight, it is also reasonable to believe that such influence in low-speed backward flight conditions would mainly be viscous-effects related, and, hence, not manageable with the Peters-He induced velocity model with a physics based approach.

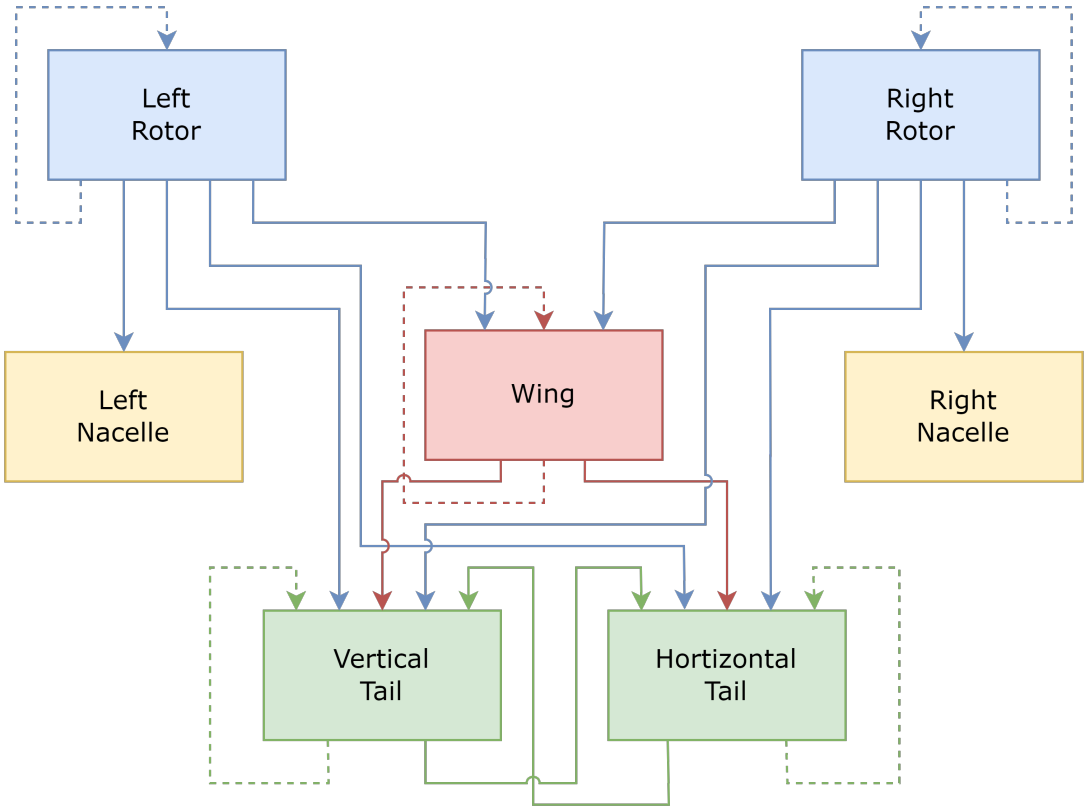


Figure 6.2: FSM aerodynamic interference outline. Active interference effects of an aerodynamic component on other parts are marked by a continuous line. On the other hand, dashed lines mark the account for self-induced velocity on a component.

cannot be directly applied to a multi-physics flight simulation models. Thus, the author came up with its own operative procedure for the solution verification of the FSM. Nevertheless, the hereby proposed approach was based on several concepts pointed out in [8] and [28], starting from the numerical error contributions highlighted in eq. (6.5).

$$\delta_{NUM} = \delta_{DE} + \delta_{IT} + \delta_{RO} \tag{6.5}$$

In the following subsections, the set of FSM model parameters considered for the assessment of δ_{DE} and δ_{IT} affecting each fidelity metrics is presented and described. The round-off error δ_{RO}, on the other hand, is neglected, since reasonably expected to be substantially smaller than the two aforementioned error contributions. Then, an outline of the numerical accuracy estimation process is provided and the results of the solution verification phase are discussed.

6.2.1. Parameters of Influence

To select the parameters of interest[4] for the solution verification, it is necessary to deepen the FSM spatial and temporal discretization, the numerical time integration scheme and trim equations solution algorithm. Indeed, only with a complete overview of the latter topics it is possible to understand what FSM parameters affect δ_{IT} and δ_{DE}.

Iterative Convergence Error Parameters of Influence

Contributions to the iterative convergence error δ_{IT} are indeed present whenever a system of equations is solved numerically with an iterative method [29]. Within the framework under consideration, such occurrence happens several times. Being the flight simulation model a time-marching, mappings between model inputs and outputs can be computed only via discrete time integration of eq. (6.2). Hence, the first contribution to δ_{IT} is associated to the implicit Newmark-Beta time integration scheme, which requires a linear system of equations to be solved at every time step. The second instance is related to the non linear system of equations that enforces the trim of the aircraft, summarized in the following equation.

$$\underline{F}(\underline{x}) = \underline{0} \tag{6.6}$$

where

$$\underline{x} = [xt_{col} \ , \ xt_{lon} \ , \ xt_{lat} \ , \ xt_{ped} \ , \ \theta \ , \ \phi]^T \tag{6.7}$$

$$\underline{F}(\underline{x}) = [\dot{v}_x \ , \ \dot{v}_y \ , \ \dot{v}_z \ , \ \dot{p} \ , \ \dot{q} \ , \ \dot{r}]^T \tag{6.8}$$

Finally, eq. (6.6) is solved exploiting a variation of the Newton-Raphson method, which, in turn, requires to solve the linear system of equations (6.9) every time the trim variables $\underline{x}$ are updated from iteration k to $k+1$.

$$\underline{\underline{J_F}}(\underline{x}_{jac}) \ (\underline{x}_{k+1} - \underline{x}_k) = -\underline{F}(\underline{x}_k) \tag{6.9}$$

where $\underline{\underline{J_F}}(\underline{x}_{jac})$ represents the numerical approximation of the Jacobian matrix of $\underline{F}(\underline{x})$ evaluated at $\underline{x}_{jac}$.

[4] In principle, it is indeed true to state that, for each fidelity metric, both δ_{IT} and δ_{DE} depend on any FSM parameter and input datum. Nevertheless, the author here uses the expressions "*parameters of influence*" or "*parameters of interest*" to refer to those parameters that are explicitly and formally bond to the numerical algorithms and discretization procedures which cause the occurrence of δ_{IT} and δ_{DE} in the first place.

Rigorously speaking, the exit criteria for the numerical solution algorithms of all the aforementioned equations systems affect δ_{IT}. However, since eq. (6.9) solver can be effectively considered to be only a provider for an update candidate solution to eq. (6.6), it may be argued that the exit criteria related to the former does not affect the bounds of δ_{IT}.
In addition, the iterative convergence error is affected by another factor: the criterium for the identification of a *steady state* condition during time integration. As aforementioned, the FSM is a time-marching system, meaning that time integration of equations of motion is needed to map model inputs (e.g. control positions) into model outputs (e.g. airframe accelerations). Thus, within the Newton-Raphson algorithm, time integration up to a steady state condition is required to evaluate the mapping $\underline{F}(\underline{x})$ and its Jacobian matrix $\underline{\underline{J_F}}$ at each iteration. As a consequence, the numerical criterion for the definition of the steady state condition itself affects the residual of eq. (6.6) during the solution process and, therefore, impacts δ_{IT}.

By default, FLIGHTLAB allows the user to set exit criteria for both the linear and the non linear equations solvers. Concerning the former, the maximum number of iterations and a scalar, relative convergence tolerance value can be specified. As soon as one of the two is reached, the iterative method is arrested. In regards of the latter, independent tolerances for the residual linear and angular acceleration components in the body frame at steady state can be set. Thus, once all user defined acceleration tolerances are met, the solution is considered to be converged.
Considering the identification of the steady state condition, FLIGHTLAB relies once more on the values of the body frame acceleration components. In particular, their average values on the last N_{avg} rotor revolutions are compared with the averages on the previous N_{avg} revolutions. Once the differences between the averages fall within the user specified tolerances for all linear and angular acceleration components, the steady state is considered to be achieved. However, no check concerning the variation of the accelerations within the last N_{avg} revolutions is carried out. By default, in FLIGHTLAB, the linear and angular acceleration tolerances for the steady state definition are bound to a half of the trim equation tolerances. Hence, whenever the user asks for tighter trim equation tolerances, the criterium for the steady state identification is tightened accordingly. In the present work, the author made use of this FLIGHTLAB default option, which ultimately led to a reduction of the independent model parameters which affect δ_{IT}. Moreover, N_{avg} was maintained fixed and equal to 1.

A summary of all the independent FSM parameters which affect δ_{IT} and are involved in the solution verification is presented in table 6.3.

Solver	**Parameter**	
Linear system iterative solver	Maximum iterations	N_{lin}
	Convergence tolerance	ϵ_{lin}
Trim equation solver	Longitudinal acceleration tolerance	$\dot{v}_{x\ tol}$
	Lateral acceleration tolerance	$\dot{v}_{y\ tol}$
	Vertical acceleration tolerance	$\dot{v}_{z\ tol}$
	Roll rate derivative tolerance	$\dot{p}_{tol}$
	Pitch rate derivative tolerance	$\dot{q}_{tol}$
	Yaw rate derivative tolerance	$\dot{r}_{tol}$

Table 6.3: List of the FSM parameters affecting the model prediction iterative convergence error δ_{IT} and involved in the solution verification process

Discretization Error Parameters of Influence

Concerning δ_{DE}, on the other hand, contributions arise whenever a temporal or spatial discretization is carried out. Considering the FSM outline presented in section 6.1, this occurs in several instances. First of all, a time discretization is performed to allow for numerical time integration with the Newmark-Beta scheme. The integration step $\mathrm{d}t$ is defined according to the following equation as a function of n_{az}.

$$\mathrm{d}t = \frac{2\pi}{\Omega\ n_{az}} \tag{6.10}$$

Secondly, all aerodynamic surfaces (including the rotor blades, the wing and the horizontal and vertical tail) are divided into n_{seg} sections. As a consequence, both the sampling of the dynamic pressure spanwise variation and the resulting integral aerodynamic loads are affected by the choice of this discretization, as exemplified in eq. (6.11).

$$\underline{\mathrm{F}} = \sum_{j=1}^{n_{seg}} \underline{f}_j\ \mathrm{d}l_j \tag{6.11}$$

where $\underline{f}_j$ and $\mathrm{d}l_j$ are the load per unit length and the width of the j-th aerodynamic segment of a lifting surface, respectively. In addition, as anticipated in section 6.1, the aerodynamic interference model [20] included within the FSM requires to compute the pressure distribution integrals in eq. (6.4) in order to quantify the induced velocity at the interference point of interest. However, some of these integrals do not come with a closed-form analytical expression. Thus, the composite Simpson integration rule with n_{simps} intervals is used for their approximation. Eventually, the last contribution to δ_{DE} comes from the finite state representation of the lifting surfaces inflow, since the approximation

resulting from the truncation of states of a dynamic system may be interpreted as a discretization error. Within FLIGHTLAB, the number of induced velocity states can be set by prescribing the inflow maximum power of radial variation n_{rad} and number of harmonics n_{har} (corresponding to the maximum value of indexes k and j of eq. (6.3), respectively). Then, the total number of inflow states is obtained according to the table reported in fig. 6.3.

n_{rad}	n_{har}												Total
	0	1	2	3	4	5	6	7	8	9	10	11	12	Inflow States
0	1													1
1	1	1												3
2	2	1	1											6
3	2	2	1	1										10
4	3	2	2	1	1									15
5	3	3	2	2	1	1								21
6	4	3	3	2	2	1	1							28
7	4	4	3	3	2	2	1	1						36
8	5	4	4	3	3	2	2	1	1					45
9	5	5	4	4	3	3	2	2	1	1				55
10	6	5	5	4	4	3	3	2	2	1	1			66
11	6	6	5	5	4	4	3	3	2	2	1	1		78
12	7	6	6	5	5	4	4	3	3	2	2	1	1	91

Figure 6.3: Peters-He model total number of inflow states as a function of the inflow highest power of radial variation n_{rad} and number of harmonics n_{har} (from [19])

Nevertheless, if all the discretization parameters mentioned above were to be considered, the solution verification process would become excessively costly and intricate. Hence, with the aid of some physical considerations, some parameters were left out. In low speed flight conditions, fixed aerodynamic surfaces are expected to work at low efficiencies and their wake is expected to be dominated by flow separation (e.g. wing download and vertical tail sideforce in lateral flight). Thus, the associated active interference effects, as well as their dependence on the corresponding discretization parameters, would not be dominant. Building on this idea, all induced velocity and interference modelling parameters associated to the fixed aerodynamic surfaces were excluded.
In addition, another simplification was made concerning the rotor inflow finite state representation. Building on the convergence study of the Peters-He model reported in [19], two important observations can be made.

- For a given number of harmonic, significant variations in the inflow distribution at

the rotor disk can be observed increasing the number of radial shape functions.

- For a given number of radial shape functions, provided that the number of harmonics is equal to or greater than the number of rotor blades, negligible dependence of the inflow distribution on the number of harmonics is observed.

So, relying on these evidences, the author decided to fix the number of harmonics n_{har} of the rotors inflow distributions equal to 3. As a consequence, the only discretization error contribution arising from the truncation of radial shape functions was retained in the analysis.

All the independent FSM discretization parameters considered in the solution verification are summarized in table 6.4.

Discretization type	**Parameter**	
Temporal	Time step for time integration	n_{az}
Spatial	Rotor blade aerodynamic sections	$n_{seg\ r}$
	Integral discretization for rotor(s) wake active interference	n_{simps}
	Wing aerodynamic sections	$n_{seg\ w}$
	Vertical tail aerodynamic sections	$n_{seg\ vt}$
	Horizontal tail aerodynamic sections	$n_{seg\ ht}$
Inflow states truncation	Rotor inflow highest power of radial variation	n_{rad}

Table 6.4: List of the FSM parameters affecting the model prediction discretization error δ_{DE} and involved in the solution verification process

6.2.2. Numerical Error Estimation Procedure

Considering the large amount of parameters summarized in table 6.3 and 6.4, a practical procedure to estimate δ_{IT} and δ_{DE} could not rely on the evaluation of the entire grid of possible combinations. Thus, some considerations were made to simplify the process.

First of all, since the linear system solver may stop upon either reaching the maximum number of iterations N_{lin} or the convergence tolerance ϵ_{lin}, in order to effectively reduce its contribution to δ_{IT}, N_{lin} shall be increased every time ϵ_{lin} is tightened.
A similar way of reasoning can be applied to the trim equations solver tolerances. Despite having different values, the user shall always tighten them all together in order to make sure that the solution algorithm interruption is delayed.
Finally, the author decided to bond together the linear solver tolerances and the trim

equations solver tolerances as well. As a consequence, in the framework hereby presented, δ_{IT} convergence studies can effectively be carried out in one dimension and as a function of a newly defined single equivalent parameter ϵ_{IT}. As soon as ϵ_{IT} changes, all parameters in table 6.3 update accordingly.
Concerning the discretization error parameters of influence, only $n_{seg\ ht}$ and $n_{seg\ vt}$ were bonded, resulting in a single parameter for the identification of the tail aerodynamic discretization. The final list of independent parameters explicitly involved in the error estimation algorithm is summarized in table 6.5.

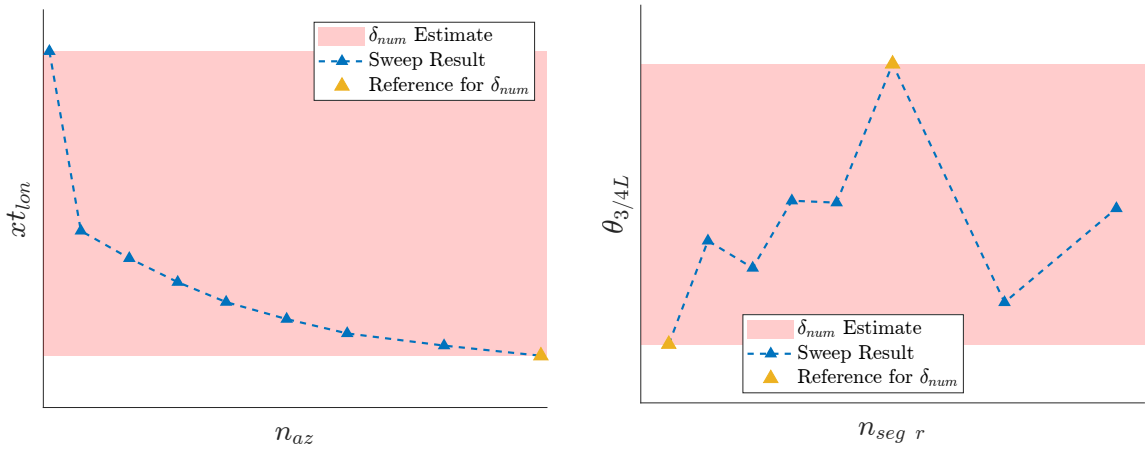
Numerical error	**Parameter**	
δ_{IT}	Equivalent iterative convergence parameter	ϵ_{IT}
δ_{DE}	Time step for time discretization	n_{az}
	Rotor blade aerodynamic sections	$n_{seg\ r}$
	Integral discretization for rotor(s) wake active interference	n_{simps}
	Wing aerodynamic sections	$n_{seg\ w}$
	Tail aerodynamic discretization	$n_{seg\ t}$
	Rotor inflow highest power of radial variation	n_{rad}

Table 6.5: List of independent FSM parameters of influence of δ_{IT} and δ_{DE} involved in the error estimate algorithm

The process used to accomplish the solution verification is reported below. Within its description, the author makes use of the expression "*convergence study*" to refer to a sweep of model predictions with respect to a numerical parameter, and to the associated error estimate. Within a sweep, which is carried out for each validation flight, the values of all fidelity metrics are collected. The sweep is interrupted upon the achievement of a *reasonable* approximation of an asymptotic value for all fidelity metrics or upon collection of sufficient evidence suggesting no asymptotic behavior will be observed[5]. If asymptotic convergence is observed, the error on each fidelity metric is defined as the distance from the representative asymptotic value (fig. 6.4a). On the other hand, if no convergence occurs, the amplitude of the oscillations in the fidelity metric is taken as the error estimate (fig. 6.4b).

However, there is a sole exception to these rules, represented by the estimation of the numerical error due to ϵ_{IT}. In [28], the suggested procedure to assess the iterative convergence error is to perform a sweep of model predictions with progressively tightened exit criteria. However, in the present application, the direct implementation of this technique

[5]At least with the values of table 6.5 model parameters that can be achieved with acceptable computational time.



(a) δ_{num} estimation on xt_{lon} sweep with respect to n_{az}

(b) δ_{num} estimation on $\theta_{3/4L}$ sweep with respect to $n_{seg\ r}$

Figure 6.4: Example of δ_{num} estimation from sweep result (validation flight $\psi_{gs} = 0°$

may result in an underestimation of δ_{IT}.
Indeed, whenever a non linear, algebraic system of equations, is solved with an iterative method, the numerical solution is identified by a set of points (in the neighbourhood of the true solution) in the space of the equations unknowns. The points within the set all result in acceptable residuals of the system of equations. The size of this set in the unknowns space is then directly associated to the tolerance set on the residuals: the tighter the tolerance, the smaller the distance between the true solution and the set boundary. Hence, in the attempt of providing an accurate estimate of δ_{IT}, an assessment of the size of the aforementioned numerical solutions set shall be made. However, such task cannot be achieved by mean of a single model prediction for each tolerance value. On the other hand, a sufficiently large amount of predictions scattered within the set might be considered a good proxy of its size, thus leading to a more accurate estimate of δ_{IT}.
A practical exemplification of this view is displayed in fig. 6.5. In the right-hand side plot of the figure, 50 realizations of $\theta_{3/4L}$ on $\psi_{gs} = 0°$ obtained with the same value of ϵ_{IT} and different initialization of the solution algorithm are reported. Considering the maximum and minimum sampled values, it is possible to identify a *convergence band* for the SRQ. Such *convergence band* should represent the image of the numerical solutions set through the FSM. Basing on its amplitude, δ_{IT} is then estimated, as in the left-hand side plot of fig. 6.5.

However, it is reasonable to say that this procedure leads to a more robust estimate of the iterative convergence error, with respect to considering only a single, arbitrary realization of the SRQ for a given value of ϵ_{IT}.

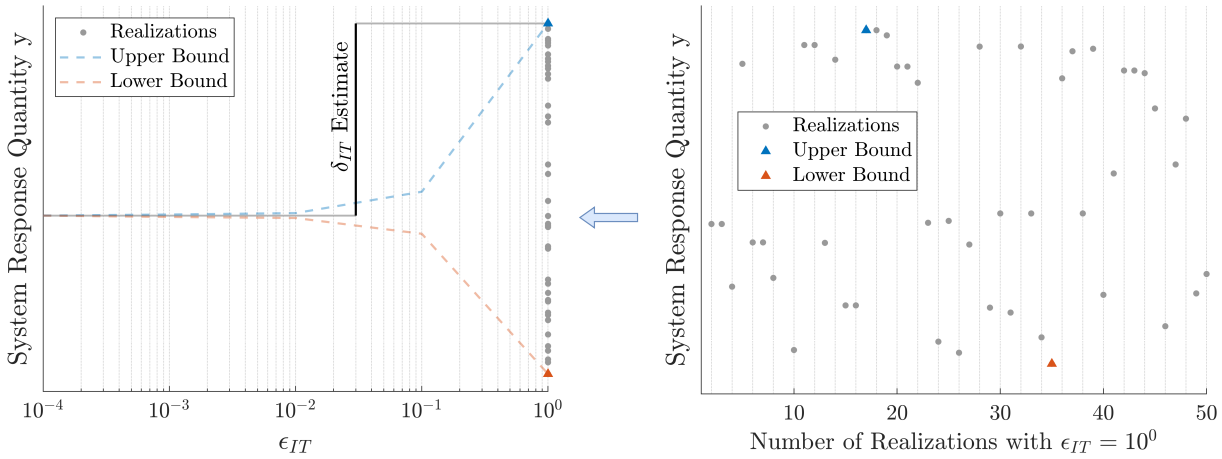


Figure 6.5: Example of the estimation of δ_{IT} for a system response quantity y. Several predictions of y are performed for each value of ϵ_{IT}, and are then used to construct an *iterative convergence band*

The error estimation procedure was built using the following arguments. First of all, building on some preliminary investigations carried out on the baseline FSM, it was anticipated that n_{rad} generates the most dominant contribution to δ_{num}. Thus, it is natural to set the definition of this parameter in the outermost loop of the procedure.
Secondly, as stated in [28], it is good practice to ensure that whenever a discretization error assessment is performed, the iterative convergence error is at least two order of magnitude less[6]. Thus, the author decided to perform the convergence study on ϵ_{IT} immediately after the choice of n_{rad}. In this way, building on preliminary investigations of the expected order of magnitude of δ_{DE} for each fidelity metric, it is possible to choose from the start a suitable value of ϵ_{IT}.
Concerning the assessment of the time and spatial discretization error, it is indeed true that, either through active aerodynamic interference or through trim equations, interactions are present among any pair of aerodynamic components. Thus, arbitrarily, the author decided to follow the stream of active aerodynamic interference for the discretization error assessment. Checks were performed to ensure that the estimated order of magnitude of a previously assessed δ_{DE} contribution was valid also when changing other discretization parameters values.

The error estimation procedure follows.

1. Choose n_{rad}
2. Perform a convergence study on δ_{IT} varying ϵ_{IT}

[6]Otherwise, the interference of the iterative convergence error on the model predictions spoils the observation of the discretization error.

3. Choose a value of ϵ_{IT} suitable for the δ_{DE} assessment
4. Perform a convergence study with respect to n_{az}
5. For a suitable value of n_{az}, perform a convergence study with respect to $n_{seg\ r}$
6. For a suitable value of n_{az} and $n_{seg\ r}$ perform a convergence studies with respect to $n_{seg\ w}$, n_{simps}[7] and, eventually, $n_{seg\ t}$
7. Confirm the δ_{DE} estimate due to $n_{seg\ r}$ provided in step 5 with updated values of $n_{seg\ w}$, $n_{seg\ t}$ and n_{simps}
8. Confirm the δ_{DE} estimate due to n_{az} provided in step 4 with updated values of $n_{seg\ r}$, $n_{seg\ w}$, $n_{seg\ t}$ and n_{simps}
9. Confirm the δ_{IT} estimate of step 2 with updated values of n_{az}, $n_{seg\ r}$, $n_{seg\ w}$, $n_{seg\ t}$ and n_{simps}

As the procedure is carried out for several n_{rad} values, upon choice of a reference set of FSM parameters, error estimates on each fidelity metrics and associated to each parameter in table 6.5 are obtained. Then, building on this information, some observations can be drawn. First of all, the error estimate associated to n_{rad} can be isolated, providing insight on its trend. Secondly, considering the different contributions arising from ϵ_{IT} and discretization parameters, the analyst is supplied with the necessary information to reach the desired numerical error threshold on each fidelity metric acting on the dominant contributions, to limit the computational time.

The results provided in this section are effectively error estimates, not uncertainties. Their conversion into numerical uncertainties is carried out within the solution validation process in section 6.3, according to the specific validation methodology and stressing the necessary assumptions.

6.2.3. Results

The most outstanding result concerns the effect of n_{rad}. The typical output of an n_{rad} sweep on a fidelity metric is reported in fig. 6.6. In these charts, the variability of the metric with respect to the number of inflow (and interference) states and other numerical error sources can be appreciated. In addition, the error bar on each point represents the combination (through a sum) of all the remaining numerical error estimates. This

[7] The discretization errors due to $n_{seg\ w}$ and n_{simps} were assessed together, performing multiple $n_{seg\ w}$ sweeps for different values of n_{simps}. The reason behind this choice can be traced back to the strong correlation between the two parameters and to the fact that a sufficiently high value of n_{simps} was to be employed in order to spot asymptotic convergence with respect to $n_{seg\ w}$.

effective numerical error decomposition allows to isolate the inflow error behavior and, therefore, to draw several observations.

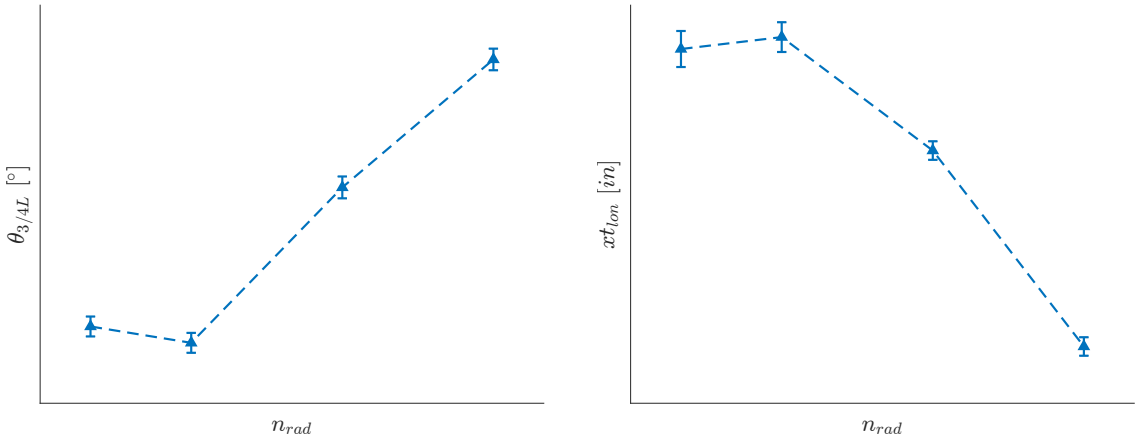
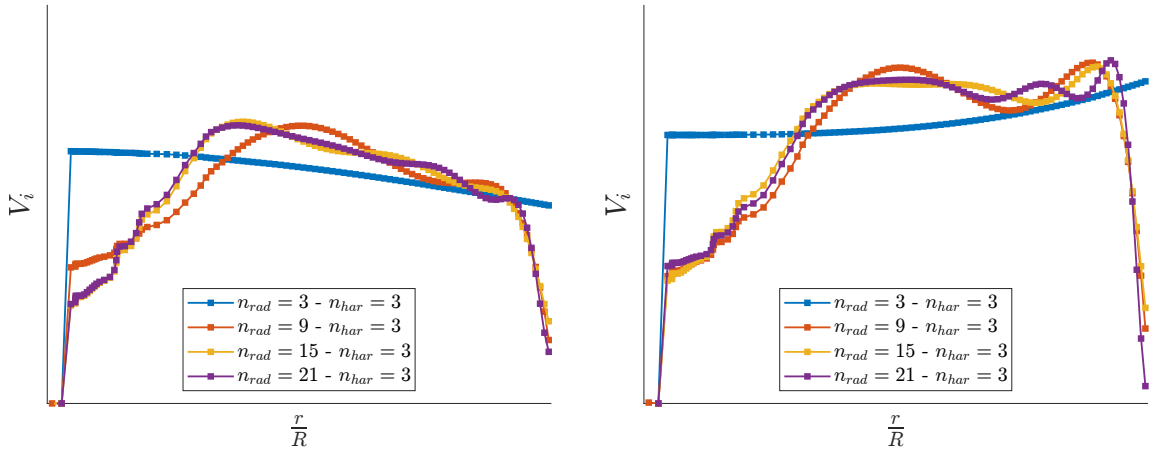


Figure 6.6: Left rotor collective and longitudinal control position numerical error estimates. The dashed line represents the variation of the SRQ with respect to n_{rad}. The errorbars on each point, conversely, represents the sum of all other numerical error contributions estimates accounted for in this study (e.g. due to ϵ_{IT}, n_{az}, etc.)

First of all, fig. 6.6 highlights the magnitude difference between the error associated to inflow discretization (dashed line) and the error due to the remaining parameters involved (error bar amplitude). Thus, for this FSM, considering the typical flight simulation execution time constraints imposed by industrial applications, the numerical error due to inflow and interference states truncation is indeed dominant.
Secondly, while the uncertainties owed to trim equation tolerances, aerodynamic segments and time step typically converge to zero as the discretization improves, the same behavior was not observed for the examined inflow states samples. Therefore, the numerical error due to inflow states truncation was not confidently assessed either.
Moreover, as previously mentioned, the inflow harmonics were kept fixed to 3 for this study (according to [19]). Nevertheless, studies carried out by LHD on articulated hovering helicopter rotors showed that increasing the number of harmonics over the number of blades results in collective control position variations of the same order of magnitude of the ones observed with respect to n_{rad}.

For the sake of better understanding the roots of this behaviour, an investigation of the rotor inflow radial distribution on an isolated rotor model was performed. The results obtained for two different collective blade pitch angles are displayed in fig. 6.7. As re-

ported, even truncating the series at a high number of states, the changes in the inflow distribution (especially at the tip region) are still evident.



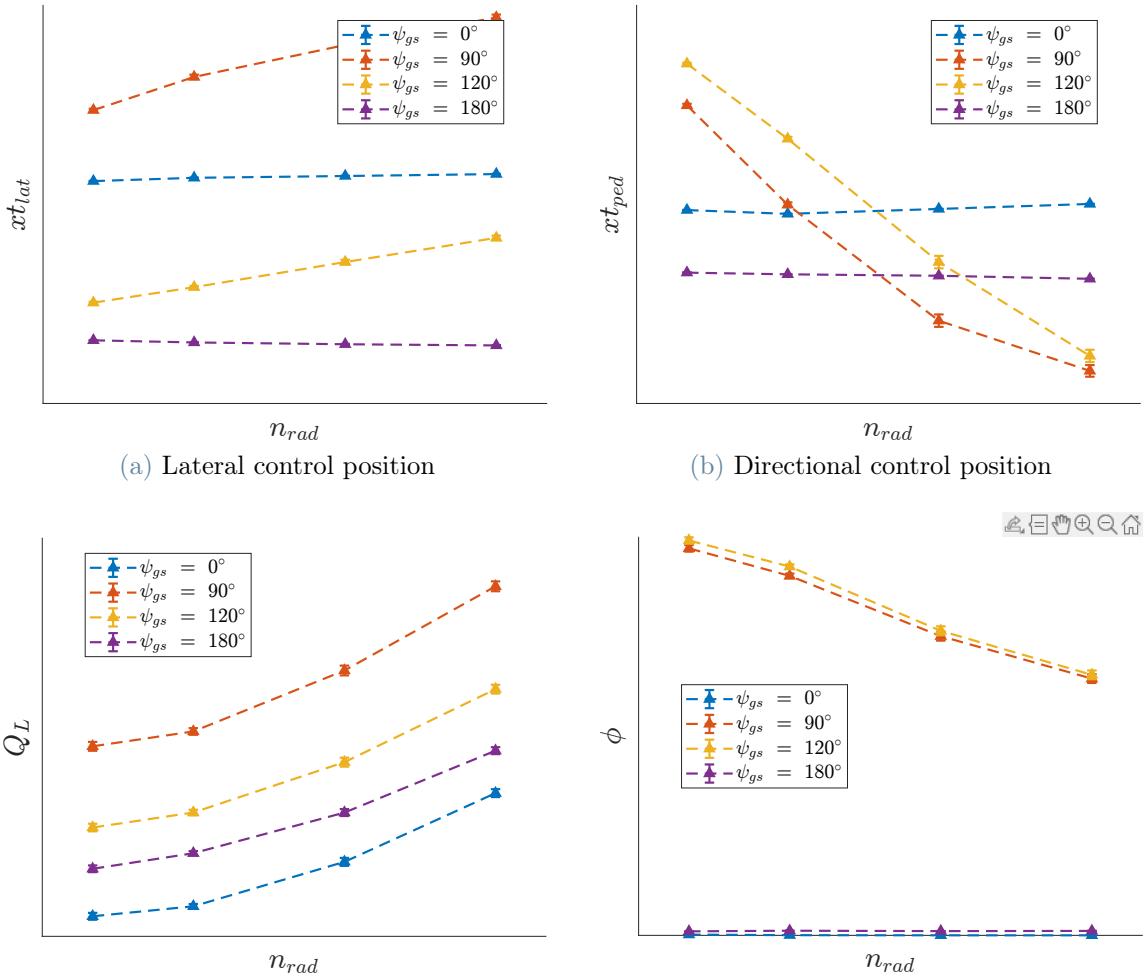
(a) *Low* collective control position (b) *High* collective control position

Figure 6.7: Radial induced velocity distribution for different values of n_{rad} on an isolated rotor model in still air

It is also interesting to observe that when the effects of n_{rad} on the validation metrics of all flights are compared, some trends emerge (fig. 6.8). Indeed, regardless ψ_{gs}, a systematic growth in trim collective and rotor torque and power figures is observed when n_{rad} is increased (fig. 6.8c). Concerning *lateral* validation flights (namely, $\psi_{gs} = 90°$ and $\psi_{gs} = 120°$), the increase in n_{rad} seem to result in an increase in the lateral control position, in a decrease in the directional control position and in a decrease in the roll angle. Thus, despite qualitative, these trends may come useful to the analyst in the interpretation of the validation results, to understand whether FSM negligence may be traced back to a model-form error associated to the inflow.

In conclusion, building on the evidence provided above, the author decided not to consider the effect of n_{rad} into the numerical error. Indeed, considering the great increase in computational expense that comes together with an increase in n_{rad}, it may be argued that even if convergence might eventually occur for sufficiently high values of n_{rad}, the associated computational time would make the model unsuitable even for offline desktop simulations. Hence, the most reasonable way to account for the error associated with Peters-He inflow model in the present application is to include it into the model-form error.

Moving on to the other parameters affecting the numerical error, asymptotic convergence was typically observed on convergence studies involving ϵ_{IT}, n_{az}, n_{simps}, $n_{seg\ w}$ and $n_{seg\ t}$.



(a) Lateral control position

(b) Directional control position

(c) Left rotor torque

(d) Roll angle

Figure 6.8: Comparison of the trends of some SRQs with respect to n_{rad} on different validation flights

On the other hand, variations of the SRQs with respect to $n_{seg\ r}$ typically exhibited an oscillating, non monotonic behaviour (fig. 6.4b). Nevertheless, their order of magnitude was well characterized and proved to be very robust with respect to both the validation flight and all other numerical parameters. Building on these considerations, several discretization levels were evaluated, eventually leading to the choice of ϵ_{IT} and discretization parameters to be exploited for the rest of the work.

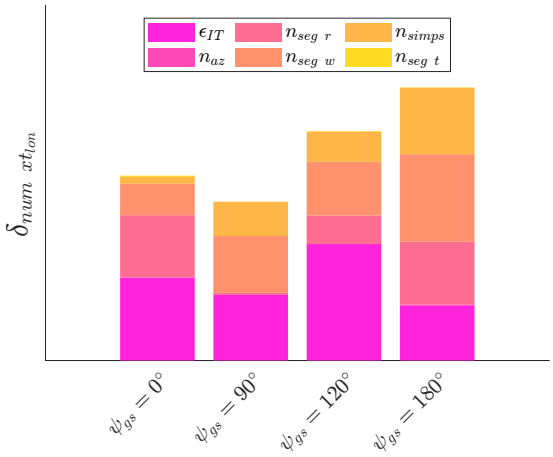
This final choice was dictated by a compromise between computational time, confidence on the error estimate[8] and error estimate magnitude. Concerning the latter, recalling the observations concerning the corollaries discussed in section 4.1.4, particular attention should be devoted to ensure that the cumulative numerical error estimate is well below

[8]Namely, whenever asymptotic convergence was observed, the author qualitatively ensured that the chosen discretization level was within the asymptotic range of convergence.

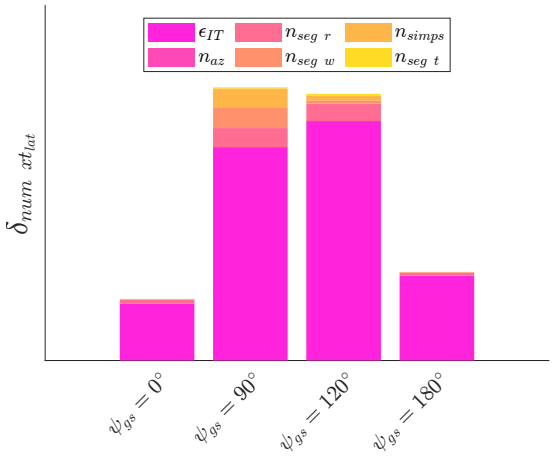
the validation acceptable margins chosen by the applicant for all fidelity metrics. Otherwise, even neglecting the uncertainties due to model input and experimental data, the applicant may not be able to exhibit evidence of model fidelity adequacy according to the selected margins.
For the present work, the margins reported in table 5.2 were used as a guideline, and not strictly applied. In fact, considering the small absolute values assumed by control positions in some validation flights (e.g. lateral control position in $\psi_{gs} = 120°$ or longitudinal control position in $\psi_{gs} = 0°$), it was not always possible to achieve compliance to the aforementioned requisite[9] with acceptable values of computational time. Thus, in regard of control positions, the author considered acceptable discretization levels which provided a sufficiently small absolute cumulative error estimate.

The numerical error estimate on some SRQs for the ultimately chosen values of the parameters in table 6.5 is reported in fig. 6.9. As it can be spotted from the histograms, typically both the total error and its contributions manifest a dependence on the flight condition. However, the reduced amount of validation flights exploited in this work did not provide sufficient material to infer reliable correlations of δ_{num} with value of model inputs (such as ground speed, density altitude, etc.). Nevertheless, this variability poses doubts regarding how to account for the numerical error in the validation process and during model extrapolation. In particular, shall each validation flight use its own estimate? or an arbitrarily chosen value should be used for all flights? Moreover, whenever a numerical error estimate shall be interpolated/extrapolated, is it correct to use the absolute value of the error? or the percentage error shall be used instead?
Despite the answers to these questions may lose relevance as soon as the numerical error becomes so small to be negligible with respect to the other sources of uncertainty, it is reasonable to say that they depend on the specific application. In this work, during the model validation phases, the choice was made to manage each validation flight with its own error estimate (appropriately turned into numerical uncertainty).

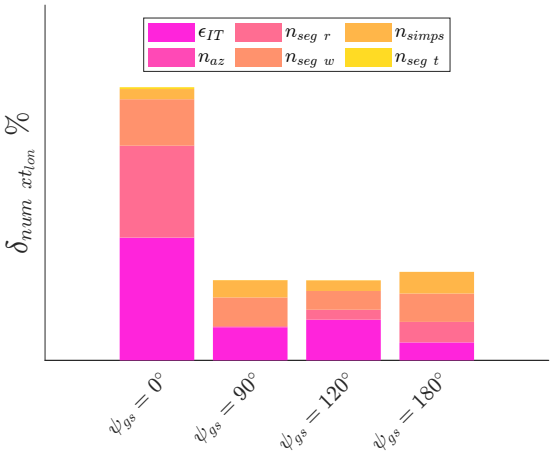
[9] namely, "*the cumulative numerical error estimate is well below the validation acceptable margins chosen by the applicant for all fidelity metrics*".



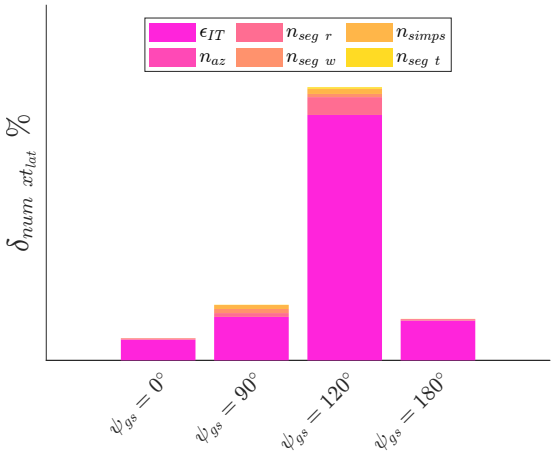
(a) Longitudinal control position δ_{num} decomposition



(b) Lateral control position δ_{num} decomposition



(c) Longitudinal control position δ_{num}% decomposition



(d) Lateral control position δ_{num}% decomposition

Figure 6.9: Numerical error estimates contributions on all validation flights

6.3. Model Validation

6.3.1. Input Uncertainty Identification and Characterization

The first step in solution validation concerns the identification and characterization of FSM input uncertainties. In the present work, together with LHD, it was decided to include within this study uncertainties due to wind speed and direction, aircraft mass and center of gravity position. Using engineering judgement, they have been considered as the most influential factors.

Concerning the wind uncertainty mathematical characterization, no series of observa-

tions/realizations was available. On the other hand, regarding aircraft mass and balance, the realizations available to the author were not sufficient to reasonably characterize the associated uncertainty. Hence, no uncertainty evaluation based on statistical analysis was possible. As a consequence, all uncertainties have been evaluated as type B [16] in ASME framework and epistemic in the Roy-Oberkampf framework[10].

Concerning the wind speed magnitude, the FAA advisory circular for FAR part 29 [5] was taken as a reference. Among the recommended procedures for general performance flight testing, the advisory circular suggests that "*hover performance testing should be conducted in winds of 3 knots or less*". Thus, in absence of additional information, the author assumed $0 - 3$ knots to be a reasonable wind speed uncertainty range for the LHD low speed flight tests selected for validation. On the other hand, regarding the wind direction, any value within the $0° - 360°$ range is possible.

Considering the aircraft weight, the author relied once more on [5]. In this regards, the advisory circular states that "*test weights should be maintained within +3% and -1% of the target weight for each data point*". Hence, assuming that the standard procedure for certification flight testing were employed for the test point chosen for validation, it is possible to state that 3% reasonably represents the maximum allowable uncertainty for the aircraft weight. Moreover, since, conversely to the wind speed, a nominal aircraft weight measure is provided in the validation flight data, the author decided to assign a Gaussian PDF to this input. This choice was dictated by the will to embed within the uncertainty characterization the belief of the author that the nominal value is the most likely, and the best available estimate of the true value. Hence, the mean of the PDF was set equal to the nominal flight data, while the standard deviation was set equal to 1% of the mean[11].

Eventually, in regard of the CG position, no applicable criteria were found. Thus, the author decided to assign a Gaussian PDF to each coordinate of the CG. Once more, the mean was set equal to the nominal CG position datum. Then, the 99.73% confidence interval was set equal to 0.35 inches, hence resulting in a standard deviation of $0.11\overline{6}$ inches.

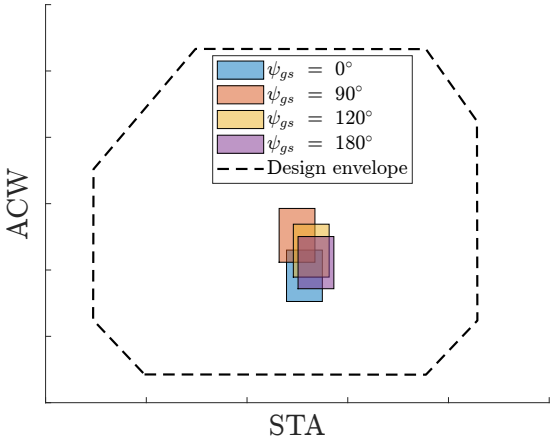
A summary of the FSM inputs uncertainties and a representation of the uncertainty boxes resulting from the assumptions above are reported in table 6.6 and fig. 6.10 respectively.

[10] Namely, the uncertainty on model inputs values reflect the belief and opinion of the author according to the available knowledge.

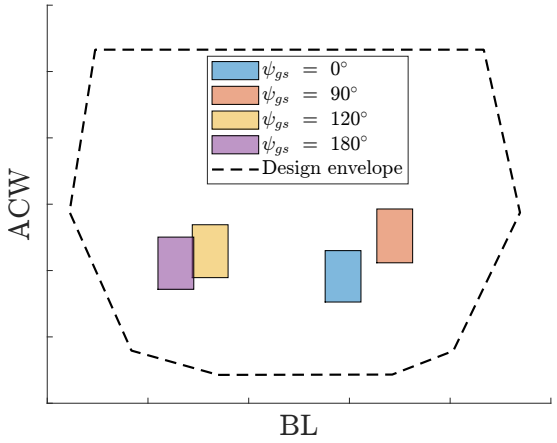
[11] In order to ensure a 99.73% confidence interval on the 3% bound derived from the FAA advisory circular [5].

Model input		PDF	Mean	Std deviation	Bounds
Wind speed	WSP	-	-	-	$0-3$ kts
Wind azimuth	WAZ	-	-	-	$0-360$ deg
Aircraft weight	ACW	Gaussian	ACW_{nom}	1% ACW_{nom}	-
CG fuselage	STA	Gaussian	STA_{nom}	$0.11\overline{6}$ in	-
CG butt line	BL	Gaussian	BL_{nom}	$0.11\overline{6}$ in	-
CG waterline	WL	Gaussian	WL_{nom}	$0.11\overline{6}$ in	-

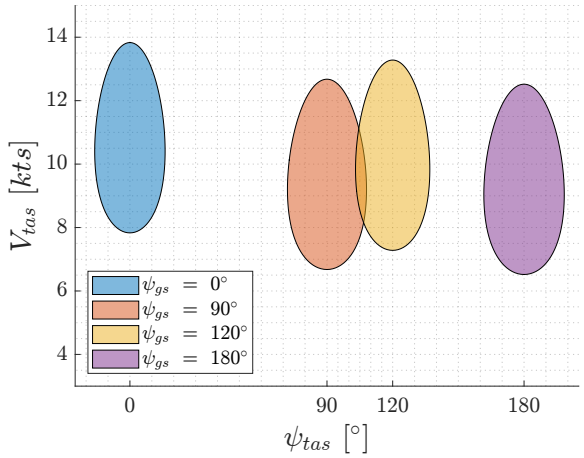
Table 6.6: Summary of FSM input uncertainty identification and characterization



(a) 99.73% confidence intervals on the CG fuselage station and aircraft weight of validation flights against design envelope



(b) 99.73% confidence intervals on the CG butt line station and aircraft weight of validation flights against design envelope



(c) True air speed magnitude and azimuth uncertainty boxes for each validation flight

Figure 6.10: FSM input uncertainty boxes

6.3.2. Sensitivity Analysis

Prior to any uncertainty propagation, sensitivity analyses should be performed with the scope of assessing the robustness of a computational model and educating the analyst on the model behaviour. Within the present work, the aim is to exploit sensitivity analyses to identify, among the candidate uncertain model inputs, those whose uncertainty results in negligible effects on all the SRQs of interest. Once identified, such inputs may be treated as deterministic, at the benefit of the uncertainty quantification computational time.

Methodology

The choice of the sensitivity analysis method usually depends on the non-linearities of the model and the computational expense of a single model run. Hence, considering the strong non-linearities exhibited by the FSM presented in section 6.1, several simple and computational inexpensive sensitivity methods (e.g. derivative-based local methods or classical one-at-a-time methods) have been excluded.
Among the global sensitivity methods, the Variance-Based-Decomposition (VBD) method stands out for its widespread adoption in the scientific literature and the reliability of its sensitivity measures: the main effect sensitivity index S_i and the total effect sensitivity index S_{T_i}, both reported in the following equations.

$$\mathrm{S}_i = \frac{Var(\mathrm{E}(\mathrm{Y}|x_i))}{Var(\mathrm{Y})} \tag{6.12}$$

$$\mathrm{S}_{T_i} = \frac{Var(\mathrm{Y}) - Var(\mathrm{E}(\mathrm{Y}|x_{-i})}{Var(\mathrm{Y})} \tag{6.13}$$

where x_i represents the i-th model input and $x_{-i} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_k)$. The main effect sensitivity S_i estimates the expected fraction of $Var(\mathrm{Y})$ that could be removed if the true value of x_i was available. However, it does not capture effects due to the interactions among different input. Conversely, the total effect sensitivity S_{T_i} accounts for all the effects on $Var(\mathrm{Y})$ involving x_i, thus both its main effect and its interactions with the other model inputs. More than one method is available for computing the VBD sensitivity measures. The method proposed by Saltelli in [2] allows to access S_i and S_{T_i} at the computational cost of $N(k+1)$, where N is the number of necessary model evaluations. As reported in [10], N is typically at least 100, and can easily become as high as tens of thousands for slow converging models.

Nevertheless, for the present work, a different approach was pursued and the Morris One-at-A-Time (MOAT) global screening method was chosen. The MOAT method was

firstly introduced by Morris in 1991 [21] and was later revised in its sensitivity metrics and sampling strategy by Campolongo, Cariboni and Saltelli [10]. The method is an extension of the local OAT derivative based approach, since it overcomes the limitations of the latter by efficiently exploring the whole model input space. In addition, it is well suited for identifying the most important factors in a model with a large number of inputs (namely, too large to apply variance-based techniques). Thus, despite this not being the case for the present application, the author decided to pursue this approach anyway, in sight of a potential larger scale application in the industry (where the competitiveness of MOAT with respect to variance-based methods is even greater). A brief outline of the MOAT method and a presentation of its sensitivity metrics is provided below, while a detailed description can be found in [9].

For each of the k model input, a number of SRQ(s) incremental ratios, referred to as Elementary Effects (EEs), are computed at different location within the input space according to eq. (6.14).

$$d_i(\mathbf{X}) = \frac{y(\mathbf{X} + \Delta \mathbf{e}_i) - y(\mathbf{X})}{\Delta} \tag{6.14}$$

where $\mathbf{X} \in \mathbb{R}^k$ represents the location within the input space, $\mathbf{e}_i$ is the i^{th} vector of the canonical basis for the model input space, Δ a scalar increment defined within the method, and y a SRQ of interest. EEs are evaluated at different locations within the input domain through an efficient sampling technique [10] based on the construction of r trajectories[12] of $k+1$ points. Each trajectory is generated from a starting point, and develops through the model input space via perturbations of one input at a time (fig. 6.11). At each perturbation, an EE is computed, and a new point belonging to the trajectory is generated from that perturbation. At last, once all k inputs have been perturbed, the trajectory is made of a collection of $k+1$ points. Eventually, upon the collection of r EEs at different locations of the input space, the sensitivity indices, referred to as μ, μ^* and σ^* in the present work, are computed for each SRQ-input pair according to eq. (6.15), (6.16) and (6.17).

$$\mu_i = \frac{1}{r} \sum_{j=1}^{r} d_i(\mathbf{X}^{(j)}) \tag{6.15}$$

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^{r} |d_i(\mathbf{X}^{(j)})| \tag{6.16}$$

[12] This particular feature of the method represents its substantial difference with traditional, local OAT methods.

$$\sigma_i^* = \sqrt{\frac{1}{r}\sum_{j=1}^{r}(d_i(\mathbf{X}^{(j)}) - \mu_i)^2} \tag{6.17}$$

The mean μ_i (originally proposed by Morris) and the modified mean μ_i^* (proposed by Campolongo et al. in [10]) give an indication of the overall effect of the i^{th} model input on the SRQ. The standard mean μ carries the information about the sign of the EEs, but, as a consequence, is prone to type II errors[13] [10]. On the other hand, the modified mean μ^* is resilient to type II error, but does not carry any information about the sign of the EEs. By default then, the Dakota implementation of the MOAT method return the modified mean. Nevertheless, since μ can be computed with no additional cost once the EEs are known, the analyst, whenever interested in the information carried by the sign, may retrieve its value as well. Still, for the purpose of the sensitivity analysis within this work, the modified mean μ^* is sufficient. σ^*, on the other hand, is an indicator of the dispersion of EEs throughout the model input domain. Thus, it effectively represents non linear effects and interactions between the different model inputs. σ^* is null for linear models.

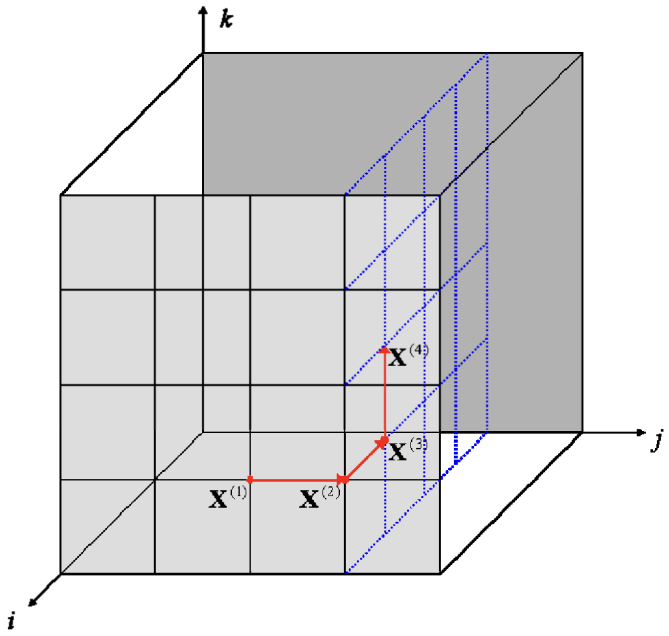


Figure 6.11: An example of trajectory in the input space when $k = 3$ (from [9])

The MOAT method requires two parameters to be selected by the user: the number of

[13] i.e., in this application, to fail in identifying an important input. Indeed, if an input significantly affects an SRQ but with opposite trends depending on the position $\mathbf{X}$ in the model input space, the resulting EEs mean may be small, despite the parameter being important.

trajectories r and the number of partitions p (e.g. 4 in fig. 6.11) for each dimension of the input space domain. Increase in both parameters results in both more samples and a greater coverage of the input space. Eventually, the number of function evaluations amounts to $r(k+1)$, with r typically being in the range of $10 \sim 20$ [10]. The typically low sample size required by the MOAT method to inform the analyst on the most important and unimportant parameters (especially in large models) represents its main strength with respect to VBD methods. Nevertheless, this advantage comes at the price of providing qualitative information only. Still, Saltelli et al. [10] compared MOAT with VBD observing great agreement between the two methods with the same sample size and concluding that MOAT μ^* at low samples sizes are indeed great proxies of the more accurate VBD indices S_{T_i}.

Implementation and Analysis Setup

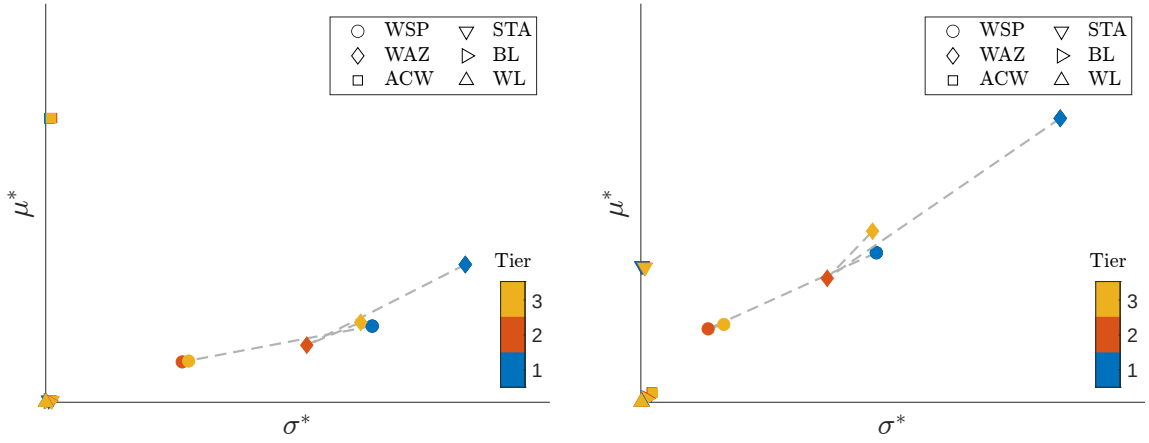
The sensitivity analysis was performed with Dakota and was carried out on all validation flights. Sensitivity indices for all fidelity metrics in section 5.1 were computed. Unbounded uncertain inputs (namely, ACW and CG coordinates) were limited in their 99.73% confidence interval, since MOAT requires all model inputs to be bounded. Moreover, since, as aforementioned, the MOAT method is parametric in r and p, a brief independence study on the sensitivity metrics was performed. Nevertheless, considering the fact that μ^* and σ^* are qualitative results, this preliminary investigation was only devoted to ensure that no doubt related to importance (or unimportance) of any model input could arise changing the sensitivity method options. The values of r and p explored within the independence assessment are summarized in table 6.7.

	Tier 1	**Tier 2**	**Tier 3**
r	10	30	60
p	9	18	27
n	70	210	420

Table 6.7: Summary of MOAT parameters values explored for the sensitivity analysis independence assessment: r represents the number of trajectories, p the number of partitions for each model input, n the total number of function evaluations (aircraft trims) required for the analysis

Results

The results of a sensitivity analysis based on the MOAT method are typically reported in scatter plots, as displayed in fig. 6.12 and 6.13. Each plot summarizes the sensitivity picture for a SRQ of interest and, within the plot, each point position carries the information of the influence, in terms of μ^* and σ^*, of a given input on the SRQ. Nevertheless, in the graphs reported in fig. 6.12 and 6.13, a triplet of points (identified by the same marker shape and linked by a dashed gray line) is reported for each input, in order to appreciate on the same plot the variability of the sensitivity indices with respect to r and p tiers of table 6.7. Moreover, the conclusions carried out in this section find direct confirmation in fig. 6.12 and 6.13, but they also extend to the other validation flights.



(a) Collective control

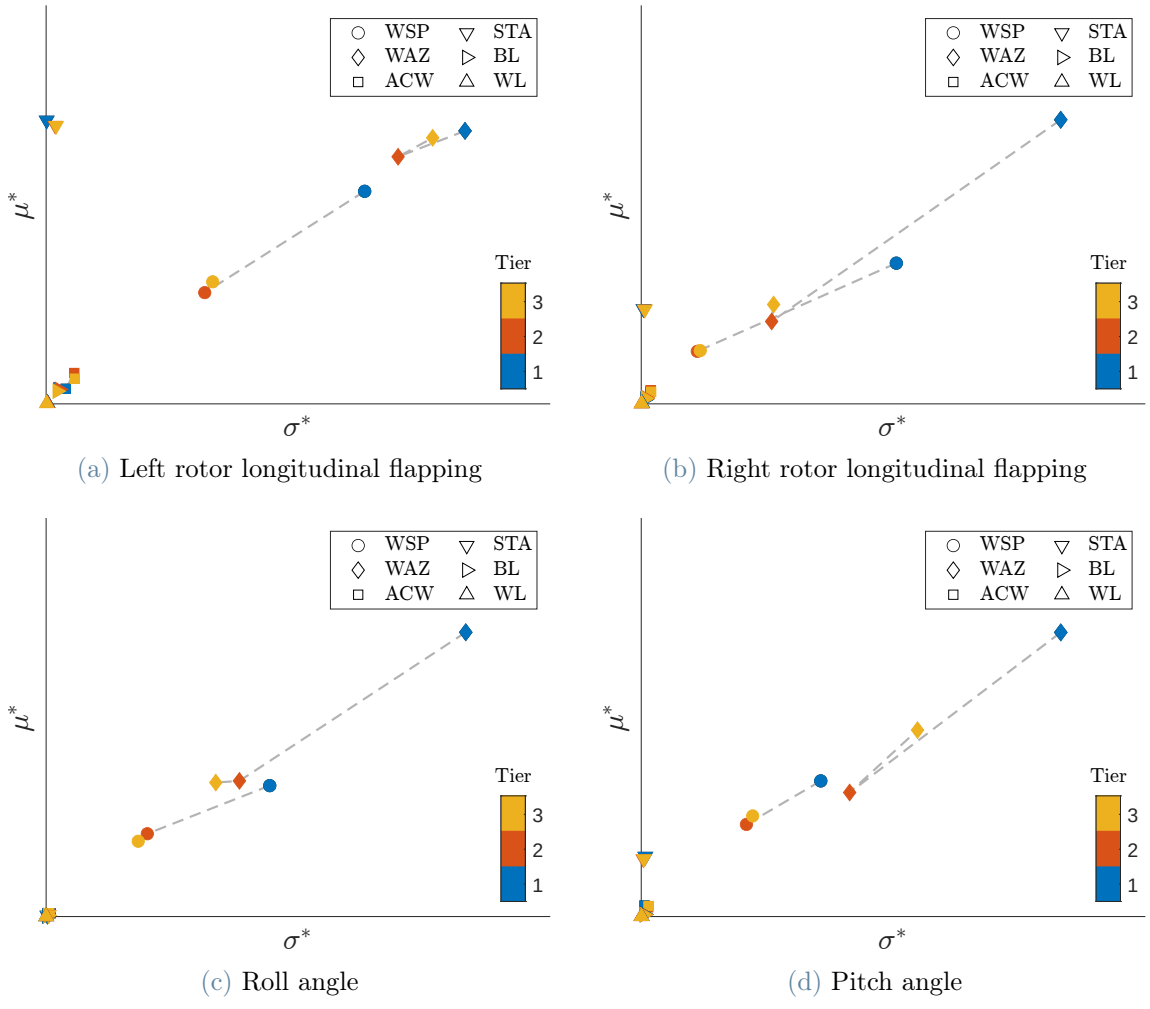
(b) Longitudinal control

(c) Lateral control

(d) Directional control

Figure 6.12: MOAT sensitivity indices of control positions at trim for the $\psi_{gs} = 120°$ validation flight. Different MOAT parameters are identified by different colors, while indices corresponding to different model inputs are identified by distinct marker shapes

First of all, it is important to observe that, for every SRQ-input pair, the different tiers do not provide significantly different μ^* values. On the other hand, significant variations in σ^* associated to WSP and WAZ occur on several SRQs. These changes can be traced back to the large number of partitions p (and trajectories r) that is intuitively necessary to characterize the whole wind envelope in terms of direction. Indeed, it may be argued that WAZ might require a much larger number of partitions with respect to all other input intervals (e.g. ACW, STA, BL) for its dispersion index σ^* to "converge". Nevertheless, being μ^* what ultimately drives the analyst in identifying an input as influential, and considering that, as previously mentioned, the method is qualitative, it is safe to conclude that the differences between the three tiers do not lead to ambiguous conclusions.



(a) Left rotor longitudinal flapping

(b) Right rotor longitudinal flapping

(c) Roll angle

(d) Pitch angle

Figure 6.13: MOAT sensitivity indices of longitudinal flapping angles and aircraft attitudes at trim for the $\psi_{gs} = 120°$ validation flight. Different MOAT parameters are identified by different colors, while indices corresponding to different model inputs are identified by distinct marker shapes

Secondly, looking at fig. 6.12 and 6.13, it is possible to observe that the points associated to WSP and WAZ typically occupy the right-hand side of the graphs (corresponding to greater σ^*). This is expected, considering that WSP and WAZ team up in characterizing the wind speed uncertainty. Thus, the impact of the former on a given SRQ strictly depend on the value of the latter (and viceversa), resulting in great variations of the associated EEs throughout the input domain. Conversely, ACW, STA, BL and WL always lay on the left hand side, meaning that their effect on the output is more or less independent on the values assumed by other inputs.

Finally, it is worth noticing that all input quantities, except WL, result in large μ^* for at least one model fidelity metric. Indeed, WSP and WAZ provide important effects on all SRQ (dominating, in particular, aircraft attitudes and directional control). ACW results in a great influence on the collective control position and, thus, rotors collective and torque figures. STA variations determine non negligible effects on longitudinal control position, pitch attitude and longitudinal rotor flapping, while BL strongly influences the lateral control position. Conversely, WL always occupies the left-bottom corner of the plots, with μ^* sensitivity values typically at least an order of magnitude smaller with respect to the aforementioned inputs. Thus, to the benefit of the computational expense, the author decided to exclude any uncertainty associated to WL from all UQ analyses, treating it as a deterministic input.

6.3.3. ASME Solution Validation

Within this section, the solution validation process according to ASME VV Standard 20, summarized in section 4.1.2, is carried out.

Numerical Uncertainty Definition

The ASME validation procedure requires the applicant to compute the standard deviation of the numerical error δ_{num} parent population. Thus, starting from the error estimates retrieved during the solution verification process, the applicant shall make some assumptions in order to combine the different contributions to the numerical error and transform the numerical error estimate to a standard uncertainty measure. Within the present work, all numerical error sources are assumed to be uncorrelated and each error estimate gathered in section 6.2 is assumed to represent the upper bound for the corresponding numerical error contribution. Then, a uniform PDF was assumed for each contribution to the numerical error, in order to convert the error estimate into a standard uncertainty, compatible with the framework proposed by ASME. The PDF amplitude was set equal

to the estimated error magnitude and the coverage factor exploited for the conversion operation was set 1.73. Eventually, relying on the hypothesis of non-correlation between the different contributions, the final numerical standard uncertainty u_{num} was obtained using the RSS, hence combining the estimated standard uncertainties associated to each contribution. An outline of the entire process is displayed in fig. 6.14.

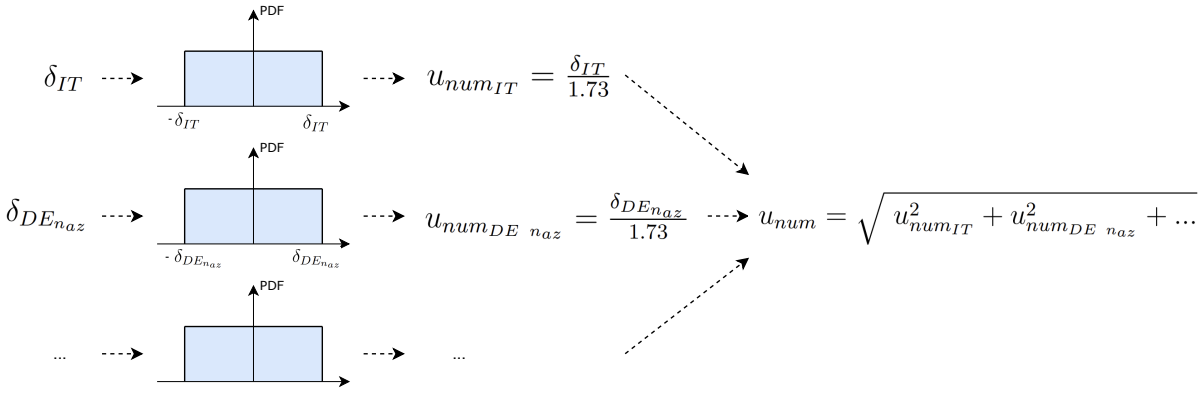


Figure 6.14: Summary of numerical uncertainty u_{num} evaluation process within the ASME validation framework

Experimental Uncertainty Definition

The uncertainty associated to the experimental data was computed according to the widespread metrology standards reported in [16]. In particular, the experimental value D of each SRQ was retrieved averaging the flight datum time history over the acquired time window, according to eq. (6.18). On the other hand, the associated standard uncertainty was computed using the formula of the standard uncertainty for the population mean, reported in eq. (6.19).

$$D = \frac{1}{N}\sum_{n=1}^{N} y_n \tag{6.18}$$

$$u_D = \frac{1}{N}\sqrt{\frac{1}{N}\sum_{n=1}^{N}(y_n - D)^2} \tag{6.19}$$

where y_n is a single realization of the measurement time history.

Input Uncertainty Definition

The uncertainty associated to the model inputs was estimated with Monte Carlo sampling using Dakota, considering the non-linearity of the FSM.

Then, building on the results of the FSM sampling, and considering the definition of δ_{input} presented in eq. (6.20), the input-related uncertainty u_{input} can be computed as the standard deviation of the sampled SRQ population, according to eq. (6.21).

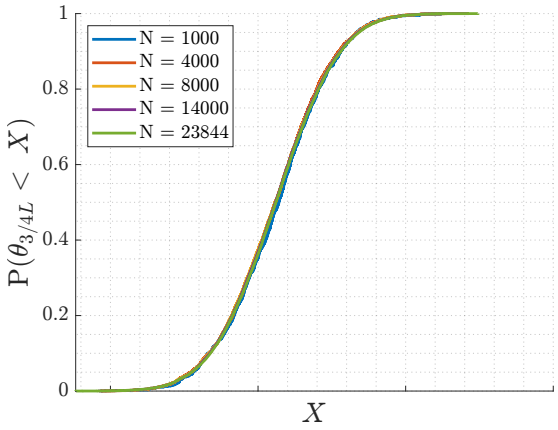
$$\delta_{input} = y_n - S \tag{6.20}$$

$$\begin{aligned} u_{input} &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}[\delta_{input} - \mu_{\delta_{input}}]^2} \\ &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}[(y_n - S) - \frac{1}{N}\sum_{n=1}^{N}(y_n - S)]^2} \\ &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}[y_n - S - \mu_y + S]^2} \\ &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}[y_n - \mu_y]^2} \\ &= \sigma_y \end{aligned} \tag{6.21}$$

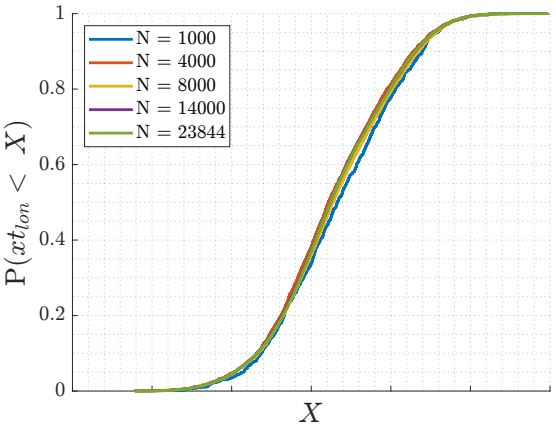
where y_n is a single SRQ prediction obtained with Monte Carlo sampling and N is the number of samples.

For each validation flight, approximately 24000 samples of each SRQ of interest were generated. In order to gain an insight on the robustness of the results with respect to the sample size, histories of the population mean μ and standard deviation σ of each SRQ were visually inspected. Nevertheless, such operation was not aimed to the quantification of the sampling uncertainty on u_{inp}, considering that ASME does not provide any guidance for such an activity. Conversely, it was carried out to collect qualitative information about what SRQs standard deviation estimates are surely not reliable and would, in principle, require more samples to improve the confidence on their values.
Some results associated to the control positions in the $\psi_{gs} = 90°$ validation flight are reported in fig. 6.15. Despite the suppression of the Y-axis ticks in the images, it is possible to conclude that the sample size of 24000 generally provides a reliable estimates of μ and σ for control positions, attitudes and rotors collective and torque figures. Nevertheless, as it can be grasped in fig. 6.16, the same conclusions cannot be always carried out for rotor flapping angles, whose variations in the CDF and statistical moments may be still considered non-negligible at this samples size. Thus, despite validation results associated

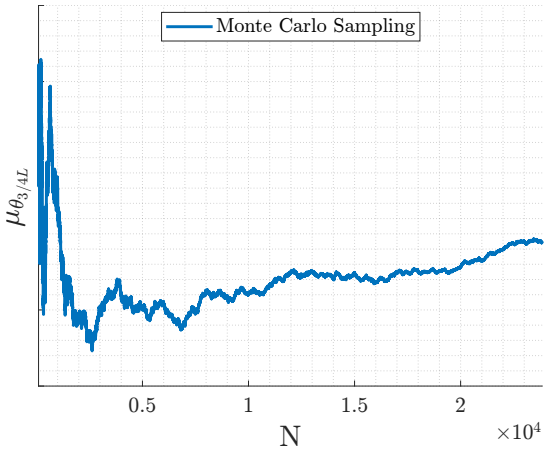
to these SRQs are reported in the following sections for the sake of completeness, it is reasonable to state that an increase in the number of samples is needed to make them more trustworthy.



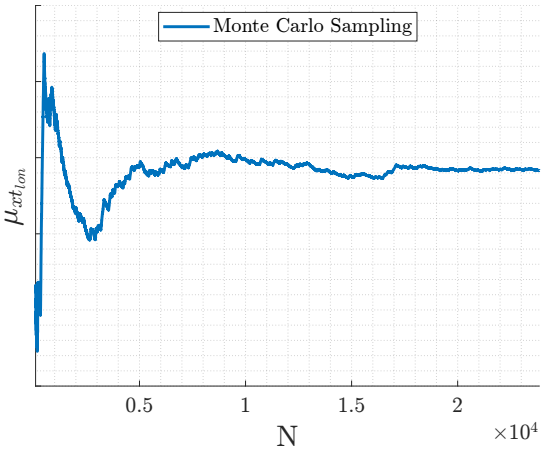
(a) Discrete CDF of the left rotor collective obtained with different samples sizes



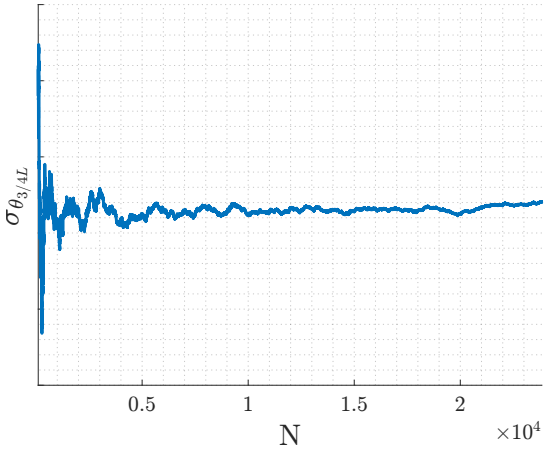
(b) Discrete CDF of the longitudinal control position obtained with different samples sizes



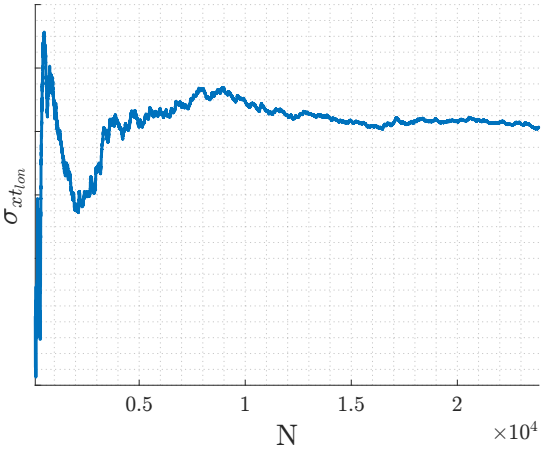
(c) Left rotor collective samples mean



(d) Longitudinal control samples mean

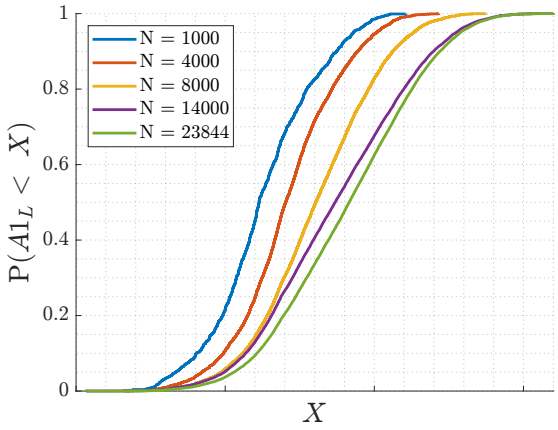


(e) Left rotor collective samples standard deviation

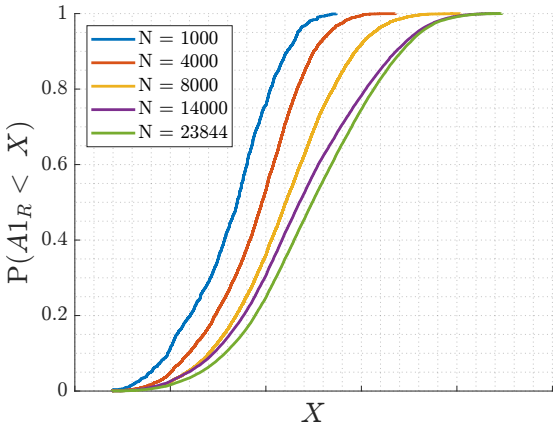


(f) Longitudinal control samples standard deviation

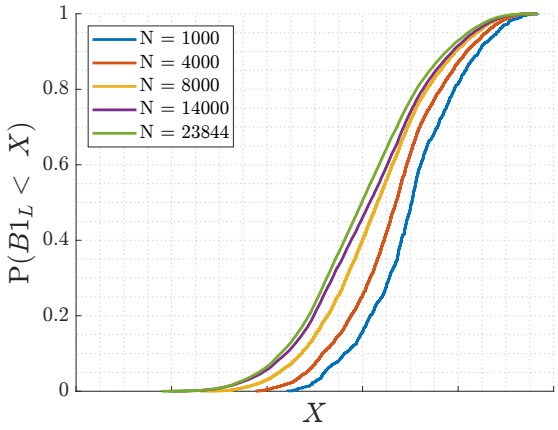
Figure 6.15: Results of the Monte Carlo sampling analysis on left rotor collective and longitudinal control position for validation flight $\psi_{gs} = 90°$



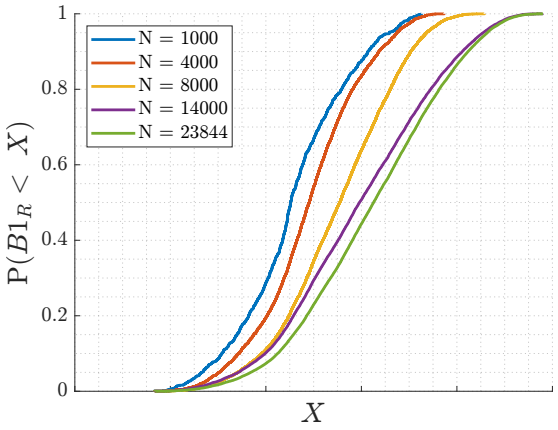
(a) Left rotor longitudinal flapping discrete CDF at different samples sizes



(b) Right rotor longitudinal flapping discrete CDF at different samples sizes



(c) Left rotor lateral flapping discrete CDF at different samples sizes



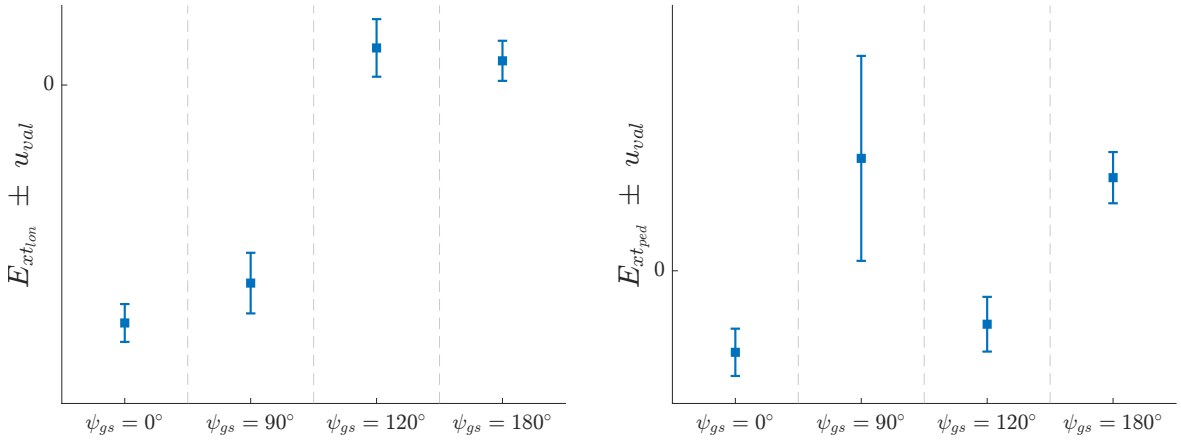
(d) Right rotor lateral flapping discrete CDF at different samples sizes

Figure 6.16: Results of the Monte Carlo sampling analysis on validation flight $\psi_{gs} = 90°$ rotors flapping angles

E and u_{val} Validation Metrics

Upon the computation of u_{num}, u_D and u_{input} as exposed in the previous sections, the validation metric u_{val} can be computed according to eq. (4.7). On the other hand, E can be easily computed according to eq. (4.3).

The results of the validation process are discussed in this section and can be consulted in fig. 6.17 and appendix B. In all these figures, the errorbars are equal to the validation standard uncertainty u_{val}, which do not represent the bounds of the comparison error E at all.



(a) Longitudinal control position (b) Directional control position

Figure 6.17: E and u_{val} validation metrics results for longitudinal and directional control positions on all validation flights. The magnitude of u_{val} is reported as an errorbar on E

As mentioned in section 4.1.2, by reporting fidelity acceptability margins onto the plots of the comparison error, the analyst is able to assess whether or not the required fidelity requirements are met. Upon negative outcome of the investigation, the applicant shall investigate the causes of noncompliance and figure out model improvements.
The application of the ASME procedure corollaries reported in section 4.1.2 leads to the conclusion that the validation uncertainty is too large (with respect to the comparison error E), to allow for a reliable quantification of the model error and measure model improvements. In fact, $E >> u_{val}$ in a handful of circumstances only: longitudinal control position at $\psi_{gs} = 0°$ (fig. 6.17a) and pitch angle at $\psi_{gs} = 180°$ (fig. B.2e). In all other instances, E is typically of the same order of magnitude of u_{val}.
The reasons behind these occurrence can be traced back to excessively large u_{input} and, less frequently, u_D. As reported in fig. 6.18 (and appendix B), u_{input} contribution to the overall validation standard uncertainty is typically at least an order of magnitude greater than the ones provided by u_D and u_{num}. Still, some exceptions are present, in which the most dominant contribution can be traced back to u_D (e.g. directional control position in $\psi_{gs} = 120°$ or lateral control position in $\psi_{gs} = 180°$). Thus, if the FSM turned out to be non compliant to fidelity requirements in the present application, a reduction of the input uncertainties and/or gathering higher quality data from validation-oriented flight testing are indeed necessary to set up well-advised model updates in these flight conditions.
Nevertheless, if the applicant were satisfied of the level of fidelity achieved, the FSM may still be suited to support the certification activity, as soon as large uncertainties (and a consequent reduction of the CR) are affordable.

(a) Validation flight $\psi_{gs} = 120°$

(b) Validation flight $\psi_{gs} = 180°$

Figure 6.18: Decomposition of $u_{val}{}^2$ into the contributions originating from u_{input}, u_D and u_{num}

Moreover, fig. 6.18 (and appendix B) highlight that the numerical error contribution to u_{val} is typically negligible with respect to the others. Thus, it may be argued that the verification of the assumptions behind the numerical error contributions combination and conversion to uncertainty, as well as the absence of *ad hoc* methods for the solution verification multi-body models, did not turn out to be of primary concern in this application.

6.3.4. Roy-Oberkampf Solution Validation

Within this section, the solution validation process proposed by Roy and Oberkampf, summarized in section 4.2, is carried out.

Numerical Uncertainty Definition

Within the Roy-Oberkampf approach, all numerical uncertainties are managed as interval-valued epistemic uncertainties. Moreover, as reported in [29] and previously mentioned in section 4.2.2, the numerical uncertainty magnitude U_{NUM} is directly obtained from the error magnitude estimate and the combination of the different contributions is carried

out with a sum. Thus, in order to derive U_{NUM} for the current application, all δ_{num} contributions estimates retrieved during solution verification (namely, δ_{IT}, $\delta_{DE\ n_{az}}$, etc.) were converted into epistemic uncertainties first. Then, all uncertainties were combined with a sum. An outline of the process is reported in fig. 6.19.

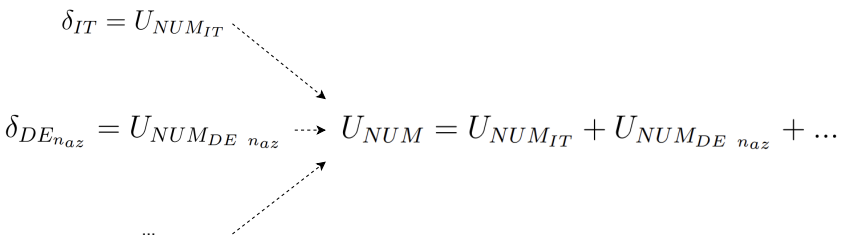


Figure 6.19: Summary of the numerical uncertainty U_{NUM} evaluation process within the Roy-Oberkampfs framework

Experimental Uncertainty Definition

In the Roy-Oberkampf framework, the uncertainty on the experimental data is managed with the computation of a discrete CDF from the experimental observations. An exemplification of the process is displayed in fig. 6.20.

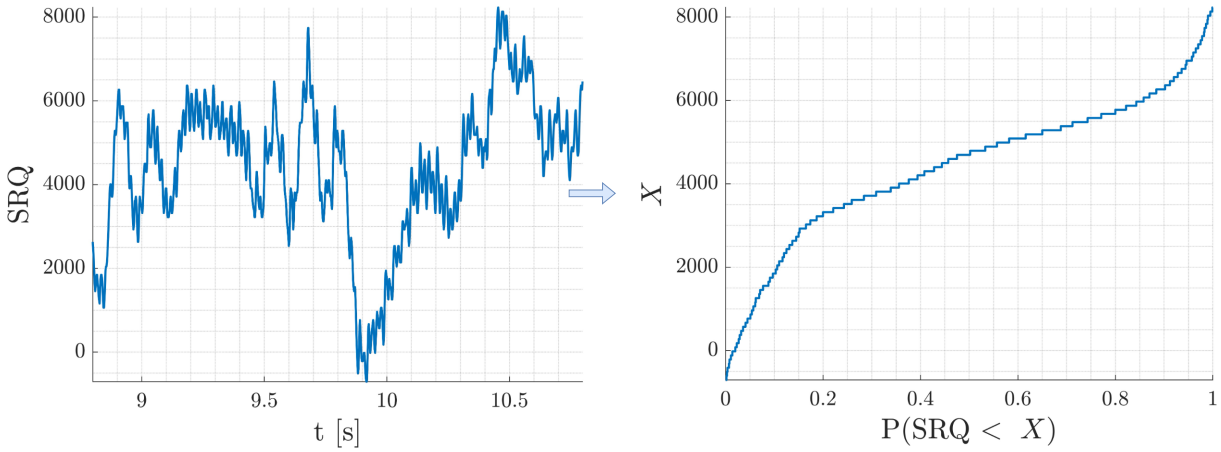


Figure 6.20: Process of creation of the experimental data CDF within Roy-Oberkampf verification and validation framework. The example in the figure is referred to the left rotor torque data of the $\psi_{gs} = 0°$ validation flight

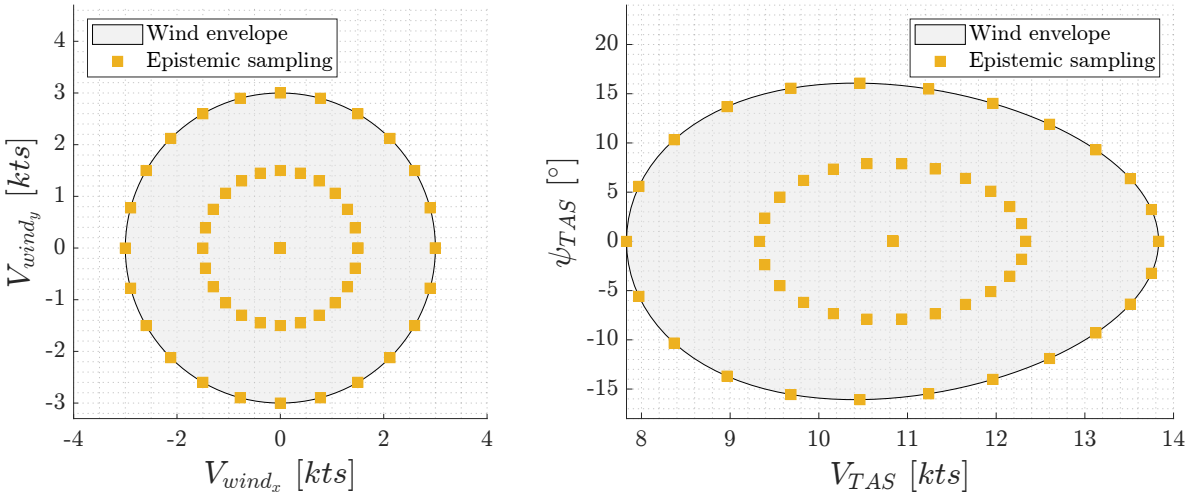
Input Uncertainty Definition

In the Roy-Oberkampf approach, the uncertainties associated to model inputs are characterized by mean of a probability box. In this work, as stated in section 6.3.1, all uncertainties associated to the FSM inputs are strictly epistemic, since certainly reducible and

derived by means other than statistical analysis. However, the decision to assign PDFs to the aircraft mass and CG position convinced the author to frame the latter model inputs uncertainties as aleatory within this validation attempt. Indeed, this choice results in the advantage of distinguishing, at the end of the UQ, the SRQ uncertainty due to the wind from the one due aircraft mass and balance. In addition, this decision paves the way to test the full algorithm for the computation of the pbox, summarized in fig. 4.3.
The first step of the procedure requires the analyst to choose the sample sizes M and N defined in fig. 4.3.

The number of samples M, related to the sampling of interval-valued inputs (namely, WSP and WAZ), was chosen to ensure sufficient coverage of the wind envelope. Two partitions were considered for the wind speed magnitude WSP, resulting in three levels of velocity samples: 0, 1.5 and 3 kts. Conversely, a 15° step was used for the definition of the wind azimuth partitions. The aforementioned wind sampling strategy, displayed in fig. 6.21, results in M equal to 47.

In order to choose the aleatory uncertainties sample size N, and to provide at least an estimate to the sampling error associated with its choice, a preliminary analysis was carried out on two validation flights (namely, $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$). For each flight, 25600 aleatory uncertainties samples were evaluated in no wind conditions ($V_{gs} = V_{TAS}$). Then, collecting samples from this database, several CDFs ensambles at lower sample sizes

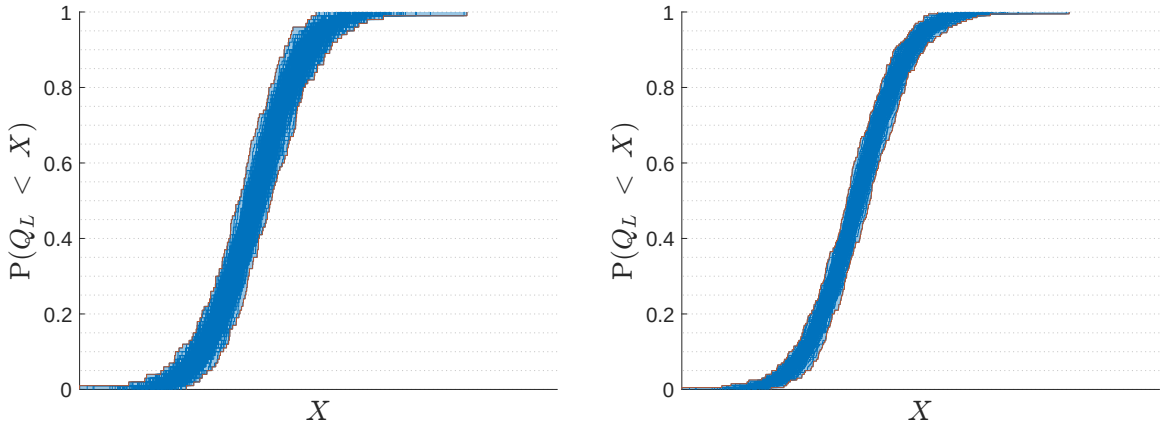


(a) Wind speed sampling

(b) Resulting TAS uncertainty sampling on $\psi_{gs} = 0°$ validation flight

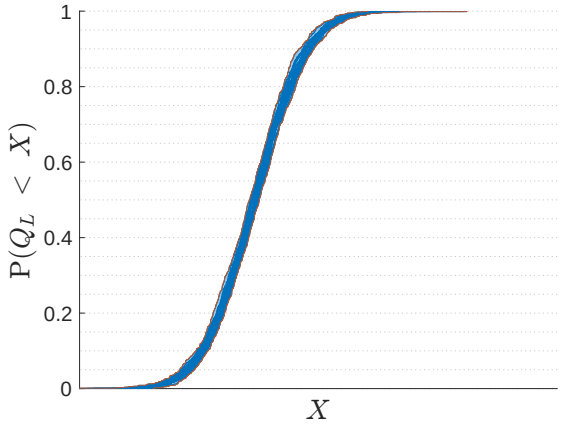
Figure 6.21: Wind sampling strategy for input uncertainty quantification within Roy-Oberkampf framework

N (and different seeds) were generated for each SRQ[14], as displayed in fig. 6.22. For each value of N, the area occupied by the ensemble (see fig. 6.22) was taken as an estimate of the sampling error δ_{SAMPL} associated to the CDF computed with N samples. Indeed, whenever a disagreement between the model and the experiment arises in the validation phase, it is the area of this ensemble that (as an epistemic uncertainty) directly affects the resulting validation metric value.

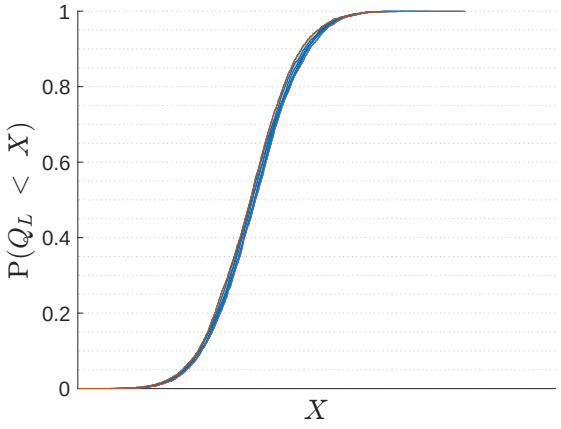


(a) $N = 100$ (256 CDFs with different seeds)

(b) $N = 200$ (128 CDFs with different seeds)



(c) $N = 800$ (32 CDFs with different seeds)



(d) $N = 3200$ (8 CDFs with different seeds)

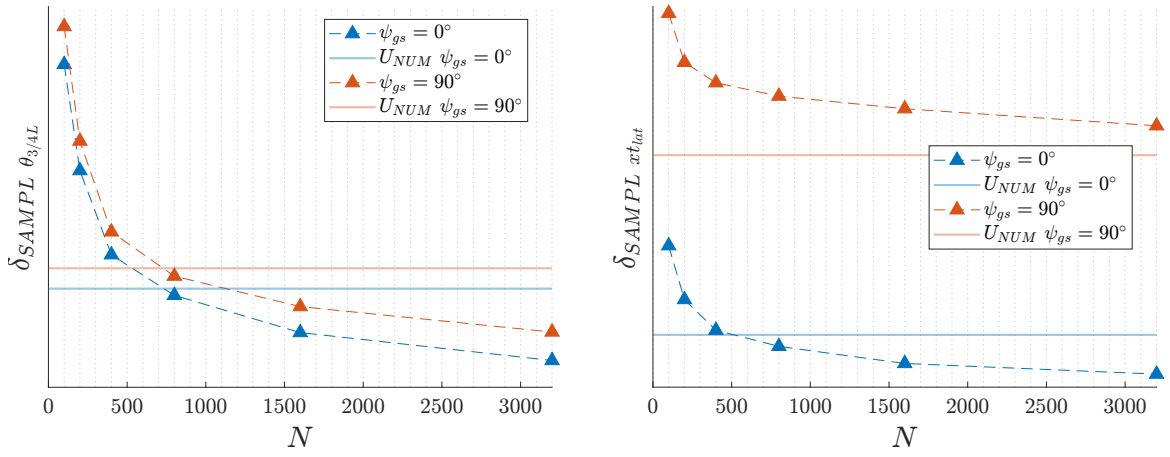
Figure 6.22: Evolution with N of the sampling error estimate on the discrete CDF of the left rotor torque ($\psi_{gs} = 90°$ validation flight)

Finally, it is important to stress that in this investigation, according to the aforementioned procedure, all ensambles are obtained exploiting all 25600 starting samples. Thus, the number of CDFs in the ensemble reduces as N increases and it can be argued that this latter occurrence might result in a bias in the estimated sampling error trend. In addition,

[14]e.g. choosing $N = 100$ allows to collect 256 discrete CDFs computed at different seeds from the starting database of 25600 SRQs evaluations.

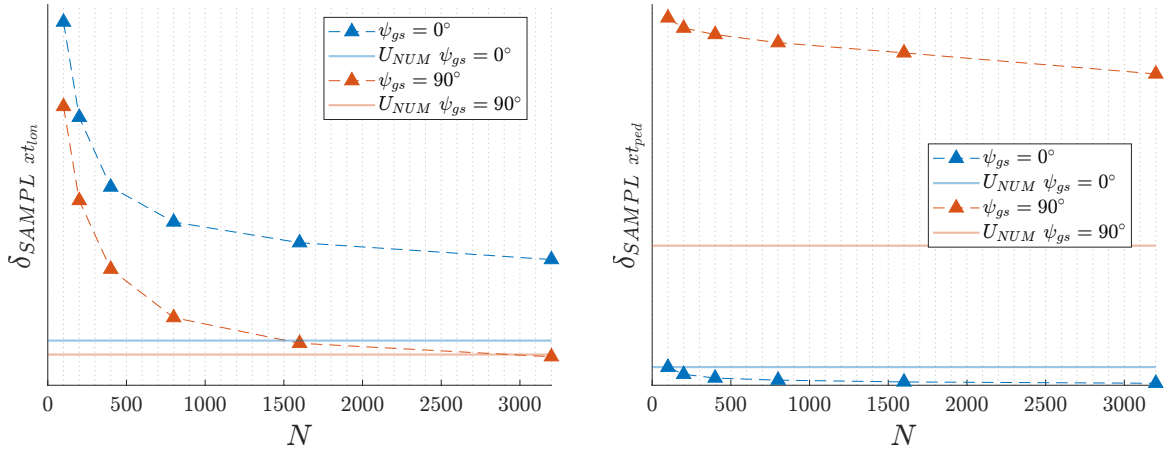
whenever δ_{SAMPL} is comparable, or smaller than U_{NUM}[15], it may also be argued that the sampling error estimation is spoiled by the numerical uncertainty and, hence, it is even less reliable. Aware of the limits of this approach, its results were only exploited as a tool to help the author make an educated choice of N and inform, whenever present, on the occurrence of situations characterized by unacceptable sampling uncertainties. As a consequence, these δ_{SAMPL} estimates are not accounted for in the validation reported in the following sections.

The results of the CDFs sampling error estimation analysis are reported in fig. 6.23 to 6.25. In each plot, the numerical uncertainty U_{NUM} is reported as well, in order to provide reference for the sampling error estimate order of magnitude.



(a) Left rotor collective

(b) Lateral control position



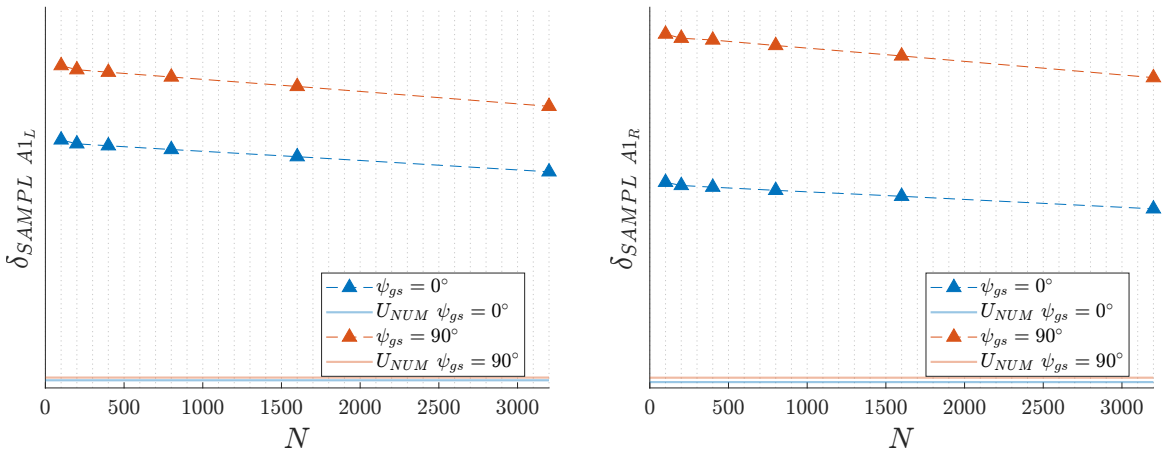
(c) Longitudinal control position

(d) Directional control position

Figure 6.23: CDF sampling error estimates for control positions and left rotor collective on $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$ validation flights

[15]Especially when $U_{NUM_{IT}}$ occupies a large portion of U_{NUM}.

Starting from control positions and rotor collective and torque figures, $N = 400 \sim 800$ typically results in sampling uncertainties on the CDF which are of the same order of magnitude of the numerical uncertainty. However, the same conclusions cannot be carried out for flapping angles, for which all samples sizes result in unacceptably large estimated sampling uncertainties on the CDFs (fig. 6.24).



(a) Left rotor longitudinal flapping angle

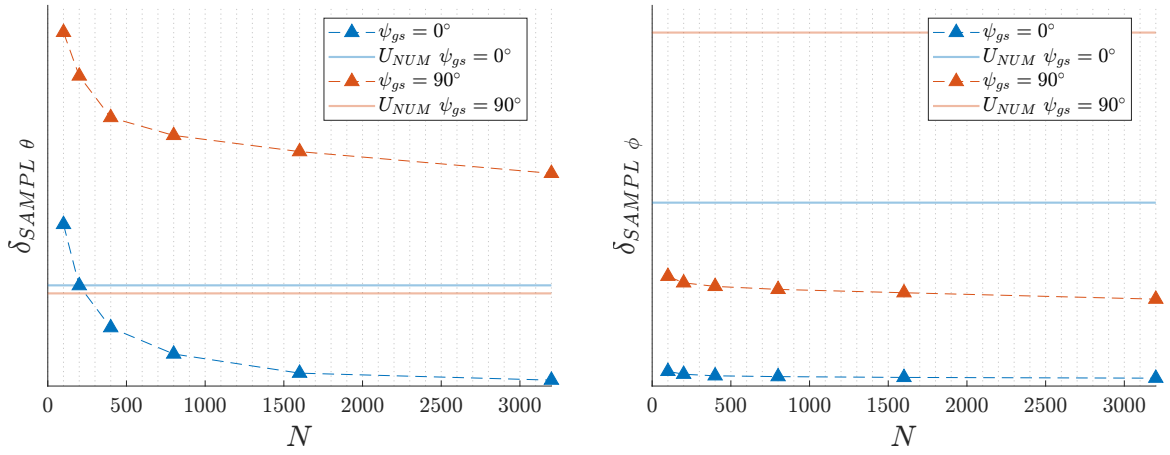
(b) Right rotor longitudinal flapping angle

(c) Left rotor lateral flapping angle

(d) Right rotor lateral flapping angle

Figure 6.24: CDF sampling error estimates for rotor flapping angles on $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$ validation flights

Finally, concerning aircraft attitude angles (fig. 6.25), the same outcome observed for controls positions and rotors collective and torque figures takes place. So, eventually, as a compromise between sampling accuracy and computational effort, the author decided to set the aleatory uncertainty sample size N of fig. 4.3 equal to 400. Nevertheless, it is clear from fig. 6.24 that using $N = 400$ while neglecting the sampling error definitely leads to unreliable outcomes regarding rotor flapping angles validation. Thus, no result



(a) Pitch angle (b) Roll angle

Figure 6.25: CDF sampling error estimates for attitude angles on $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$ validation flights

concerning these SRQs is presented within the Roy-Oberkampf validation framework.

Upon the choice of M and N, the algorithm reported in fig. 4.3 was implemented using Dakota. As a result the probability boxes for each SRQ and validation flight were obtained.
Moreover, in the pursue of reducing the computational expense of the pbox computation, another route was undertaken by the author. Within this novel approach, summarized in fig. 6.26, a direct optimization of the SRQs with respect to all the FSM uncertain input was carried out[16]. Then, the values of WSP and WAZ resulting in the maximum and minimum value of each SRQ were retrieved from the optimization results. Next, a couple of discrete CDFs (one for each optimal wind condition) are computed for the SRQ, sampling N times ACW, STA and BL. These two CDFs are assumed to represent the probability box boundaries.

A comparison of the resulting probability boxes obtained with the conventional sampling approach and the optimization-based approach is reported in fig. 6.27. For this particular application, the disagreement between the two methods always fell within the numerical uncertainty of the results and the estimated order of magnitude of the sampling error. This outcome may be in part traced back to the limited interference effects between the wind uncertainty and ACW, STA and BL uncertainties (see section 6.3.2). As a result, a decoupling between the computation of the pbox bounding epistemic inputs values and the shape of the pbox envelope can be achieved with limited approximations. However,

[16]In order for the optimization to be carried out sensibly, the values of ACW, STA and BL were bounded within their 99.73% confidence interval (three standard deviations from the mean).

the little disagreement between the two methods also testifies the sufficient coverage of the epistemic input space achieved with the conventional sampling technique and M samples. Moreover, on some instances (e.g. fig. 6.27a), were the wind conditions corresponding to the pbox limit were far from all M samples, the optimization-based approach proved its capability to reduce the sampling error associated to epistemic variables. In addition, the optimization-based method achieved the results at a fraction of the computational effort[17] and, since its cost is directly coupled to the number of SRQ of interest, this edge may extend even more if less fidelity metrics are used.
In the validation phase, pboxes obtained with both the conventional and the optimization-based method undergo the validation process and the comparison between the two propagation techniques is extended to the validation area metric.

[17] $\sim$ 10000 function evaluations, against the $\sim$ 20000 necessary for the conventional sampling method

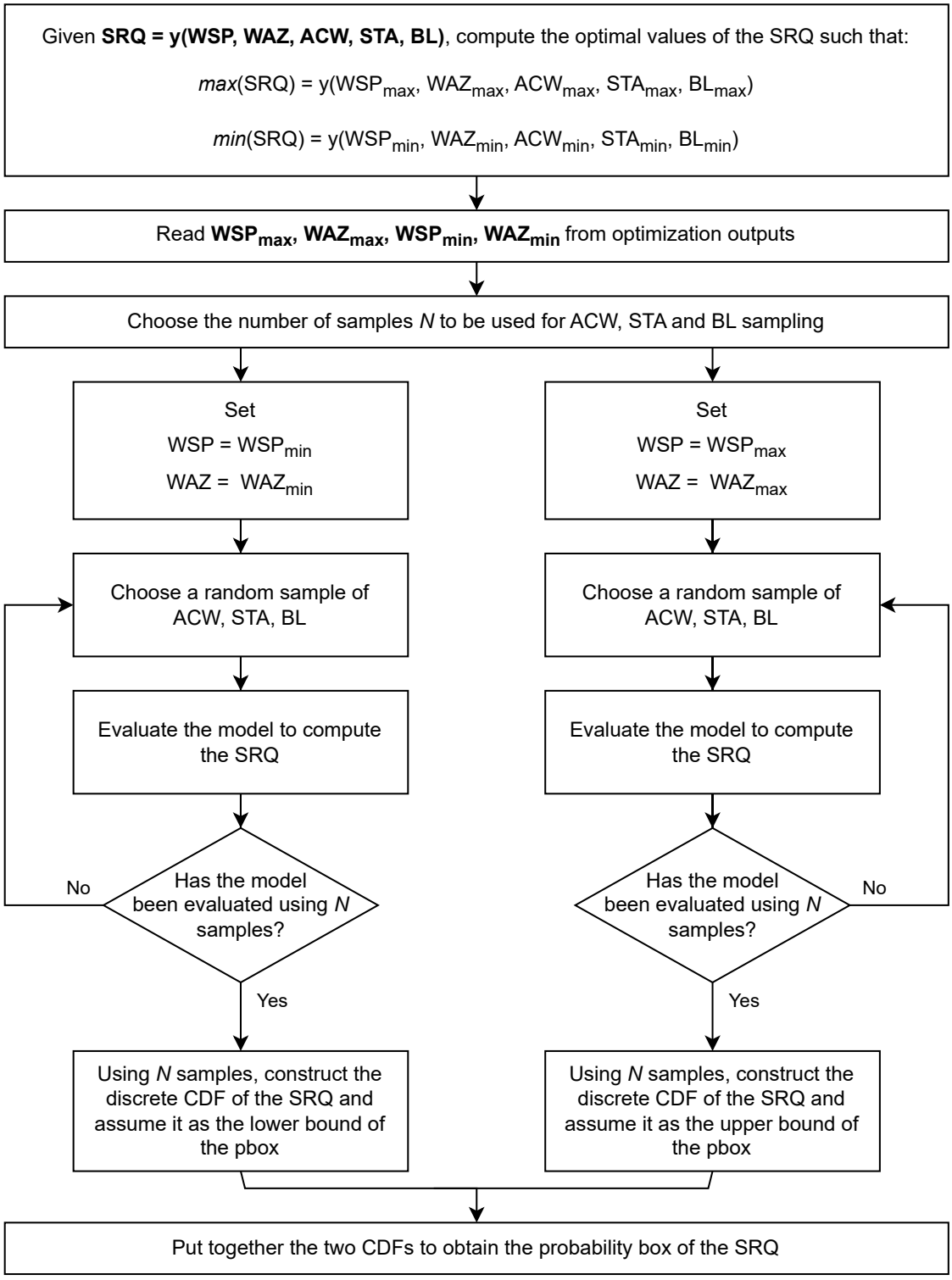
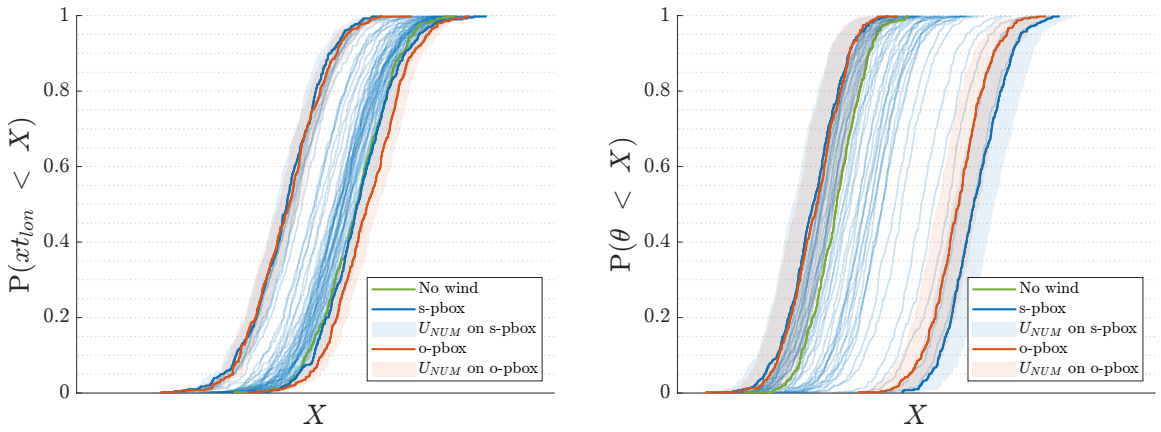


Figure 6.26: Flow chart of the optimization-based approach for the computation of the probability box



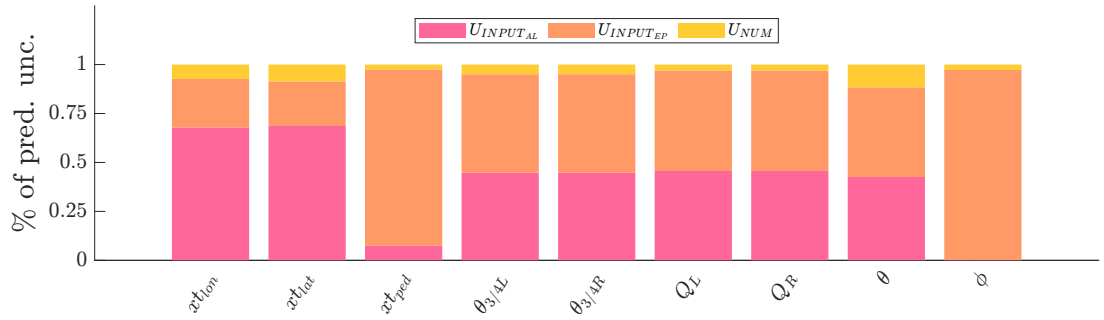
(a) Longitudinal control position (b) Pitch angle

Figure 6.27: Comparison of the probability boxes obtained with the conventional sampling (s-pbox) and optimization-based (o-pbox) approach ($\psi_{gs} = 0°$ validation flight)

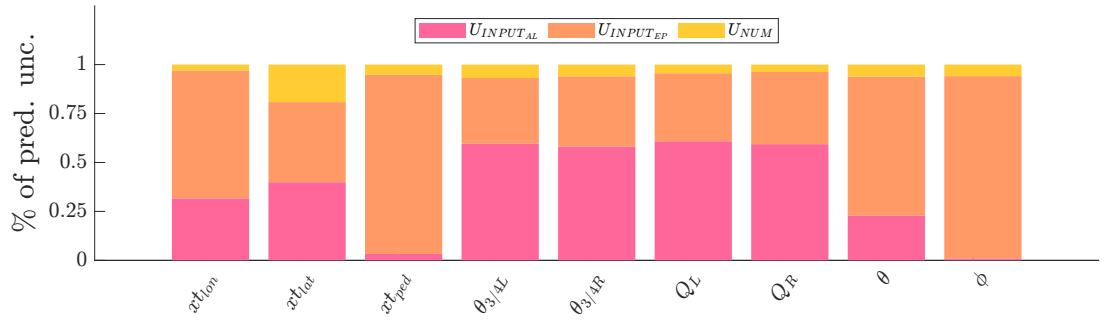
Eventually, considerations concerning the most important contributions to model prediction uncertainty can be made in this framework by looking at the size and shape of each probability box. Nevertheless, in the attempt of providing the analyst an effective tool to have access to the full picture of prediction uncertainty decomposition, the author devised an approximated histogram representation (similar to the one proposed for u_{val} in section 6.3.3). An example of this representation is reported in fig. 6.28, while the results for all validation flights can be retrieved from appendix C.

Within the histogram, the contributions from aleatory and epistemic input uncertainties are separated from the numerical ones and reported as % of the total prediction uncertainty. The epistemic input uncertainty contribution is computed as the area of the probability box before the augmentation due to U_{NUM}. Conversely, the numerical contribution in the histogram is identified by the double of U_{NUM}, since the augmentation rule mentioned in section 4.2.2 requires to append U_{NUM} on either side of the pbox. Eventually, the aleatory input uncertainty contributions has been approximated as the arithmetic average of the 99% confidence intervals of the left and right-hand side pbox boundary CDFs. Then, the total prediction uncertainty was approximated as the sum among all contributions.

Despite approximated, this representation allows the analyst to intuitively collect information associated to all fidelity metrics in the same figure, providing an effective starting point to the decision-making process aimed at reducing model predictive uncertainty.



(a) Validation flight $\psi_{gs} = 0°$



(b) Validation flight $\psi_{gs} = 90°$

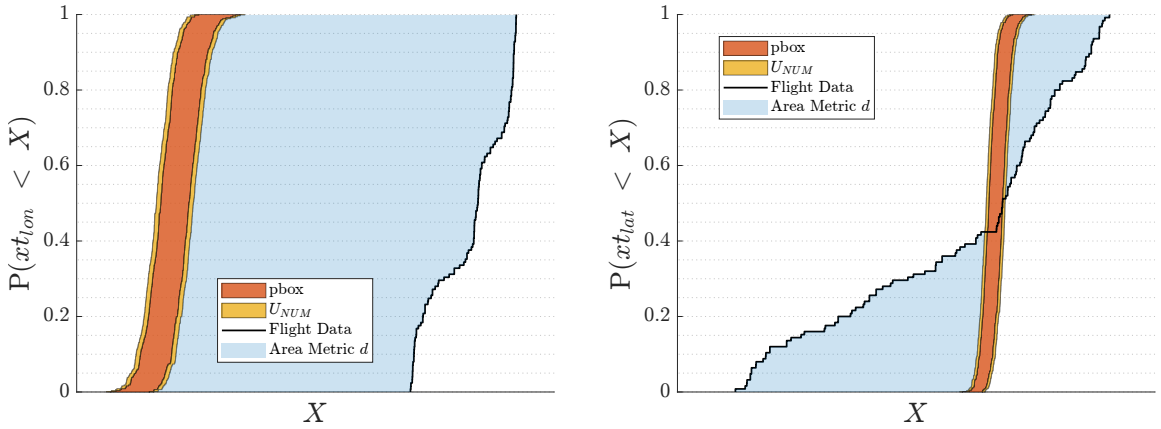
Figure 6.28: Model prediction uncertainty decomposition into the contributions from aleatory and epistemic inputs and numerical error. The results refer to the s-pbox

Building on these results, it is possible to conclude that the directional control position and the roll angle are dominated by wind uncertainty, regardless the flight condition. Conversely, longitudinal and lateral control position predictive uncertainty exhibit an important contribution due to weight and CG position uncertainty in $\psi_{gs} = 0°$ and $\psi_{gs} = 180°$, while wind becomes dominant in $\psi_{gs} = 90°$ and $\psi_{gs} = 120°$ validation flights. Similar conclusions hold for rotors collective and torque figures. However, their predictive uncertainty decomposition seems more resilient to different flight condition.
The contribution from U_{NUM} is typically the smallest, despite peaks up to $\sim 20\%$ can be spotted in the lateral control position for $\psi_{gs} = 90°$.

d Validation Metric

Upon collection of the numerical uncertainty, probability box and experimental data CDF of each SRQ, the area validation metric d was computed, as displayed in the example reported in fig. 6.29. The metric was evaluated both on the pbox obtained with the conventional sampling technique proposed by Roy and Oberkampf (from now on referred to as d-S) and on the pbox obtained with the optimization-based approach (from now on referred to as d-O). The results (reported in this section and appendix C) are hereby

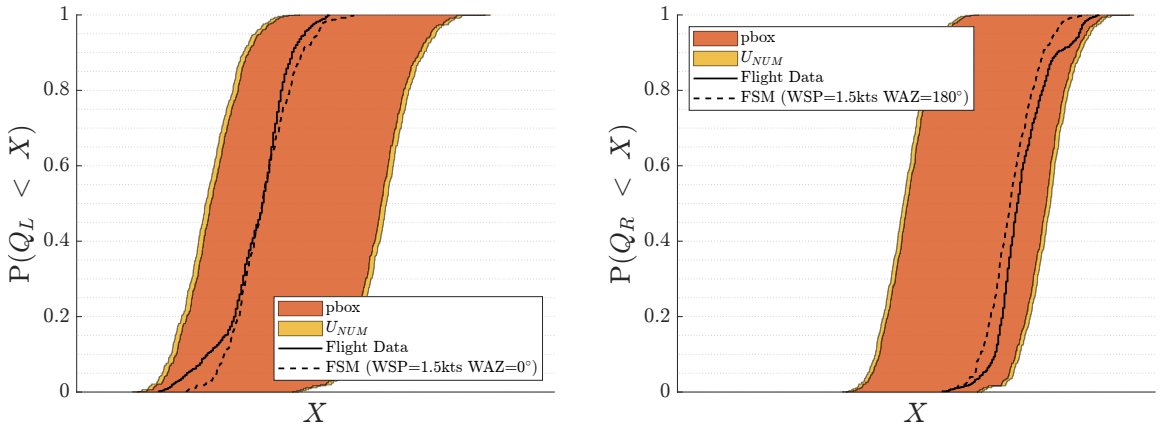
discussed.



(a) Longitudinal control position (b) Lateral control position

Figure 6.29: Graphical representation of the d area metric on the longitudinal and lateral control positions ($\psi_{gs} = 0°$ validation flight)

Since, as reported in section 4.2, the area metric measures the *evidence of disagreement* between the simulation and the experimental data, several occurrences where the area metric is identically 0 can be spotted (fig. 6.30).
On the other hand, whenever evidence of disagreement is present (namely, $d \neq 0$), model update techniques may be conceived to improve agreement even without reducing numerical and model-input associated uncertainties.

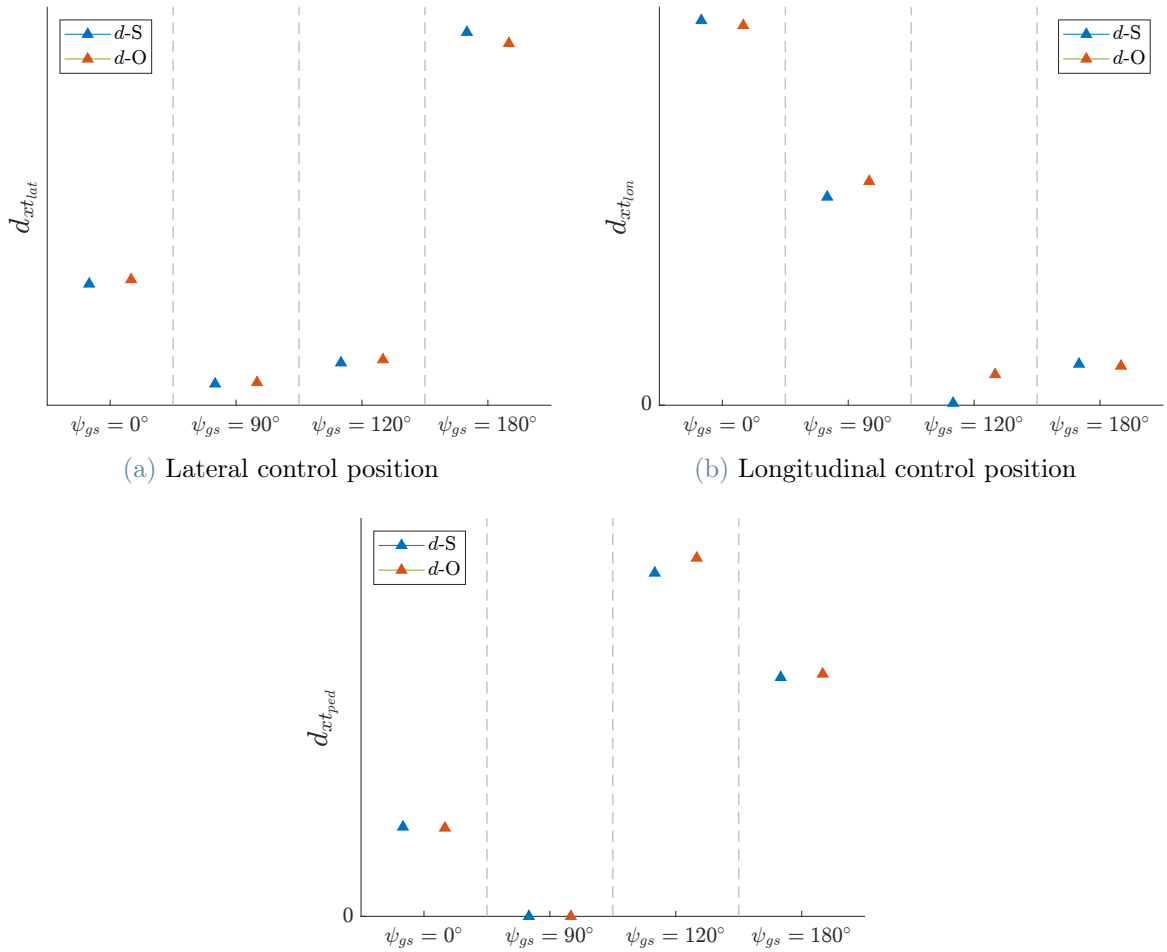


(a) Left rotor torque validation: $d = 0$ (b) Right rotor torque validation: $d = 0$

Figure 6.30: Area validation metric computation on left and right rotors torque figures ($\psi_{gs} = 0°$ validation flight)

When several different fidelity metrics are involved in the same validation attempt (as

in this work), situations in which many of them simultaneously show no disagreement with the experiment may arise. However, it might happen that the agreement occurs in regions of the probability box associated with very different interval-valued inputs values. An example of this occurrence is reported in fig. 6.30. Indeed, according to the area metric, no disagreement between the FSM and the experiment exist for the two rotors torque figures. However, going more in detail, it is possible to observe that, for the left rotor (fig. 6.30a), the experimental datum sits in the left-hand side portion of the pbox, corresponding to a simulation output with wind of ~ 1.5 kts from the 0° azimuth angle.



(a) Lateral control position

(b) Longitudinal control position

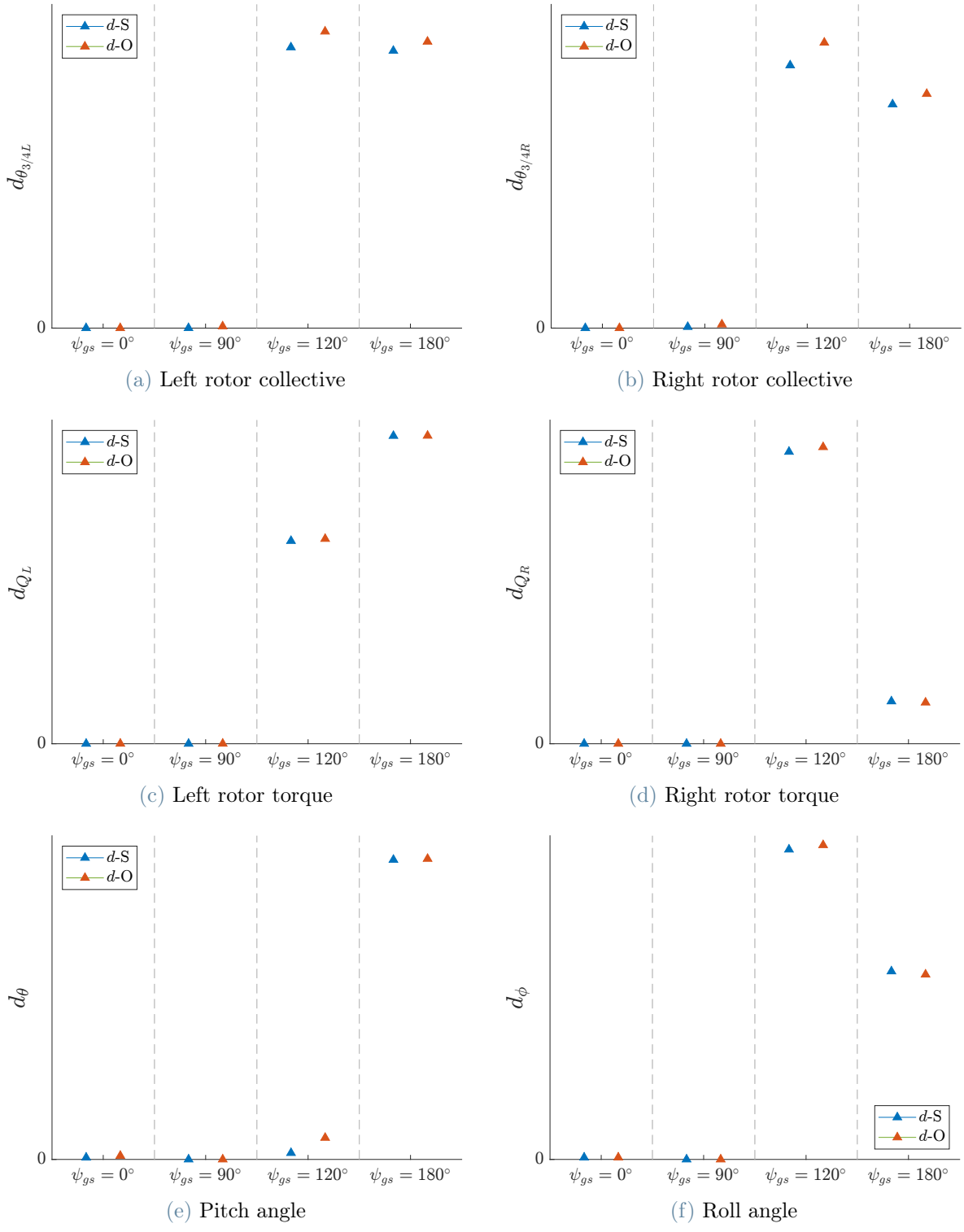
(c) Directional control position

Figure 6.31: Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for control positions on all validation flights

On the other hand, in fig. 6.30b, the experiment sits on the far right hand side of the pbox (corresponding to ~ 1.5 kts wind from an azimuth angle of 180°). Hence, evidence

of disagreement between the FSM prediction and the experiment can be identified. Still, the area metric, which accounts for a single SRQ (and provide no uncertainty or range associated to its value, unlike E of ASME) fails to draw the attention of the analyst to this evidence.

The differences between the two approaches for the computation of the pbox, as reported in fig. 6.31 and fig. 6.32, little disagreement is generally observed. Thus, for the sake of the advantages provided by the optimization-based approach in this particular application, only its results are reported in section 6.3.6, where the different validation metrics are compared.



(a) Left rotor collective

(b) Right rotor collective

(c) Left rotor torque

(d) Right rotor torque

(e) Pitch angle

(f) Roll angle

Figure 6.32: Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for aircraft attitudes rotors collective and torque figures on all validation flights

6.3.5. Interval Analysis Solution Validation

Considering the validation procedures presented in chapter 4, it is clear that the computation of u_{val} (in ASME framework) and the construction of the probability box (in Roy-Oberkampf framework) are costly tasks. Indeed, as reported in [8], Latin Hypercube Sampling (LHS) may be a lower cost alternative to the Monte Carlo sampling for the computation of u_{val}, but this advantage comes at the expense of a reduced confidence on the final value of the validation uncertainty itself. Moreover, an additional feature of both these UQ procedures is that they rely on sampling algorithms which, regardless the number of samples, carry a sampling error which in practice may be difficult (or expensive) to assess.
At the same time, building on the final observations carried out in sections 4.1.4 and 4.2.4, it may be argued that the computation of SRQs statistics[18] may not be strictly necessary for some possible applications of the validation process within RoCS framework. Indeed, whenever model validation is exploited within the model update and tuning loop or it is used, in preliminary phases, to quantify the maximum allowable input and experimental uncertainties to achieve a prescribed margin of model fidelity with sufficient confidence, the low computational expense and the simplicity of the procedure becomes critical. It is in this scenario, hence, that interval analysis might stand out as a viable tool to overcome the aforementioned limitations of sampling techniques and to support the analysis in the course of the validation campaign planning and model update iterations.

Methodology

Formally speaking, such a framework incorporates elements of both the ASME standard (validation metrics and corollaries) and the Roy-Oberkampf approach (interval-valued uncertain quantities). All model inputs affected by uncertainty are characterized as intervals[19]. Then, the propagation of input uncertainties through the model is performed via global optimization (i.e. interval analysis) techniques. As a result, bounding minimum and maximum values for S are obtained. Then, the resulting interval of possible values of S is treated as an epistemic uncertainty on the model output itself, referred to as U_{INPUT}. Concerning the numerical uncertainty, no difference with respect to Roy-Oberkampf framework is present. Thus, it is treated as an epistemic uncertainty whose derivation follows the guidelines discussed in section 6.3.4. Finally, concerning the experiment, the realization(s) of the SRQs of interest are used to derive an interval U_D

[18] and comparison error statistics, in the case of ASME approach.
[19] thus, with no difference with respect to the mathematical description characterizing epistemic, interval-valued input uncertainties within the Roy-Oberkampf framework.

which bounds all possible contemplated values of the measurand D. If the time history of the SRQ is available, the bounds of the interval may be defined basing upon a certain confidence range on the measurand (e.g. 95%). On the other hand, if the analyst believes that the number of realization (or the quality of the experimental datum) is insufficient to estimate a reasonably accurate confidence interval, the minimum and maximum values among the realizations at disposal can be used.

Then, upon the quantification of U_{INPUT}, U_{NUM} and U_D, the validation metrics can be evaluated. An outline of the process is reported in fig. 6.33.

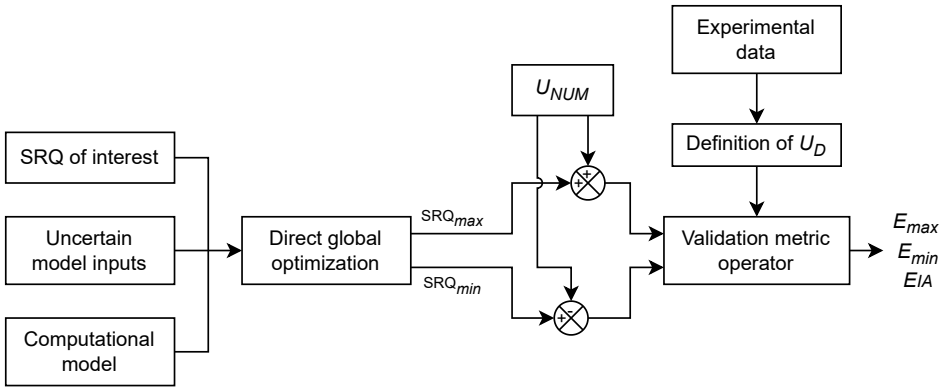


Figure 6.33: Flow chart of the validation metric quantification process within the framework of interval analysis

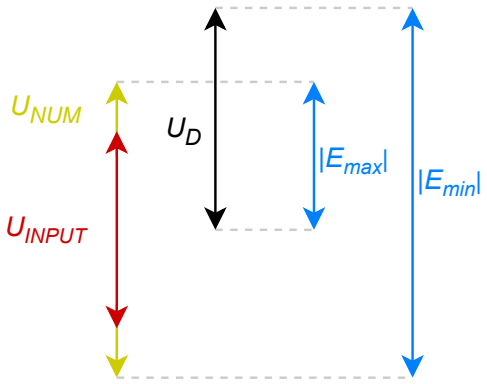


Figure 6.34: Graphical exemplification of the computation of E_{min} and E_{max} validation metrics

$$E_{min} = \min(S\text{-}D) \tag{6.22}$$

$$E_{max} = \max(S\text{-}D) \tag{6.23}$$

$$E_{IA} = E_{max} - E_{min} \tag{6.24}$$

At first, the numerical uncertainty U_{NUM} is used to augment the interval associated to

input uncertainties, as displayed in fig. 6.33. Then, the augmented interval is combined with U_D according to the rules of interval analysis, namely eq. (6.22) and (6.23), in order to derive the upper and lower bound of the comparison error. These latter quantities, together with their difference reported in eq. (6.24), constitute the validation metrics for this framework.

Upon computation of the validation metrics, corollaries based on the ones discussed in section 4.1.2 can be derived.

- If both $|E_{max}| >> E_{IA}$ and $|E_{min}| >> E_{IA}$, then it is safe to assume that δ_{model} has the same order of magnitude of E_{max} and E_{min} and that its estimation is reliable;
- if either $|E_{max}| \lesssim E_{IA}$ or $|E_{min}| \lesssim E_{IA}$, then the comparison error is dominated by either numerical, input and/or experimental uncertainties and little information can be retrieved and used about the modelling error value.

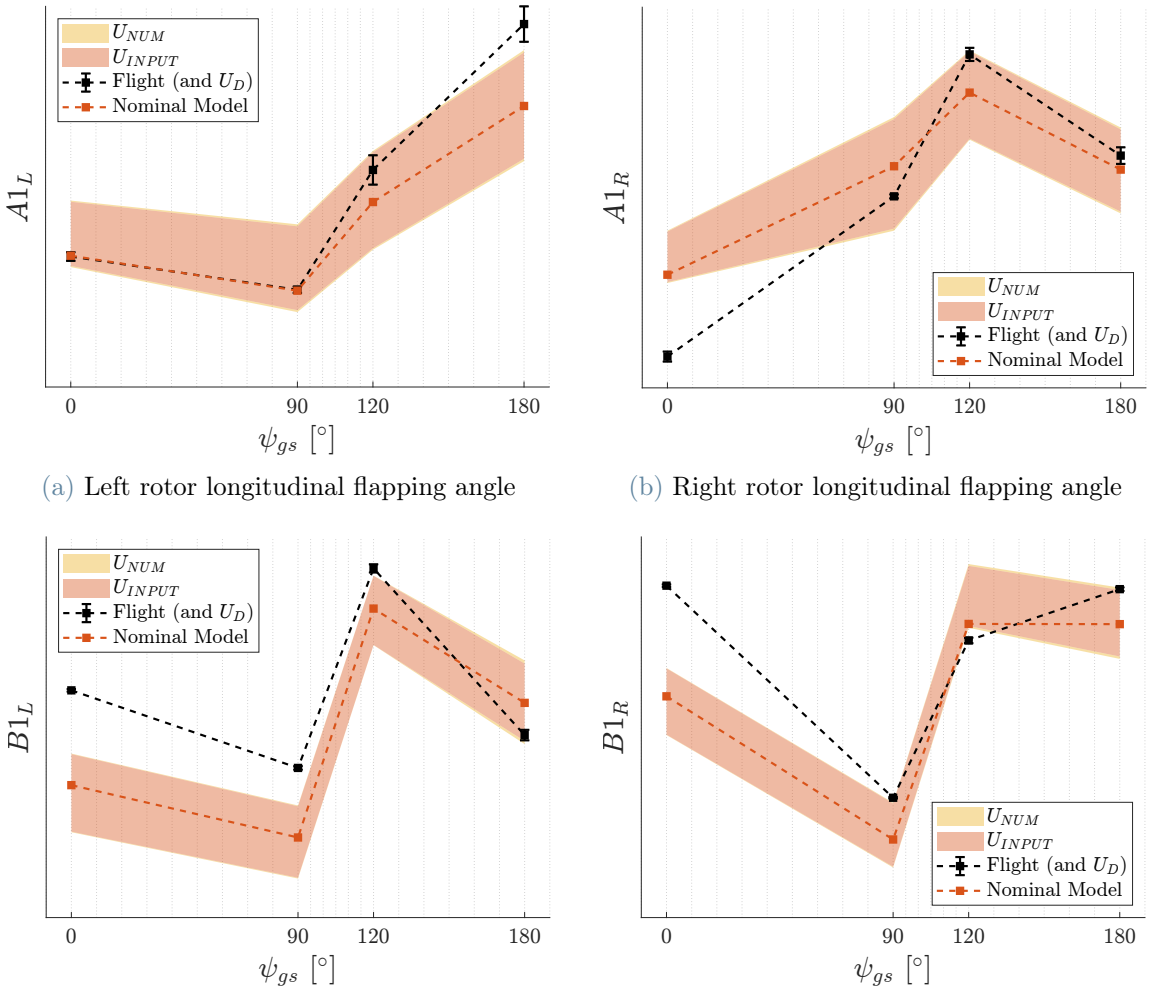
Fidelity assessments and model improvements can be measured directly on E_{max} and E_{min}.

Concerning model adequacy assessment, the uncertainty on the model predictive capability in the DoP can be approximated by combining the appropriately extrapolated model error $\delta_{model} \subset [E_{min}; E_{max}]$ together with U_{INPUT} and U_{NUM} evaluated at the prediction point of interest, according to eq. (6.25).

$$U_{PRED} = (SRQ_{max} + U_{NUM} + E_{max}) - (SRQ_{min} - U_{NUM} + E_{min}) \tag{6.25}$$

Results

The validation procedure presented above was carried out on the present test case. The global optimization of each SRQ of interest with respect to the uncertain inputs was performed with Dakota. Moreover, the unbounded model inputs (namely, ACW, STA and BL) were bounded within three standard deviations from their mean, thus identifying the corresponding intervals with the 99.73% confidence interval. Concerning the experimental uncertainty, in this instance, U_D was centered in D (computed according to eq. (6.18)), with an amplitude equal to three standard deviations of the measurand. However, as previously mentioned, the analyst may come up with other techniques for its definition basing on the number of samples available and its knowledge about the presence of systematic errors.



(a) Left rotor longitudinal flapping angle

(b) Right rotor longitudinal flapping angle

(c) Left rotor lateral flapping angle

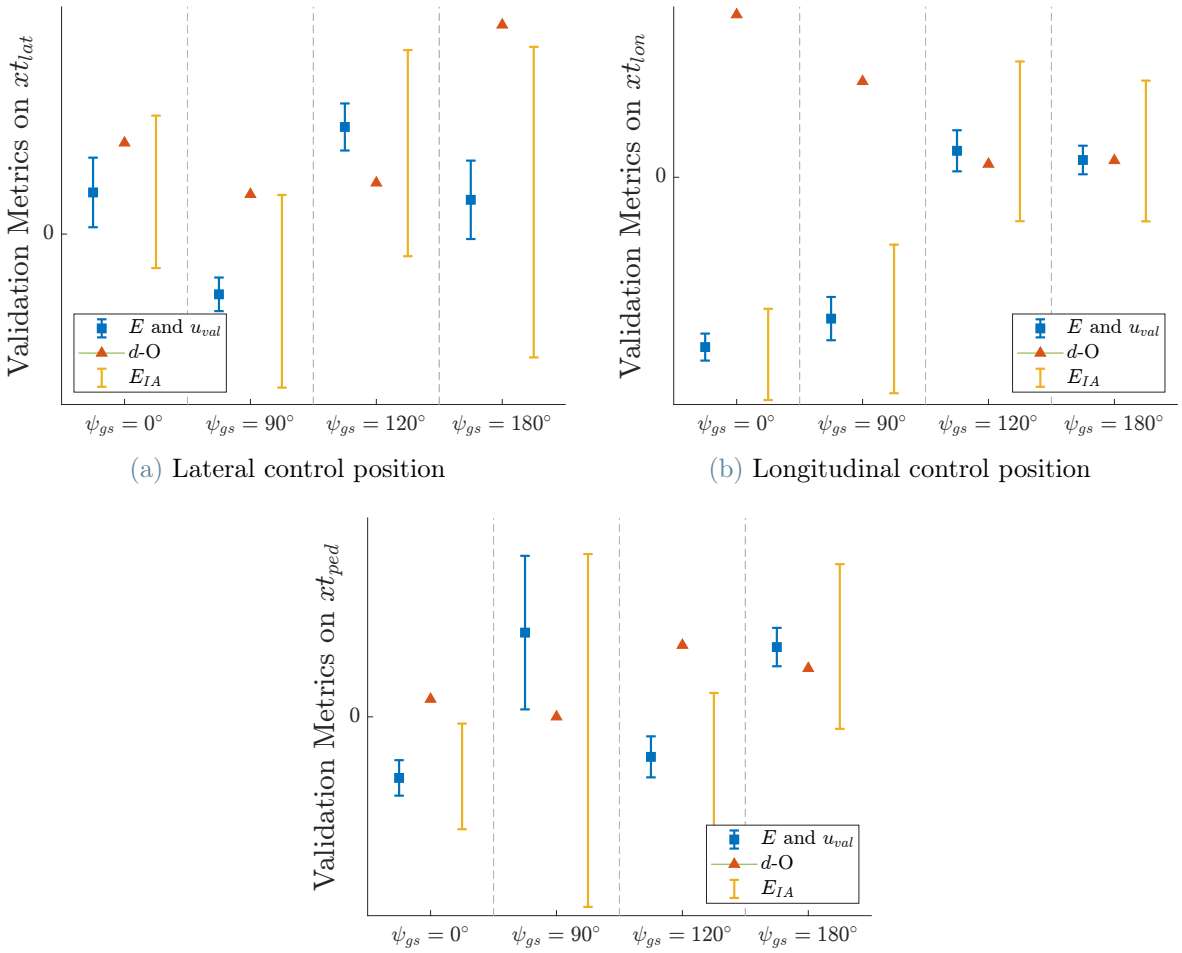
(d) Right rotor lateral flapping angle

Figure 6.35: Results of the uncertainty propagation and aggregation on FSM prediction and flight data for rotors flapping angles

The result of input uncertainty propagation are reported in appendix D, together with its aggregation with the numerical uncertainty and experimental data. Fig 6.35 reports the results of the procedure presented in section 6.3.5 on rotors flapping angles. Straightaway, this result highlights that this approach makes it possible, in the present application, to carry out sensible conclusions regarding the validation of rotor flapping angles. Moreover, as already stressed in section 6.3.3, input uncertainty U_{INPUT} turns out to be the dominant contribution to E_{IA} by far, with exceptions in a handful of instances only, where U_D dominates instead. The values of E_{max}, E_{min} and E_{IA} are reported in section 6.3.6, where they are compared with the other validation metrics.

6.3.6. Validation Metrics Comparison

In this section, the results obtained with the three aforementioned VV frameworks and metrics are compared. Moreover, observations regarding their comparison are drawn to propose when and why one method may be advantageous with respect to another. Comparison of the rotors flapping angles is proposed between ASME framework and interval analysis only, since in 6.3.4 evidence was provided that the finite-size sample on which d rely is not trustworthy for these fidelity metrics.



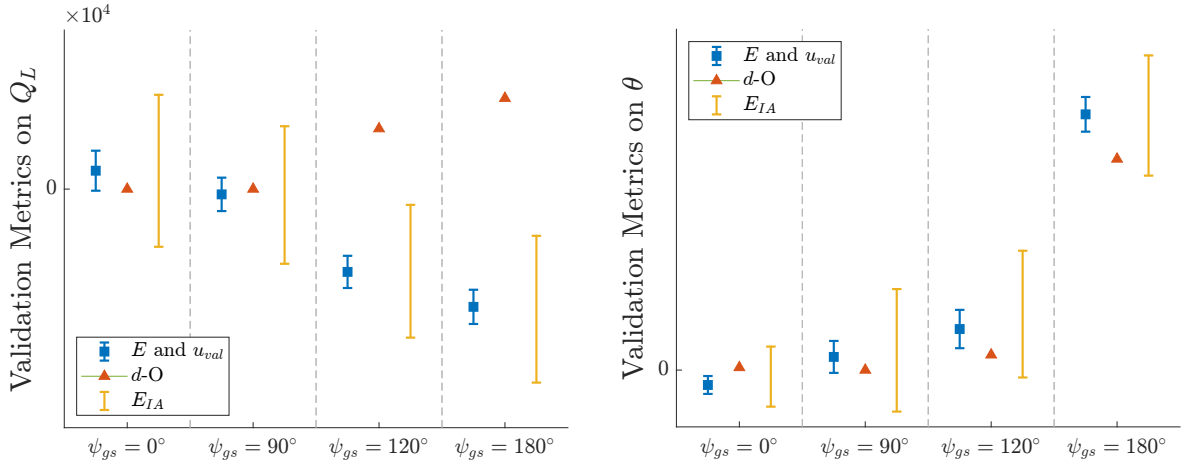
(a) Lateral control position

(b) Longitudinal control position

(c) Directional control position

Figure 6.36: Validation metrics comparison for control positions on all validation flights

Decent agreement in the trends and values of the metrics can be observed in fig. 6.36 and 6.37. Being both based on the comparison error, E_{IA} and $E \pm u_{val}$ are indeed very similar. The yellow band of E_{IA} always include the nominal comparison error E of ASME, as expected. Moreover, the amplitude of the yellow bar is typically of the same order of magnitude of u_{val} and, frequently, 2 to 3 times greater. This is expected considering



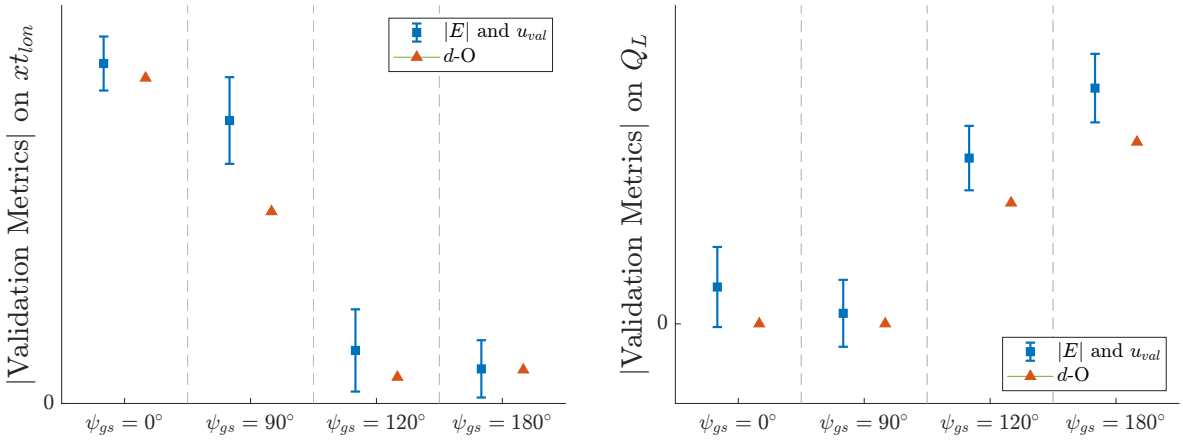
(a) Left rotor torque (b) Pitch angle

Figure 6.37: Validation metrics comparison for left rotor torque and aircraft pitch angle

that u_{val} is a standard uncertainty, while the E_{IA} is extended and directly represents the interval in which the comparison error is expected to fall. The area metric d assumes values which are comparable to the other metrics as well. As a consequence, it is reasonable that the fidelity acceptability margins conceived for the comparison error may also apply sensibly to this validation metrics. In addition, it also typically follows the same trends of ASME and interval analysis, with two types of exceptions. Whenever the comparison error assumes negative values, considering that the area metric is a positive definite quantity, a difference in the trends emerges, characterized by a pseudo-symmetry with respect to the X-axis of the plot. Thus, whenever E carries the information about the sign of the error, the area metric doesn't. However, if there is any[20], this can always be retrieved by the analyst through a visual inspection of the pbox and the area metric. Indeed, this condition is the one that occurs in this study every time d and E-based metrics follow diverging trends (e.g. fig. 6.37a). To support this conclusion, the comparison error from ASME and the area metric from Roy-Oberkampf are compared in magnitude in fig. 6.38.

Alike E and u_{val}, also $|E_{max}|$, $|E_{min}|$ and E_{IA}, are typically of the same order of magnitude. Hence, when the corollaries are applied to interval analysis validation metrics, the analyst is supplied with compelled evidence that the uncertainty due to input and flight data is too big to isolate the value of the model error. As a consequence, a reduction of the input and experimental uncertainty is necessary in order to set up well-advised model improvements based on these data, unless the applicant is willing to accept a large uncertainty.

[20] e.g. the experiment CDF is always on the same side of the FSM pbox, and so a "sign" can be intuitively associated to the error.



(a) Longitudinal control position (b) Left rotor torque

Figure 6.38: Comparison of $|E| \pm u_{val}$ and d

Concerning the area metric, it is straightaway clear that no model improvement can be formulated basing on the instances where $d = 0$. On the other hand, depending on the quality of flight data (as mentioned in section 6.3.4), model update procedures may be conceived for the cases where $d \neq 0$. Nevertheless, as observed in section 4.2.4, d does not come with any associated uncertainty (unlike E in the other two frameworks). When the model is extrapolated and the model-form error estimate is exploited in the DoP for certification-aimed prediction or credibility assessments, any information concerning the uncertainties involved in the validation phase is lost. As a consequence, another fundamental difference among the frameworks emerges in regard of model adequacy assessment.

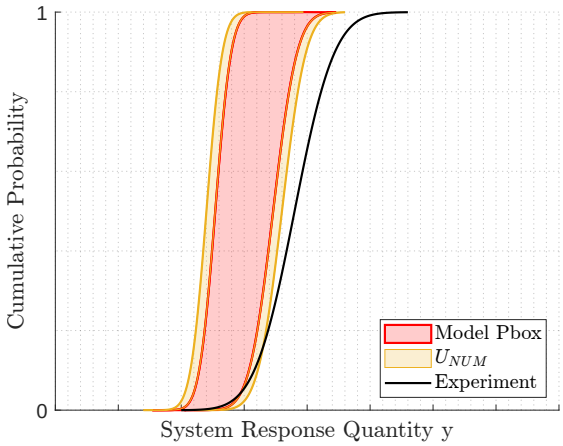
Within the validation methodologies based either on ASME standard or interval analysis, any increase in the uncertainties involved in the model validation phase ends up worsening model credibility. The validation metrics u_{val} and E_{IA} not only draw the attention of the analyst to large uncertainties involved in the validation process but also, and more importantly, are able to transport that information (related to the *goodness* of the model validation) into the DoP. Essentially, any uncertainty associated to model-inputs (u_{input} or U_{INPUT}), numerical approximations (u_{num} or U_{NUM}) and flight measurements (u_d or U_D) in the validation points indirectly affects the model credibility in the DoP through the model-form error uncertainty (either represented by u_{val} or E_{IA}). Hence, the credibility of a FSM is not only affected by the comparison error E and the fidelity of the model itself, but also by the *quality* of the validation which has been carried out to assess its fidelity. As a result, the very same model validated with reduced uncertainties in the validation data can lead to a smaller margin with the same CR when used to support certification (thanks to a reduction on u_{val} or E_{IA}).

Conversely, whenever d is used, this connection does not always emerge. Indeed, as the analyst inadvertently exaggerate either the epistemic input uncertainty or U_{NUM} during model validation, the value of d reduces and, as a consequence, recalling eq. (4.9), model credibility improves. An exemplification of this view is presented and discussed in fig. 6.39, where the increase of model credibility as a consequence of larger input uncertainties considered during the validation phase is displayed.
Thus, apparently, it is possible to conclude that whenever d is used, it becomes critical not to overestimate epistemic uncertainties. However, this might not be trivial at all, considering the fact that epistemic input uncertainties may include inputs of which little information is available and accounting for the difficulties associated to the estimation of U_{NUM} (as mentioned in section 6.3 and [29]). On the other hand, $E \pm u_{val}$ and interval analysis methods may be considered more resilient in this regard, since, at least, they make available to the analyst an indication of the uncertainty associated to the model-form error.

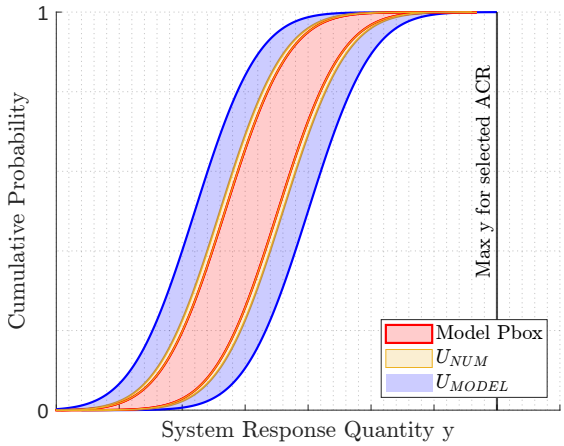
In regard of uncertainty contributions decomposition, all validation frameworks are able to inform the analyst about what uncertainties dominate model predictions and/or validation metrics. Indeed, within the ASME framework, u_{input}, u_{num} and u_D can be compared to figure out the greatest contribution to u_{val}, as shown in section 6.3.3, and to the simulation error δ_S in general. On the other hand, within the approach proposed by Roy and Oberkampf, the size and shape of the pbox, together with the magnitude of U_{NUM} can be exploited to decide what is the most effective way to reduce model prediction uncertainty, as highlighted in section 6.3.4. Finally, in the framework of interval analysis, U_{INPUT}, U_{NUM} and U_D can be compared, as mentioned in section 6.3.5.

Another way of comparing the validation metrics relies on their computational cost. Indeed, in the present study, u_{val} and d turned out to be the most computationally intensive metrics to compute, due to the sampling techniques required for the estimation of u_{input} and of the model probability box, respectively. On the other hand, at the cost of sacrificing the information on SRQs statistics, interval analysis proved to be much cheaper, resulting in at least an order of magnitude reduction in terms of required function evaluations per fidelity metrics. As it can be deduced from the flowchart presented in fig. 6.33, the computational cost of E_{max}, E_{min} and E_{IA} scales up with the number of fidelity metrics involved in the study (since a dedicated global optimization is required for each SRQ individually). On the other hand, the cost of u_{val} and d, in principle, is independent from the number of fidelity metrics involved in model validation. However, as observed in sections 6.3.3 and 6.3.4 with rotor flapping angles, their cost is bounded above the number of function evaluations required to obtained satisfactory convergence of the statistics of

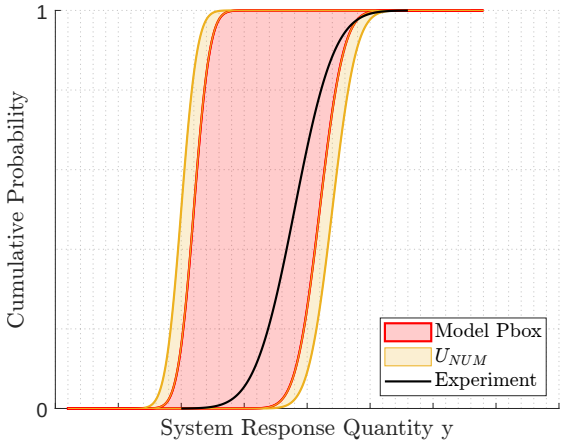
the *latest-converging* fidelity metric. Thus, in a practical application, the analyst may find that the validation cost depends on what fidelity metrics are involved in the study as well.



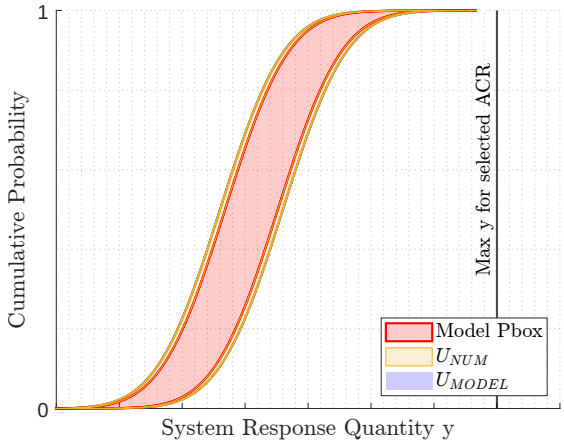
(a) Example of model validation outcome in a validation point V ∈ DoV



(b) Example of credibility assessment outcome in the prediction point P ∈ DoP. The model-form uncertainty on the model prediction is retrieved from the model validation summarized in fig. 6.39a



(c) Example of model validation outcome in a validation point V ∈ DoV. The reducible model input uncertainty and the numerical uncertainty have been inadvertently overestimated with respect to fig. 6.39a



(d) Example of credibility assessment outcome in the prediction point P ∈ DoP. The model-form uncertainty on the model prediction is retrieved from the model validation summarized in fig. 6.39c. The overestimation of the model prediction epistemic uncertainty in the validation phase (resulting in $d = 0$) causes an improvement in model credibility (measured according to the CR)

Figure 6.39: Examples of model validation and credibility assessment activities within the Roy-Oberkampf framework. No extrapolation uncertainty is considered affecting U_{MODEL}

7 | Conclusions

In the present work, the ASME VV 20 [8] and Roy-Oberkampf [29] approaches to VV of computational models are revised. Then, emulating the first phases of a partial credit demonstration of compliance to EASA low speed controllability requirements [7], the aforementioned VV methodologies have been applied to a state-of-the-art FSM developed by LHD. In the course of the application, the CbS guidelines developed by RoCS are strictly followed. FLIGHTLAB and Dakota are employed as tools for flight dynamics modelling and SA/UQ, respectively.

The solution verification process is carried out accounting for several model discretization and solution algorithm parameters. At the time of writing, there is no accepted solution verification procedure suited for the quantification of the numerical error of flight mechanics multi-body models. Hence, the author applies a verification algorithm specifically conceived for the present application, stressing the assumptions, challenges and limits of the currently adopted procedures when applied to complex state-of-the-art multi-body systems. Valuable insight is gained on the behaviour of the Peters-He finite state dynamic inflow model with different number of states. In addition, reliable estimates of the numerical error associated to the FSM solution are obtained.

In the framework of model validation, a preliminary SA with respect to FSM inputs is carried out with the MOAT [9] method implemented in Dakota. Despite a direct comparison with more widely adopted SA methods (e.g. VBD) is not provided, it is predicted that the reduced cost of MOAT may result in a significant reduction of the computational expense of SAs in FSMs with a large number of uncertain input parameters. Model validation is then carried out with both ASME VV 20 and Roy-Oberkampf approaches. Moreover, a validation methodology based on interval analysis is proposed and applied to the FSM. The three VV procedures are eventually compared in terms of fundamental assumptions, computational expense, validation metric values and suitability to the CbS process developed by RoCS.

ASME and Roy-Oberkampf validation procedures, as expected, proved to be significantly more expensive than interval analysis. Despite this, all validation methodologies lead to

similar values of model validation metrics for all SRQs. The VV framework proposed by ASME for CFD and CHT [8] may be best suited for the application to FSMs in RoCS. Indeed, its corollaries provide a powerful tool at disposal of the analyst to critically interpret the outcomes of model validation. Moreover, the standard validation uncertainty u_{val} allows model credibility in the DoP to be affected by *quality* of the FSM model validation (namely, the magnitude of the uncertainties accounted for during validation). This is a powerful feature, especially when, as in the present work, FSM input uncertainties cannot be accurately estimated. Nevertheless, when computational expense is of paramount importance, e.g. when a preliminary guess of affordable FSM input uncertainties shall be provided, interval analysis might turn out to be best suited, thanks to its capability of delivering great proxies of the comparison error E uncertainty band obtained with the ASME approach at a fraction of the cost.

Finally, it is important to stress that, despite the VV methodology, a solid estimation of FSM input parameters is necessary to set up a well advised validation process and reinforce model credibility. Moreover, all validation procedures, as they are presented in chapter 4 and 6.3.5, are well suited for scalar, single valued validation metrics (i.e. SRQs at trim). Nevertheless, their practical implementation presented in this work shall at least be revised in order to make it suitable to different kinds of fidelity metrics, such as non-linear time responses or frequency responses.

Bibliography

[1] *https://www.flightlab.com/*. ADVANCED ROTORCRAFT TECHNOLOGY.

[2] S. A. *Making best use of model valuations to compute sensitivity indices*. Computer Physics Communications, 2002.

[3] B. M. Adams, W. J. Bohnhoff, K. R. Dalbey, M. S. Ebeida, J. P. Eddy, M. S. Eldred, R. W. Hooper, P. D. Hough, K. T. Hu, J. D. J. M. Khalil, K. A. Maupin, J. A. Monschke, E. M. Ridgway, A. A. Rushdi, D. ThomasSeidl, J. A. Stephens, L. P. Swiler, A. Tran, and J. G. Winokur. *Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.16 User's Manual*. Sandia National Laboratories, 2022.

[4] AIAA. *Recommended Practice: When Flight Testing Modelling is Used to Reduce Flight Testing Supporting Aircraft Certification*. 2021.

[5] Anonymous. *FAR Part 29 Advisory Circular 29-2C*. U.S. Department of Transportation - Federal Aviation Administration (FAA), 2018.

[6] Anonymoys. *Certification Specifications for Helicopter Flight Simulation Training Devices 'CS-FSTD(H)'*. European Aviation Safety Agency (EASA), 2012.

[7] Anonymoys. *Certification Specifications, Acceptable Means of Compliance and Guidance Material for Large Rotorcraft (CS-29) - Amendment 11*. European Aviation Safety Agency (EASA), 2023.

[8] ASME. *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*. 2009.

[9] F. Campolongo and J. Cariboni. *Sensitivity Analysis: How to Detect Important Factors in Large Models*. Institute for the Protection and Security of the Citizen (IPSC), EASF European Commission, 2007.

[10] F. Campolongo, J. Cariboni, and A. Saltelli. *An Effective Screening Design for*

Sensitivity Analysis of Large Models. Institute for the Protection and Security of the Citizen (IPSC), EASF European Commission, 2006.

[11] K. R. Dalbey, M. S. Eldred, G. Geraci, J. D. Jakeman, K. A. Maupin, J. A. Monschke, D. ThomasSeidl, L. P. Swiler, A. Tran, F. Menhorn, and X. Zeng. *Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.16 Theory Manual.* Sandia National Laboratories, 2022.

[12] R. DuVal and C. He. *Validation of the FLIGHTLAB virtual engineering toolset.* The Aeronautical Journal, 2018.

[13] S. Ferson and L. Ginzburg. *Different Methods are Needed to Propagate Ignorance and Variability.* Reliability Engineering and System Safety, 1996.

[14] S. Ferson and J. G. Hajagos. *Arithmetic with Uncertain Numbers: Rigorous and (Often) Best Possible Answers.* Reliability Engineering and System Safety, 2004.

[15] S. Ferson, W. L. Oberkampf, and L. Ginzburg. *Model Validation and Predictive Capability for the Thermal Challenge Problem.* Computer Methods in Applied Mechanics and Engineering, 2008.

[16] J. C. for Guides in Metrology (JCGM). *Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement.* 2010.

[17] I. O. for Standardization (ISO). *International Vocabulary of Basic and General Terms in Metrology - 2nd Edition.* 1993.

[18] V. Gaioni. *A Practical Verification and Validation Procedure for Computational Fluid Dynamics.* Universidad Politecnica di Madrid, 2020.

[19] C. He. *Development and Application of Generalized Dynamic Wake Theory for Lifting Rotors.* PhD thesis, Georgia Institute of Technology, 1989.

[20] C. He. *Finite State Dynamic Wake Interference Modeling for Rotorcraft Simulation.* PhD thesis, Advanced Rotorcraft Technology (ART), 1997.

[21] M. D. Morris. *Factorial Sampling Plans for Preliminary Computational Experiments.* Technometrics, 1991.

[22] G. D. Padfield, S. van't Hoff, L. Lu, P. Podzus, M. White, and G. Quaranta. *Preliminary Guidelines for the Rotorcraft Certification by Simulation Process.* Politecnico di Milano, 2023.

[23] D. A. Peters, C. He, and A. Su. *A Closed-Form, Finite-State Model for the Unsteady Aerodynamics of Rotors.* Computational Mechanics, 1988.

[24] L. F. Richardson. *The Approximate Arithmetical Solution by Finite Differences of Physical Problems Involving Differential Equations, with an Application to the Stresses in a Masonry Dam.* Philosophical Transacrions of the Royal Society A. Mathematical, Physical and Engineering Sciences, 1911.

[25] P. J. Roache. *Perspective: A Method for Uniform Reporting of Grid Refinement Studies.* Journal of Fluids Engineering, 1994.

[26] P. J. Roache. *Conservativism of the Grid Convergence Index in Finite Volume Computations on Steady-State Fluid Flow and Heat Transfer.* Journal of Fluids Engineering, 2003.

[27] P. J. Roache and P. M. Knupp. *Completed Richardson Extrapolation.* Communications in Numerical Methods in Engineering, 1993.

[28] C. J. Roy and W. L. Oberkampf. *Verification and Validation in Scientific Computing.* Cambridge University Press, 2010.

[29] C. J. Roy and W. L. Oberkampf. *A complete Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing.* 48th AIAA Aerospace Sciences Meeting, 2010.

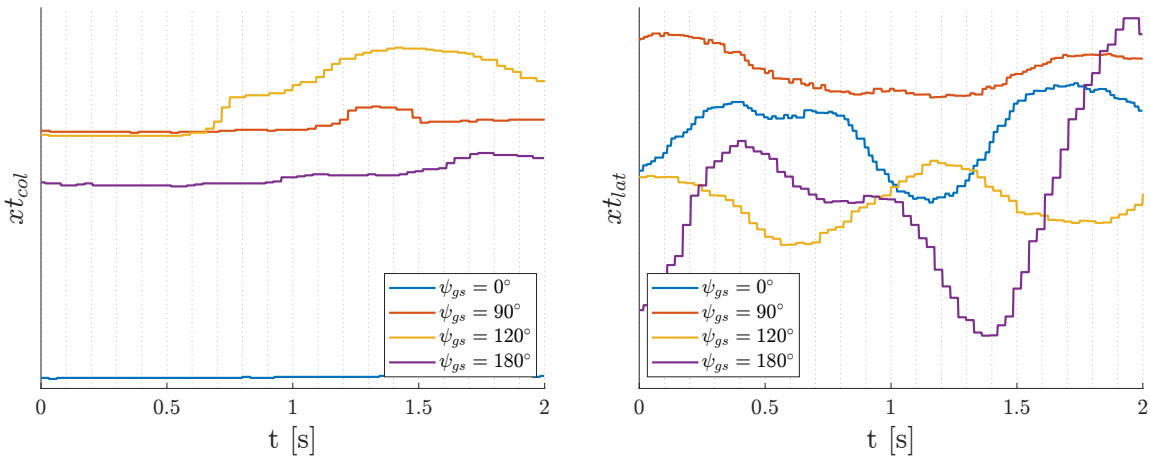
[30] J. Schaefer, V. Romero, S. Schafer, B. Leyde, and C. Denham. *Approaches for Quantifying Uncertainties in Computational Modeling for Aerospace Applications.* AIAA Scitech 2020 Forum, Orlando, FL, 2020.

[31] F. Stern and H. W. Coleman. *Uncertainties and CFD Code Validation.* Journal of Fluids Engineering, 1997.

[32] S. van't Hoff, L. Lu, M. Jones, M. White, and G. Quaranta. *Preliminary Set of Metrics for Quantifying Flight-Model Fidelity.* Politecnico di Milano, 2020.

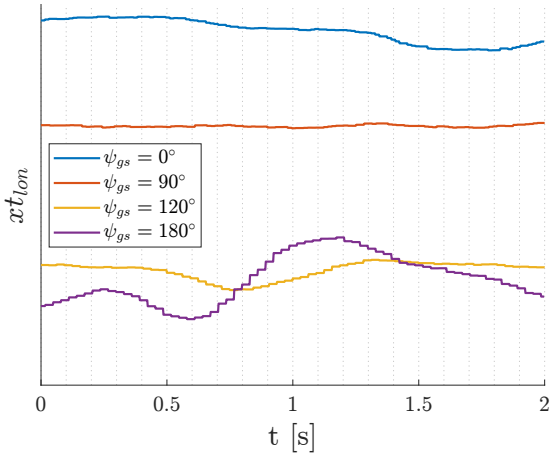
[33] S. vant't Hoff. *ACR: Low-Speed Controllability.* RoCS Workshop, 2022.

A | Flight Data Time Histories

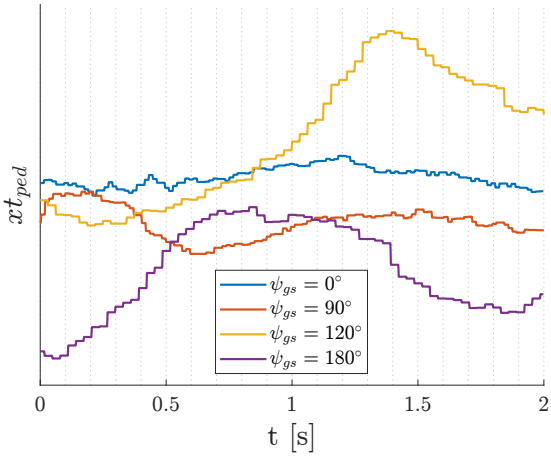


(a) Collective control position

(b) Lateral control position

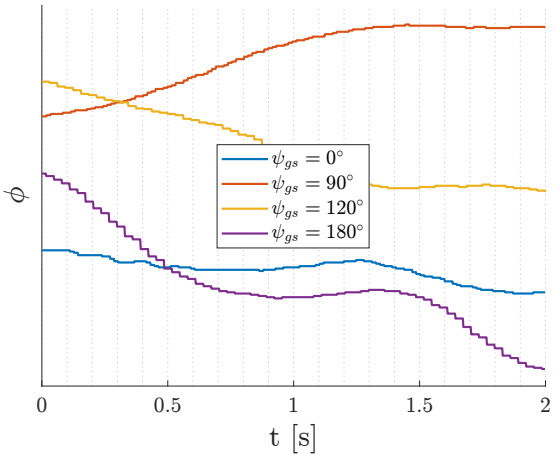


(c) Longitudinal control position

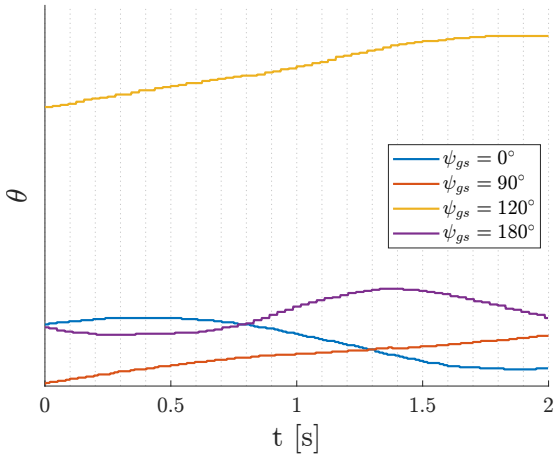


(d) Directional control position

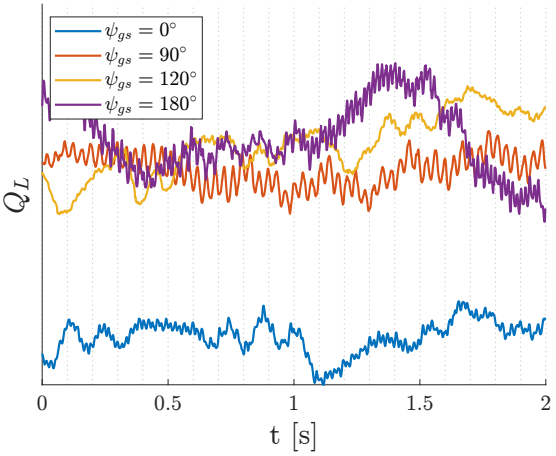
Figure A.1: Total controls time histories from flight tests



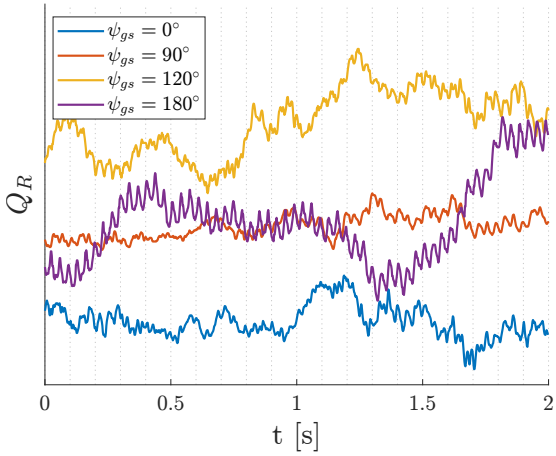
(a) Roll angle



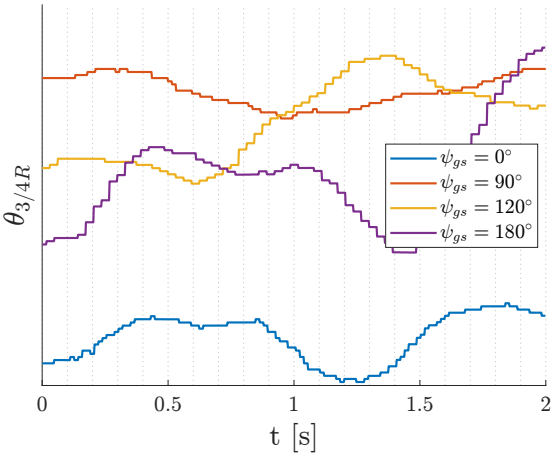
(b) Pitch angle



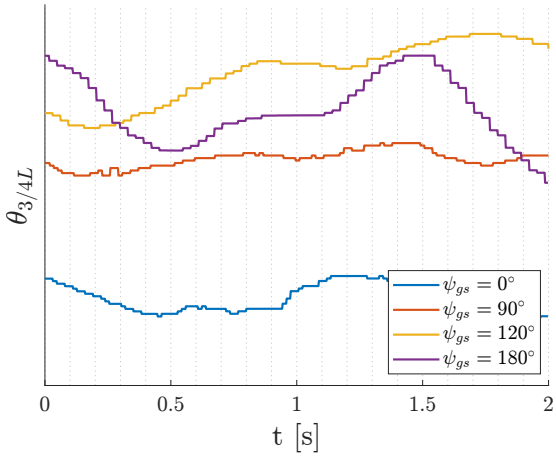
(c) Left rotor torque



(d) Right rotor torque

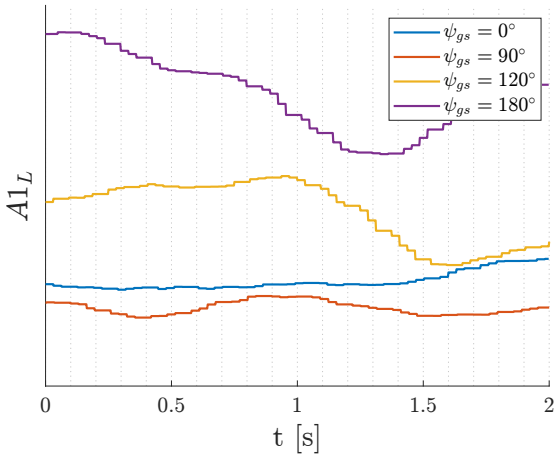


(e) Left rotor collective

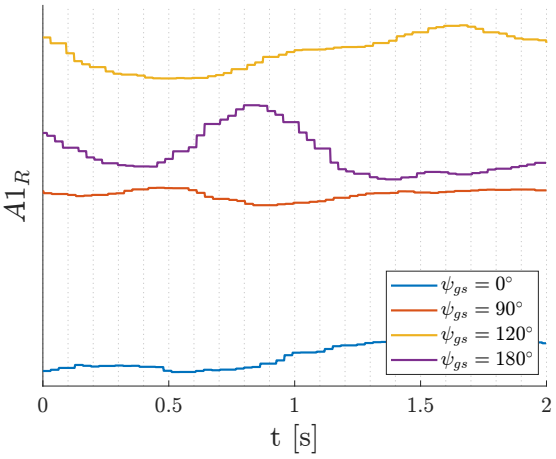


(f) Right rotor collective time history

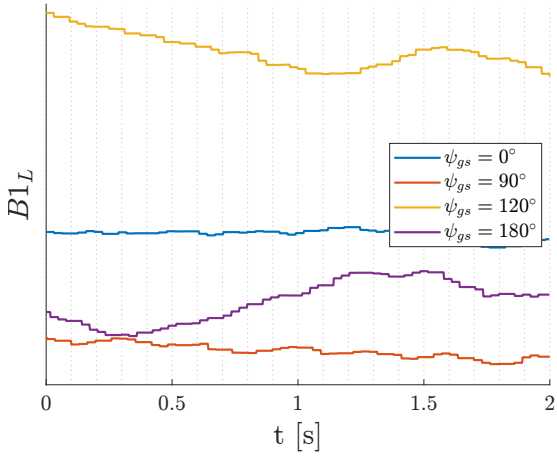
Figure A.2: Aircraft attitudes and rotors collectives and torque figures time histories from flight tests



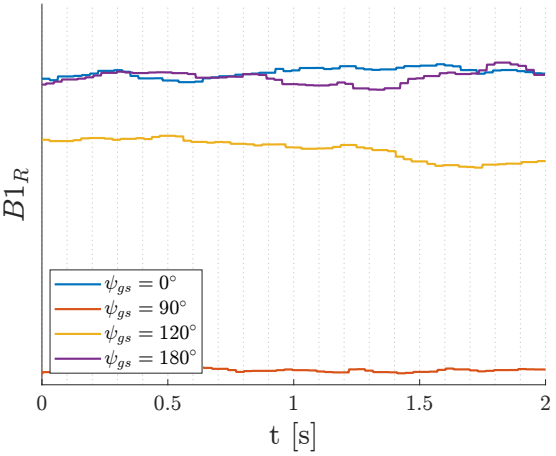
(a) Left rotor longitudinal flapping angle



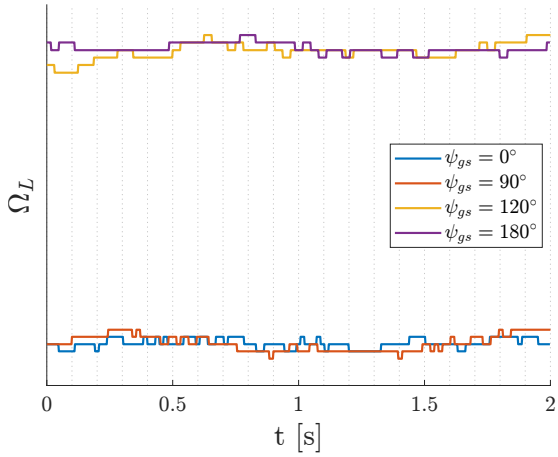
(b) Right rotor longitudinal flapping angle



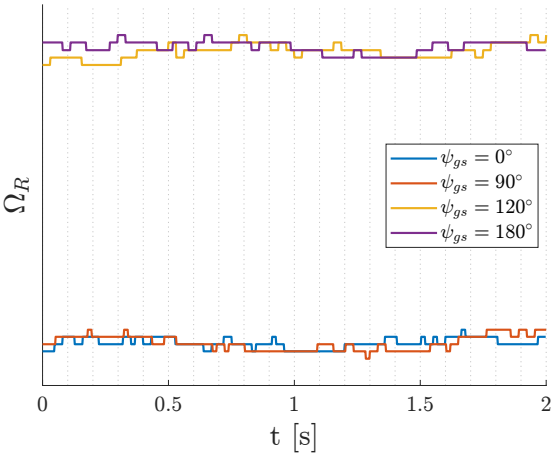
(c) Left rotor lateral flapping angle



(d) Right rotor lateral flapping angle

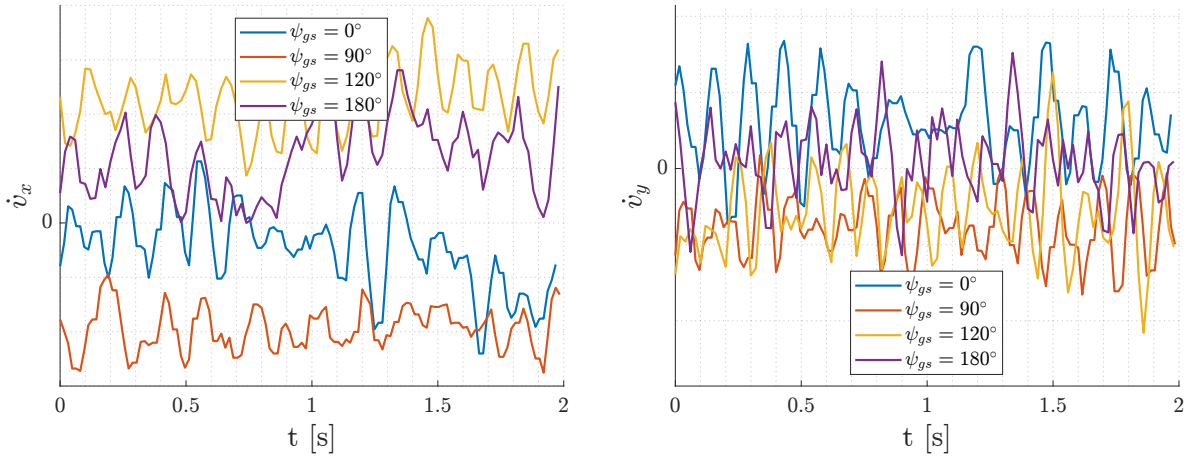


(e) Left rotor RPM



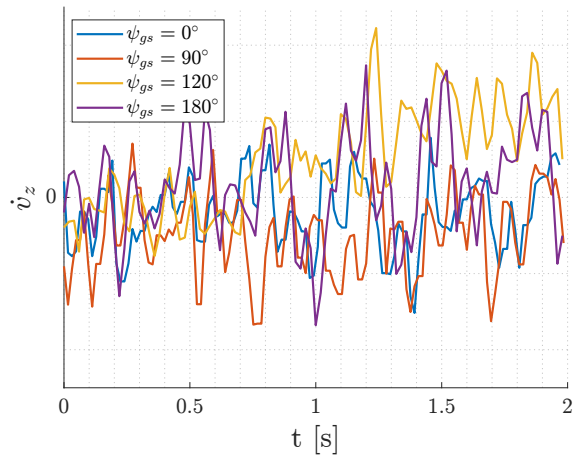
(f) Right rotor RPM

Figure A.3: Rotors flapping angles and RPM time histories from flight tests

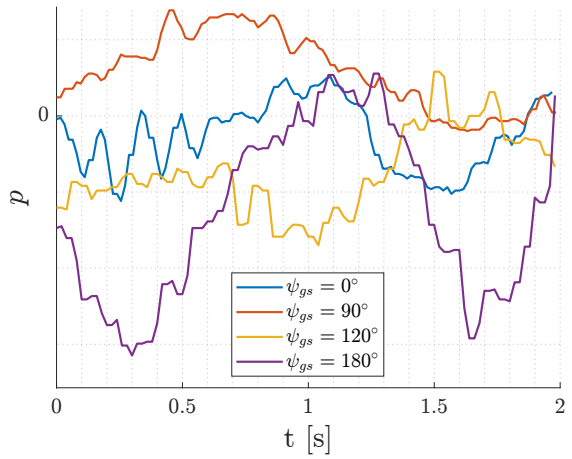


(a) Longitudinal acceleration

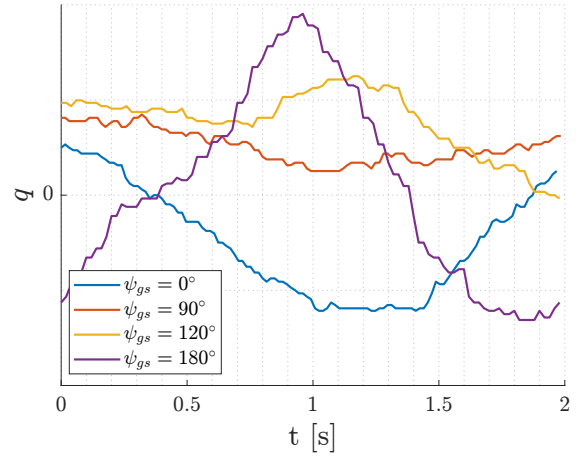
(b) Lateral acceleration



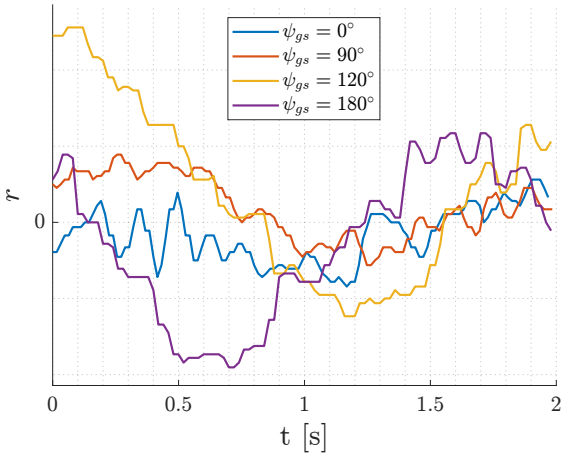
(c) Vertical acceleration



(d) Roll rate



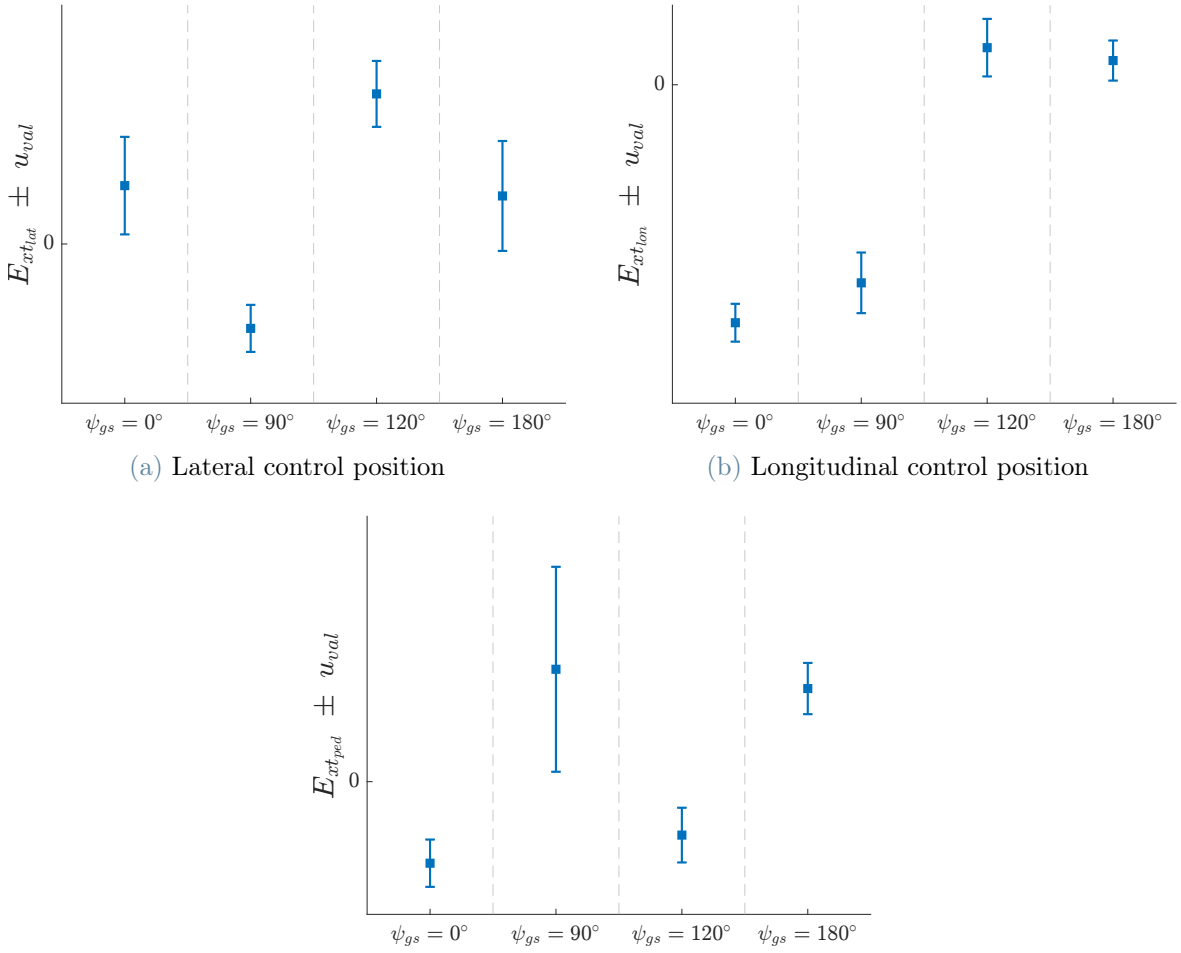
(e) Pitch rate



(f) Yaw rate

Figure A.4: Aircraft (body-axes) linear accelerations and angular rates time histories

B | ASME VV Standard 20 - Validation Results

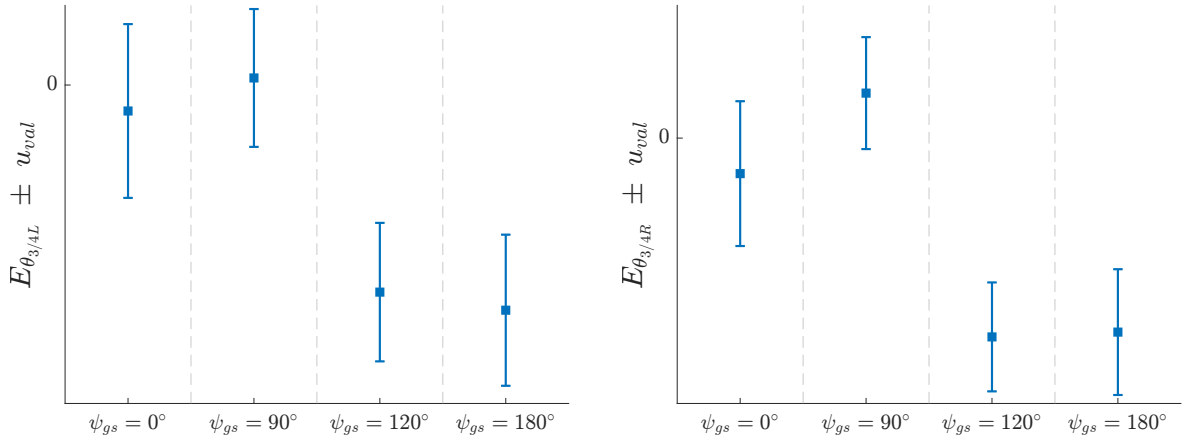


(a) Lateral control position

(b) Longitudinal control position

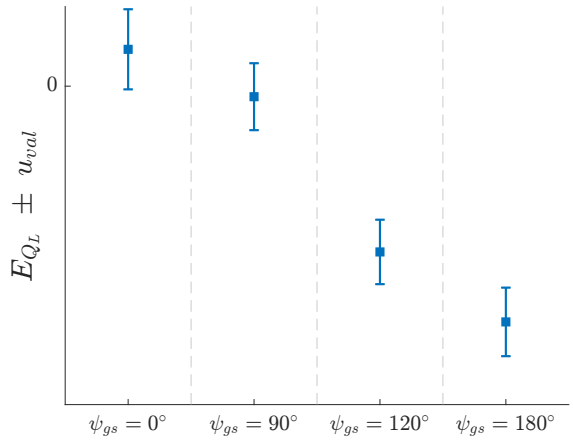
(c) Directional control position

Figure B.1: E and u_{val} validation metrics results for control positions on all validation flights. The magnitude of u_{val} is reported as an errorbar on E

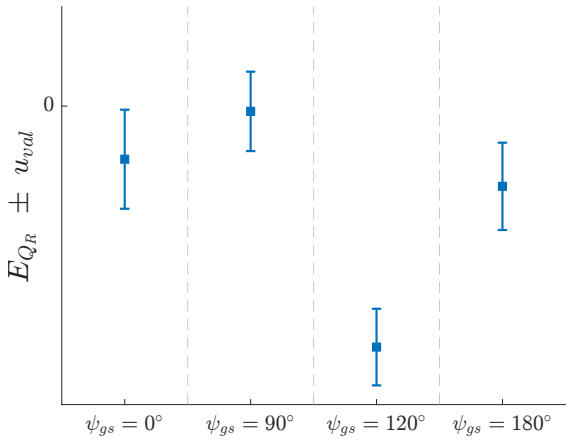


(a) Left rotor collective

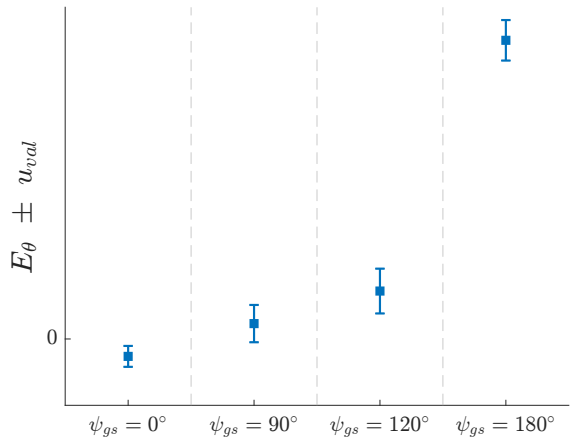
(b) Right rotor collective



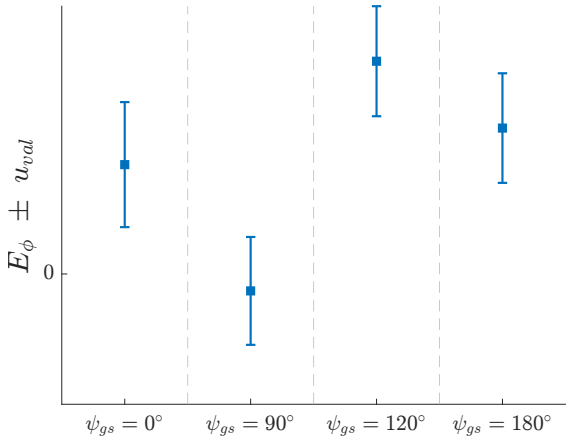
(c) Left rotor torque



(d) Right rotor torque

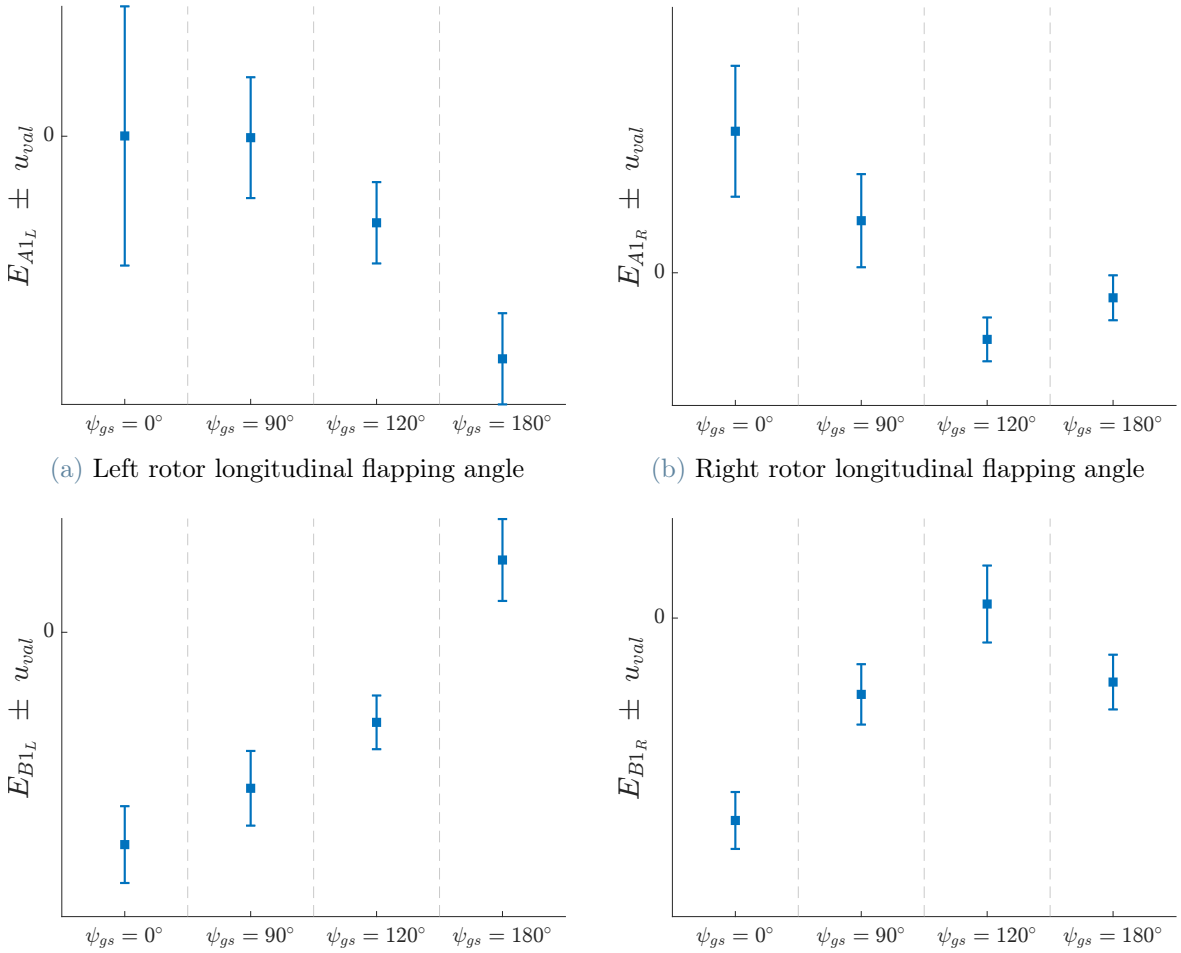


(e) Pitch angle



(f) Roll angle

Figure B.2: E and u_{val} validation metrics results for aircraft attitudes and rotors collective and torque figures on all validation flights. The magnitude of u_{val} is reported as an errorbar on E



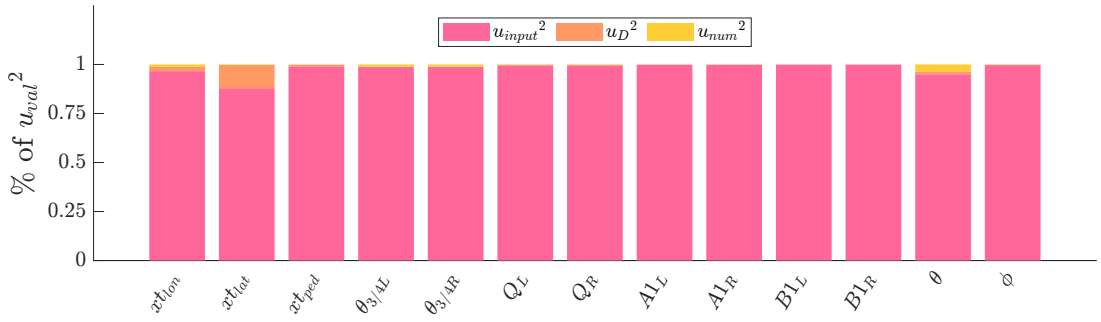
(a) Left rotor longitudinal flapping angle

(b) Right rotor longitudinal flapping angle

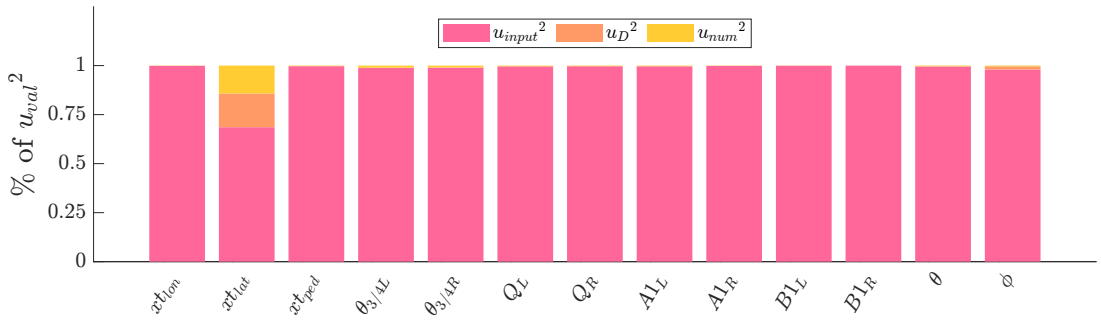
(c) Left rotor lateral flapping angle

(d) Right rotor lateral flappung angle

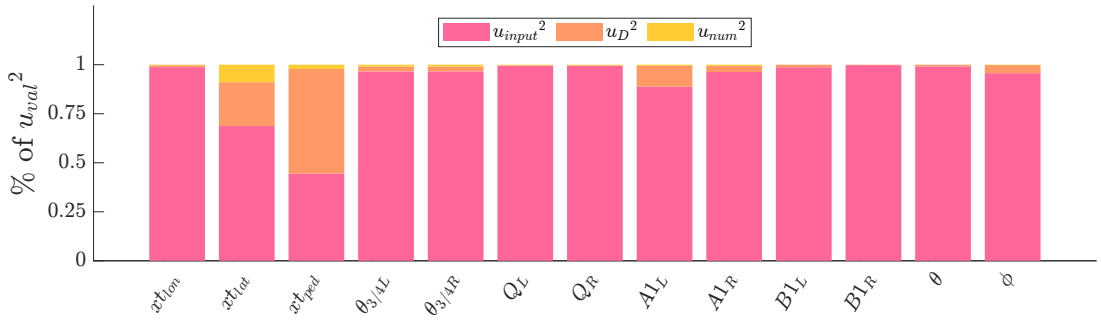
Figure B.3: E and u_{val} validation metrics results for rotors flapping angles on all validation flights. The magnitude of u_{val} is reported as an errorbar on E



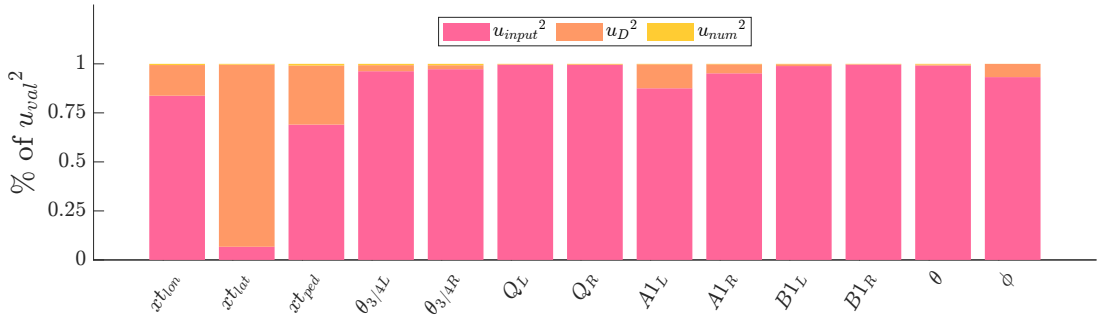
(a) Validation flight $\psi_{gs} = 0°$



(b) Validation flight $\psi_{gs} = 90°$



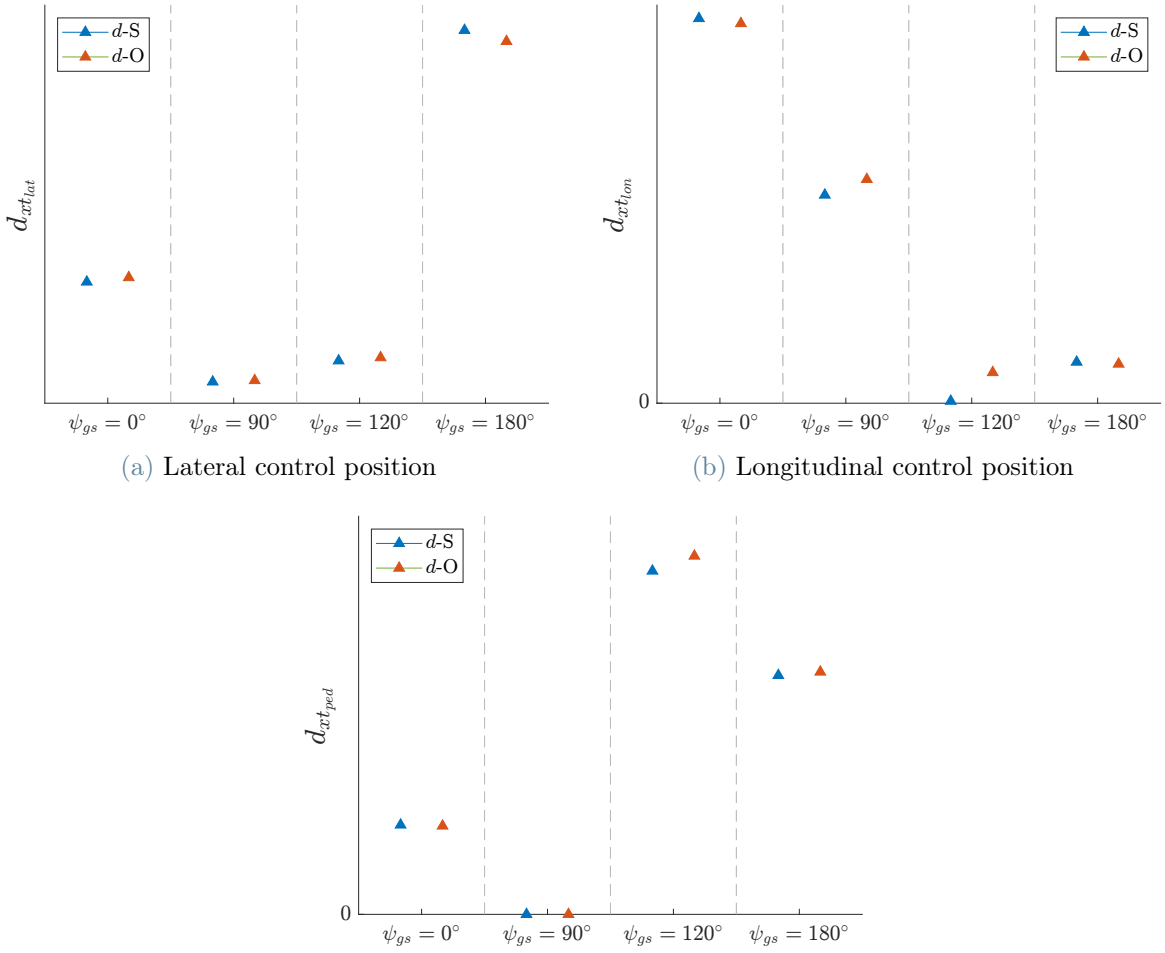
(c) Validation flight $\psi_{gs} = 120°$



(d) Validation flight $\psi_{gs} = 180°$

Figure B.4: Decomposition of $u_{val}{}^2$ into the contributions from u_{input}, u_D and u_{num}

C | Roy-Oberkampf Approach - Validation Results

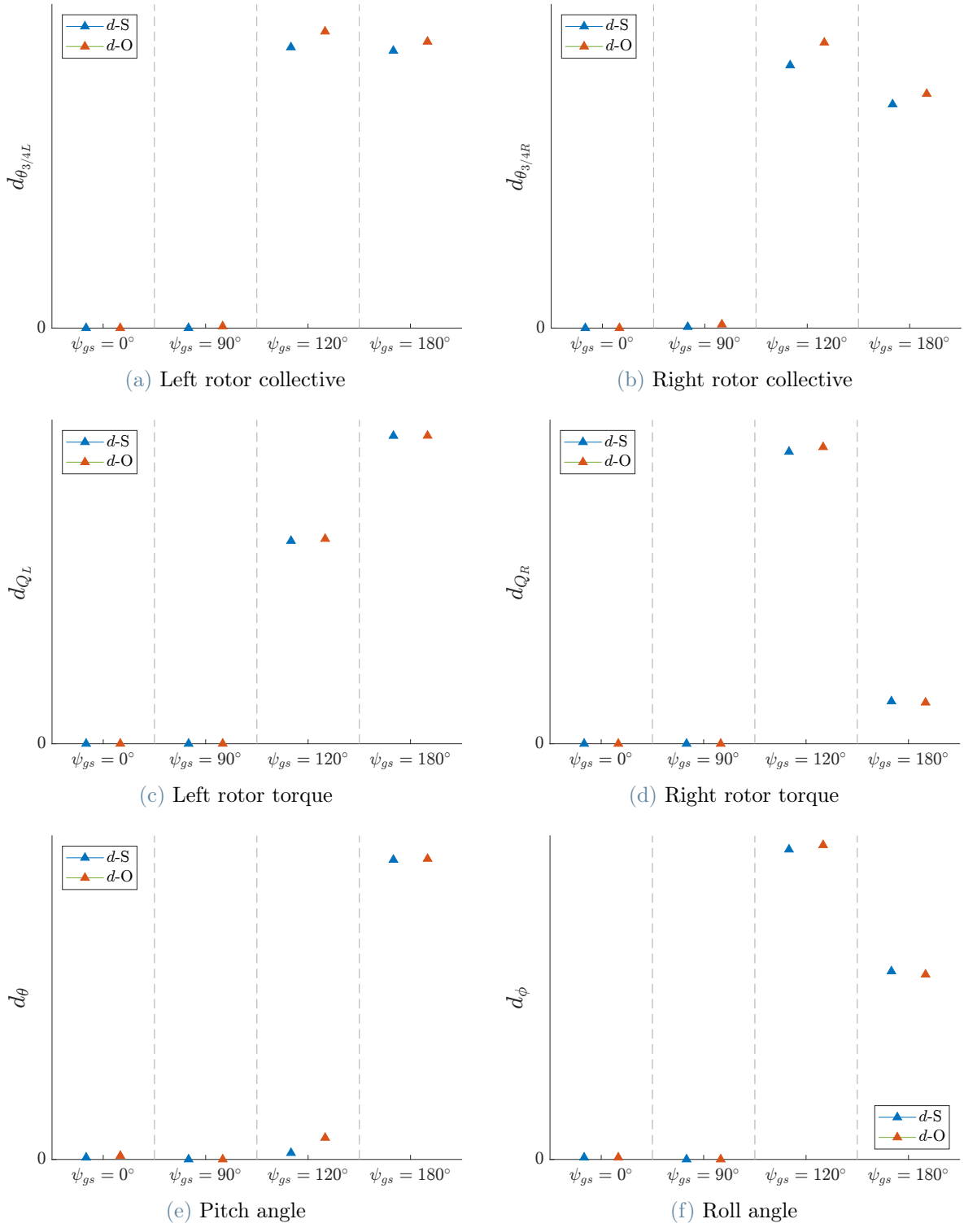


(a) Lateral control position

(b) Longitudinal control position

(c) Directional control position

Figure C.1: Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for control positions on all validation flights



(a) Left rotor collective

(b) Right rotor collective

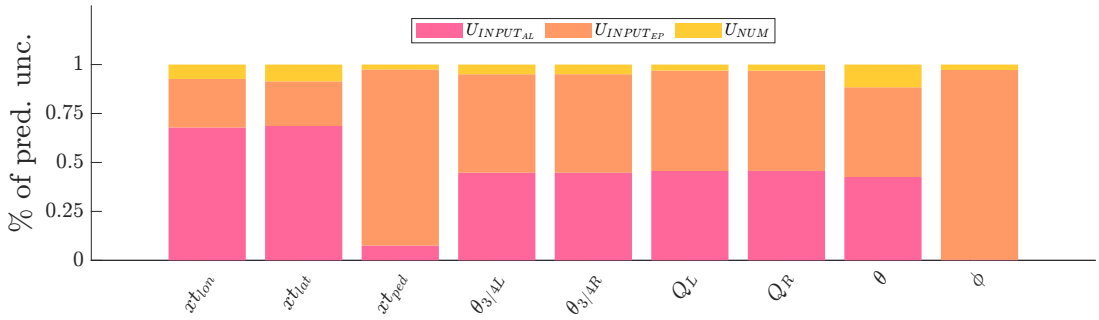
(c) Left rotor torque

(d) Right rotor torque

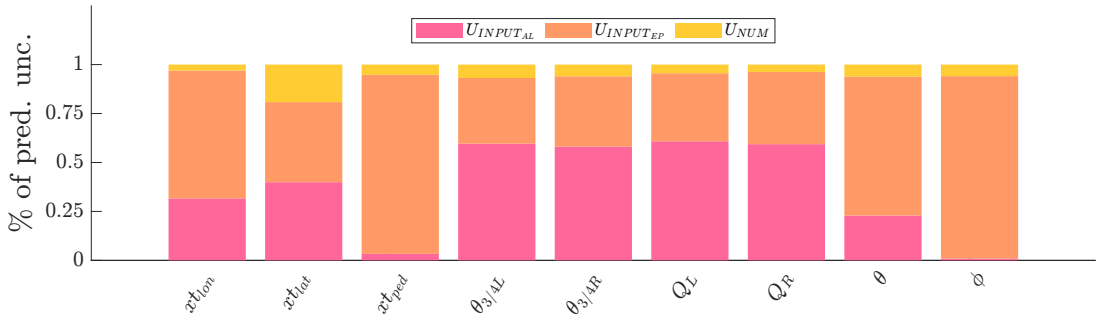
(e) Pitch angle

(f) Roll angle

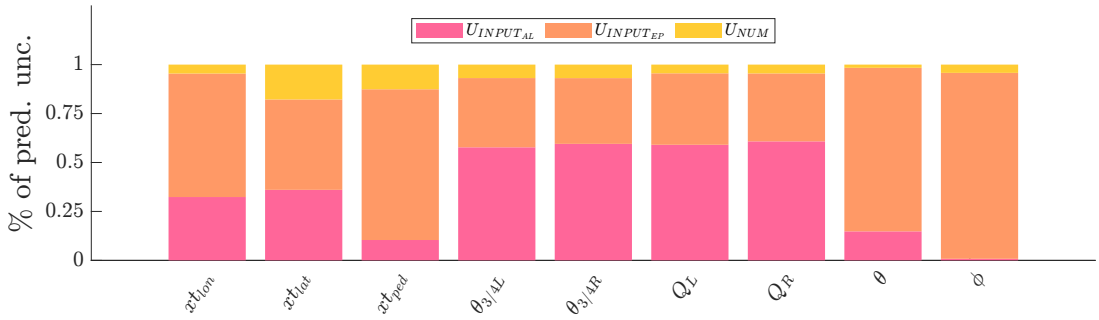
Figure C.2: Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for aircraft attitudes rotors collective and torque figures on all validation flights



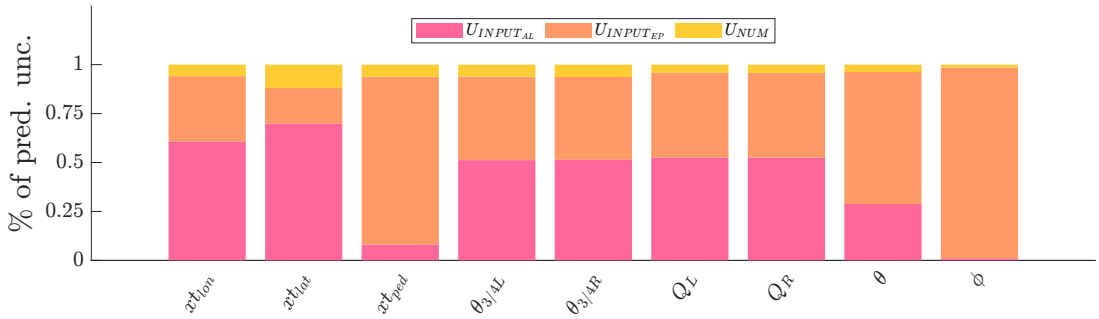
(a) Validation flight $\psi_{gs} = 0°$



(b) Validation flight $\psi_{gs} = 90°$



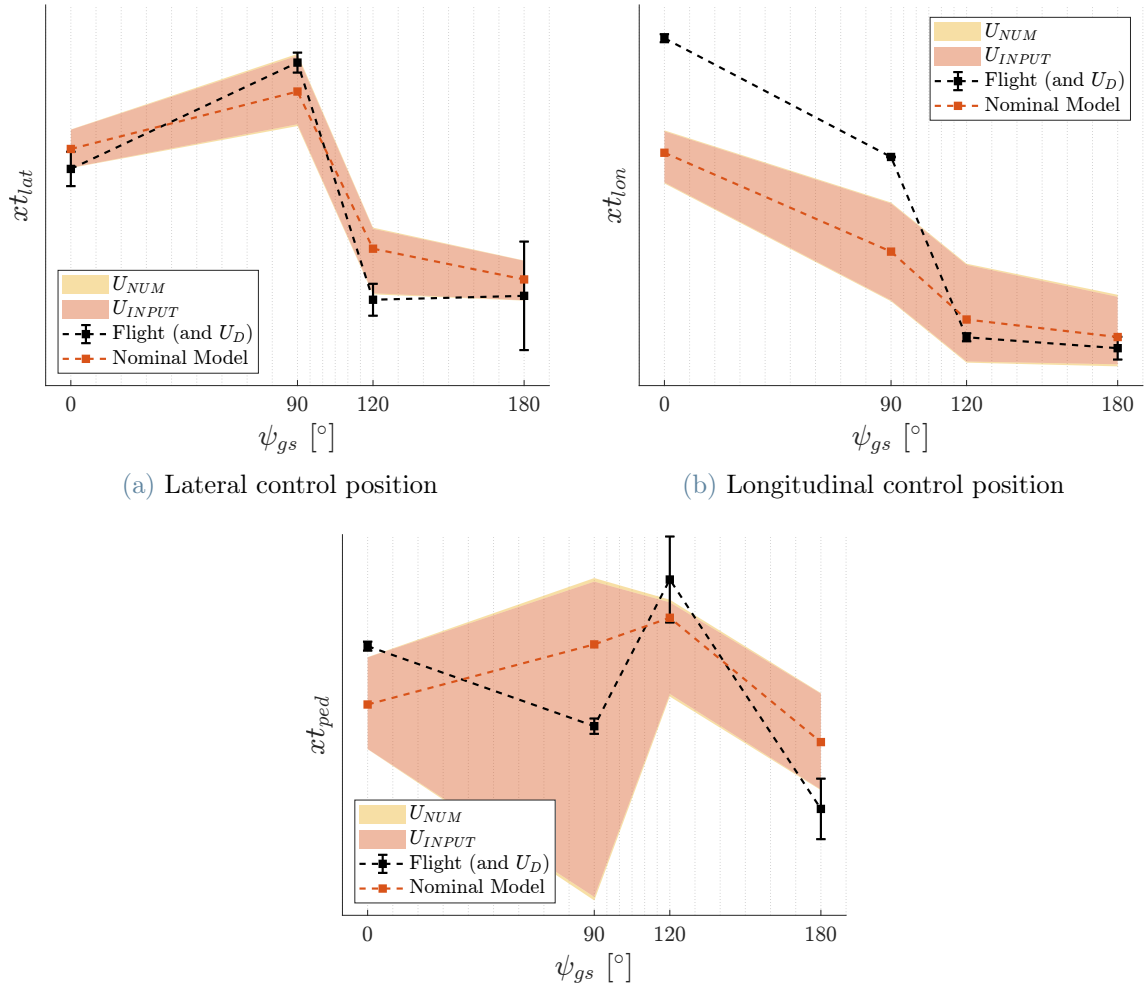
(c) Validation flight $\psi_{gs} = 120°$



(d) Validation flight $\psi_{gs} = 180°$

Figure C.3: Model prediction uncertainty decomposition into the contributions from aleatory and epistemic inputs and numerical error. The results are computed on the s-pbox

D | Interval Analysis - Validation Results

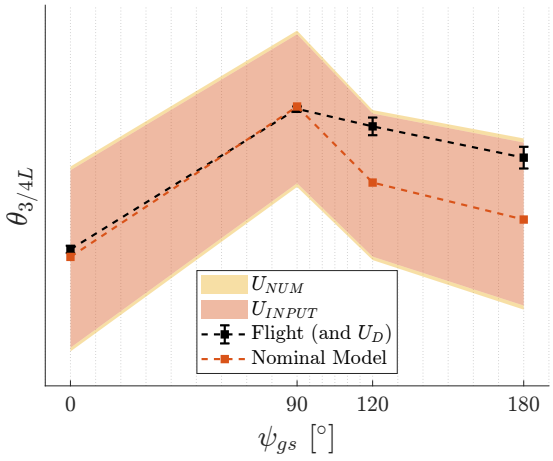


(a) Lateral control position

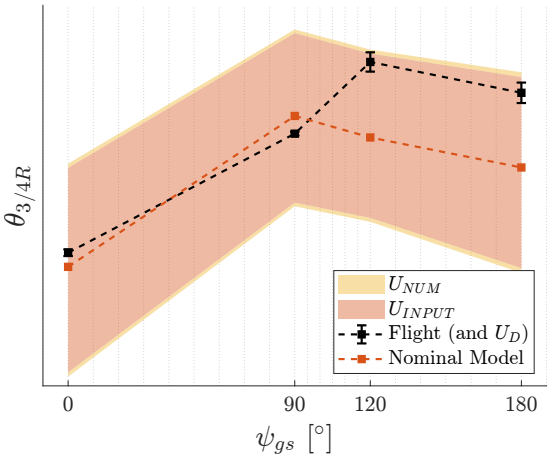
(b) Longitudinal control position

(c) Directional control position

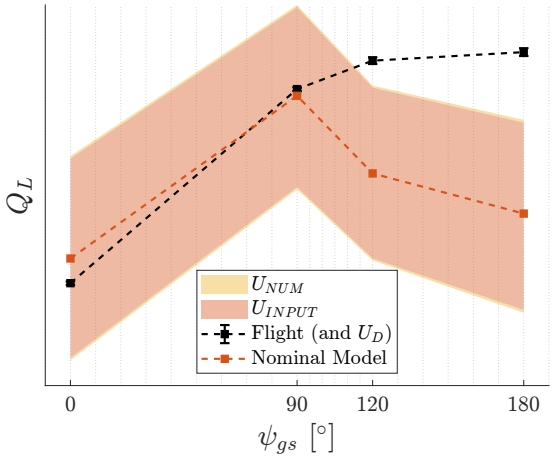
Figure D.1: Results of interval analysis uncertainty propagation and aggregation on FSM prediction and flight data for control positions



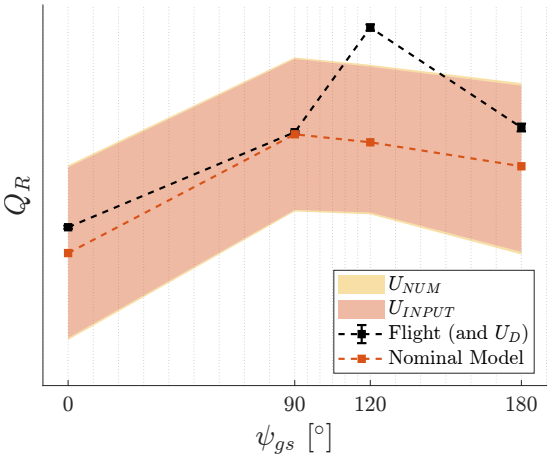
(a) Left rotor collective



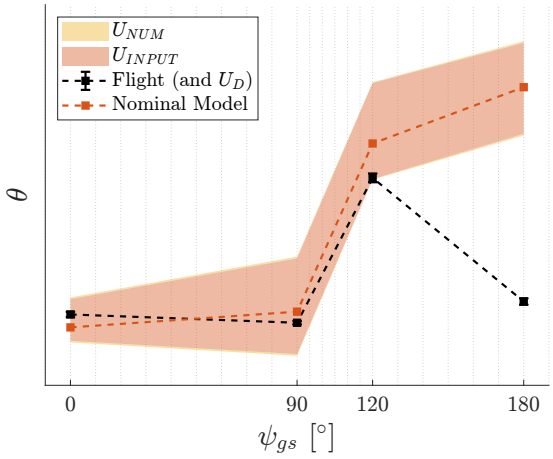
(b) Right rotor collective



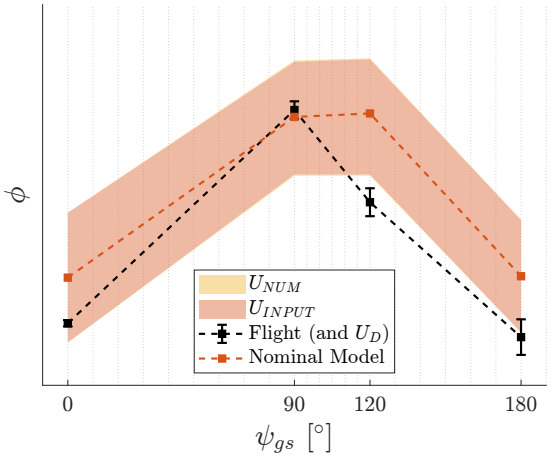
(c) Left rotor torque



(d) Right rotor torque

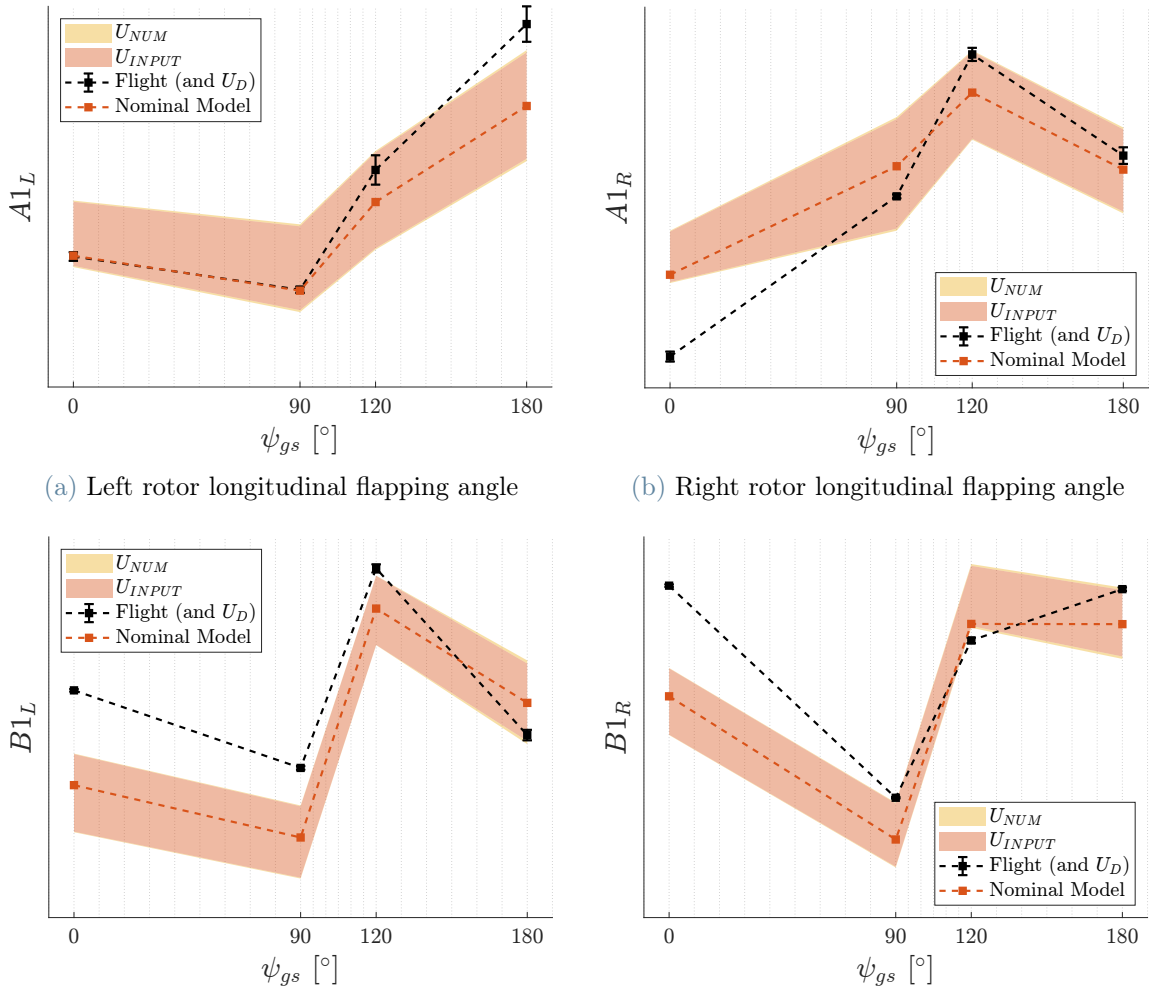


(e) Pitch angle



(f) Roll angle

Figure D.2: Results of interval analysis uncertainty propagation and aggregation on FSM prediction and flight data for aircraft attitudes and rotor collective and torque figures



(a) Left rotor longitudinal flapping angle

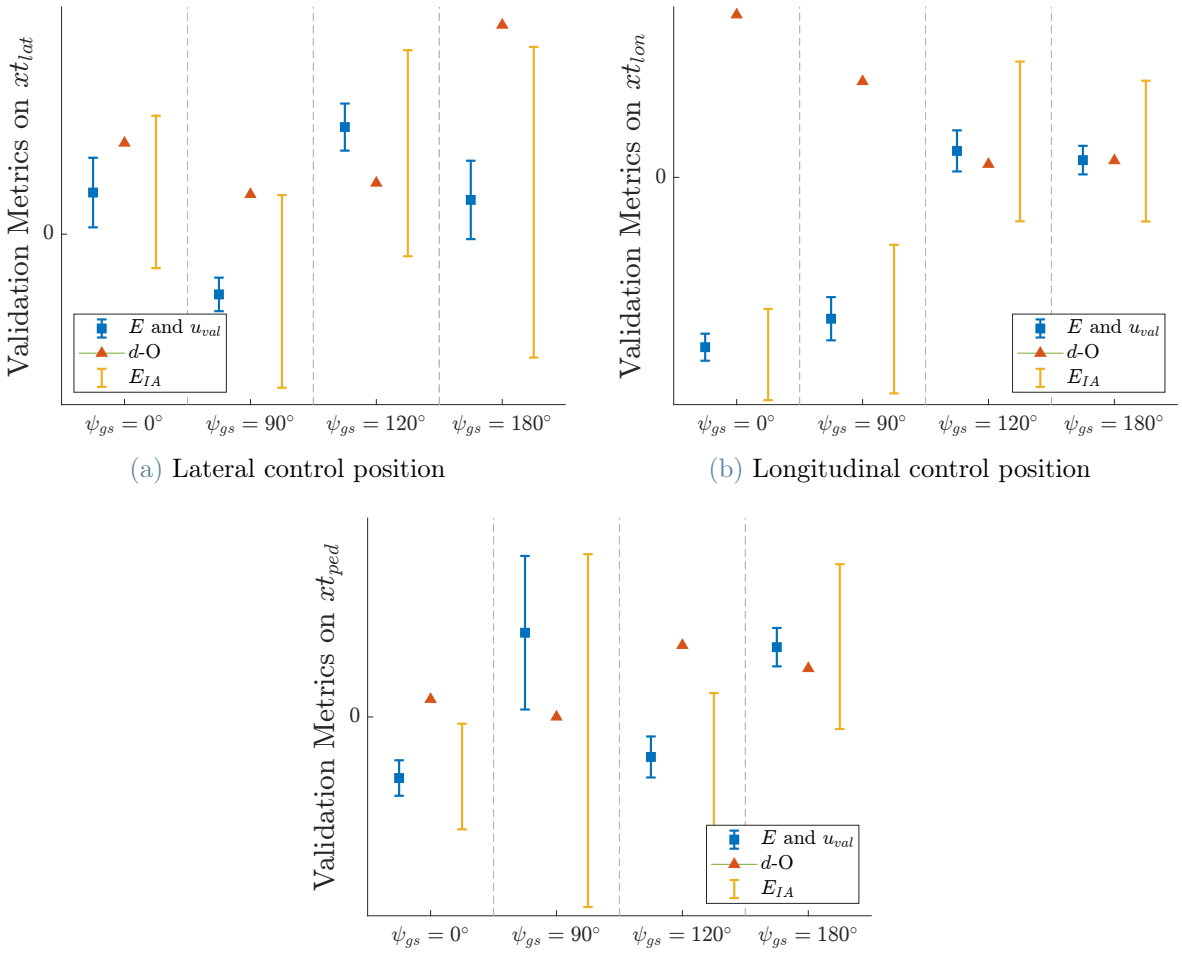
(b) Right rotor longitudinal flapping angle

(c) Left rotor lateral flapping angle

(d) Right rotor lateral flapping angle

Figure D.3: Results of the uncertainty propagation and aggregation on FSM prediction and flight data for rotors flapping angles

E | Validation Metrics Comparison

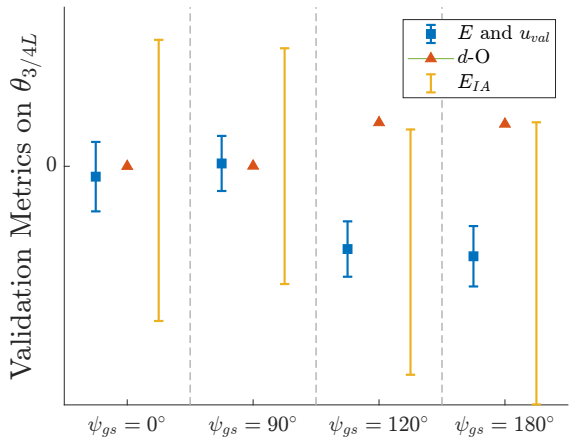


(a) Lateral control position

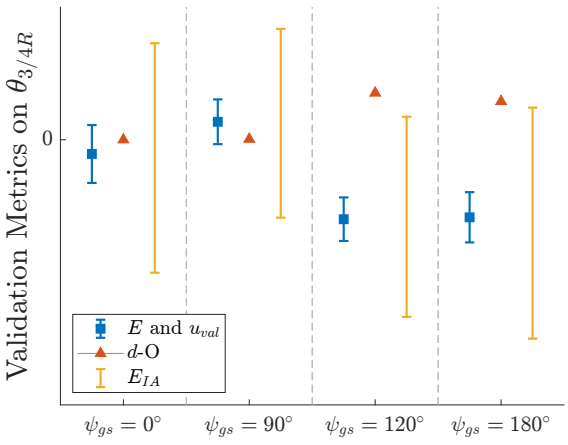
(b) Longitudinal control position

(c) Directional control position

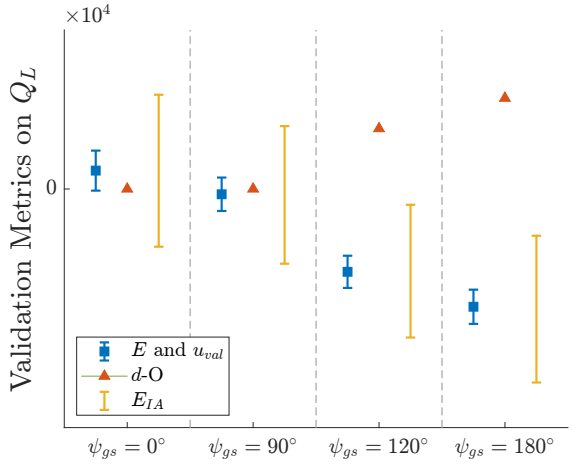
Figure E.1: Validation metrics comparison for control positions on all validation flights



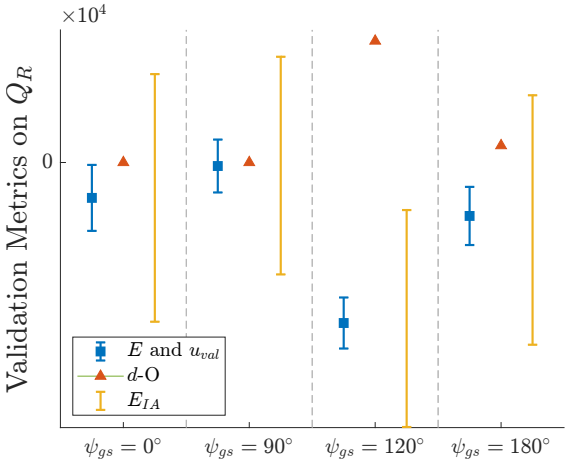
(a) Left rotor collective



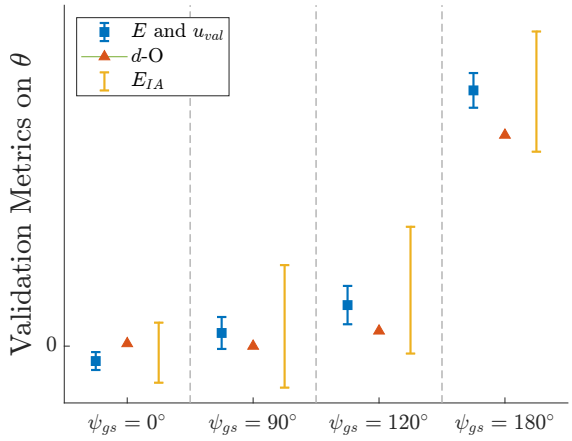
(b) Right rotor collective



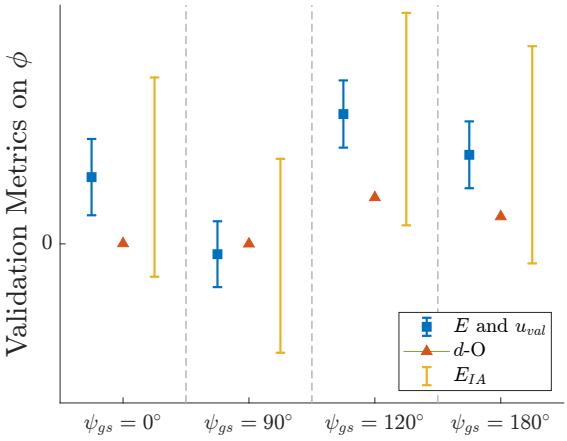
(c) Left rotor torque



(d) Right rotor torque

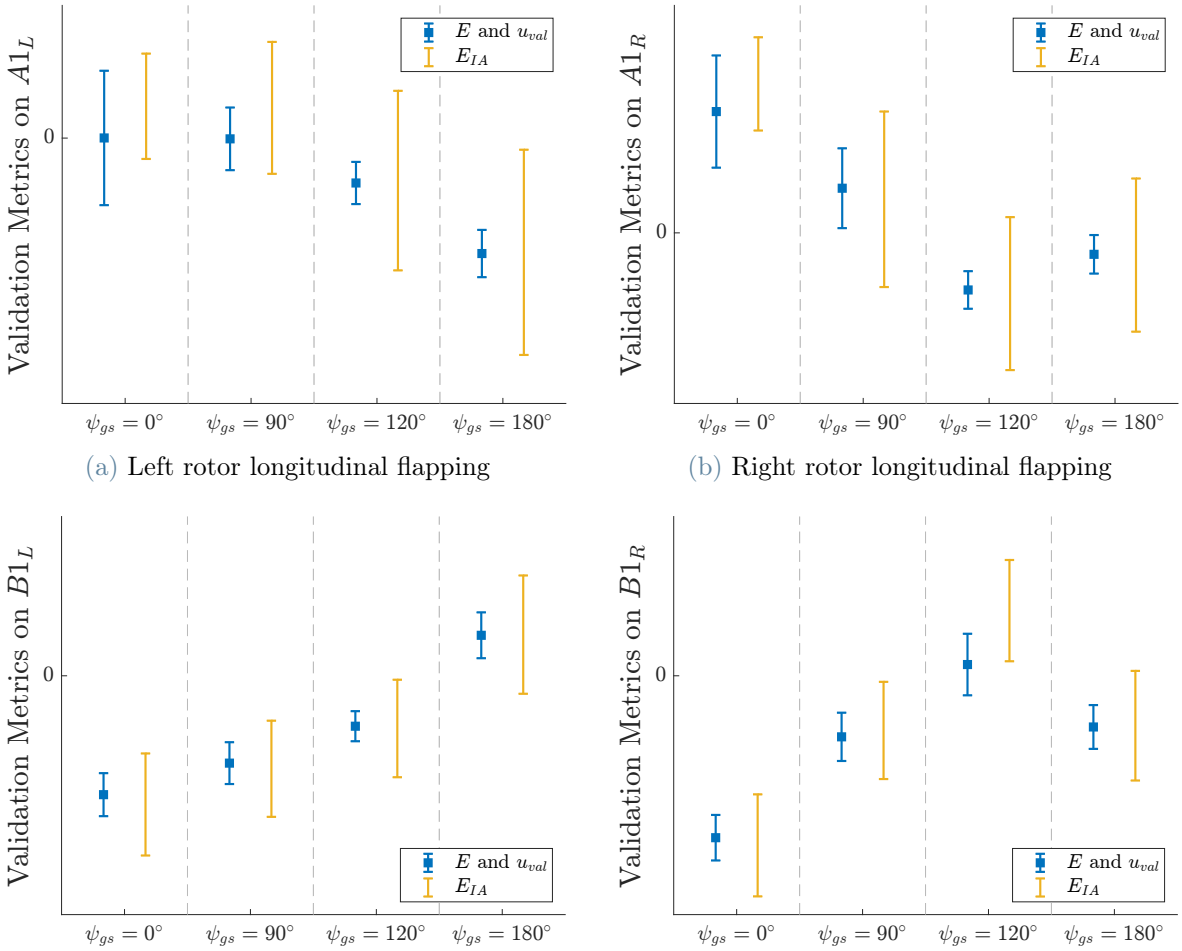


(e) Pitch angle



(f) Roll angle

Figure E.2: Validation metrics comparison for aircraft attitudes and rotors collective and torque figures on all validation flights



(a) Left rotor longitudinal flapping

(b) Right rotor longitudinal flapping

(c) Left rotor lateral flapping

(d) Right rotor lateral flapping

Figure E.3: Validation metrics comparison for rotor flapping angles on all validation flights

List of Figures

3.1 CbS process summary (from [22]) . . . 11
3.2 Sketch of RoCS application domains and their mutual relationship (from [22]) . . . 12
3.3 Process to create a simulation model (from [22]) . . . 16
3.4 Outline of a component-based tilt-rotor flight simulation model (from [22]) 16
3.5 Overview of the sources of uncertainties and model validation process proposed in the ASME VV Standard 20 [8] . . . 18
3.6 Overview of the sources of uncertainty and model validation process proposed by Roy and Oberkampf in [28] . . . 18

4.1 Partial nomenclature schematic for ASME VV 20 approach (from [8]) . . . 22
4.2 Example of model adequacy assessment within ASME validation framework 26
4.3 Flow chart of nested sampling loops for propagation of both aleatory and epistemic uncertainties (from [28]) . . . 32
4.4 Example of the p-box obtained for a SRQ y (from [29]) . . . 33
4.5 Area validation metric example with four experimental measurements available for the SRQ (from [29]). The uncertain simulation prediction is represented by a single CDF, meaning that there is no interval-valued uncertainty in model inputs and that numerical uncertainties have been neglected 34
4.6 Increase in epistemic uncertainty associated to model prediction due the addition of numerical and model-form uncertainty . . . 35
4.7 Examples of model adequacy assessments within the Roy-Oberkampf framework . . . 35
4.8 Example of application of the area metric to *extreme* conditions . . . 36
4.9 Example area validation metric computation and model-form uncertainty definition and application . . . 38

5.1 Examples of flight data time histories for the selected validation flights . . 42
5.2 Prediction, validation and extrapolation domains . . . 44
5.3 Validation points position within the weight-balance envelope of the rotorcraft 45

6.1 FSM components breakdown 48
6.2 FSM aerodynamic interference outline. Active interference effects of an aerodynamic component on other parts are marked by a continuous line. On the other hand, dashed lines mark the account for self-induced velocity on a component. 52
6.3 Peters-He model total number of inflow states as a function of the inflow highest power of radial variation n_{rad} and number of harmonics n_{har} (from [19]) 56
6.4 Example of δ_{num} estimation from sweep result (validation flight $\psi_{gs} = 0°$. 59
6.5 Example of the estimation of δ_{IT} for a system response quantity y. Several predictions of y are performed for each value of ϵ_{IT}, and are then used to construct an *iterative convergence band* 60
6.6 Left rotor collective and longitudinal control position numerical error estimates. The dashed line represents the variation of the SRQ with respect to n_{rad}. The errorbars on each point, conversely, represents the sum of all other numerical error contributions estimates accounted for in this study (e.g. due to ϵ_{IT}, n_{az}, etc.) 62
6.7 Radial induced velocity distribution for different values of n_{rad} on an isolated rotor model in still air 63
6.8 Comparison of the trends of some SRQs with respect to n_{rad} on different validation flights 64
6.9 Numerical error estimates contributions on all validation flights 66
6.10 FSM input uncertainty boxes 68
6.11 An example of trajectory in the input space when $k = 3$ (from [9]) 71
6.12 MOAT sensitivity indices of control positions at trim for the $\psi_{gs} = 120°$ validation flight. Different MOAT parameters are identified by different colors, while indices corresponding to different model inputs are identified by distinct marker shapes 73
6.13 MOAT sensitivity indices of longitudinal flapping angles and aircraft attitudes at trim for the $\psi_{gs} = 120°$ validation flight. Different MOAT parameters are identified by different colors, while indices corresponding to different model inputs are identified by distinct marker shapes 74
6.14 Summary of numerical uncertainty u_{num} evaluation process within the ASME validation framework 76
6.15 Results of the Monte Carlo sampling analysis on left rotor collective and longitudinal control position for validation flight $\psi_{gs} = 90°$ 79

6.16 Results of the Monte Carlo sampling analysis on validation flight $\psi_{gs} = 90°$ rotors flapping angles . . . 80
6.17 E and u_{val} validation metrics results for longitudinal and directional control positions on all validation flights. The magnitude of u_{val} is reported as an errorbar on E . . . 81
6.18 Decomposition of ${u_{val}}^2$ into the contributions originating from u_{input}, u_D and u_{num} . . . 82
6.19 Summary of the numerical uncertainty U_{NUM} evaluation process within the Roy-Oberkampfs framework . . . 83
6.20 Process of creation of the experimental data CDF within Roy-Oberkampf verification and validation framework. The example in the figure is referred to the left rotor torque data of the $\psi_{gs} = 0°$ validation flight . . . 83
6.21 Wind sampling strategy for input uncertainty quantification within Roy-Oberkampf framework . . . 84
6.22 Evolution with N of the sampling error estimate on the discrete CDF of the left rotor torque ($\psi_{gs} = 90°$ validation flight) . . . 85
6.23 CDF sampling error estimates for control positions and left rotor collective on $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$ validation flights . . . 86
6.24 CDF sampling error estimates for rotor flapping angles on $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$ validation flights . . . 87
6.25 CDF sampling error estimates for attitude angles on $\psi_{gs} = 0°$ and $\psi_{gs} = 90°$ validation flights . . . 88
6.26 Flow chart of the optimization-based approach for the computation of the probability box . . . 90
6.27 Comparison of the probability boxes obtained with the conventional sampling (s-pbox) and optimization-based (o-pbox) approach ($\psi_{gs} = 0°$ validation flight) . . . 91
6.28 Model prediction uncertainty decomposition into the contributions from aleatory and epistemic inputs and numerical error. The results refer to the s-pbox . . . 92
6.29 Graphical representation of the d area metric on the longitudinal and lateral control positions ($\psi_{gs} = 0°$ validation flight) . . . 93
6.30 Area validation metric computation on left and right rotors torque figures ($\psi_{gs} = 0°$ validation flight) . . . 93
6.31 Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for control positions on all validation flights . . . 94

6.32 Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for aircraft attitudes rotors collective and torque figures on all validation flights 96
6.33 Flow chart of the validation metric quantification process within the framework of interval analysis 98
6.34 Graphical exemplification of the computation of E_{min} and E_{max} validation metrics 98
6.35 Results of the uncertainty propagation and aggregation on FSM prediction and flight data for rotors flapping angles 100
6.36 Validation metrics comparison for control positions on all validation flights 101
6.37 Validation metrics comparison for left rotor torque and aircraft pitch angle 102
6.38 Comparison of $|E| \pm u_{val}$ and d 103
6.39 Examples of model validation and credibility assessment activities within the Roy-Oberkampf framework. No extrapolation uncertainty is considered affecting U_{MODEL} 106

A.1 Total controls time histories from flight tests 113
A.2 Aircraft attitudes and rotors collectives and torque figures time histories from flight tests 114
A.3 Rotors flapping angles and RPM time histories from flight tests 115
A.4 Aircraft (body-axes) linear accelerations and angular rates time histories . 116

B.1 E and u_{val} validation metrics results for control positions on all validation flights. The magnitude of u_{val} is reported as an errorbar on E 117
B.2 E and u_{val} validation metrics results for aircraft attitudes and rotors collective and torque figures on all validation flights. The magnitude of u_{val} is reported as an errorbar on E 118
B.3 E and u_{val} validation metrics results for rotors flapping angles on all validation flights. The magnitude of u_{val} is reported as an errorbar on E . . . 119
B.4 Decomposition of ${u_{val}}^2$ into the contributions from u_{input}, u_D and u_{num} . . 120

C.1 Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for control positions on all validation flights 121
C.2 Comparison of the values of the area validation metric obtained with the sampling (d-S) and optimization-based (d-O) approach for aircraft attitudes rotors collective and torque figures on all validation flights 122

C.3 Model prediction uncertainty decomposition into the contributions from aleatory and epistemic inputs and numerical error. The results are computed on the s-pbox 123

D.1 Results of interval analysis uncertainty propagation and aggregation on FSM prediction and flight data for control positions 125
D.2 Results of interval analysis uncertainty propagation and aggregation on FSM prediction and flight data for aircraft attitudes and rotor collective and torque figures 126
D.3 Results of the uncertainty propagation and aggregation on FSM prediction and flight data for rotors flapping angles 127

E.1 Validation metrics comparison for control positions on all validation flights 129
E.2 Validation metrics comparison for aircraft attitudes and rotors collective and torque figures on all validation flights 130
E.3 Validation metrics comparison for rotor flapping angles on all validation flights 131

List of Tables

3.1 Influence levels within RoCS framework 13
3.2 Predictability levels within RoCS framework 14
3.3 Influence-predictability levels matrix with minimum confidence levels (from [22]) 20

5.1 Summary of time domain fidelity metrics (and associated acceptable margins) proposed by RoCS and CS-FSTD(H) for trim fidelity assessment . . . 40
5.2 Summary of time domain fidelity metrics exploited in this work 41

6.1 Summary of the SRQs of the FSM 47
6.2 FSM aerodynamic components features and models 49
6.3 List of the FSM parameters affecting the model prediction iterative convergence error δ_{IT} and involved in the solution verification process 55
6.4 List of the FSM parameters affecting the model prediction discretization error δ_{DE} and involved in the solution verification process 57
6.5 List of independent FSM parameters of influence of δ_{IT} and δ_{DE} involved in the error estimate algorithm 58
6.6 Summary of FSM input uncertainty identification and characterization . . 68
6.7 Summary of MOAT parameters values explored for the sensitivity analysis independence assessment: r represents the number of trajectories, p the number of partitions for each model input, n the total number of function evaluations (aircraft trims) required for the analysis 72