Ref. Ares(2020)1595408 - 16/03/2020

RoCS D3.1 - PRELIMINARY SET OF METRICS FOR QUANTIFYING FLIGHT-MODEL FIDELITY

Project title	Rotorcraft Certification by Simulation
Project acronym	RoCS
Grant Agreement No.	831969
Project call	H2020-CS2-CFP08-2018-01
Topic	JTI-CS2-2018-CfP08-FRC-01-19
Work Package	WP3
Lead Partner	NLR
Contributing Partner(s)	LIV, POLIMI
Security classification	CO
Contractual delivery date	30/10/2019
Actual delivery date	16/03/2020
Version	0.4
Reviewers	Michael Jones (DLR), Linghai Lu (LJMU)

History of changes

Version	Date	Comments	Main Authors
0.1	27/09/2019	First version, establishing document structure and content	S. van 't Hoff (NLR)
0.2	18/11/2019	Update with major contributions from NLR and POLIMI.	S. van 't Hoff (NLR), G. Quaranta (POLIMI)
0.3	03/12/2019	Included contributions from UoL on HQ metrics and P-V models. Switch to new report template.	S. van 't Hoff (NLR), G. Quaranta (POLIMI) M. White (UoL)
0.4	30/01/2020	Expanded preliminary fidelity metrics for selected CSRFA.	S. van 't Hoff (NLR), G. Quaranta (POLIMI) M. White (UoL)
1.0	16/03/2020	Final v1.0 release.	S. van 't Hoff (NLR)

Disclaimer

This document contains confidential information in the form of the RoCS project findings, work and products and its use is strictly regulated by the RoCS Consortium Agreement and by Contract no. 831969.

Neither the RoCS Consortium nor any of its officers, employees or agents shall be responsible, liable in negligence, or otherwise howsoever in respect of any inaccuracy or omission herein.

The contents of this document are the sole responsibility of the RoCS consortium and can in no way be taken to reflect the views of the Clean Sky 2 Joint Undertaking.

This project has received funding from the Clean Sky 2 Joint Undertaking (JU) under grant agreement No 831969. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Clean Sky 2 JU members other than the Union

TABLE OF CONTENTS

Executive Summary 5
1 Introduction 6
2 Overall simulation fidelity assessment: predictive and perceptual fidelity 7
3 Acceptable use of simulation in certification 9
4 Unique challenges for predictive fidelity metrics for certification by simulation 10
4.1 Availability of validation flight test data 10
4.2 Definition of acceptable tolerances 10
4.3 Reliance on objective metrics 11
5 Existing predictive fidelity metrics 12
5.1 Metrics currently employed at LH 12
5.2 Training simulator certification standards CS-FSTD(H) 17
5.2.1 Context 17
5.2.2 Parameter tolerances as a metric 17
5.2.3 Validation data 18
5.2.4 Subjective pilot evaluation 18
5.2.5 Discussion 18
5.3 Objective HQ metrics 19
5.3.1 HQ metrics for simulation fidelity assessment 19
5.3.2 Lifting standards project: two-stage approach to simulation fidelity assessment 19
5.3.3 HQ metrics for predictive fidelity assessment 20
5.4 Statistical methods to quantify fidelity 21
5.5 Fidelity bounds based on unnoticeable dynamics 22
5.6 Metrics derived from pilot-vehicle models 23
6 Simulation validation with partial test data set 25
6.1 Introduction 25
6.2 Extrapolation beyond tested envelope 25
6.3 Uncertainty quantification 26
6.3.1 Background 26
6.3.2 Sources of uncertainty 26
6.3.3 Methods for uncertainty quantification 27
7 Metrics selected for preliminary fidelity assessment of LH models 30

7.1 Critical azimuth: CS-29.143(c) 30
7.1.1 Compliance demonstration parameters 30
7.1.2 Preliminary model predictive fidelity metrics 31
7.2 Cat-A take-off: CS-29.53(a) 32
7.2.1 Compliance demonstration parameters 33
7.2.2 Preliminary model predictive fidelity metrics 34
8 References 37

EXECUTIVE SUMMARY

D3.1 - Preliminary set of metrics for quantifying model fidelity addresses the topic of fidelity assessment for flight simulation used in lieu of flight testing in the certification of helicopters and tiltrotors. The report is aimed specifically at the definition and application of quantitative metrics to evaluate the predictive fidelity of the simulation model. The relation to the assessment of perceptual and overall simulation fidelity is discussed in Chapter 2 of this report.

Whereas flight simulation is a well-established practice for pilot training purposes, the application in the context of aircraft certification is currently much more restricted. Today, the acceptable use of simulation in certification is limited to specific scenarios in which flight demonstration is, e.g., too hazardous or the required environmental or aircraft conditions are too difficult to attain. Moreover, specific guidance on the acceptable simulation fidelity or, indeed, the means or methods by which to assess the fidelity of the simulation is missing.

The application of flight simulation and the definition of the associated predictive fidelity metrics in the context of aircraft certification are faced by several unique challenges. The first and foremost challenge is inherent to the application and involves the (partial) absence of validation flight test data. The demonstration of the credibility of the simulation outside of the validated envelope is a particular concern. Another challenge lies in the definition of acceptable tolerances; how good is good enough?

The assessment of flight simulation predictive fidelity has been the subject of significant research effort over the years, particularly in the context of pilot training devices. Chapter 5 of this report provides an extensive review of the relevant literature on the topic, distinguishing in particular the metrics currently in use at Leonardo, those defined in the certification standards for pilot training flight simulators, as well as metrics derived from rotorcraft handling qualities design standards.

The fidelity metrics presented inherently rely on the availability of flight test data for their evaluation. In case the validation data set is missing or incomplete, there will be a need for extrapolation beyond the tested envelope. In this case, questions arise regarding, e.g., the evolution of the prediction error of physics-based model, the minimum data set required to demonstrate simulation credibility, and the limits on the extrapolation distance. These questions will be addressed in the continuation of RoCS WP3 by relying on flight simulation models and validation flight test data of existing aircraft made available by Leonardo. In addition, subject to the availability of suitable data, an effort will be made to investigate the application of Uncertainty Quantification as one means to quantify simulation credibility.

The continued simulation modelling and fidelity assessment activities in RoCS WP3 will focus on four specific paragraphs of the rotorcraft/tiltrotor certification standards as identified in WP2. The preliminary set of flight simulation model predictive fidelity metrics associated with each of the paragraphs is presented in Chapter 7.

1 INTRODUCTION

Flight simulation fidelity assessment has been the topic of much research in recent years. The predictive fidelity of the flight-model is distinguished from the perceptual fidelity of the flight simulation experience by the pilot as addressed in Chapter 2. In both cases, fidelity is intended to reflect and quantify simulation quality relative to the real world. Typically, fidelity research has been in the context of flight simulation for training purposes. The application of flight simulation in lieu of flight testing for certification has historically been limited, with authorities restricting its use as discussed in Chapter 3 of this report. The certification context inherently introduces some unique challenges in terms of simulation fidelity assessment that will be addressed in Chapter 4. This is followed in Chapter 5 by a review of the simulation model predictive fidelity metrics proposed in open literature and their potential for application in the context of certification. The particular challenge of simulation fidelity assessment in the absence of flight test validation data is discussed in detail in Chapter 6. Finally, Chapter 7 presents the preliminary metrics defined for the sections of the certification standards selected in WP1 of the RoCS project.

2 OVERALL SIMULATION FIDELITY ASSESSMENT: PREDICTIVE AND PERCEPTUAL FIDELITY

The comparison of results from predictive and perceptual assessments forms a key component of the overall fidelity assessment process. This is required to establish that the predicted and perceptual results are consistent; for the same reasons in the simulator as they are in flight and to understand better any differences. A flow diagram representing the process for the assessment of predicted and perceptual simulator fidelity is shown in Figure 1.

The process begins with a definition of the required purpose of the flight simulator, and hence the tasks that will be trained (Blocks 1-3 in Figure 1), which will set the required level of fidelity. Once the purpose of the simulator has been defined, testing using the simulator and the simulated aircraft can be conducted (Block 4). This leads to the assessment of the predicted fidelity (Block 5), using a set of metrics described later. The results for each simulator component in the predicted fidelity stage can then be analysed to arrive at an overall level of predicted fidelity for a particular task. The results from these tests feed into the first decision point. The question is; do the individual predictive fidelity metrics show a sufficiently good match between flight and simulation? (Block 6). This stage highlights the quality of individual components of the simulation. Subject to a satisfactory result at this stage, further flight and simulator testing can be conducted to examine the perceptual fidelity of the simulation (Block 7). As with the predictive fidelity, metrics are computed for each test point, in this case relying predominantly on subjective fidelity ratings, (Block 8), and a decision made as to the suitability of the resultant Level of perceptual fidelity for the intended purpose (Block 9).

A third decision point addresses the acceptability of the comparison between predictive and perceptual fidelity (Block 10). This stage is analogous to the comparison between predictive and assigned HQs, as an assessment of the validity of the testing. If the test results are valid, it would be expected that the predictive level of fidelity for the simulator would agree with that from the perceptual assessment processes. In addition, the analysis at this point provides a further indicator as to the source of discrepancies between flight and simulation. For example, if the predictive metrics for the flight model show a good match, while the perceptual metrics do not, then the indication is that the fidelity issues lie within the generation of the task-dependent cues and not the flight model. If all questions (Blocks 6, 9 and 10) can be answered positively then a decision can be made that the simulator is fit for its designed purpose and can be accepted for service (Blocks 11 and 12). If, however, one of the fidelity requirements is not met, this would be an indicator that the simulator is not fit for purpose, and an upgrade, either to the cueing or the flight model or both, is required (Block 13). It should be recognised that a simulator may be fit for some purposes but not others and thus have limited fidelity.

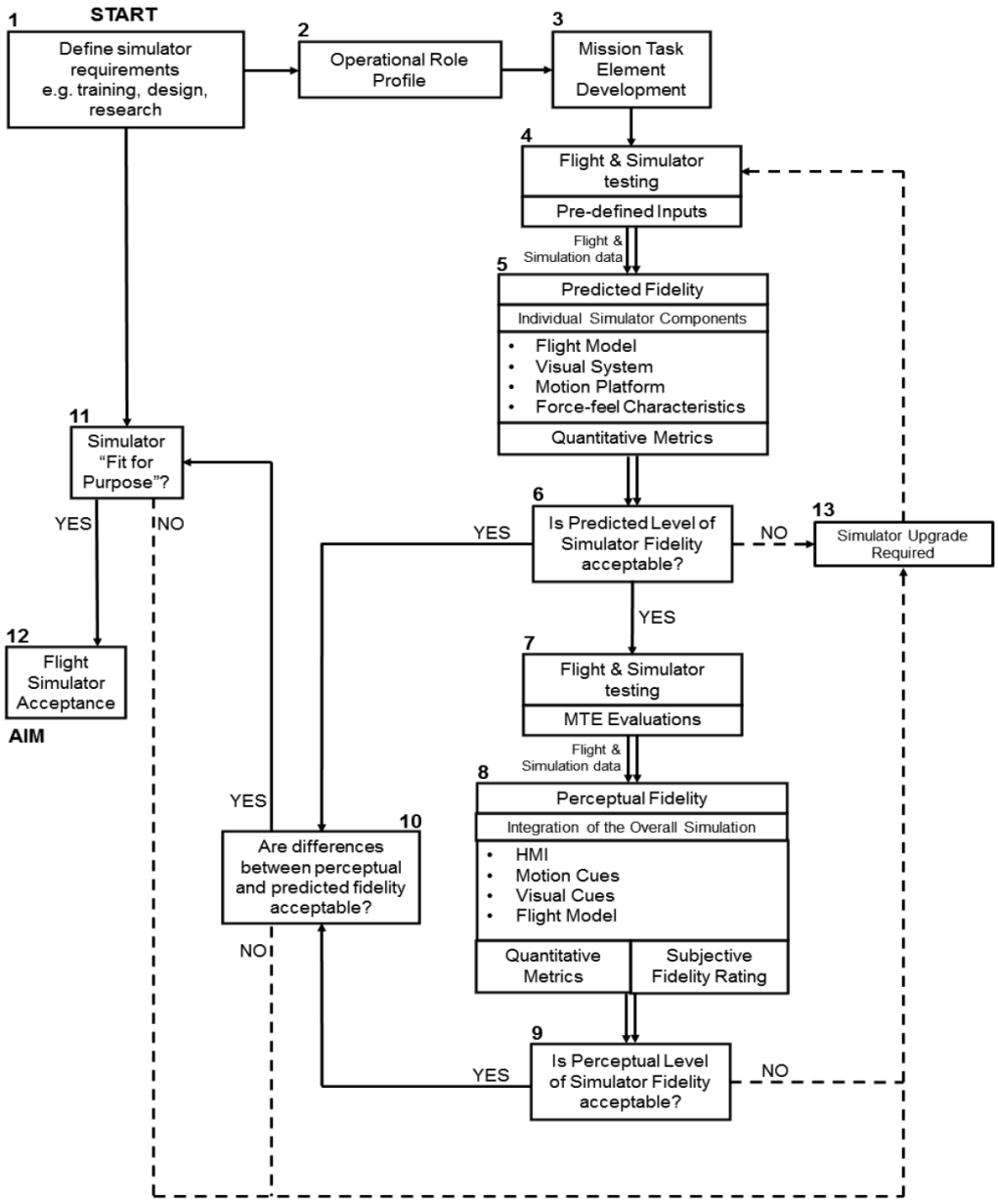


Figure 1: Methodology for integrated predicted and perceptual simulator fidelity assessment

3 ACCEPTABLE USE OF SIMULATION IN CERTIFICATION

Flight simulation is an attractive alternative to flight testing from many perspectives, including cost, flexibility, planning, safety,etc. However, despite the continuous advancements in the field, flight testing for validation purposes will remain a necessity for the foreseeable future. For the same reason, airworthiness certification authorities continue to be understandably cautious about accepting flight simulation as a sole means of compliance.

The FAA has published guidance in AC25-7D[8] regarding the acceptable use of simulation in lieu of flight testing for performance and handling qualities of fixed-wing aircraft. The conditions under which simulation *may* be an acceptable means of compliance are defined therein as:

- Flight demonstration is too risky.
- Required environmental or aircraft conditions are too difficult to attain.
- Simulation is used to augment a reasonably broad flight test program.
- The objective is to demonstrate repeatability, or performance for a range of pilots.

Furthermore, AC25-7D states that the simulation should be *suitably validated* for the conditions of interest, where the level of substantiation of the simulator to flight correlation should be commensurate with the level of compliance. In other words, the closer the case is to being non-compliant, the higher the required fidelity of the simulation. Metrics to measure the fidelity are not defined. Nor is a distinction made between the use of simulation to interpolate or extrapolate flight test data.

The guidance for CFR-29/CS-29[34] does not contain similar verbiage, but §29.21 does refer to calculation and inference (incl. simulation and extrapolation) as potential proof of compliance under the condition that equal accuracy can be demonstrated. The only other references to flight simulation are in the context of tail rotor failure, where flight testing is too hazardous, and altitude extrapolation of hV data. In both cases, the guidance does not provide any insight into the type of simulation and whether it should be performed in a simulator or not.

4 UNIQUE CHALLENGES FOR PREDICTIVE FIDELITY METRICS FOR CERTIFICATION BY SIMULATION

There are three main challenges in the definition of predictive fidelity metrics for certification by simulation:

- Availability (and reliability) of validation flight test data
- Definition of acceptable tolerances
- Reliance on objective metrics

More generally, the selected metrics shall represent the distinguishing flight characteristics of the aircraft across the trim conditions and/or amplitude and frequency range of pilot control relevant for the simulation task. The legitimacy of the metrics and tolerances are to be substantiated by evidence, i.e., test data, and suitably documented.

4.1 AVAILABILITY OF VALIDATION FLIGHT TEST DATA

Flight simulation predictive fidelity metrics are evaluated against validation flight test data. In the case of certification, the validation flight test data may not be available, or the data set may be incomplete. For instance, the available test data potentially does not cover the full certification flight envelope, is not available for a particular aircraft configuration, or is missing entirely because the conditions or manoeuvres have not been flight tested, e.g., for practical or flight safety reasons. Nevertheless, the selected fidelity metrics shall enable quantification of the credibility of the simulation based on dedicated or available flight test validation data to demonstrate that the simulation model is fit for purpose. It is expected that it will be necessary to rely on physics-based simulation approaches in order to justify the application of the simulation outside of the validated envelope.

4.2 DEFINITION OF ACCEPTABLE TOLERANCES

A number of the flight simulation fidelity metrics that have been proposed in the past do enable the quantification of simulation accuracy relative to flight, but without providing a clear foundation for defining acceptable tolerances and, therefore, fidelity. To enable certification compliance demonstration by simulation, the definition of the tolerances as a minimum standard is critical.

It is proposed herein that, potentially in addition to generally applicable fidelity metrics such as those defined in ADS-33E, specific fidelity metrics are defined in direct relation to requirements in the paragraph of the certification standard that is being addressed. The acceptable tolerance on simulation error can then be directly related to the demonstration requirement. For instance, when considering controllability as per CS 29.143, the prediction tolerance on pilot control position may be defined as a function of the distance between the predicted control position and the minimum accepted control margin (see §6.2). In this way, the requirement on simulation fidelity is linked to the proximity to noncompliance.

On the other hand, whereas for training simulator applications the aim is to achieve the highest level of fidelity, the application of simulation for certification compliance demonstration potentially allows for a

lower level of fidelity to be accepted as long as conservativeness can be demonstrated. The worst-case fidelity that may be accepted needs to be explored further in the remainder of the project.

4.3 RELIANCE ON OBJECTIVE METRICS

In the case of piloted flight simulation, the simulation fidelity is the net result of the predictive fidelity of the simulation flight-model and the perceptual fidelity achieved by the flight simulator as described in Chapter 2. The simulator perceptual fidelity may skew the pilot assessment of the overall fidelity of the simulation, relative to the predictive fidelity of the flight-model. In the context of certification, flight simulation may be both piloted and off-line depending on the compliance demonstration objective. In either case, the flight-model simulation fidelity must be demonstrated as much as possible using objective metrics. The application of subjective metrics based on pilot assessment will be addressed in RoCS WP4.

5 EXISTING PREDICTIVE FIDELITY METRICS

5.1 METRICS CURRENTLY EMPLOYED AT LH

Leonardo Helicopters (LH) employs a set of quality criteria for rotorcraft models (including both helicopters and tiltrotors) based on a comparison of model results with experimental data. Consequently, the metrics used are not applicable to predictive models where a comparison with flight data is not available.

The metrics used have been established using CS-FSTD(H)[1], the certification standards for training flight simulators, as a reference and follows the same rationale. The quality of the model is defined based on a comparison between simulation and flight test data for a specific set of manoeuvres. The predicted aircraft response is required to fall within a specified error band, defined either as absolute values or in percentage of each considered parameter. The selection and range of parameters are dependent on the flight manoeuvre considered.

The rules defined in CS-FSTD(H) have been extended in order to account for aircraft that cannot fly with AFCS OFF, as is the case for tiltrotors. Additionally, parameters have been included that are deemed by LH to be relevant in the context of design and certification (e.g., flapping angles) that are not included in CS-FSTD(H).

The following modifications to rules set in CS-FSTD(H) have been applied based on company experience (without discussion with the relevant certification authority):

1. Elimination of the dependence of error boundaries on flight condition. The only distinction that remains is the one between trim and dynamic conditions. Consequently, whenever in CS-FSTD(H) different tolerances are defined for different manoeuvres; the tighter bound has been applied.

2. All parameters are taken into account for all manoeuvres and not only those considered as more relevant. Therefore, no distinction is made between on-axis and off-axis response.

3. Additional parameters have been added such as: rotor torque, flapping angles, vertical speed and lateral load factor, AFCS downstream command to control surfaces, etc.

4. The tolerance on sideslip has been tightened from 2 to 1 degree and an alternate 10% error option has been added based on the experience gained on high-speed conditions for tiltrotors.

5. The tolerance on height above ground has been tightened from 20 to 10ft for simulation of OEI conditions for helicopters.

6. A percentage tolerance has been added for altitude (for up-and-away conditions) equal to 7.5% of distance from trim (similar to the tolerance for horizontal distance).

7. The tolerance on roll rate has been tightened from 3 to 2 deg/s to align it with the one used on pitch and yaw axes.

8. A check on the relative phase between roll and yaw rate has been added for lateral-directional oscillations.

9. For static and manoeuvring tests, the absolute error for inceptor positions is defined in % of the total control travel, following experience gained with simulation of tiltrotors.

All metrics defined here are related to time-domain simulation and are consequently applicable to that domain. No frequency domain criteria have been defined.

The following Table 1 summarizes all parameters considered. When two tolerance values are given for a parameter, the less restrictive one is used unless indicated otherwise.

Table 1: Simulation model metrics (acceptable +/- maximum errors on measures) used by LH

NAME	TYPE	ABSOLUTE ERROR			% ERROR	
		UNITS	TRIM	DYNAMICS	TRIM	DYNAMICS
Calibrated Air Speed	Performance	kts	.	3	-	-
Vertical Speed	Performance	ft/min	100	100	10	10
Height above ground	Performance	ft	.	10	-	
Altitude (up-and-away)	Performance	ft	.	100	-	7.5
Horizontal distance (near-earth)	Performance	ft	.	100	-	7.5
Rotor Speed	Performance	% of nominal RPM	1.5	1.5	-	-
Engine Torque	Performance	% of engine max continuous torque	3	3	-	-
Longitudinal stick position (cockpit)	Longitudinal dynamics	%	5	10	-	-
Collective lever position (cockpit)	Longitudinal dynamics	%	5	10	-	-
Longitudinal stick position (cockpit)	Longitudinal dynamics for stability test	%	2.5	-	10 (variation wrt flight test trim value)	-
Longitudinal stick position (AFCS)	Longitudinal dynamics	%	5	10	-	-
Pitch	Longitudinal dynamics	deg	1.5	1.5	-	10
Pitch rate	Longitudinal dynamics	deg/s	-	2	-	10
Normal load factor	Longitudinal dynamics	g	-	0.1	-	-
Angle of attack	Longitudinal dynamics	deg	1	1	10	10
Longitudinal oscillation period	Longitudinal dynamics	s	-	-	-	10

Table 1: Simulation model metrics (acceptable +/- maximum errors on measures) used by LH

NAME	TYPE	ABSOLUTE ERROR			% ERROR	
		UNITS	TRIM	DYNAMICS	TRIM	DYNAMICS
Longitudinal oscillation time to halve/double	Longitudinal dynamics	s	-	-	-	10
Longitudinal Oscillation damping ratio	Longitudinal dynamics	% of critical damping	-	2	-	-
Lateral stick position (Cockpit)	Lateral-directional dynamics	%	5	10	-	-
Directional control position (Cockpit)	Lateral-directional dynamics	%	5	10	-	-
Lateral stick position (Cockpit)	Lateral-directional static stability	%	2.5	-	10 (variation wrt flight test trim)	-
Directional control position (Cockpit)	Lateral-directional static stability	%	2.5	-	10 (variation wrt flight test trim)	-
Lateral stick position (AFCS)	Lateral-directional dynamics	%	5	10	-	-
Directional control position (AFCS)	Lateral-directional dynamics	%	5	10	-	-
Roll	Lateral-directional dynamics	deg	1.5	1.5	-	10
Roll rate	Lateral-directional dynamics	deg/s	-	2	-	10
Heading	Lateral-directional dynamics	deg	-	2	-	10
Yaw rate	Lateral-directional dynamics	deg/s	-	2	-	10
Lateral load factor	Lateral-directional dynamics	g	-	0.1	-	-
Sideslip angle	Lateral-directional dynamics	deg	1	1	10	10

Table 1: Simulation model metrics (acceptable +/- maximum errors on measures) used by LH

NAME	TYPE	ABSOLUTE ERROR			% ERROR	
		UNITS	TRIM	DYNAMICS	TRIM	DYNAMICS
Lat. Dir. Oscillation period	Lateral-directional dynamics	s	-	0.5	-	10
Lat. Dir. Oscillation time to halve/double	Lateral-directional dynamics	s	-	-	-	10
Lat. Dir. Oscillation time damping ratio	Lateral-directional dynamics	% of critical damping	-	2	-	-
Relative phase roll-sideslip	Lateral-directional dynamics	s	-	1	-	20
Relative phase roll-yaw rate	Lateral-directional dynamics	s	-	1	-	20
Rotors mast torque	Internal parameter	% of rotor max. continuous torque	3	3	-	-
Rotors longitudinal flapping	Internal parameter	deg	1	1	10	10
Rotors lateral flapping	Internal parameter	deg	1	1	10	10

The models are evaluated through a two-step process.

STEP 1 - parameter level:

For each validation dataset, for each parameter included in Table 1, the quality of the match of each parameter is classified as:

1. GREEN: lower than thresholds;
2. YELLOW: lower than **two** times the thresholds;
3. RED: lower than **four** times the thresholds;
4. BLACK: higher than four times the thresholds.

STEP 2 - flight condition level:

Given a flight condition, for each group of model parameters (performance, longitudinal dynamics, lateral-directional dynamics and internal parameters), a colour code is assigned as follows:

1. GREEN: 90% of the parameters of the group are classified GREEN and 10% are not worse than YELLOW.

2. YELLOW: 90% of the parameters of the group are classified not worse than YELLOW and 10% are not worse than RED.
3. RED: 90% of the parameters of the group are classified not worse than RED and 10% are classified as BLACK.
4. BLACK: All other cases.

Depending on the simulation task/objective, the relevant groups of parameters to be evaluated are selected and the minimum level required for each of the groups is specified. The remaining parameters are required to be realistic in terms of trend and magnitude based on engineering judgment.

It is noted that the tolerances specified in Table 1 are not considered hard thresholds of acceptability since exceedances are allowed up to a certain degree. Moreover, there is no link to the proximity to noncompliance and all parameters that are considered in the assessment are weighted equally (i.e., no consideration of on-axis vs off-axis response). Finally, no information has been provided on how groups of parameters are selected, what determines the minimum match level for a given simulation task, or how realistic trend and magnitude is defined.

5.2 TRAINING SIMULATOR CERTIFICATION STANDARDS CS-FSTD(H)

5.2.1 Context

The CS-FSTD(H)[1] has been established as a standard to qualify flight simulators used for training purposes. The ultimate aim of a training simulator is to provide a representative training environment in which the pilot/crew skills and strategy can be learned, developed or assessed and where the level of adaptation of the piloting skills and strategy, e.g., due to inadequate cueing or differences in aircraft dynamics, is minimal (i.e., does not lead to negative transfer of training).

The means by which the latter are achieved in terms of underlying physical modelling is of secondary concern from a simulator qualification point of view. The CS-FSTD(H) standards merely reflect general requirements regarding the physical aircraft features and phenomena that shall be represented in the flight simulator. For instance, qualitative requirements are posed regarding the level of aerodynamic modelling, including phenomena such as ground effect, vortex ring state and retreating blade stall. There is, however, no constraint on the level of non-physical tuning, which would not be acceptable for a model used to extrapolate or otherwise replace flight test data.

The context of flight simulation for training purposes is distinctly different from the application of simulation to the certification of the aircraft itself. Nevertheless, the simulation to flight test error tolerances that form the basis of the Acceptable Means of Compliance of CS-FSTD(H) may be of merit.

5.2.2 Parameter tolerances as a metric

The Acceptable Means of Compliance of CS-FSTD(H) are defined in terms of parameter-level tolerances applied to the error between the flight simulation and the validation data, thereby providing a measure of the quality of the match between the two. Tolerances are specified for specific manoeuvres and flight conditions, emphasizing the control and aircraft parameters that are of primary interest for a given scenario. No consideration is made of the type of helicopter or rotorhead in the definition of the tolerances.

In terms of the simulation model, the tolerances are divided into two categories: 1) performance, and 2) handling qualities. The performance section includes requirements for take-off, hover climb, descent, autorotation and landing. The handling qualities section generally only considers small amplitudes and is based on trims, step control responses (w/o clear definition of the control input) and short-period stability tests in various conditions. There are no standards or tolerances defined for flight dynamics aspects in the frequency domain, apart from the modal frequencies and damping under the handling qualities tests.

The time-domain parameter tolerances adopted in CS-FSTD(H) provide an intuitive understanding of the differences between the simulation and the validation data, even though the cause for a particular deviation may remain obscured. However, no background is provided as to how the parameters were selected, or how the magnitudes of the tolerances have been defined. In fact, it is believed that the requirements have been inherited from fixed-wing standards with the original substantiation lost in history. The suitability of the tolerances, at least for most applications, is evidenced by the acceptance of the flight simulators for training purposes by qualified test pilots, but this does not guarantee suitability for other purposes, including flight simulation for certification.

The guidance material leaves room for exceedance of the specified tolerances in that “Validation tests that do not meet the test criteria should be addressed to the satisfaction of the competent authority.”, and “When it is difficult or impossible to match FSTD to helicopter data or approved validation data throughout a time history, differences should be justified by providing a comparison of other related variables for the condition being assessed.” Also, the quality of the match to validation data for less-than-critical parameters is in certain cases defined as *correct trend and magnitude* instead of specifying quantified tolerances.

5.2.3 Validation data

The validation data, in the context of CS-FSTD(H), are typically comprised of flight test data, but also data generated using the OEM *engineering simulator* is used in some cases. When non-flight test data is used as an allowable form of reference validation data, a near-exact match is required (the guideline is 20% of the corresponding flight test tolerance). In either case, the validation tests are largely based on time histories with the published tolerances being applied throughout.

5.2.4 Subjective pilot evaluation

Although the intent of CS-FSTD(H) is to evaluate the flight simulation as objectively as possible, pilot acceptance is acknowledged to be an important consideration. In particular, for specific flight conditions that cannot be validated by flight test (typically failure scenarios) or fall outside of the validated envelope, subjective pilot assessment, e.g., through the FAA handling qualities rating method[8], is accepted as the only means of compliance. However, the methodology for subjective pilot assessment is not well-defined and there is no formal process for resolving cases where the objective and subjective assessments are in disagreement.

5.2.5 Discussion

The tolerance-based flight simulation fidelity assessment adopted in CS-FSTD(H) has merits in terms of the intuitive implementation and the fact that the tolerances are tailored to specific flight conditions and manoeuvers. The suitability of the magnitudes of the tolerances when applied in the context of simulation-based compliance demonstration to a specific paragraph of CS-29 remains to be investigated in RoCS WP3. To the extent possible within the project, this effort should also consider the dependence on helicopter

type. Furthermore, the inclusion of large-amplitude time-domain (where applicable) as well as frequency domain metrics to complement the time-domain handling qualities assessment in CS-FSTD(H) should be considered.

5.3 OBJECTIVE HQ METRICS

5.3.1 HQ metrics for simulation fidelity assessment

GARTEUR Action Group (AG) HC/AG-12 was formed to conduct a critical examination[9]-[11] of the then-existing simulator standards in JAR-STD 1H[12] (consolidated along with JAR-STD 2H[13] and JAR-STD 3H[14] into JAR-FSTD(H)[15] in 2008, and subsequently adopted as CS-FSTD(H) in 2012), including correlation of handling qualities and fidelity metrics. The work revealed a range of shortcomings. For example, AG-12 showed that the relationship between fidelity and the JAR-STD 1H tolerances is sensitive to the nature of the manoeuvre being flown and, more significantly, that matching tolerances does not always lead to matching handling qualities. The use of HQ metrics for fidelity assessments was one of HC/AG-12's key recommendations.

The use of HQs engineering as a framework to quantify overall simulation fidelity has developed in several forms. Hess and colleagues[16]-[18] introduced the handling qualities sensitivity function as a quality metric. Padfield et al.[19], and McCallum and Charlton[20] first proposed the handling qualities standard, ADS-33E-PRF 'Handling Qualities Requirements for Military Rotorcraft'[21], as the basis for deriving metrics; since the HQ parameters define the flying characteristics, they ought also to be suitable to judge the fidelity. Within the JSHIP project, Advani and Wilkinson[22] and Roscoe and Thompson[23] used comparative performance and control activity, and handling qualities ratings given for the same tasks flown in simulation and flight. In each of these approaches, the adopted philosophy has been to develop a rational and systematic approach to the identification of differences between simulation and flight, and hence the areas where the simulator is deficient. These methods have met with partial success, but this only highlights the need for new fidelity criteria for use in design, development and product qualification. Development of new criteria was the focus of the Lifting Standards research project at the University of Liverpool (UoL)[24] and the new approach is presented in the following sections.

5.3.2 Lifting standards project: two-stage approach to simulation fidelity assessment

It is the need to have objective measures of predictive fidelity, complemented by subjective measures of perceptual fidelity, that was the main focus of the UoL 'Lifting Standards: A Novel Approach to the Development of Fidelity Criteria for Rotorcraft Flight Simulators'[25]-[27]. A two-stage approach for defining fidelity criteria for simulator qualification was developed in Lifting Standards. The first stage involved the development of a quantitative basis for prediction of fidelity using metrics derived, in part, from HQ engineering. The second stage consists of perceptual fidelity metrics supplemented by a simulator fidelity rating scale[28], used to assign the perceptual fidelity of the simulator.

In the area of HQ engineering, two assessment processes, prediction and assignment, are integrated, and combine to give the overall HQs of an aircraft. The practices adopted in the Lifting Standards project draw on this integrative process and the HQ performance specification, ADS-33E-PRF[21]. For both processes, the test aircraft is assessed to be in one of three HQ 'levels' which are defined by HQ ratings and reflect the workload and task performance of the pilot controlling the aircraft.

Much of this methodology can be directly applied to the fidelity assessment of a flight simulator – both handling qualities and fidelity are intimately related to pilot control strategy and task performance. If a pilot makes the same control inputs in the simulator as they are required to in flight, then correct behavioural patterns have been learned, and the simulator training will have been effective; meaning that the pilot will have benefitted from their time in the simulator. The goal in fidelity assessment is to establish the 'quality' of the simulator in replicating the behaviour of the real aircraft and its pilot, rather than purely to assess the handling qualities of the simulated aircraft. In the case of the prediction metrics, the fidelity assessment is focussed on the simulator components, e.g. the flight model, the image generation system and so forth, with the fidelity predictions for each component contributing towards a prediction for the overall simulation. For the assignments, the pilot's impression of the behavioural accuracy of the model is closely linked with the experienced cues. The primary generators of task cues within the simulator are the visual, motion, audio and inceptor force-feel systems. In fidelity, we describe the pilot's experience as the perceptual fidelity.

5.3.3 HQ metrics for predictive fidelity assessment

The first part of the fidelity assessment process involves an analysis of the individual simulator components. For the flight model, the predicted HQs of the test aircraft and flight model are computed, with dynamic response criteria drawn from the response to clinical tests such as pulse, step, doublet and frequency sweep control inputs. HQ metrics define the level of performance/compensation achievable/required to fly the defined missions. Metrics have been developed in both the time and frequency domains, to assess the full range of aircraft response, from low to high frequency and from small to large amplitude and are presented as the dynamo construct shown in Figure 2[29].

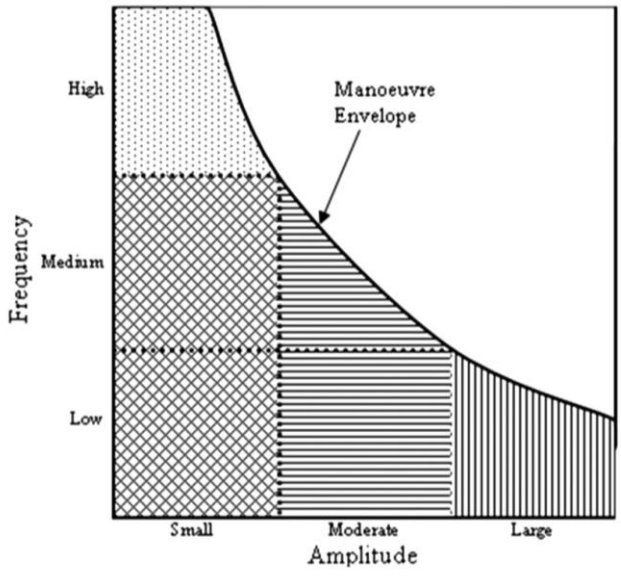


Figure 2: Dynamo construct for dynamic response criteria[29]

Dynamic response criteria are often displayed in two-parameter charts, for example damping and frequency, or bandwidth and phase delay, with boundary lines demarking regions of Level 1, 2 and 3 performance. As measures of dynamic performance, these metrics provide a sound basis for quantifying

predictive simulation fidelity, since they are referenced to missions and pilot control strategies in operational tasks. The fidelity of the simulated flight model is defined as the goodness of the match between flight and simulation for each of the predicted HQ metrics.

Moving through the dynamo construct, the stability and agility criteria adopted in the predicted HQs section of ADS-33E-PRF to assess each region are:

1. Small amplitude, high frequency – bandwidth and phase delay.
2. Small amplitude, low to medium frequency – open-loop stability.
3. Moderate amplitudes – quickness.
4. Large amplitudes – maximum response.

A further set of HQ metrics is required that specify the required level of handling for the cross-coupled, off-axis responses, e.g. pitch response to roll control inputs (and vice versa) and the yaw response to collective control inputs. Additionally, for forward flight the magnitude of the pitch response to a collective input is assessed:

The comparison of these HQ metrics between flight and simulation provides an indication of the fidelity of the model as shown in [24] and [29]. An open question at this point in time is how closely the HQ metrics of the aircraft need to be matched in simulation in order for the simulation to be considered adequately representative for application in a certification context.

5.4 STATISTICAL METHODS TO QUANTIFY FIDELITY

GARTEUR Actions Groups AG06 and AD09 investigated, amongst other things, statistical methods to quantify the fidelity of helicopter simulation models[2],[3]. Multiple formulations were proposed including time and frequency domain cost functions driven by the simulation-to-test error in aircraft response and control parameters, and phase and amplitude of the frequency response, respectively.

The statistical methods aimed to derive cost functions that reflect the goodness-of-fit, or the significance of the deviation between the flight simulation and flight test data based on simulation-to-test deviations of the control and aircraft states. The time-variable cost functions were evaluated per parameter, or the parameters were combined into a single cost function using weighting matrices based on the acceptable error (tolerance) for individual parameters.

The methods proposed in AG06/09 consist of statistical evaluations that indicate whether the simulation "deviates significantly" from the validation data, statistically speaking. The parameters introduced do not have a physical interpretation that promotes intuitive understanding regarding the deficiencies of the simulation. Moreover, most of the formulations rely on the a priori definition of acceptable error, thereby providing little benefit over the tolerance-based approach adopted in CS-FSTD(H).

Finally, in addition to open questions regarding how to define acceptable ranges of fidelity cost function values, it is unclear how this approach would be extended to a case where validation flight test data is not available or incomplete. The authors of this report did not find statistical methods in open literature that are promising in terms of their suitability for quantifying simulation fidelity in a certification context.

5.5 FIDELITY BOUNDS BASED ON UNNOTICEABLE DYNAMICS

The frequency-domain envelopes defining Maximum Unnoticeable Added Dynamics (MUAD), first developed in the 1980s, define regions of acceptable levels of mismatch between a flight simulation and the real aircraft in terms of magnitude and phase of the simulation error response based on the pilot's ability to perceive the difference in the dynamics[4]. The envelopes were originally defined through a series of flight simulator trials in which a pilot were asked to provide a Cooper-Harper rating of the handling qualities of a baseline low-order model of fixed-wing aircraft pitch dynamics, as well as various modifications of it obtained by adding various degrees of dynamics. The envelopes are adopted in MIL-STD-1797A as criteria for fixed-wing handling qualities simulation.

The MUAD envelopes were originally defined for fixed-wing, but have since found application in the field of rotorcraft and tiltrotor simulation, typically in the realm of flight dynamics simulation[5]. In most cases, the envelopes are adopted without rigorous verification of their suitability for the intended application. However, research by Mitchell et al[6], in which the method was applied to a helicopter in hover, revealed that the envelopes were too stringent for this application due to the higher bandwidth of the baseline aircraft, while the envelopes were not stringent enough at lower frequencies. The specific piloting task and simulator facility may further affect the bounds of the noticeable dynamics that form the basis for the MUAD concept.

Following the work presented in [6], Mitchell defined a new tentative set of envelopes, termed Allowable Error (AE) envelopes, that are a function of the dynamics of the baseline aircraft. However, as noted by the authors, the proposed envelopes are a rough approximation and application of the methodology is inherently impractical as the envelopes are dependent on the bandwidth of the aircraft. To date, universally accepted unnoticeable dynamics envelopes for rotorcraft are still lacking, although the concept continues to be applied in certain fields, particularly in control law design in the US.

The concept of MUAD and AE envelopes has been revised in recent attempts aiming towards the objective evaluation of the influence of altered dynamics on pilot control behaviour using a cybernetic approach [7][8]. In fact, Matamoros et. al. point out that one strong limiting factor of the AE envelopes proposed by Mitchell is related to the fact that they were determined only considering subjective ratings from a single pilot and without objective metrics to back up the ratings. Using several pilots and applying rigorous statistical methods to evaluate their control effort, Matamoros et al found that the dependence on the bandwidth of baseline model dynamics of MUAD envelopes can, at least under the tested conditions (pitch attitude control task, pure dipole added dynamics), be discarded.

In the context of certification by simulation, there are three main deficiencies to the concept of unnoticeable dynamics; 1) there are no universal unnoticeable dynamics envelopes that are proven to be applicable to all rotorcraft, piloting tasks and simulation facilities, 2) similar to other metrics discussed herein the concept only reflects the simulation fidelity in terms of aircraft dynamics, providing no validation of other physical parameters, and 3) system identification validation flight test data that is required for evaluation may not be easy to obtain in critical parts of the flight envelope (e.g. at VNE).

Despite the practical limits described above, frequency domain correlation between flight test data and the simulation model, or different models of varying fidelity, provides valuable insight in the discrepancies of the simulation, even if bounds of allowable error are hard to define.

5.6 METRICS DERIVED FROM PILOT-VEHICLE MODELS

The Adaptive Pilot Model (APM) methodology has been developed in the context of training flight simulators to quantify overall simulation fidelity based on an analysis of pilot visual guidance strategy - identifying the control loops utilised, levels of abruptness and the cues available to support anticipation. The premise is that if the control strategy adopted to perform a flying task is 'equivalent' in flight and simulation, then the fidelity is good and the flight simulator fit for purpose with regards to handling qualities aspects.

The meaning of equivalent is developed in terms of what can be described as the Adaptive Pilot Model, whereby the combined pilot and aircraft is modelled as an equivalent system[30]. Comparisons made of model parameters identified from the same curve fitting process applied to data from flight and simulation tests then form the basis of system fidelity. In the APM approach, the changing pilot gains relating to velocity and distance control, for example, are tracked through the manoeuvre.

The general hypothesis behind the APM approach follows from earlier representations[31]-[33], whereby, in flying a manoeuvre, the pilot acts to transform the coupled aircraft-pilot system to a simple relationship between command and output. In the acceleration-deceleration manoeuvre, for example, the pilot initiates the manoeuvre from a hover in response to the command R_c and finishes in a new hover when the error (R_c-R) is reduced to zero (Figure 3). The distance to stop is defined as -X. Figure 4 shows the corresponding closed-loop control scheme, where θ_c is the commanded pitch attitude and θ is the current pitch attitude.

By applying several simplifying assumptions the vehicle-pilot model can be reduced to extract, e.g., pilot gains, and the frequencies and damping of the closed-loop dynamics. The comparison of the identified parameters between flight test and simulation provides a means to quantify and assess the fidelity of the simulation, with the identified parameters acting as fidelity metrics.

The described methodology obviates the complexity of a direct comparison of control inputs generated by a test pilot and a virtual pilot model. The approach may be of value in the certification simulations for complex manoeuvres such as Category A take-off and landing procedures. However, as is the case for many of the fidelity metrics discussed herein, the definition of appropriate tolerances remains a challenge.

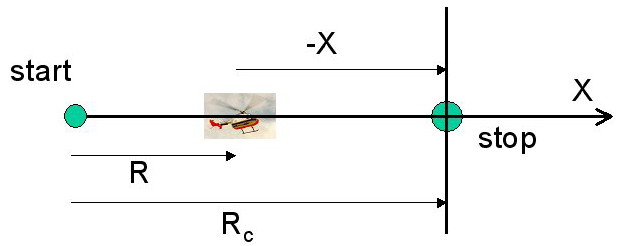


Figure 3: Illustration of accel-decel manoeuvre logic

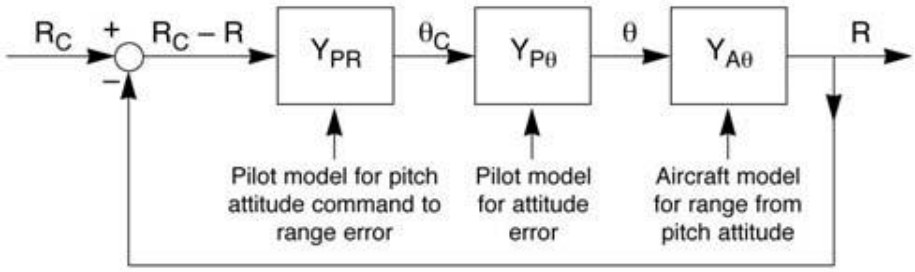


Figure 4: Closed-loop control of aircraft range

6 SIMULATION VALIDATION WITH PARTIAL TEST DATA SET

6.1 INTRODUCTION

The application of simulation for the purpose of certification compliance demonstration inherently implies a situation where the available validation data is incomplete, requiring either interpolation or extrapolation. Data from particularly high-risk flight conditions may be missing entirely (e.g., power-off landing), whereas for other certification requirements flight test data may only be available in a part of the envelope. In both scenarios, flight simulation using a physics-based approach is a promising means to acquire the compliance data at reduced cost and/or risk. However, how does one demonstrate the credibility of the simulation in the (partial) absence of validation data?

6.2 EXTRAPOLATION BEYOND TESTED ENVELOPE

Figure 5 illustrates the problem at hand using the altitude-extrapolation of trim control margin as an example. The non-hashed area indicates the demonstrated worst-case simulation error bandwidth within the tested envelope, taking into account the scatter in the measurement data. The hashed area reflects the unknown evolution in the prediction error over the extrapolation 'distance'. Finally, the blue error bar indicates the allowable prediction error at the corner of the envelope in relation to the distance to the compliance limit (in this case the minimum control margin required for adequate gust control). Questions that arise are:

- What is an acceptable error tolerance in the validated part of the envelope given the allowable error at the extrapolated conditions?
- What is the minimum number of validation data points that is needed to gain confidence in the simulation and accept extrapolation using a physics-based model?
- Can we define limits for the allowable extrapolation distance for a given parameter (e.g., altitude, weight, airspeed) and flight condition?

The evolution of the prediction error along the dimension of extrapolation is not known a priori and, even for physics-based models, will depend on the flight condition, the extrapolation parameter and distance, the aircraft configuration, and the simulation methods used. To determine appropriate guidelines, it will be necessary in the continuation of RoCS WP3 to review these trends for the selected paragraphs of the CS using available test data from past certification efforts (spanning the complete certification envelope) and modern physics-based simulation tools, preferably considering multiple dissimilar aircraft. In practice, the available test data and models is expected to be restricted to the AW109.

It is noted that extrapolation can only be justified if, in addition to the tolerance on the prediction magnitude, the simulation is able to replicate the correct trend. In other words, correct magnitude and trend must be demonstrated.

In terms of extrapolation, AC 29-2C contains guidance that limits straightforward extrapolation for altitude to ±4,000 ft for performance in hover, take-off and landing, whereas a ±2,000 ft extrapolation limit is specified for IGE handling qualities, h-V testing, and engine operating characteristics. In contrast, controllability and stability shall be flight demonstrated at least at the lowest practical altitude and the

highest cruise altitude (with interpolation between). Weight may be extrapolated for certain aspects, but only along an established W/σ line within the allowable altitude extrapolation range up to the maximum gross weight of the rotorcraft. The proposed application of physics-based simulation in lieu of flight testing will require these limits to be re-evaluated.

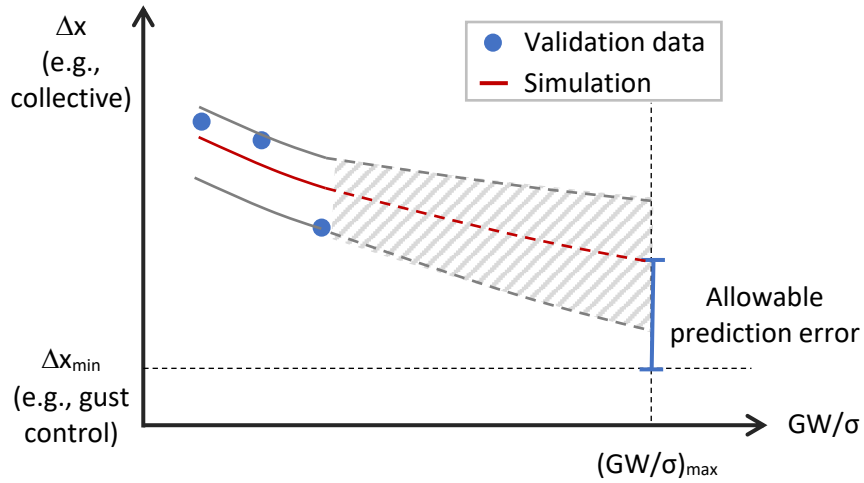


Figure 5: Illustration of trim control margin simulation prediction error in an altitude extrapolation scenario.

6.3 UNCERTAINTY QUANTIFICATION

6.3.1 Background

Uncertainty Quantification (UQ) is an umbrella term for various methods of the quantitative characterization of uncertainties in various applications. Two types of uncertainty quantification can be distinguished: forward uncertainty propagation and inverse uncertainty quantification.

Forward uncertainty propagation is the statistical effect of variable uncertainty on the model's output that is based on these variables. This can be performed prior to any experiments. Inverse uncertainty quantification estimates the discrepancy between experimental data and the model. This requires experiment results to be available so that the discrepancies between the experiment and model may be identified.

6.3.2 Sources of uncertainty

Within simulation-based modelling three types of uncertainty can be identified: parametric uncertainty, model-form uncertainty and predictive uncertainty.

6.3.2.1 Parametric uncertainty

Parametric uncertainty consists of the uncertainty in the inputs of the model and the used parameters in the model. Many models require deterministic values as input values or parameters although it is in most cases not possible to get the exact value with infinite precision. This uncertainty will create differences between the output of the model and the physical representation. Parametric uncertainty can be categorized into two types: aleatoric and epistemic uncertainty. Aleatoric uncertainties are unknown

random outcomes. It is caused by the inherent variability of the process and is therefore irreducible. Epistemic uncertainties are caused by a lack of knowledge or data. These uncertainties can be reduced by more data, analysis or research.

6.3.2.2 Model-form uncertainty

Model-form uncertainty consists of the uncertainty in representations of physical processes which is influenced by, for example, varying boundary conditions, different theories of behaviours or response, or linearity assumptions. These different models could produce different outputs while representing or simulating the same physical process. This uncertainty in the outputs is caused by model-form uncertainty.

6.3.2.3 Predictive uncertainty

Predictive uncertainty is the uncertainty in the difference between the output of a single model and the actual physical process. This difference is caused by simplifications or negligence of physical phenomena.

6.3.3 Methods for uncertainty quantification

There are several methods available to apply uncertainty quantification on the three aforementioned uncertainty types. A limiting factor in these methods is the availability of experimental data. While uncertainty quantification can be applied to parametric and model-form uncertainty in the absence of experimental data, predictive uncertainty cannot be performed when experimental data is not available.

6.3.3.1 Parametric uncertainty quantification

When quantifying parametric uncertainty, sampling methods can be used. The advantage of using sampling methods is that they are relatively easy to implement in complex simulations because of their non-intrusive nature. The disadvantage is that these methods are quite exhaustive and are computationally intensive. A popular sampling method is Monte Carlo Simulation.

Monte Carlo Simulation

Monte Carlo Simulation is a sampling technique where a physical process is simulated multiple times, each with different start conditions. The results of these individual simulations are combined into a final result. This result could be a probability distribution of the possible outcomes (figure 2). The advantage is that Monte Carlo simulation both shows what could happen and how likely it could happen. Furthermore, Monte Carlo can be used to determine the probability that a certain event will occur. An example would be the amount of simulation output values that exceed a certain threshold for given initial value distributions.

The set of initial conditions can be determined by random sampling. The downside is that it is not guaranteed that the samples will be picked from the extreme cases of the simulation, while these extreme cases are of interest in certification. To account for this downside, Latin Hypercube sampling could be used. With Latin Hypercube sampling first a fixed number of sample points is determined, in contrast to random sampling. After that sample points are picked in such a way that every parameter is guaranteed to be uniformly represented.

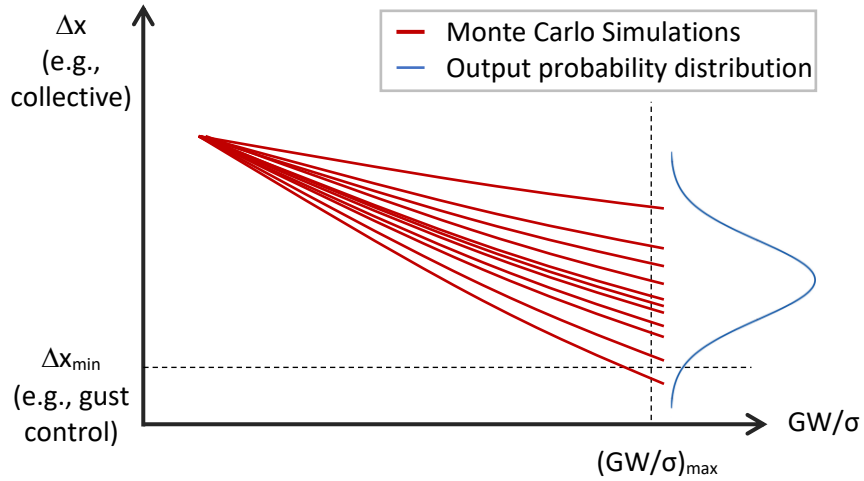


Figure 6- Illustation of Monte Carlo Simulations of trim control margin in an altitude extrapolation scenario

6.3.3.2 Model-form uncertainty quantification

The methods used in quantifying model-form uncertainty differ from each other in whether experimental data is available or not. A frequently used method that relies on experimental data is Bayesian Model Averaging. When no experimental data is available the adjustment factors approach could be used.

Bayesian Model Averaging

This method is used to calculate posterior model distributions (for predictive uncertainty quantification) and to update model probabilities (for model-form uncertainty quantification). With use of a set of experimental data points, Bayesian Model Averaging consists of two steps. First the predictive uncertainty of the individual models is determined using a Maximum Likelihood Estimate. After that, the model probabilities are updated using the model likelihoods. The updated model probabilities can be determined by Equation 5.3.1 where $P(y|M_i, D)$ is the predictive distribution of the model and $P(M_i|D)$ the posterior model probability. In this equation: *y* is the model output, *D* is the given experimental data, *M* is the model, and *N* is the amount of models.

$$P(y|D) = \sum_{i=1}^{N} P(y|M_i, D)P(M_i|D) \tag{5.3.1}$$

To determine the posterior model probability of a model equation 5.3.2 is used. For this determination an equal probability is assumed among all models (equation 5.3.3).

$$P(M_i|D) = \frac{P(M_i)P(D|M_i)}{\sum_{j=1}^{N} P(M_j)P(D|M_j)} \tag{5.3.2}$$

$$P(M_i) = \frac{1}{N} \tag{5.3.3}$$

Adjustment Factors Approach

The Adjustment Factors Approach is a method that uses expert opinions to quantify model-form uncertainty using an adaption of Bayes' Theorem. A model probability is assigned to each model which is based on expert judgement regarding for each individual accuracy relative to the other models. This will result in a probability of being the best model $P(M_i)$ for each model. With this probability value a factor can be determine to adjust the best model.

6.3.3.3 Predictive uncertainty quantification

To quantify predictive uncertainty experimental data must be available. When these data is available Bayesian Model Averaging can be used to calculate posterior model distributions in order to quantify predictive uncertainty.

7 METRICS SELECTED FOR PRELIMINARY FIDELITY ASSESSMENT OF LH MODELS

7.1 INTRODUCTION

The following sections address the preliminary flight simulation fidelity metrics for the paragraphs of CS-29 selected in WP2 of the RoCS project. Each section starts with a description of the requirement, the related compliance demonstration and the presumed role/scope of flight simulation therein. Next, the proposed preliminary predictive fidelity metrics are presented along with a description of the validation flight test data that is presumed to be available for their evaluation. The proposed metrics consider only objective criteria, in both frequency and time-domain, that can be evaluated through desktop simulation (overall simulation fidelity assesses through subjective pilot assessment is addressed in D4.1). It is emphasized that the proposed metrics are preliminary and will be subject to continued evaluation within RoCS WP3 based on available AW109 models and flight test data. No distinction is yet made between the application to helicopters and tiltrotors; it is presumed that the same metrics (or at least a similar approach) can be applied to both types of aircraft.

7.2 CRITICAL AZIMUTH: CS-29.143(C)

7.2.1 Compliance demonstration

CS-29.143(c) states that:

> "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** on or near the ground in **any manoeuvre appropriate to the type** (such as crosswind take-offs, sideward flight, and rearward flight), with:
>
> (1) Weight selected by the applicant;
> (2) Critical centre 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 requirement refers to any manoeuvre appropriate to the type. In order to reduce the scope, only IGE hover will be considered herein. The same approach and models are expected to be suitable also for OGE conditions.

Flight test demonstration of cross-wind hover is typically achieved using a pace car rather than relying on steady cross-wind atmospheric conditions. The flight test compliance demonstration involves the substantiation of:

- Control and power margin
- Directional control authority (recognizable yaw response)
- Control inceptor forces and freedom of motion
- Handling qualities

The guidance does not provide a definition of what constitutes a recognizable yaw response, but the compliance flight testing does specifically address this part of the requirement by demonstrating yaw response to directional control inputs from stable cross-wind hover.

In the following, no metrics are defined for control inceptor forces and motion or other HMI aspects (an irreversible control system is assumed). It is assumed that such aspects will be addressed through flight or ground testing.

Handling qualities assessment is expected to require a piloted simulation with a subjective pilot rating methodology looking at all aspects that affect handling qualities. The performance parameters (control/power margin and yaw response), on the other hand, can be addressed in a desktop simulation.

7.2.2 Preliminary model predictive fidelity metrics

It is assumed that time-history flight test validation data are available as a basis for the evaluation of fidelity metrics for wind from all azimuths (or pace-car equivalent) at low-altitude conditions and unspecified (i.e., not necessarily critical) loading and rotor speed, at least for hover. Flight testing at high altitude will likely continue to be required to spot-check the simulation predictions and to address aspects not captured by the simulation such as exhaust gas re-ingestion and compressor stall.

The proposed preliminary simulation model predictive fidelity metrics for CS-29.143(c) are summarized in Table 2 and Table 3. The time-domain metrics are evaluated based on two separate trajectory simulations (time-histories), one that replicates the flight test compliance demonstration for the requirement and the other that reproduces the ADS-33E-PRF flight test to determine the collective-to-yaw response in hover. As a reference for tolerance definition, Figure 7 presents the time-domain collective-to-yaw response coupling requirements from ADS-33E-PRF.

The parameters identified in Table 2 and Table 3 with a tolerance equal to “Null” are those that are presumed to be driving the simulation. In other words, the simulation shall be set up to match the validation flight test data nearly exactly for those parameters. The remaining free parameters form the basis for the fidelity assessment. CT&M specifies a requirement to match the validation data in terms of Correct Trend & Magnitude.

The tolerances on the attitudes have been adopted from CS-FSTD(H). The reasoning is that these parameters are not of primary concern for aircraft performance, but are relevant for the piloting task in the simulator (i.e., for handling qualities assessment). The application of tolerances accepted for pilot training applications, therefore, appears reasonable, at least as initial baseline. In the event that the hover pitch/bank attitude approaches the maximum attitudes for landing, the error tolerance may need to be re-evaluated.

To verify that the control dynamics and handling qualities are adequately simulated, the application of frequency-domain metrics is considered of added value. Given the fact that the (hover) control task in cross-wind involves small-amplitude high-frequency closed-loop control inputs in all axes, the low-speed/hover pedal-to-yaw bandwidth and phase-delay[1] as defined in ADS-33E-PRF are interesting

[1] Bandwidth represents the range of frequencies in which control inputs do not threaten the closed-loop stability of the system. Phase delay is the slope of the phase between attitude and control at ‘high’ frequency and is affected by,

parameters to compare against flight test. The related definition for handling quality levels is shown in Figure 8 for reference. The bandwidth and phase delay for the other axes may be of interest as well.

Note that the frequency response validation data is not expected to include wind due to the difficulty of obtaining such data, especially near the ground.

Interaxis pitch-roll coupling is another relevant factor affecting handling qualities in hover. As such, the related ADS-33E-PRF criterion (for target acquisition and tracking tasks) is also a potential basis for assessing simulation fidelity. The criterion for yaw rate response to lateral gusts may be of interest as well, but dedicated validation flight test data is not likely to be available.

Table 2: Proposed time-domain model predictive fidelity metrics for CS-29.143(c)

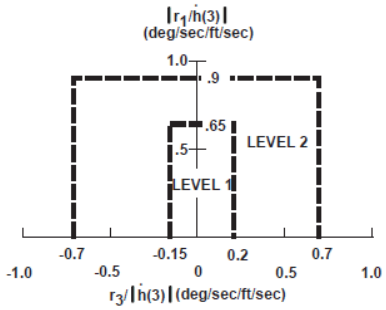
Type	Parameter	Error tolerance	Remarks
Time-history: pace car or cross-wind hover with pedal inputs	Height above ground	Null	
	Rate of climb	Null	
	Control position (all axes)	Allowable error*	
	Engine torque	Allowable error*	
	Peak yaw rate (on-axis)	Allowable error*	
	Pitch attitude	±1.5°	CS-FSTD(H)
	Bank angle	±2°	CS-FSTD(H)
	Wind	Null	
Time-history: collective to yaw response (from hover)	Rate of climb ($\dot{h}$)	TBD	ADS-33E-PRF §3.3.9.1
	Peak yaw rate (r_1 and r_3)	TBD	ADS-33E-PRF §3.3.9.1
	Control position (all axes)	±10%	
	Engine torque	CT&M	
	Pitch attitude	CT&M	
	Bank angle	CT&M	
	Wind	Null	No wind

*: Predicted distance to compliance limit at certification condition.

Table 3: Proposed frequency-domain model predictive fidelity metrics for CS-29.143(c)

Type	Parameter	Error tolerance	Remarks
Frequency response: yaw control in hover w/o wind	Pedal-to-yaw bandwidth ω_{BW}	TBD	ADS-33E-PRF §3.3.5.1
	Pedal-to-yaw phase delay τ_p	TBD	
Frequency response: pitch-to-roll and roll-to-pitch in hover	Average p/q	TBD	ADS-33E-PRF §3.3.9.3
	Average q/p	TBD	

e.g., the equivalent time delay between attitude response and pilot control input. Both parameters need to be included to uniquely define the handling qualities of the aircraft.[7]



Where:

r_1 = first peak (before 3 seconds) or r(1) if no peak occurs before 3 seconds

r_3 = r(3) - r_1 for $r_1 > 0$, or r_1 - r(3) for $r_1 < 0$

r(1) and r(3) are yaw rate responses measured at t = 1 and 3 seconds, respectively, following a step collective input at t = 0

In the unlikely event of more than one peak before 3 seconds, the largest peak (by magnitude) should be designated as r_1

Figure 7: ADS-33E-PRF time-domain requirement for yaw-due-to-collective coupling.

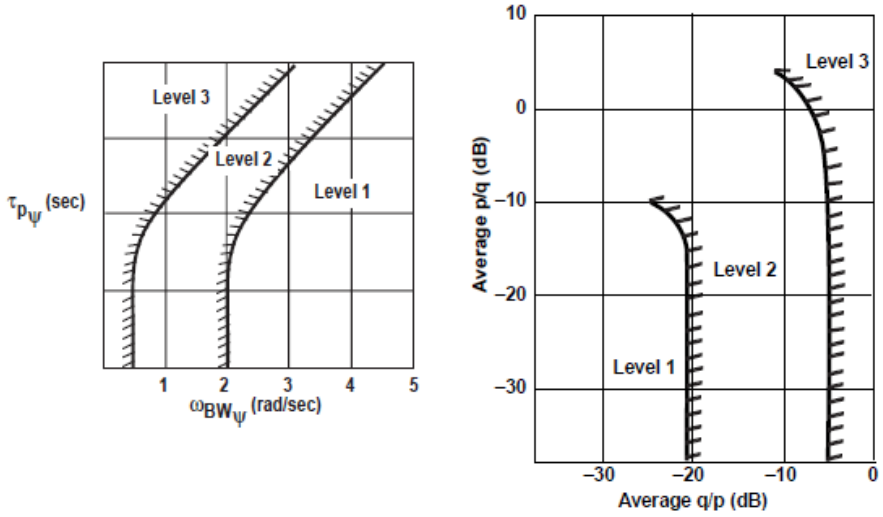


Figure 8: ADS-33E-PRF frequency-domain requirements for hover/low-speed yaw control bandwidth and phase delay (left) and pitch-due-to-roll/roll-due-to-pitch interaxis coupling (right).

7.3 CAT-A TAKE-OFF: CS-29.53(A)

7.3.1 Compliance demonstration

CS-29.53(a) states that:

> "The take-off performance must be determined and scheduled so that, if **one engine fails** at any time after the start of take-off, the rotorcraft can:
>
> (a) **Return to and stop safely on, the take-off area**;"

In relation to this requirement, CS-29.51 states that:

> "(a) The take-off data required by CS 29.53, 29.55, 29.59, 29.60, 29.61, 29.62, 29.63 and 29.67 must be determined:
>
> (1) At each weight, altitude, and temperature **selected by the applicant**; and

(2) With the operating engines within approved operating limitations.

(b) Take-off data must:

(1) Be determined on a **smooth, dry, hard surface**; and

(2) Be corrected to assume a level take-off surface.

(c) No take-off made to determine the data required by this paragraph may require **exceptional piloting skill or alertness**, or exceptionally favourable conditions."

The flight test compliance demonstration aims to show that the performance and handling qualities of the aircraft are adequate for an average pilot to perform the specified take-off procedure in a safe and repeatable manner up to the maximum take-off weight specified by the applicant. This includes an evaluation of the robustness against variations in piloting procedure through so-called Abuse Case Testing.

In practice, the following compliance demonstration parameters are involved:

- Power/NR margin
- Control margins
- Landing and drive system loads
- Helicopter position / height above ground
- Handling qualities
- Pilot workload

Adherence to landing gear and drive system load limitations is not demonstrated by a direct load measurement, but through specified limits for touchdown sink rate and transient rotor speed. As such, fidelity assessment of in-flight loads predictions is not required.

It is presumed that performance aspects can be adequately demonstrated through desktop simulation. The assessment of handling qualities, and potentially pilot workload, require a piloted simulation (using the same flight-model) with a representative cockpit HMI.

7.3.2 Preliminary model predictive fidelity metrics

It is assumed that AEO/OEI trim and Category A take-off time-history flight test data is available for validation (instead of typical control response data as proposed for CS-29.143(a)). In addition, system identification flight test data will be required to evaluate the fidelity in the frequency-domain.

The proposed time-domain fidelity metrics are presented in Table 4. The metrics are derived from AEO/OEI trim conditions and the rejected take-off trajectory from engine failure at/before the TDP down to a height of 10 ft AGL[2]. In addition, IGE hover and level fly-by simulations are defined to assess the fidelity of the ground effect modelling. For certification purposes the metrics shall be evaluated for a representative range of GW/σ and head/cross-wind.

The tolerances on the attitudes and the control positions other than collective, mostly relevant for the piloted handling qualities simulations, have been adopted from CS-FSTD(H). The tolerances on collective,

[2] Simulation of touchdown not necessary if entry conditions at 10 ft AGL are within safe envelope (ref. AW189 certification approach).

torque and rotor speed are specified as a (arguably arbitrary) fraction of the specified compliance limit. In case of AEO climb, the error tolerance on torque is not specified as a fixed percentage, but is related to the allowable impact of this error on the fly-away height loss following single-engine failure during take-off. This approach enables a tolerance specification that can be compared to the compliance limit that states a minimum obstacle clearance of 35 ft.

Note that, considering the fact that the Category A maximum take-off weight is typically defined to achieve maximum performance and, thus, minimum allowable margins with the compliance limits (e.g., touchdown sink rate), no reference is made to "allowable error". Instead, it is proposed that the compliance limits be adjusted to account for the maximum demonstrated prediction error (based on available flight test data). For instance, if the OEI touchdown sink speed is underpredicted by 100 ft/min (within the validated part of the envelope), the maximum allowable touchdown sink speed shall be reduced by *at least* 100 ft/min when demonstrating compliance using simulation.

The tolerances specified in Table 4, unless defined as "±...", do not necessarily need not to be applied symmetrically around the validation flight test data. Prediction error that is larger than the indicated tolerance but in the conservative direction, perhaps intentionally so, may be acceptable for certification purposes.

The OEI trajectory simulation is presumed to be set up to match the OEI rotor speed (hence the tolerance on rotor speed is "null"). This allows the height (loss) and rate of descent to be exploited as fidelity metrics, both of which can be related to compliance limits relevant for the certification task.

The frequency-domain metrics in Table 5 are aimed at assessing the fidelity of the simulation of the aircraft handling qualities, and by extension the pilot workload, prior to compliance demonstration of these aspects through piloted simulations. The metrics emphasize pitch control and pitch-roll interaxis coupling as the most relevant to the longitudinal control task. Conceivably, other off-axis frequency responses (e.g., collective to pitch/yaw) could be added as well. Height and rotor speed control, both fairly low-frequency, are assessed through the time-domain metrics in Table 4.

Table 4: Proposed time-domain model predictive fidelity metrics for CS-29.53(a) – rejected take-off

Type	Parameter	Error tolerance	Remarks
Trim: AEO climb	Rate of climb	Null	
	Airspeed	Null	
	Torque	5 ft OEI height loss	Obstacle clearance: 35 ft
	Rotor speed	Null	CS-FSTD(H)
	Pitch attitude	±1.5°	CS-FSTD(H)
	Bank angle	±1.5°	CS-FSTD(H)
	Collective position	Eq. to torque tolerance	
	Cyclic & pedal position	±10%	CS-FSTD(H)

Type	Parameter	Error tolerance	Remarks
Trim: 30s OEI descent	Rate of descent	Null	
	Airspeed	Null	
	Torque (operative engine)	Null	Transmission limit
	Rotor speed (NR)	Energy: 10% TD RoD limit	
	Pitch attitude	±1.5°	CS-FSTD(H)
	Bank angle	±1.5°	CS-FSTD(H)
	Heading	±2°	CS-FSTD(H)
	Collective position	Eq. to NR tolerance	
	Cyclic & pedal position	±10%	CS-FSTD(H)
Trim: IGE hover & fly-by	Height above ground	Null	
	Airspeed	Null	
	Torque	Climb/descent power: 10% TD V/S limit	
	Rotor speed	Null	
	Pitch attitude	±1.5°	CS-FSTD(H)
	Bank angle	±1.5°	CS-FSTD(H)
	Collective position	Eq. to torque tolerance	
	Cyclic & pedal position	±10%	CS-FSTD(H)
Trajectory: rejected take-off from TDP down to 10 ft AGL	Height above TO surface	10% min. obstacle clearance	Obstacle clearance: 35ft
	Rate of descent (OEI)	10% TD RoD limit	
	Airspeed	Null	
	Torque	Null	Transmission limit
	Rotor speed	Null	
	Pitch attitude	±1.5°	CS-FSTD(H)
	Bank angle	±1.5°	CS-FSTD(H)
	Heading	±2°	CS-FSTD(H)
	Collective position	Eq. to RoD tolerance	
	Cyclic & pedal position	±10%	CS-FSTD(H)

Table 5: Proposed frequency-domain model predictive fidelity metrics for CS-29.53(a) – rejected take-off

Type	Parameter	Error tolerance	Remarks
Frequency response: pitch control in hover/low-speed	Cyclic-to-pitch bandwidth ω_{BW} Cyclic-to-pitch phase delay τ_p	TBD TBD	ADS-33E-PRF §3.3.2.1
Frequency response: pitch-to-roll and roll-to-pitch in hover	Average p/q Average q/p	TBD TBD	ADS-33E-PRF §3.3.9.3

7.4 DYNAMIC STABILITY IFR: CS-29 APP B PAR VI

7.4.1 Compliance demonstration

CS-29 App B Par VI states that:

> "Dynamic stability
>
> (a) Any oscillation having a period of less than 5 seconds must damp to ½ amplitude in not more than one cycle.
>
> (b) Any oscillation having a period of 5 seconds or more but less than 10 seconds must damp to ½ amplitude in not more than two cycles.
>
> (c) Any oscillation having a period of 10 seconds or more but less than 20 seconds must be damped.
>
> (d) Any oscillation having a period of 20 seconds or more may not achieve double amplitude in less than 20 seconds.
>
> (e) Any aperiodic response may not achieve double amplitude in less than 9 seconds."

Figure 9 from the associated guidance material in AC 29-2C translates the requirement into a chart of damping ratio against oscillation period including a defined boundary of acceptable dynamic stability for IFR. The compliance demonstration parameters are, therefore, the oscillation period and damping ratio (or equivalently the time to half/double amplitude) of the dynamic modes of the stabilized aircraft.

The requirement must be demonstrated in the envelope approved for IFR flight, from V_{MINI} to $VNE_{IFR,}$ over the complete altitude range. Flight simulation is presumed to be applied to extrapolate test data to the corners of the envelope.

An advantage of the application of simulation for compliance demonstration lies in the flexibility of defining control inputs and external disturbances (e.g., turbulent gusts) to excite the dynamic modes of the aircraft, which is a particularly difficult task for the flight test pilot.

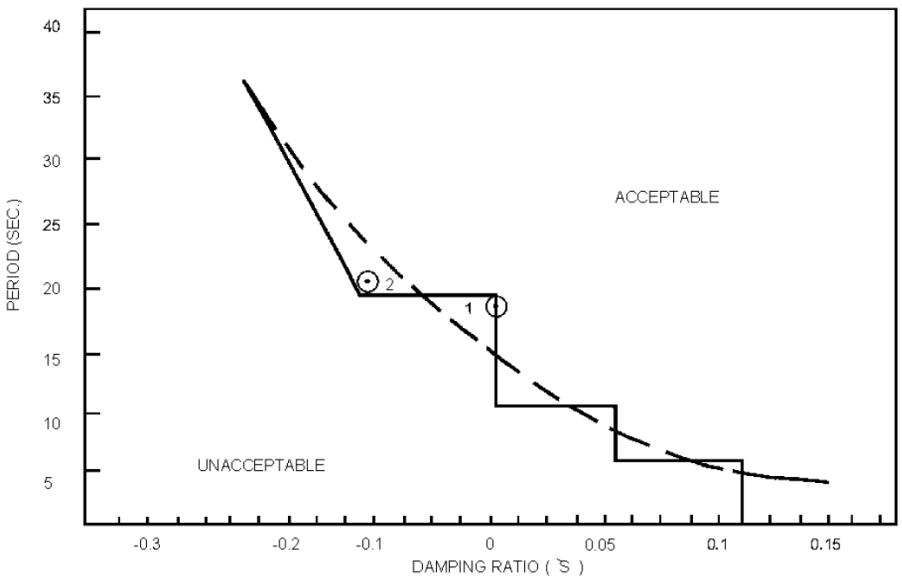


Figure 9: Rotorcraft dynamic stability requirements for IFR from AC 29-2C Appendix B.

7.4.2 Preliminary model predictive fidelity metrics

The available validation flight test data is assumed to include time-history compliance demonstration flight test data obtained at low altitude from V_{MNI} to VNE_{IFR} (i.e., requiring altitude-extrapolation). Use of flight test data in turbulent gusts would be of interest for validation purposes, but the measurement of the gust field is expected to be problematic.

Table 6 presents the fidelity metrics for periodic and aperiodic responses (in any axes). In both cases, the control inputs used to excite the dynamic modes of the aircraft in flight test shall be replicated in the simulation to ensure that there is no undue effect of the size of the input on the predicted damping ratio or time to half/double amplitude. The simulation shall display correct trend and magnitude in the time-domain to ensure that not only the period and damping is matched, but also the multi-axis motion itself.

The identified tolerances on period, time to double/half amplitude, and damping ratio are based on CS-FSTD(H). In practice, an alternative is to adjust the compliance limit defined by the boundary in Figure AC 29.APX B-1 to account for the maximum demonstrated error in predicted oscillation period and damping ratio.

Table 6: Proposed time-domain model predictive fidelity metrics for CS-29 App B Par VI – Dynamic stability IFR

Type	Parameter	Error tolerance	Remarks
Time-history: aperiodic response	Airspeed Height Torque Control positions (from trim) Pitch/bank angle	CT&M CT&M CT&M Null ±2° or ±10% in 20 sec	 CS-FSTD(H)
Time-history: periodic response	Airspeed Height Torque Control positions (from trim) Attitude angles Oscillation period Time to ½ or double amplitude Damping ratio	CT&M CT&M CT&M Null Equiv. to 10% time to ½ or double amplitude ±10% ±10% ±0.02	 CS-FSTD(H) CS-FSTD(H) CS-FSTD(H)

7.5 STABLITY AUGMENTATION SYSTEM (SAS) FAILURE: CS-29 APP B PAR VII(A)(1)

7.5.1 Compliance demonstration

CS-29 App B Par VII(a)(1) states that:

> "Stability augmentation system (SAS)
>
> (a) If a SAS is used, the reliability of the SAS must be related to the effects of its failure. Any **SAS failure condition** that would prevent continued safe flight and landing must be extremely improbable. It must be shown that, for any failure condition of the SAS which is not shown to be extremely improbable:
>
> (1) The helicopter is **safely controllable** when the **failure or malfunction** occurs at **any speed or altitude within the approved IFR operating limitations**;"

The requirement for the aircraft remaining safely controllable after SAS failure implies demonstration of adequate:

- control and power margin, and
- handling qualities in degraded SAS state

to achieve safe recovery and establish continued safe flight within a practical flight envelope and respecting structural load limits.

7.5.2 Preliminary model predictive fidelity metrics

A key benefit of flight simulation for the compliance demonstration of this requirement is that it allows the demonstration of failure modes that cannot be tested in flight (for practical or flight safety reasons). A

consequence of this is that validation flight test data of the actual failure cannot generally be expected to be available. Instead, metrics need to be defined based on more generic flight test data such as control input response data.

Table 7 presents the proposed time and frequency-domain fidelity metrics. Time-domain metrics are derived from the aircraft response following a control stick step input representative in magnitude of a full-stroke SAS actuator displacement (as may occur due to a failure). The result is expected to be moderate to large-amplitude motion in the failed axis. As such, the tolerances specified refer to the on-axis response. The off-axis response is required to display correct trend and magnitude. Note that related metrics may equivalently be defined in terms of attitude quickness and maximum response as discussed in §5.3.

The frequency-domain metrics are restricted to the (nominal SAS) on-axis bandwidth and phase delay, as relevant for closed-loop (small-amplitude) control tasks. Metrics based on open-loop stability (quantified by the frequencies and damping of the aircraft dynamic modes) may be of interest, but validation data is unlikely to be available for the aircraft in degraded SAS state.

Table 7: Proposed time and frequency-domain model predictive fidelity metrics for CS-29 App B Par VII – SAS failures

Type	Parameter	Error tolerance	Remarks
Time-history: Full-stroke step input control response (single-axis)	Airspeed Height Torque Control positions (from trim) Pitch angle change Bank angle change Heading change	CT&M CT&M CT&M Null ±10% or ±1.5° ±10% or ±3° ±10% or ±2°	 CS-FSTD(H) CS-FSTD(H) CS-FSTD(H)
Frequency-domain: on-axis response	Bandwidth ω_{BW} Phase delay τ_p	TBD TBD	

8 REFERENCES

[1] "Certification Specifications for Helicopter Flight Simulation Training Devices", CS-FSTD(H), European Aviation Safety Agency, Initial Issues, June, 2012.

[2] A.M. Dequin et al, "Final Report of Action Group HC AG06: Mathematical Modelling for the Prediction of Helicopter Flying Qualities", GARTEUR/TP-075/2, April, 1997.

[3] H. Haverdings et al, "Final Report of Action Group HC AG09: Mathematical Modelling for the Prediction of Helicopter Flying Qualities- Phase 3", GARTEUR HC/AG09 TP-116, November, 1999.

[4] R.E. Smith, J. Hodgkinson, and R.C. Snyder, "Equivalent System Verification and Evaluation of Augmentation Effects on Fighter Approach and Landing Flying Qualities" U. S. Air Force Wright Aeronautical Lab., AFWAL-TR-81-3116, Vol. 2, Wright-Patterson Air Force Base, Ohio, 1981.

[5] K. Strope, C. Borden and J. Harding, "Verification & Validation of a UH-60 FLIGHTLAB Model in Support of the UH-60M Limited User Test", American Helicopter Society 60th Annual Forum, Maryland, June, 2004.

[6] D.G. Mitchell, C. He, and K. Strope, "Determination of Maximum Unnoticeable Added Dynamics", Atmospheric Flight Mechanics Conference and Exhibit, Colorado, August, 2006.

[7] G.D. Padfield, "Helicopter Flight Dynamics; including a treatment of tilt rotor aircraft", 3rd edition, Wiley Publishing, 2018.

[8] M.C. Romanowski, "Flight Test Guide for Certification of Transport Category Airplanes", AC 25-7D, initiated by AIR-670, April, 2018.

[9] G.D. Padfield, M. Pavel, D. Casolaro, G. Roth, M. Hamers and A. Taghizad, "Validation Criteria for Helicopter Real-Time Simulation Models, Sketches From the Eork of GARTEUR HC/AG-12", 30th European Rotorcraft Forum, Marseilles, France, 2004.

[10] G.D. Padfield, M. Pavel, D. Casolaro, G. Roth, M. Hamers and A. Taghizad, "Fidelity of Helicopter Real-Time Models", 61st Annual Forum of the American Helicopter Society, Grapevine, TX, USA, June, 2005.

[11] M. Pavel, M.D. White, G.D. Padfield, G. Roth, M. Hamers and A. Taghizad, "Validation of Mthematical Models for Helicopter Flight Simulators Current and Future Challenges ", The Aeronautical Journal, Royal Aeronautical Society, Volume 117, Number 1190, pp. 343 – 388 April 2013.

[12] Anon., "Helicopter Flight Simulators", JAR-STD 1H, Joint Aviation Authorities, April, 2001.

[13] Anon., “Helicopter Flight Training Devices”, JAR-STD 2H, Joint Aviation Authorities, September, 2003.

[14] Anon., “Helicopter Flight Simulators”, JAR-STD 3H, Joint Aviation Authorities, May, 2002.

[15] Anon., “Helicopter Flight Simulation Training Devices”, JAR-FSTD H, Joint Aviation Authorities, May, 2008.

[16] R.A. Hess and T. Malsbury, “Closed Loop Assessment of Flight Simulator Fidelity”, Journal of Guidance, Control and Dynamics, 14, (1), pp 191–197, 1991.

[17] R. A. Hess and W. Siwakosit, “Assessment of Flight Simulator Fidelity in Multi-Axis Tasks Including Visual Cue Quality”, Journal of Aircraft, 38, (4), pp 607–614, 2001.

[18] J.A. Schroeder, W.W.Y. Chung and R.A. Hess, “Evaluation of Motion Fidelity Criterion with Visual Scene Changes”, Journal of Aircraft, 37, (4), pp 580–587, 2000.

[19] G.D. Padfield, M.T. Charlton and A.T. McCallum,” The Fidelity of Hi-Fi Lynx on the DERA Advanced Flight Simulator using ADS-33 Handling Qualities Metrics”, DRA/AS/FDS/TR96103/1, pp 1-152, December, 1996.

[20] A.T. McCallum and M.T. Charlton, “Structured Approach to Helicopter Simulator Acceptance, The Challenge of Realistic Rotorcraft Simulation”, November RAeS conference, London, UK, 2001.

[21] Anon., “Handling Qualities Requirements for Military Rotorcraft”, ADS-33E-PRF, US Army, March, 2000.

[22] S.K. Advani and C.H. Wilkinson, “Dynamic Interface Modelling and simulation – a Unique Challenge”, Royal Aeronautical Society Conference on Helicopter Flight Simulation, London, UK, 2001.

[23] M.F. Roscoe and J.H. Thompson, “JSHIP’s Dynamic Interface Modelling and Simulation Systems: A Simulation of the UH-60A Helicopter/LHA Shipboard Environment Task”, 59th Annual Forum of the American Helicopter Society, Phoenix, AZ, USA, May, 2003.

[24] P. Perfect, M.D. White, G.D. Padfield and A.W. Gubbels, “Rotorcraft Simulation Fidelity: New Methods for Quantification and Assessment”, The Aeronautical Journal, Vol. 117, Issue 1189 pp. 235-282, March, 2013.

[25] M.D. White, P. Perfect, G.D. Padfield, A.W. Gubbels and A.C. Berryman, “Acceptance Testing and Commissioning of a Flight Simulator for Rotorcraft Simulation Fidelity Research”, Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, Volume 227 Issue 4, pp. 663 – 686, April, 2013.

[26] M.D. White, P. Perfect, G.D. Padfield, A.W. Gubbels and A.C. Berryman, “Progress in the Development of Unified Fidelity Metrics for Rotorcraft Flight Simulators”, 66th Annual Forum of the American Helicopter Society, Phoenix, AZ, USA, May, 2010.

[27] M.D. White, P. Perfect, G.D. Padfield, A.W. Gubbels and A.C. Berryman, “Integrating Predicted and Perceived Fidelity for Flight Simulators”, 36th European Rotorcraft Forum, Paris, France, September, 2010.

[28] P. Perfect, E. Timson, M.D. White, G.D. Padfield, R. Erdos and A.W. Gubbels, “A Rating Scale for the Subjective Assessment of Simulation Fidelity”, The Aeronautical Journal, August, Volume 11, No 1206, pp. 953 – 974, 2014

[29] G.D. Padfield, “Helicopter Flight Dynamics: Including a Treatment of Tiltrotor Aircraft”, Third Ed. John Wiley & Sons, 2018.

[30] G.D. Padfield and M.D. White, “Measuring Simulation Fidelity through an Adaptive Pilot Model”, Aerospace Science and Technology, Vol. 9, Issue 5, pp. 400 – 408, July, 2005.

[31] G.D. Padfield amd M.D. White, “Simulation Fidelity Criteria using an Adaptive Pilot Model”, 29th European Rotorcarft Forum, Friedrichshafen, Germany, September, 2003.

[32] D.T. McRuer and E.S. Krendel, “Mathematical Models of Human Pilot Behaviour”, AGARD AG 188, 1974.

[33] R.K. Heffley, “A Pilot-in-the-Loop Analysis of Several Kinds of Helicopter Acceleration-Deceleration Manoeuvres”, NASA CP 2216, Helicopter Handling Qualities, 1982.

[34] “Certification Specifications and Acceptable Means of Compliance for Large Rotorcraft”, CS-29, European Union Aviation Safety Agency, Amendment 6, December, 2018.