



HAL
open science

Simulation Fidelity Assessment for Rotorcraft Methods and Metrics Sketches from the Work of NATO AVT-296

Marilena Pavel, Mark Tischler, Mark White, Olaf Stroosma, Michael Jones,
David Miller, Vincent Myrand-Lapierre, Michel Nadeau-Beaulieu, Armin
Taghizad

► To cite this version:

Marilena Pavel, Mark Tischler, Mark White, Olaf Stroosma, Michael Jones, et al.. Simulation Fidelity Assessment for Rotorcraft Methods and Metrics Sketches from the Work of NATO AVT-296. 77th Annual Vertical Flight Society Forum and Technology Display, FORUM 2021: The Future of Vertical Flight, May 2021, Virtuel, United States. hal-03592354

HAL Id: hal-03592354

<https://hal.science/hal-03592354>

Submitted on 1 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Simulation Fidelity Assessment for Rotorcraft – Methods and Metrics

Sketches from the Work of NATO AVT-296

Marilena D. Pavel
Delft University of Technology
Delft, The Netherlands

Mark B. Tischler¹
US Army DEVCOM
AvMC
Moffett Field, U.S.A.

Mark White
The University of Liverpool
Liverpool, U.K.

Olaf Stroosma
Delft University of Technology
Delft, The Netherlands

Michael Jones
German Aerospace Center
Braunschweig, Germany

David Miller
The Boeing Company
Philadelphia, U.S.A.

Vincent Myrand-Lapierre
CAE, Global Engineering
St-Laurent, Québec, Canada

Michel Nadeau-Beaulieu
CAE, Global Engineering
St-Laurent, Québec, Canada

Armin Taghizad
ONERA
Salon-de-Provence, France

ABSTRACT

Fidelity is a fundamental concept for crucial decision-making processes during engineering design, test and evaluation of aircraft and rotorcraft. Fidelity, as applied to rotorcraft, proves to be a prerequisite for most human factors research and simulator development programs. This paper reviews the concept of fidelity, its metrics and methods concentrating on rotorcraft applications. The paper stems from the Applied Vehicle Technology (AVT) Panel of the NATO Science and Technology Organization (STO) work of NATO AV-296 concentrating on recent advancements on rotorcraft flight simulation model fidelity prediction and assessment. Starting from the question of what is meant with the term simulation fidelity, the paper will demonstrate that even though the term fidelity can be loosely translated as simulation goodness or faithfulness to reality, when it comes to fidelity assessment, this concept has a multidimensional and multifaceted character depending on the configuration considered. A key message of the paper is that simulation fidelity should be decomposed in different dimensions and attributes with their own metrics. Using multiple fidelity metrics is therefore meaningful for assessing simulation fidelity in general and in particular for the case of rotorcraft.

1. INTRODUCTION

Throughout the life cycle of any rotorcraft, flight dynamics simulation models are developed and used for different purposes such as: aircraft design analysis, control system design and development, and simulator qualification. Simulation fidelity of such models remains the key for success in application as this enables: 1) increased confidence in predictions of aircraft behavior to inform the design processes; 2) increased efficiency and safety of flight envelope expansion during development process (including through life support); 3) cost and time savings during qualification and certification processes and 4) more effective and efficient pilot, crew, and mission training.

The Applied Vehicle Technology (AVT) Panel of the NATO Science and Technology Organization (STO) has recently engaged the rotorcraft industry, research laboratories and universities through the NATO AV-296 Research Task Group (Tischler et al., 2021) on the topic of “Rotorcraft Flight Simulation Model Fidelity Improvement and Assessment”. The focus of this group was mainly on developing and documenting various system identification methods for updating flight dynamics mathematical models with flight data. One of the activities performed within NATO AVT-296 consisted of understanding the methods and metrics used for simulation fidelity assessment as suitable for the final intent of the model. The goal of the present paper is to review the concept of fidelity and define the methods and metrics to be

¹ Retired from US Army; currently, Tischler Aeronautics; mrtischler@aol.com

DISTRIBUTION STATEMENT A. Approved for public release.

Presented at the Vertical Flight Society’s 77th Annual Forum & Technology Display, Virtual, May 10-14, 2021. Copyright © 2021 by the Vertical Flight Society. All rights reserved.

used for rotorcraft simulation fidelity. Questions will be answered such as:

- what is model fidelity and why is this needed,
- how to measure simulation fidelity and what are common metrics,
- what is the process of qualifying the fidelity of helicopter simulators,
- how to assess the simulator motion cueing fidelity,
- what are the connotations for the term fidelity in the context of flight simulation and how can these be assessed?

Section 1.1 introduces the concept of simulation fidelity showing its multifaceted significance. Section 2 introduces different metrics and methods that can be used to quantify simulation fidelity. In this sense, the concept of integrated cost functions as metrics in the frequency and time domains for assessing the fidelity of a simulation model are introduced. Also, the ValCrit-T and ValCrit-F validation criteria for assessing the simulation model fidelity are presented in this section. Section 3 presents the concept of Maximum Unnoticeable Added Dynamics (MUAD) and Allowable Error Envelopes (AEE) are introduced for determining 1) whether the math model is adequate or 2) if a simulation facility is adequate to accomplish proposed evaluations. Section 4 moves towards the simulation fidelity as related to a simulator device, concentrating on the definition, cueing and motion cueing fidelity as part of simulator device acceptance. Motion cueing metrics are herein reviewed, e.g. the Sinacori/Schroeder boundaries and the Objective Motion Cueing Test (OMCT). Section 5 discusses simulator predicted fidelity reviewing the Qualification Test Guide Performance Standards (QTG) for simulator qualification as a training device. Section 6 discusses perceptual fidelity presenting the concept of Simulation Fidelity Rating (SFR) scale for evaluating the fidelity of a simulation device for flight training. Section 7 presents methods for assessing the overall simulator fidelity based on handling qualities engineering. Finally, concluding remarks on the concept of fidelity and its metrics as used for rotorcraft applications end the paper.

1.1 Fidelity: A Multifaceted and Multi-dimensional Concept

Today, simulation is becoming the primary tool for crucial decision-making processes during engineering design, test and evaluation of new systems, safety-critical systems, and in training of people operating these systems. With increasing reliance on simulation results, it is important to know the validity and credibility of simulation results, or in other words, the simulation's fidelity. Indeed, fidelity is one of the most important concepts of any model or simulation development, and it is also one of the main cost-drivers of aircraft development. As a general rule, the higher the fidelity the more time and resource consuming the simulation development is. Thus, being able to state what level of fidelity is exactly required avoids unnecessary investments,

superfluous simulation components, and unusable simulations.

When trying to define fidelity in the context of aerospace, it appears that this concept has a multidimensional and multifaceted character depending on the application considered. A wide variety of interpretations and connotations are encountered in the literature, a representative sample of terms include such things as: validity, accuracy, error, granularity, resolution, precision, tolerance, verification and quality.

Firstly, the notion of fidelity can be related to the dynamics of the aircraft/rotorcraft i.e. predictive fidelity. In this sense, simulation fidelity belongs to the area of validity (accuracy) of a system's modeling.

Secondly, the notion of fidelity can be related to the flying qualities and handling qualities of the aircraft/rotorcraft. Historically, the first conversations about fidelity in aviation started in this area. Herein, NASA's Ad Hoc Advisory Subcommittee on Avionics, Controls, and Human Factors (1979) defines fidelity as "*the degree to which characteristics of perceivable states induce realistic pilot psychomotor and cognitive behavior for a given task and environment.*" Fidelity in this sense relates primarily to the effect upon the pilot—not to the effective aircraft dynamics—although both aspects are necessarily involved.

Thirdly, the notion of fidelity can be related to the ground-based flight simulation. Herein, the definition of fidelity is even more challenging: the terms of fidelity and cueing are often used together, their meaning interweaving with each other and onto other simulator characteristics. Definitions characteristics for the term fidelity in the context of a simulator include terms such as: equipment fidelity, environmental fidelity, task fidelity, functional fidelity, attribute fidelity, concrete fidelity, abstract fidelity, objective fidelity (also known as predictive fidelity, physical fidelity), subjective fidelity (perceptual fidelity, psychological/physiological fidelity). This variety of terms illustrates that fidelity is a difficult concept to attain agreement on, particularly when one considers the various levels of fidelity that may be present within the same simulator depending on the task performed.

2. FIDELITY METRICS FOR ROTORCRAFT FLIGHT DYNAMICS MODELLING

The most straightforward description of simulation fidelity is the one related to the system's dynamics. Fidelity (in the viewpoint of verification and validation) of vehicle physics is achieved in a straight-forward manner – by directly comparing simulation predictions with acquired flight test data from the real aircraft. Herein, verification can be related to the question whether the model do what it intends to do. As opposed to verification, validation asks whether the model is

correct (Mihram, 1972). For fidelity as related to flight dynamics applications, single integrated metrics are useful measures of the overall fidelity (precision) of the simulation model and quantify the simulation fidelity of a model. The frequency-domain metric J_{ave} indicates the overall integrated cost function based on the comparison of the simulation model and flight-test frequency responses. The time-domain metric J_{rms} indicates the overall integrated cost function for responses to a control input.

2.1 Model/Flight Data Mismatch and Integrated Cost Functions

2.1.1 Frequency-Domain Integrated Cost Function, J_{ave}

A frequency-domain metric useful for assessing the fidelity of a simulation model frequency response as compared to flight test data for a single input/single-output (SISO) frequency response pair (e.g. p/δ_a , q/δ_e , etc.) at a particular flight condition was originally proposed by Hodgkinson and fully covered in Hodgkinson (1998) based on the weighted sum of magnitude and phase squared errors:

$$J = \frac{20}{n_\omega} \sum_{\omega_1}^{\omega_{n_\omega}} W_\gamma [W_g (|\hat{T}_c| - |T|)^2 + W_p (\angle \hat{T}_c - \angle T)^2] \quad (1)$$

where:

$||$ = magnitude (dB) at each frequency ω

\angle = phase (deg) at each frequency ω

n_ω = number of frequency points (typically selected as $n_\omega = 20$)

ω_1 and ω_{n_ω} = starting and ending frequencies of fidelity assessment (typically covering 1–2 decades).

By selecting the n_ω frequency points $\omega_1, \omega_2, \dots, \omega_{n_\omega}$ with a uniform spacing on a log-frequency scale (rad/s), the fidelity metric or cost (J) well reflects the error as displayed on the Bode plot.

W_g and W_p are the relative weights for magnitude and phase squared-errors. The normal convention from USAF MIL-STD-1797B (2006) is to use the values $W_g = 1.0$ and $W_p = 0.01745$. This choice of weighting means that a 1-dB magnitude error is comparable with a 7.57-deg phase error. However, the fidelity metric is largely insensitive to the exact choice of these weighting values.

Tischler and Remple (2012) also included function W_γ to weight the fidelity metric more heavily when the flight data are

more reliable as determined from the coherence function at each frequency ω :

$$W_\gamma(\omega) = [1.58(1 - e^{-\gamma_{xy}^2})]^2 \quad (2)$$

For a coherence of $\gamma_{xy}^2 = 0.6$, this function reduces the weight on the squared errors by 50%. As a guideline for simulation fidelity and based on extensive experience, Tischler and Remple (2012) propose that a cost function of

$$\text{Guideline: } J \leq 100 \quad (3)$$

generally reflects an acceptable level of accuracy for flight-dynamics modelling and reflects a good simulation model response for rotorcraft. A cost function of $J \leq 50$ can be expected to produce a model that is nearly indistinguishable from the flight data in the frequency domain and time domain.

Tischler and Remple (2012) generalized the SISO cost function for a multi-input/multi-output (MIMO) matrix of output/input frequency response pairs for the simulation model $T(s)$ and flight data \hat{T}_c . The associated overall fidelity metric at the same flight condition is a direct extension of the SISO formulation of Equation (1) and is now simply the summed cost for the n_{TF} transfer functions:

$$J = \sum_{l=1}^{n_{\text{TF}}} \left\{ \frac{20}{n_\omega} \sum_{\omega_1}^{\omega_{n_\omega}} W_\gamma [W_g (|\hat{T}_c| - |T|)^2 + W_p (\angle \hat{T}_c - \angle T)^2] \right\}_l \quad (4)$$

In most cases, the matrix of flight-test responses will not have good data for several of the theoretically possible input-to-output combinations, as indicated by poor coherence for the entire frequency range of interest. Such responses are dropped entirely from the cost function. So only a subset n_{TF} of the frequency-response pairs will be included in the cost function. The frequency-response pairs retained in the identification are denoted by T_l , $l = 1, 2, 3, \dots, n_{\text{TF}}$. The choice of frequency-range ($\omega_1, \omega_{n_\omega}$) is made separately for each response pair T_l , corresponding to that pair's range of acceptable coherence.

The accuracy of a system identified model is best characterized by the average overall cost function or *integrated cost function*:

$$J_{\text{ave}} = \frac{J}{n_{\text{TF}}} \quad (5)$$

The weighting functions W_γ , W_g , and W_p all retain the same definitions as in the SISO formulation and are evaluated at each

frequency point $(\omega_1, \omega_2, \dots, \omega_{n_\omega})$ for each frequency-response pair T_l .

The interpretation of the MIMO fidelity metric extends directly from the SISO case, where an overall average cost function that achieves

$$\text{Guideline: } J_{ave} \leq 100 \quad (6)$$

generally reflects an acceptable level of accuracy for flight-dynamics modelling and is typically used rotorcraft model fidelity. Some of the individual cost functions, especially for the off-axis responses, can use the guideline of

$$\text{Guideline: } J_l \leq 150 \text{ to } 200 \quad (7)$$

without resulting in a noticeable loss of overall predictive accuracy.

2.1.2 Time-Domain Integrated Cost Function, J_{rms}

The time-domain integrated cost function J_{rms} is useful for assessing the predictive accuracy for a short-term doublet input, typically about 5-10 sec. The simulation response is determined from direct numerical integration of the equations of motion of the simulation model using measured control inputs from the flight data. The simulation model outputs are compared with the flight-data measurements and should not include reconstructed signals. The integrated time-domain cost function is determined from Tischler and Remple (2012):

$$J_{rms} = \sqrt{\left(\frac{1}{n_t \cdot n_o}\right) \sum_{i=1}^{n_t} [(y_{data} - y)^T (y_{data} - y)]} \quad (8)$$

where

y_{data} = time-history measurement vector from the flight data

y = simulation model time-history response vector

n_t = number of time-history points in the time history data record

n_o = number of outputs (measurement signals) in the time history vector, y

A good rule of thumb is to select the units of the measurement vector for SI units as

$$\text{deg, deg/sec, m/sec, m/sec}^2 \quad (9)$$

and for Imperial units as

$$\text{deg, deg/sec, ft/sec, ft/sec}^2 \quad (10)$$

Experience shows that a J_{rms} value in the range of Tischler and Remple (2012)

$$\text{Guideline: } J_{rms} \leq 1.0 \text{ to } 2.0 \quad (11)$$

generally reflects an acceptable level of accuracy for flight-dynamics modelling when the error function is calculated based on the units of Equation (9) or Equation (10).

2.2 ValCrit-T and ValCrit-F Criteria

The Group for Aeronautical Research and Technology in Europe (GARTEUR) has supported a number of collaborative activities aimed at improving the predictive capability of rotorcraft modelling. One of the main objectives of GARTEUR HC/AG-06 on “*Mathematical Modelling for the Prediction of Helicopter Flying Qualities*” [Padfield et al. (1996)] and GARTEUR HC/AG-09 on “*Mathematical Modelling For The Prediction Of Helicopter Flying Qualities*” [Haverdings et al. (2000)] was to derive new criteria and validation methods to better quantify modelling fidelity (<https://garteur.org/technical-reports/>). *ValCrit-T* and *ValCrit-F* metrics in the time domain and respectively frequency domain were developed within GARTEUR HC/AG-09 to detect statistically the differences between simulation model and flight test data that might be missed by engineering judgment alone. These metrics stem from the Chi-square statistics (pronounced with a hard ch as in “character”) [Snedecor and Cochran, 1989]. The chi-square statistic is a goodness-of-fit test and gives how well the actual results match the expected results. This is done by comparing the size of any discrepancies existing between the expected results (flight test data) and the actual results (model data). GARTEUR HC/AG-21 on “*Rotorcraft Simulation Fidelity Assessment Predicted and Perceived Measures of Fidelity*” [White et al. (2020)] applied *ValCrit-T* parameter to the problem of helicopter simulator validation according to the Qualification Test Guide (QTG) tolerances (see also section 4 on QTG).

Note that there is a correlation between the frequency-domain and time-domain integrated cost functions as presented in the previous section and the ValCrit-F and ValCrit-T metrics defined in the chi-square statistic. The cost function is also a measure of how well the simulation model is performing and this is achieved by measuring the distance between the predicted value y and the actual value y_{data} . In this sense, J_{ave} , Eq. (4) and J_{rms} , Eq. (8) can correlate with ValCrit-F and respectively ValCrit-T through their statistical p-value.

2.2.1 ValCrit-T Parameter in the Time Domain

Consider a single response measurement and simulation data to be compared over N discrete time samples. The model error at each sample can be defined as:

$$\varepsilon_i = x_{data}(t_i) - x_{model}(t_i) - x_{bias}(t_i) \quad (12)$$

wherein ε_i is the model error at each sample, $x_{data}(t_i)$ is the data of the measurement (flight test data), $x_{model}(t_i)$ is the output state from the model prediction and $x_{bias}(t_i)$ is a bias correction that can be applied to the response. A weighted average of squared errors over all samples can be defined in the variable Z_T corresponding to the allowed inaccuracy of the data [Haverdings et al. (2000)]:

$$Z_T^2(t_i) = \frac{1}{p} [\bar{x}(t_i) - \bar{x}_{model}(t_i) - \bar{x}_{bias}]^T \cdot X_T^{-1}(t_i) \cdot [\bar{x}(t_i) - \bar{x}_{model}(t_i) - \bar{x}_{bias}]$$

↑

p-dimensional state vector (measurements)

↑

output state vector from the model prediction

↑

bias correction that may be applied

↑

Matrix allowing for the weighting of the difference between model and measurements

$$X_i = \text{diag}[\sigma_{x_i}^2]$$

X is set equal to the measurement error covariance

(13)

Herein \bar{x} is a p-dimensional state vector (obtained from the measurements or from the model) and X_T allows for the time-varying weighting of the difference between model and measurements. One choice for the X_T matrix is to set it equal to the measurement error covariance, matrix X_i in (13). The variable Z_T can be regarded as a normal variate, with zero mean and unit variance, and may be considered as a closeness-of-fit parameter.

A validation criterion in the time domain can be defined which consists of a “performance” scalar J_T , the so-called *ValCrit-T* parameter, with the following function:

$$ValCrit - T = J_T = \frac{1}{N} \sum_{i=1}^N Z_T^2(t_i) \quad (14)$$

where N is the number of time samples considered. For the response of a k parameter of a simulation model the *ValCrit-T* value corresponds then to:

$$J_k = \frac{1}{N} \sum_{i=1}^N \varepsilon_i X_i \varepsilon_i \quad (15)$$

X_i matrix becomes the identity matrix in the simpler case when unity weight is given to each sample. The single *ValCrit-T* value of eq. (15) reduces then to the average square error:

$$J_k = \frac{1}{N} \sum_{i=1}^N \varepsilon_i^2 \quad (16)$$

When p responses of the simulation model are considered, the *ValCrit-T* value becomes the aggregate value over all these responses:

$$ValCrit - T = J_T = \frac{1}{p} \sum_{k=1}^p J_k = \frac{1}{pN} \sum_{k=1}^p \sum_{i=1}^N \varepsilon_{ik} X_{ik} \varepsilon_{ik} \quad (17)$$

Consider next the set error samples ε_i as a set of N normal random variables. It follows that the term $J_T pN$ can be interpreted as a simulation error statistic that can be tested against a chi-squared distribution with pN degrees of freedom, denoted as χ_{pN}^2 . The statistical significance of that error can be assessed by comparing it with a χ_{pN}^2 function’s corresponding p-value, Figure 1.

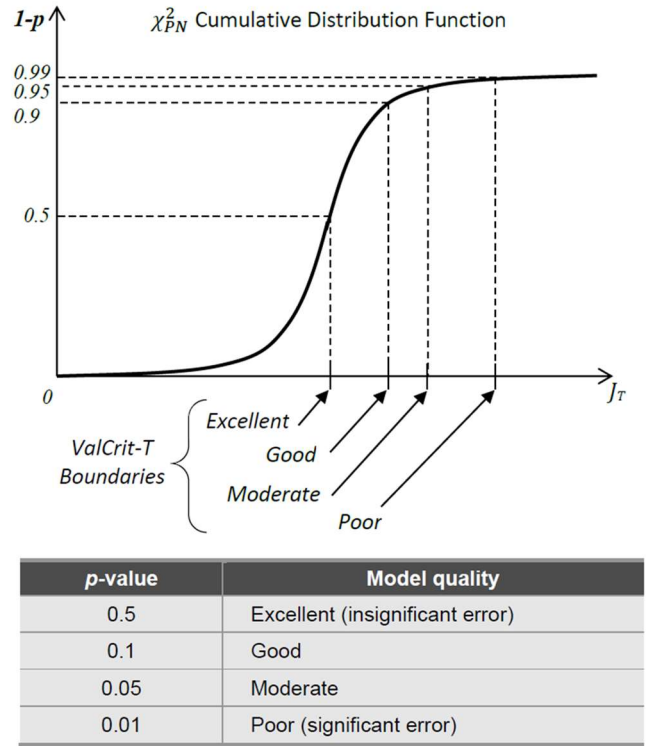


Figure 1. Test for statistical significance of *ValCrit-T* metric and their p-value for levels of simulation error as proposed in GARTEUR HC/AG-09 [Haverdings et al. (2000)].

Hypothesizing that the test statistic $J_T pN$ represents an insignificant simulation error, then this hypothesis is falsified if $J_T pN$ exceeds the value of z corresponding to a given threshold $1 - p$ of the χ_{pN}^2 cumulative distribution function. The values of p for four qualitative levels of simulation error as recommended in GARTEUR HC/AG-09 [Haverdings et al. (2000)] are listed in Figure 1. GARTEUR proposed boundaries

for the J_T (cumulative) and Z (instantaneous) functions based on statistical tables for Chi-square distribution. In this sense, $ValCrit-T$ can be seen as a generalization of the integrated cost function.

2.2.2 ValCrit-F Parameter in the Frequency Domain

The same as $ValCrit-T$ in the time domain, in the frequency domain, GARTEUR HC/AG-09 introduced a very similar validation criterion, the so-called $ValCrit-F$ parameter:

$$ValCrit - F = J_F = \frac{1}{N} \sum_{i=1}^N Z_F^2(\omega_i) \quad (18)$$

where the variable to be summed in the frequency domain is:

$$Z_F^2(\omega_i) = \frac{1}{2p} [\bar{g}(\omega_i) - \bar{g}_{model}(\omega_i) - \bar{g}_{bias}]^T \cdot X_g^{-1}(\omega_i) \cdot [\bar{g}(\omega_i) - \bar{g}_{model}(\omega_i) - \bar{g}_{bias}] + \frac{1}{2p} [\bar{\phi}(\omega_i) - \bar{\phi}_{model}(\omega_i) - \bar{\phi}_{bias}]^T \cdot X_\phi^{-1}(\omega_i) \cdot [\bar{\phi}(\omega_i) - \bar{\phi}_{model}(\omega_i) - \bar{\phi}_{bias}]$$

←

p-dimensional state vector for magnitude and phase (measurements)

←

output state vector in magnitude and phase from the model prediction

←

bias correction for magnitude and phase that may be applied

←

Matrix allowing for the weighting of the difference between model and measurements magnitude and phase

$X_i = diag[\sigma_{xi}^2]$

X is set equal to the measurement error covariance

(19)

As such one may specify statistical boundaries for J_F to exceed a certain value, and hence to have a certain model quality, because J_F exactly as J_T relates to the Chi-squared statistic as follows:

$$ValCrit - F = J_F = \frac{1}{\nu} \sum_{i=1}^N \chi_F^2(\omega_i) \quad (20)$$

where $\nu = pN$ is the number of degrees of freedom. Exactly as for the J_T , the corresponding values for J_F can be derived for defined values of probabilities of exceedance. Quality-of-fit indicators related to exceedance values of α are excellent ($\alpha=0.5$), good ($\alpha=0.1$), moderate ($\alpha=0.5$) or poor ($\alpha=0.01$), see Figure 1. The values for Z_F are based on a two-tailed probability, as Z_F may be both positive and negative, i.e. the null-hypothesis tested is:

$$H_0 \{Z_F > g \text{ or } Z_F < -g\} = H_0 \{|Z_F| > g\} \quad (21)$$

$$H_0 \{Z_F > \phi \text{ or } Z_F < -\phi\} = H_0 \{|Z_F| > \phi\}$$

For the frequency domain, GARTEUR HC/AG-09 used for each frequency a maximum allowable error of 10 percent (95th

percentile) of the gain and 20 degrees error (also 95th percentile) in phase angle, i.e.:

$$X_{g_i}(\omega) = (0.1 \times 1.645) \text{diag} \left[\bar{g}_{model}(\omega_i) \bar{g}_{model}^T(\omega_i) \right]$$

$$X_\phi(\omega) = 3.183 \quad (22)$$

The factor of 1.645 in eq. (22) stems from conversion of the 95th percentile value to standard deviation.

3. FIDELITY METRICS FOR ROTORCRAFT HANDLING QUALITIES

The development of the Maximum Unnoticeable Added Dynamics (MUAD) stems from the concept of equivalent system approach [Hodgkinson (2005)]. With the introduction of highly-augmented aircraft in the late 80s, difficulties arose in the flying qualities analysis. Classical flying qualities standards defined for classical flight dynamics could not be used anymore due the higher-order effects of the augmented aircraft. To cope with these higher-order dynamics, lower-order equivalent systems (LOES) were developed to approximate the higher-order system (HOS). The equivalent system approach meant that the actual aircraft dynamics had to satisfy flying qualities requirements in terms of “equivalent classical systems which have responses most closely matching those of the actual aircraft” (Mitchell et al., 2004). One of the challenges of using the lower order systems was to determine what acceptable levels of mismatches were between the LOES and HOS.

3.1 Maximum Unnoticeable Added Dynamics (MUAD)

The overlay of flight-test and simulation frequency responses is a direct and efficient means to validate model fidelity and assess model improvements. After making modifications to the simulation model, the comparison is repeated to determine whether the validity of the model has been improved. The simulation model accuracy for each frequency response pair can also be characterized in terms of the error response function $\epsilon_{model}(f)$ defined as:

$$\epsilon_{model}(f) \equiv T(f)/\hat{T}_c(f) \quad (23)$$

where the frequency response for simulation denoted as T , and the associated flight-test data frequency response is denoted as \hat{T}_c [Tischler and Remple (2012)]. In terms of the magnitude (dB) and phase (deg) responses,

$$Mag_{err}(f) = (|T| - |\hat{T}_c|) \quad (24)$$

$$Phase_{err}(f) = (\angle T - \angle \hat{T}_c) \quad (25)$$

where an error function with 0-dB magnitude and 0-deg phase indicates perfect tracking of the flight and simulation results. The magnitude and phase of the error response functions for the XV-15 Generic Tilt Rotor simulation model in cruise are shown as the dashed curves in Figure 2 from Tischler and Remple (2012).

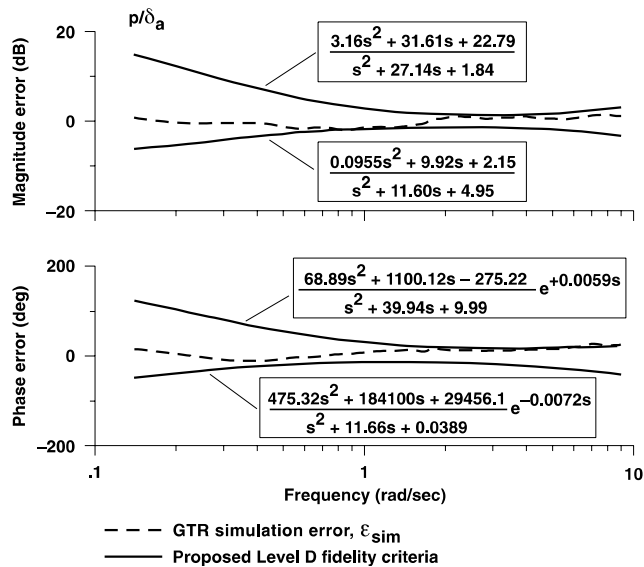


Figure 2. XV-15 Cruise Error Functions and MUAD Bounds [Hodgkinson (1998)].

Also shown in the figure as the solid curves are boundaries that correspond to limits on MUAD, beyond which a pilot will detect a deviation in the aircraft model compared to flight [Hodgkinson (1998) and Smith et al. (1981)]. These boundaries are used in the fixed-wing handling-qualities criteria of the USAF MIL-STD-1797B (2006) to evaluate the mismatch between an actual aircraft response and a lower-order equivalent system (LOES) model. The equations for the MUAD boundaries [Hodgkinson (1998)] are shown in the figure. If the error functions fall within these boundaries, then the simulation model response would be judged by a pilot as being indistinguishable from the actual flight response, thereby providing a good basis for simulation model fidelity assessment. Tischler (1995) first proposed the use of the MUAD boundaries for simulation model fidelity assessment and FAA Level D simulation fidelity criteria. The same approach of mismatch boundaries in the frequency domain was also independently proposed and applied by DLR researchers to detect the effects of unnoticeable dynamics in the case of helicopters [Hamel and Jategaonkar (1996)] and for evaluating the fidelity of in-flight simulation [Buchholz et al. (1996)]. A good overview of the low-order equivalent systems approach was presented by Hodgkinson (2005). More recent research by Mitchell et al. (2009) supports these boundaries as useful for evaluating rotorcraft simulation fidelity as well.

3.2 Allowable Error Envelopes (AEE)

In 2006 Mitchell et al. (2006b) verified the application of MUAD to military simulators qualification. The MUAD envelopes, as described previously, were developed to define limits on unnoticeable added dynamics from a fixed-wing airplane database. Given that the envelopes have been proposed for simulation validation, Mitchell et al. (2006b) applied MUAD for validation of military helicopter simulators. It was concluded that the concept of MUAD and added dynamics should be carefully implemented in the case of helicopter simulator validation. The envelopes describing Maximum Unnoticeable Added Dynamics although still generally ubiquitous when applied to simulation validation, in the case of helicopter roll tasks, added dynamics were still unnoticeable to the pilot. The new goal of Mitchell et al. (2006b) was thus to identify a set of frequency-domain envelopes defining the boundary between unnoticeable and noticeable dynamics. These envelopes are referred to as ‘Allowable Error Envelopes’ (AEE), to distinguish them from the MUAD envelopes, and to more accurately reflect their ultimate purpose, i.e. to define the allowable errors in simulation validation.

Based on MUAD envelopes for simulation validation, Mitchell et al. (2006b) developed the so-called ‘Allowable Error Envelopes (AEE)’ defined as ‘boundaries between unnoticeable and noticeable dynamics.’ The idea was that, as pilots cannot evaluate what they cannot ‘see’, variations of parameters that caused change in frequency response within envelope could not be meaningfully evaluated.

The AEE envelopes could be used to:

- determine whether the mathematical simulation model is adequate, or
- determine whether a simulation facility is adequate to accomplish proposed evaluations

It was foreseen that the AEE envelopes would probably be larger for fixed-base simulators than for moving-base simulators and that the smallest envelopes were to be expected for an in-flight simulator. Therefore, AEE should be developed separately for fixed-base piloted simulations, moving-base piloted simulations and for in-flight piloted simulations.

The experiment of Mitchell et al. (2006b) involved a helicopter hovering task as described in the ADS-33E (2000) performed with a simple helicopter model, i.e. the helicopter translated along a 45-degree angle ground track at low altitude and decelerated to stop at a predefined hover position. This task was appropriate for AEE design as it showed effective conveying of high bandwidth closed-loop pilot activity. This experiment was extended by Penn (2013) using more pilots and more simulators.

Figure 3 represents the AEE proposed boundaries plotted against the MUAD boundaries as presented in different sources.

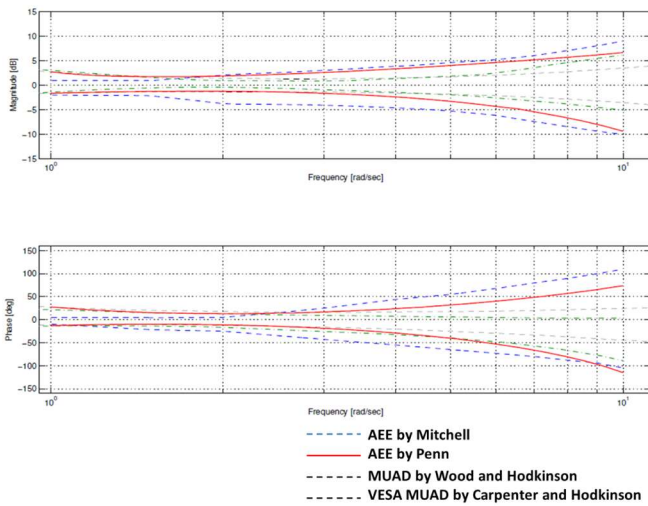


Figure 3. Allowable Error Envelope as proposed by Mitchell (2006b) (blue) and Penn (2013) (red). MUAD envelope by Wood and Hodgkinson (1980) (black dotted line) and V/STOL Equivalent Systems Analysis (VESA) MUAD envelope by Carpenter and Hodgkinson (1980) (green).

4. SIMULATOR FIDELITY

The fidelity concept has been analyzed in the context of ground-based simulators since the 1980s. At that time there had been concerns that ground-based simulators had insufficient visual cue fidelity for pilots to adopt equivalent control strategies to the “real world.” (Padfield, 2013) It was soon understood that the fidelity in the simulator world should refer to: 1) objective fidelity (also called predictive fidelity and physical fidelity) involving a series of ‘predictive’ metrics for flight model fidelity and also for other simulator’s components and 2) subjective fidelity (also called perceptual or behavioral fidelity) involving the subjective opinion of the assessing pilot, which is always a combination of the complete system.

Figure 4 presents the concept of fidelity as related to the simulator. In the real world, the pilot assesses the vehicle orientation by using his/her visual, vestibular, tactile, and aural sensors, integrating all the signals into his/her central nervous system and using his/her hands, feet, posture, and voice as motor mechanisms to finally control the vehicle. In the simulated world, likewise, for a certain task assigned, the pilot used the simulator cueing elements (e.g. visual, motion, control loading cueing) and the vehicle model to mimic the vehicle’s flight and achieve a task performance.

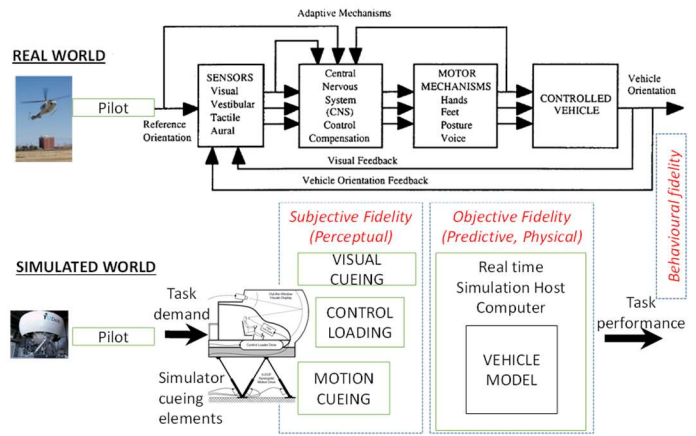


Figure 4. Fidelity concept as related to a simulator device.

Morgan and Baillie (1993) exemplified this difference between the simulated and the real world for the bandwidth-phase delay criteria. In their experiment, similar tasks and vehicle configurations were flown on the NASA Vertical Motion Simulator (VMS) and the NRC’s Flight Research Laboratory (FRL) Bell 205 in-flight simulator. Figure 5 presents the roll attitude bandwidth according to the ADS-33 requirements obtained in the VMS and FRL. One can see that while in flight the attitude bandwidth required for Level 1 handling was 2 rad/s, in ground-based simulation this was 3 rad/s. This difference was largely attributed to the pilot needing more bandwidth in the simulator to compensate for poorer visual cues, particularly the sparseness of fine-grained surface texture.

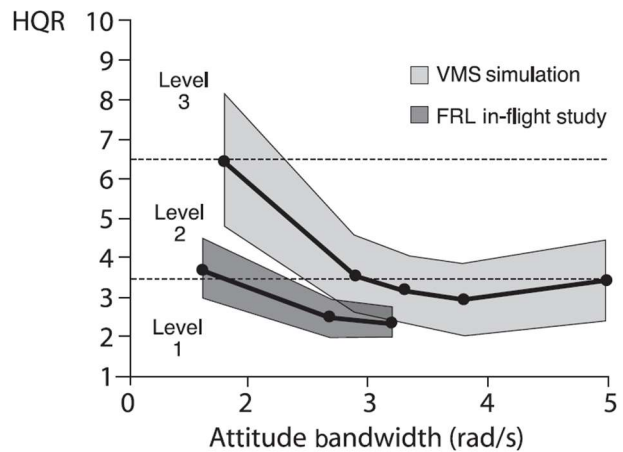


Figure 5. Flight/ground-based simulation comparison; roll attitude bandwidth with attitude response type (Morgan and Baillie (1993)).

Wang, White and Owen et al. (2012) demonstrated in a different example that, even with high fidelity flight modelling and visual motion cues, the pilot’s experience can be spoiled by poor vestibular cueing. Therefore, for simulators, fidelity involves both predictive and perceptual fidelity.

4.1 A systematic review of fidelity definition as related to ground-based simulators

The Advisory Group for Aerospace Research & Design (AGARD, 1980) represented the division of a simulator into two classes depending on the nature of the cues they provide:

- 1) Equipment cues provide a duplication of the appearance and feel of the operational equipment (the aircraft), i.e., the static and internal dynamic characteristics such as the size, shape, location, and color of controls and displays, including controller force and displacement characteristics.
- 2) Environment cues provide a duplication of environment and motion through the environment.

In this sense, one can discuss about equipment fidelity and environmental fidelity, the fidelity herein being *“the degree to which the equipment and environmental cues relate to those of the real airplane.”* (AGARD, 1980). Furthermore, AGARD work made a distinction between the real cues, measured objectively, and the cues the pilot subjectively can experience. As a result, one can discuss two types of fidelity (AGARD, 1980):

- 1) Objective Fidelity is the degree to which a simulator can reproduce the real-life counterpart aircraft, in flight, if its form, substance, and behavior were sensed and recorded by a non-physiological instrumentation system onboard the simulator.
- 2) Perceptual Fidelity is the degree to which the pilot/flight crew subjectively perceives the simulator to reproduce its real-life counterpart aircraft, in flight, in the operational task situation. The requirement that the operational equipment be considered in the context of the task situation ensures that not only cue timing and synchronization, but also cue priority effects, are taken into account. In this sense, perceptual fidelity provides a psychological/ physiological standard.

The philosophical discussion surrounding verification and validity for computerized simulation models was carefully reviewed by Mihram (1972).

In the context of simulator device, Heffley (1981) defined fidelity as *“the simulator's ability to induce the pilot trainee to output those behaviors known to be essential to control and operation of the actual aircraft in performance of a specific task.”*

According to Hays and Singer (1989), fidelity should be defined in terms of domain of interest, relative to something and in a measurable form. A training simulation domain-oriented definition for fidelity is *“the degree of similarity between the training situation and the operational situation, which is simulated”*. Fidelity is presented here as a 2-

dimensional measurement in physical and functional characteristics. Physical characteristics address aspects like look and feel while functional characteristics address aspects such as the informational, operational knowledge and stimulus and response options. Fidelity is thus characterized as a summarizing descriptor of the overall training device characteristics and trained scenarios.

Lane and Alluissi (1992) define fidelity as an engineering concept referring to *“the physical correspondence of the simulator's hardware and software to that of the actual equipment being simulated.”* Lane remarks that implementing *“all you can afford”* levels of fidelity in a simulation is not always the most correct and cost-effective approach to address the problem of how much fidelity is required for a specific purpose.

Rehmann (1995) believes that fidelity is a function of the degree to which the equipment and environmental cues relate to those of the real airplane. He underlines that at least 22 different definitions have been used in the literature to refer to different kinds of fidelity. Each of these kinds of fidelity defined during the years could be appropriate for a particular application, but each are not individually applicable to overall aircraft simulation in general.

The Pace (2015) definition for fidelity dates back to his work on fidelity from 1992. According to him fidelity is *“the degree of exactness of a model or simulation representation when compared to the real world.”* According to Pace, simulation fidelity is an absolute concept while simulation validity is a relative concept, dependent upon the simulation application.

Fidelity is defined by Gross and Freeman (1997) as *“the extent to which a model reproduces the referent along one or more of its interests”* (N.B. referent is an authoritative description of reality in the context of high-level architecture simulation). To characterize fidelity, Gross uses three classifications: 1) existence (which object of the referent exists in the model); 2) attributes (which object attributes of the referent exist in the model) and 3) behavior (what object behavior of the referent is included in the model).

McDonald (1998) defines fidelity as *“the accuracy of representation when compared to the real world.”* This fidelity has two major parts: the extent to which the simulation models each aspect of the real world and the agreement between the performance of each modeled aspect and the real performance.

Meyer's (1998) perspective on fidelity definition is to outline four terms for describing simulation goodness, i.e.: 1) Detail describing the measure of the completeness/complexity of the model w.r.t. the observable characteristics of the physical entity; 2) Accuracy as related to the exactness of the model w.r.t. the observable characteristics and behaviors of the physical entity; 3) Resolution defined as a measure of the minimum degree to which the accuracy and detail of the

constituent models must coincide with the required level of fidelity of the simulation and 4) Fidelity as related to the agreement of a simulation with perceived reality.

Roza (2005), based on the above-definitions proposed a unified framework for simulator fidelity in which a formal mathematical definition for fidelity is formulated as “*the inverse difference between the reality and simulated reality.*”

Havighurst, Fields, and Fields (2003) define *high fidelity* as the required equipment and materials necessary to adequately simulate the task the learner is expected to perform. They define *low fidelity* as equipment and materials that are less similar to what task the learner is expected to perform.

Myers, Starr and Mullins (2018) revisiting the fidelity as a fundamental concept in simulator design, concentrated on the relation between *fidelity* and *training transfer*. These two concepts are closely related. Many simulator designers, operators, technicians, and behavioral scientists believe that the simulator should be designed with the maximum fidelity possible since that is postulated to provide the most training transfer. However, doing so results in higher costs that may not be feasible for some organizations.

Finally, it can be stated that fidelity is connected with safety as safety outcomes are among the key factors in desiring appropriate fidelity within the simulator environment.

4.2 Simulator Fidelity and Necessary Cueing

The fidelity of a simulator arises from the comparison of the simulated piloting experience and the actual piloting experience, as perceived by the pilot. The closer one gets to the sensation of piloting a real vehicle the higher the simulator fidelity. However, it is often difficult to assign an objective value to simulator fidelity, given its subjective nature. Additionally, because the ground-based simulator is in fact, by design, confined to a laboratory, the engineer needs to create a controlled environment in the simulator that “forces” the pilot’s experience to be as realistic as possible. Environmental feedback, flight conditions, optical flows, vehicle physics – all of these details are purposefully distorted before being presented to the pilot. This is achieved through cueing, which, in simple terms, is similar to post-processing all the aspects of the calculated reality.

It follows that the “good fidelity” in a simulator can be associated with “proper cueing”. One can say that fidelity and cueing are both terms to describe the design of a simulator. Cueing relies upon an understanding of human factors, and upon the acceptance that what we experience at any given moment is being processed by our senses. In a well-appointed pilot-in-the-loop simulator, the engineer is feeding the pilot’s senses as much as possible with realistic experiences. Nonetheless, the goal of all these sensory feeds may not be to make pilots believe they are in a real aircraft. The goal may be to make pilots behave as if they are in a real aircraft.

Cueing is thus a fundamental concept for a ground-based simulator. Sometimes, mathematical model fidelity deficiencies might be masked or exacerbated by the simulator’s cueing systems. Evidently, all cues contribute to the piloting experience in a simulator; however, some cues play a key role in the perceived realism: 1) Visual cues, since they are the main cues used to orient an aircraft in the air. 2) Motion cues, since they map the modelled aircraft motions into feedback to the pilot’s body (vestibular system and kinaesthetic sense) 3) control force feel cues, since they provide the pilot with information about the static control force levels and dynamic feel perceived by the pilot in control of an aircraft.

While outside visual, instrument, and force feel cueing systems usually distort the aircraft’s dynamics with little more than a time delay (latency), a simulator’s motion system can introduce more elaborate distortions of the perceived aircraft’s dynamics. This is caused by the inherently limited motion space of a ground-based simulator and the motion cueing strategies (algorithms) that are necessarily deployed to transform the model’s motion cues into physical motions of the simulator. Such strategies allow tuning of simulator motion to improve the fidelity of the motions of interest. The next section explores especially the motion cueing strategies to be applied in the rotorcraft simulator and how these can influence the overall perceived fidelity of the simulation.

4.3 Motion cueing

When assessing rotorcraft dynamics performance, motion feedback is a key component. In general, a human pilot’s vestibular system is sensitive to the specific forces and rotational rates or accelerations of the helicopter. Similar to an accelerometer, a human directly perceives the aerodynamic, engine, and landing gear forces acting on the rotorcraft. These cues play a number of roles in a piloted simulation:

- They provide the pilot with additional signals to stabilize and control their aircraft. Especially for unstable vehicles such as (unaugmented) helicopters, the phase lead present in the motion cues over the visual position and attitude cues help to increase phase margin in the closed loop.
- They signal to the pilot that certain important events or state changes are occurring in the vehicle. An example are the vibrations associated with Effective Translational Lift (ETL), providing timing signals to help in performing the landing flare [Miller et al. (2009)].

Motion cues help create an immersive virtual environment that elicit realistic pilot behavior. High levels of behavioral fidelity improve training effectiveness and engineering or scientific simulator validity.

Apart from providing cues that help the pilot perform their task, physical motion can also be detrimental to task performance, e.g. vibrations impacting touchscreen effectiveness. Such effects should be replicated in the simulator, if relevant.

Multiple techniques exist to provide physical motion cues to the simulator pilot. The most common type of motion system moves the simulator cab as a whole in six degrees of freedom. In helicopter simulators sometimes a dedicated in-cabin motion platform moves part of the cabin, such as the seat. Additionally, dynamic seats can move parts of the pilot seat independently to give an illusion of motion, e.g. moving the lumbar support to signal sustained specific force cues through independent actuation of seat backpad surge and sway and seat pan cushion and seat bucket heave degrees of freedom. A common whole-cabin motion system is based on the Gough-Stewart 6-degree-of-freedom mechanism [Stewart (1965)].

The motion base’s cueing algorithm converts the simulated vehicle’s cues to simulator motions. It must achieve two competing goals: provide realistic cues to the pilot and keep the simulator in the available motion space. Many motion cueing algorithms exist, most of them proprietary, but the underlying concepts are usually derived from the UTIAS Classical Washout Algorithm described by Reid and Nahon (1985). In this algorithm, the translational specific forces and rotational accelerations go through their own high-pass filter channels, and a special cross-coupling is made (“tilt coordination”) by which sustained, low-frequency surge and sway cues can be simulated by tilting the simulator and using gravity to cue the specific force.

In both the rotational and translational channels, it is important to note that in essence the algorithm employs high-pass filters. This means that for frequencies that the filter (starts to) attenuate the motion, a phase lead is present. At higher frequencies (i.e. greater than 10 rad/s) phase lag will mostly be present due to the motion base’s inertia and control system.

The parameters used in the motion cueing algorithm (which essentially define the phase and gain cueing), can significantly affect the simulator’s motion fidelity. As such the simulator engineer has some freedom to shape or tune the motion response to the requirements for a certain application. The freedom is of course limited by the available motion envelope of the platform (governed by design and size), and the detrimental effects of false cues inherent to the algorithm.

4.3.1 Motion cueing criteria

The first objective motion cueing criteria was developed by Sinacori (1977), who identified a metric of the cueing response that specifies a level of fidelity of the resulting cues. This well-established metric examines the phase and magnitude of the cueing algorithm at 1 rad/s and categorizes it in one of three fidelity levels (low, medium, and high). Schroeder (1999) later refined the limits based on helicopter tests in NASA Ames Vertical Motion Simulator (VMS), see Figure 6. The VMS'

large vertical and horizontal travel allowed Schroeder to develop baseline tasks incorporating one-to-one motion without encountering motion envelope limits. Schroeder then developed the Modified Sinacori Criteria [Schroeder (1999)] for rotational and translational motion. To date, this method is the most commonly applied to rotorcraft simulation devices. Reardon et al. (2014) for example developed an indirect motion fidelity criteria varying the motion filter parameters from Schroeder’s experiment and using the subjective Simulation Fidelity Rating (SFR) scale (for more details on SFR scale the reader should consult section 6).

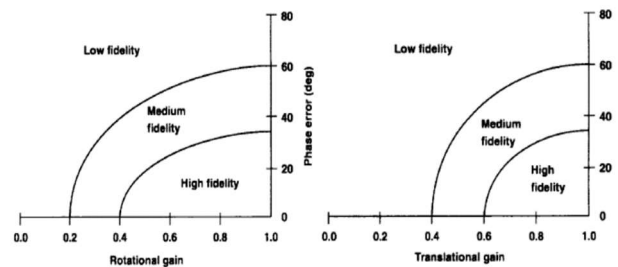


Figure 6. Sinacori/Schroeder motion fidelity criteria [Schroeder (1999)].

It is acknowledged that the assessment method as developed by Sinacori and Schroeder, has limitations. To increase the (frequency) range over which the motion cueing is evaluated and thereby creating a more comprehensive evaluation than the Sinacori/Schroeder criteria, the International Civil Aviation Organization (ICAO) has proposed the Objective Motion Cueing Test (OMCT) [Advani and Hosman (2006), Advani and Hosman (2007), Hosman and Advani (2012)]. This test works by independently exciting each degree of freedom of the motion base at frequencies varying from 0.1 to almost 16 rad/s and examining the corresponding motion outputs (main effect), as well as some unwanted cross-couplings (false cues). It examines these responses in the frequency domain to effectively generate a Frequency Response Function of the combination of the motion cueing algorithm’s software and the motion base’s hardware. As such it fits well with the common practice of investigating the helicopter’s model as a transfer function in the frequency domain. An example of fidelity envelopes, derived from a study involving 10 simulators used for fixed-wing aircraft (both training and research) is shown in Figure 7. The fidelity envelope was defined by taking the highest and lowest settings from the group of simulators (following extraction of outliers).

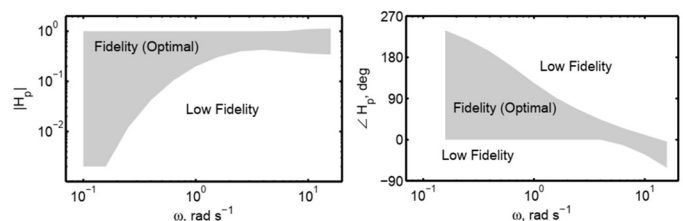


Figure 7. Example of OMCT fidelity boundaries, roll motion gain and phase [Jones (2018)].

Studies reported in Hodge et al. [(2015a), (2015b)] and Jones et al. (2017) suggested large differences between perceived motion fidelity and objective fidelity using Sinacori and Schroeder boundaries. Both studies found beneficial and representative motion was attained when performing typical mission task elements performed by rotorcraft. The tests were conducted in two simulation facilities, with independent pilots and simulation settings. The reservations are confirmed when one compares the acceptable required fidelity for current OMCT and Schroeder metrics. Figure 8 shows the difference between rotational requirements, plotted on the gain – phase plot. Using the Schroeder approach, the same boundaries are used to assess pitch, roll, and yaw dynamics. For the OMCT boundaries, a range of boundaries are presented. Both acceptable roll and yaw requirements do not reflect Schroeder boundaries. Meanwhile, pitch boundaries are more stringent. As a result of these discrepancies, further research is required in this area prior to the adoption of OMCT as an objective method to assess rotorcraft training simulators.

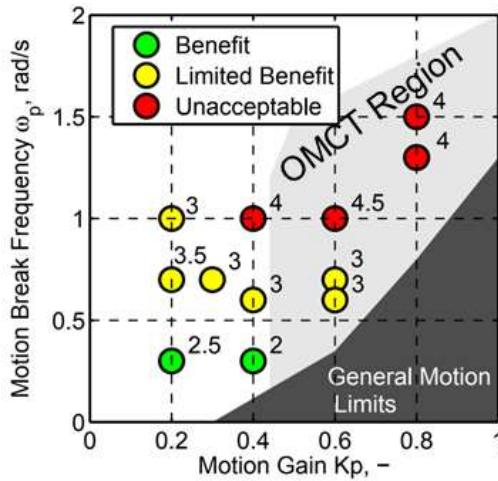


Figure 8. Comparison of current boundaries for OMCT and Schroeder metrics [Jones et al. (2017)].

In additional research investigations conducted by Jones et al. (2017), further motion cases were tested to determine initial OMCT fidelity boundaries specifically for rotorcraft simulation. During these investigations, the differences between current OMCT boundaries and perceived fidelity was again confirmed. Like results from previous studies, the investigation showed a clear preference for low phase errors in the motion system. This was even for cases with very low motion gain. Three pilots participated in the study, and stated that motion cueing with high phase errors led to significant false cues, sickness, and adversely affected performance. Pilots were clearly able to consistently recognize motion configurations. Adverse motion cues led to a reduction in task aggression (thereby suppressing false cues). During completion of typical ADS-33 maneuvers, this led to an increase in task completion time.

The linear Frequency Response Function determined in OMCT uses uncoupled and simplified (sinusoidal) input signals.

During normal flight the cueing algorithm’s non-linear characteristics can interact with inherent couplings in the helicopter model’s cues, to produce false cues and responses that are not fully captured in the general OCMT method.

Recently, to overcome some of these limitations, Miletovic (2020) and Miletovic et al. (2021) proposed the “Eigenmode distortion” (EMD) method as a novel quantitative methodology to objectively evaluate motion cueing fidelity in flight simulation. This method relies on combining the vehicle dynamics modal analysis/ mode participation factor with the Motion Cueing Algorithm (MCA) dynamics, this in order to understand the human-perceived motion quantities.

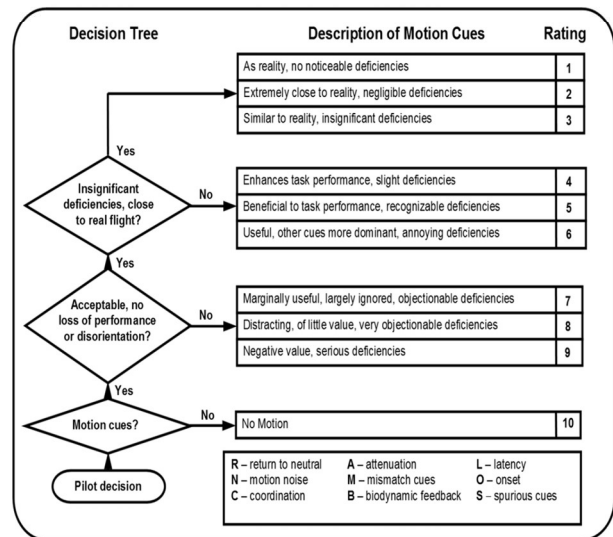
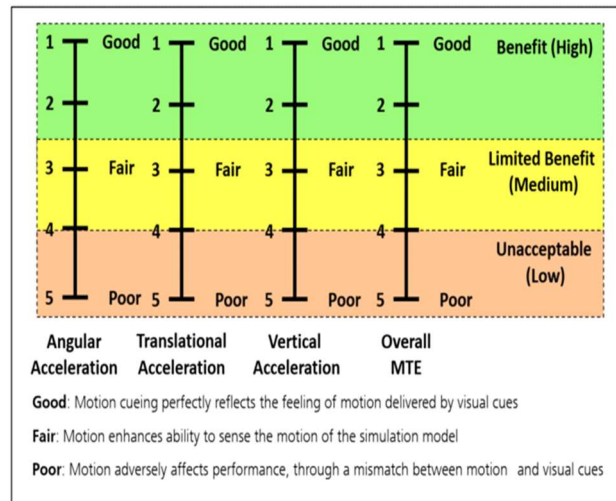


Figure 9. Motion Fidelity Rating scale (right hand-side figure from Hodge et al. (2015a/b), left hand-side figure from Jones et al. (2017)).

To standardize and streamline subjective pilot assessments a number of techniques have been used in the literature. Some have already found use outside their academic origins, while others are still being developed. The most common technique takes the form of a rating scale, similar to the Cooper-Harper

Handling Qualities Rating (HQR) Scale [Cooper and Harper (1969)], in combination with one or more standardized maneuvers, e.g. from the ADS-33E standard (2000).

The University of Liverpool and the German Aerospace Center (DLR) have both worked on defining a Motion Fidelity Rating scales, see Figure 9. In industry, Miller et al. (2009) added numerical ratings to the Simulator Motion Fidelity Scale developed by Schroeder (1999) to allow finer grading of motion fidelity within the High, Medium, and Low fidelity rating levels as shown in Table 1 during motion system tuning.

Table 1 Modified Simulator Motion Fidelity Scale Used in Industry [Miller et al. (2009)]

Fidelity Rating	Definition
High (7-9)	Motion sensations like those of flight.
Medium (4-6)	Motion Sensations are noticeably different from flight, but not objectionable.
Low (1-3)	Motion sensations are noticeably different from flight and objectionable.

Concluding, particularly for training simulators and from the view of the regulator, there is a reliance upon subjective opinion. To assure acceptability of a flight simulator for either engineering or training purposes, the end user’s subjective assessment of its fidelity cannot be ignored. In addition, many of the objective motion cueing quality metrics have been validated by expert pilot opinion with varying levels of success. Pilot opinion still remains the standard against which motion cueing quality is measured.

5. PREDICTIVE FIDELITY – METRICS AND METHODS

5.1 Predictive Fidelity

The evaluation of the fidelity of a simulation device for flight training is performed within the simulator qualification documents: Federal Aviation Administration (FAA), FAA 14 Part 60 (2016), and European Aviation Safety Agency (EASA), EASA CS-FSTD-H (2012). They are the regulatory authorities responsible for the acceptance of full flight simulators (FFS). The highest fidelity level defined by FAA/EASA corresponds to the Level-D qualification. A Level-D qualification process allows the replacement of most of the flight hours required for a pilot's type rating or recurrent training by simulator hours. A Level-D training simulator is made of many sub-system models related to the vehicle dynamics (flight dynamics, engines autopilot, flight controls), vehicles systems (avionics, ancillaries, etc.) and simulator immersive cueing environment (motion sound, visual, weather, airport environment, etc.). Each of these sub-systems must meet qualitative and quantitative validation criteria for the specific aircraft type to meet Level-D simulator requirements. A complete background and history on the qualification of helicopter training simulators over the years can be found in Pavel et al. (2013).

Regarding the quantitative requirements, these examine the response or behavior of the individual elements of a simulation device – the visual system, the motion platform (if so equipped), the flight dynamics model etc. – to a set of predetermined inputs. The results of these tests are typically termed “*predictive/objective fidelity*”. Objective fidelity only partially serves to characterize the utility of a simulator.

5.2 Qualification Test Guide (QTG) Performance Standards

Both the FAA and EASA are using a functional performance standard called “Qualification Test Guide” (QTG). The QTG is a document designed to assess and validate that the performance and handling qualities of a simulator are within prescribed limits of those of the aircraft and that all applicable regulatory requirements have been met. QTG maneuvers can be separated in 3 test categories: snapshot test, dynamics test and trajectories tests. Snapshot tests are used when a steady state condition exists in the flight test data at the instant of time captured [Myrand-Lapierre et al. (2020)]. Dynamics tests involved a pre-defined control input perturbation at a trim condition. Trajectories tests are highly non-linear maneuvers that will go through multiple flight regimes.

The QTG includes both the helicopter flight test data and simulator data used to support the validation. A flight test data package must contain more than one hundred individual events to meet the minimum Level-D validation requirements. The qualifying criteria of the mathematical model are formulated by using ‘*tolerances*’ and it includes an evaluation based on the comparison between reference flight test data and results of identical tests computed on a simulator. For example, in hover FAA (Table C2A in FAA 14 Part 60 (2016) and SUBPART C in EASA CS-FSTD(H) require for longitudinal cyclic input cases a tolerance of $\pm 10\%$ or 2 deg/sec (whichever is the highest) on the pitch rate response and of ± 1.5 degrees on the pitch attitude change following a control input. For lateral cyclic input cases, a tolerance of $\pm 10\%$ or 3 deg/sec (whichever is the highest) on the roll rate response and of ± 3 degrees on the roll attitude change following a control input are required. Also, for all cases, the Off-axis response must show correct trend for unaugmented cases.

Regarding the simulator tolerances as used in the QTG, GARTEUR Action Group (AG) HC/AG-12 [Pavel et al. (2013)] conducted a critical examination of their background. The work revealed a range of shortcomings. For example, GARTEUR HC/AG-12 showed that the relationship between fidelity and the EASA CS-FSTD-H (previous JAR-STD 1H) tolerances is sensitive to the nature of the maneuver being flown and, more significantly, that matching tolerances does not always lead to matching handling qualities.

6. PERCEPTUAL FIDELITY WITH PILOT-IN-THE-LOOP

The implicit assumption in tests of objective fidelity is that a strong quantitative match of simulator component systems with the flight vehicle will assure a high degree of simulator utility. Experience has shown that this assumption is not always valid, and that tests of objective fidelity are insufficient to guarantee a sufficiently accurate simulation. Hence, the FAA/EASA qualification standards require a piloted, subjective assessment of the simulation in addition to the quantitative elements. Subjective validation requirements are comprised of a series of training tasks and abnormal conditions that are normally spot-checked during the final assessment to ensure that there are no discontinuities between simulated flight regimes.

These subjective test “*arises from the need to confirm that the simulation has produced a totally integrated and acceptable replication of the helicopter*” [EASA CS-FSTD(H) (2012)]. However, the guidance provided in the qualification documents regarding the approach taken to subjective evaluations is very limited. Section AMC1 of CS-FSTD(H) (2000) Terminology and abbreviations, Paragraph C (2) Test Requirements (iv) of Book 2 EASA CS-FSTD(H) (2012) states:

“When evaluating functions and subjective tests, the fidelity of simulation required for the highest level of qualification should be very close to the helicopter. However, for the lower levels of qualification the degree of fidelity may be reduced in accordance with the criteria (within the document).”

This requirement is poorly defined, and potentially open to interpretation by the operator and qualifying body. The work undertaken in GARTEUR HC/AG12 [Pavel et al. (2013)] suggested that the existing requirement for the subjective aspect of simulator qualification is unsatisfactory and should be improved.

6.1 Simulation Fidelity Rating Scale

The Simulation Fidelity Rating (SFR) [Perfect et al. (2014)] was proposed to provide a repeatable, prescriptive method for the subjective assessment of fidelity into the overall qualification process. The SFR scale should be used to complement and augment the existing simulator evaluation processes of CS-FSTD(H) and other applicable simulator qualification processes. It is proposed that the SFR scale may be used as part of a fidelity evaluation methodology based on the use of engineering metrics for both the prediction of the fidelity of the individual simulator components (flight model, motion platform, visual system etc.) [Perfect et al. (2014)] and the assessment of the perceptual fidelity of the integrated simulation system, as experienced by the pilot.

The SFR scale employs several key concepts that are considered fundamental to the utility of a simulation device. They are as follows:

- Transfer of Training (ToT) – the degree to which behaviors learned in a simulator are appropriate to flight.
- Comparative Task Performance (CTP) – comparison of the precision with which a task is completed in flight and simulator.
- Task Strategy Adaptation (TSA) – the degree to which the pilot is required to modify their behaviors when transferring from simulator to flight and vice versa.

The relationship between task performance and strategy adaptation is similar to that between performance and compensation in a handling qualities evaluation. In the HQR scale, the expectation is that the pilot’s perception of deteriorating performance will stimulate higher levels of compensation, indicative of worsening Handling Qualities (HQ). While this correlation can be expected in measuring HQ, in the context of fidelity assessment task performance and adaptation will not necessarily change in correlation with each other but will instead depend on the nature of the fidelity deficiencies present in a simulator.

A matrix presenting all possible combinations of comparative performance and task strategy adaptation was constructed (Figure A1 in Appendix A); this was used to form the basic structure of the SFR scale (Figure A2 in Appendix A).

Each of the ratings SFR=1 to SFR=9 corresponds to a region in the fidelity matrix. An SFR=10 rating indicates a simulation that is entirely inappropriate for the purpose, and so comparisons with flight cannot be made. As with the HQR scale, boundaries have been defined between the potential combinations of comparative performance and adaptation, reflecting value judgements on levels of fidelity. As the SFR worsens through each level, it can be seen from Figure A2 that the individual comparative performance and adaptation measures may not degrade in a progressive manner. However, the intention is that the overall ‘experience’ of the simulation fidelity degrades progressively as the SFR worsens.

7. ASSESING SIMULATOR FIDELITY BASED ON HANDLING QUALITIES ENGINEERING

While previous sections highlighted on the one hand the QTG as metrics and tolerances for assessing simulator predictive capability and on the other side the SFR (pilot opinion) as a measure for simulator perceived fidelity, unified metrics connecting predictive and perceived fidelities needed to be developed. In the 1990s it was suggested that, since the HQ parameters define the flying characteristics, they might also be suitable to judge the fidelity. Indeed, much of this HQ 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 behavioral 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 behavior of the real aircraft and its pilot, rather than purely to assess the handling qualities of the simulated aircraft.

Examples of references that used the HQ engineering as a framework to quantify overall simulator fidelity are as followings: Padfield et al. (1996) and McCallum and Charlton (2001) first proposed the handling qualities standard, ADS-33E (2000) ‘Handling Qualities Requirements for Military Rotorcraft’, as the basis for deriving metrics. Hess and colleagues [Hess and Malsbury (1991), Hess and Siwakosit (2001), Schroeder et al. (2000)] introduced the handling qualities sensitivity function (HQSF) as a quality metric. Advani and Wilkinson (2001) and Roscoe and Thompson (2003) used comparative performance and control activity, and handling qualities ratings given for the same tasks flown in simulation and flight.

The Lifting Standards project [White et al. (2010)] focused on developing a metric-based engineering framework for the assessment of fidelity of a simulator. In this sense, a two-stage approach was used for defining fidelity criteria for simulator qualification. 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 the simulator fidelity rating scale SFR [Perfect et al. (2014)].

Recalling the area of HQ engineering, two assessment processes, prediction and assignment, are integrated, and combine to give the overall HQ of an aircraft. This is also the case for the integrative process and the HQ performance specification, [ADS-33E (2000)]. For both processes, the test aircraft is assessed to be in one of three handling qualities ‘levels’. Level 1 HQ (HQ ratings (HQR) of 1, 2, and 3) indicate that there is no requirement for improvement to the aircraft, and that all operational tasks can be accomplished with low workload. In Level 2 (with HQR of 4, 5, 6), the workload will be higher, and the level of precision reduced, but the safety of the aircraft is not significantly at risk. If Level 3 HQs are found (HQR > 6) then the level of workload has increased to the extent that task performance is no longer achievable. At the higher end of level 3 (HQR 9, 10; sometimes defined as Level 4), flight safety is compromised as the risk of loss of control increases.

In the Lifting Standards project, as in the HQ assessment process, a comparison of results from predictive and perceptual assessments formed a key component of the overall fidelity assessment process. This was required to establish that the predicted and perceptual results were consistent. A flow diagram from [Perfect et al. (2013)] representing the process

for the assessment of predicted and perceptual simulator fidelity is shown in Figure 10.

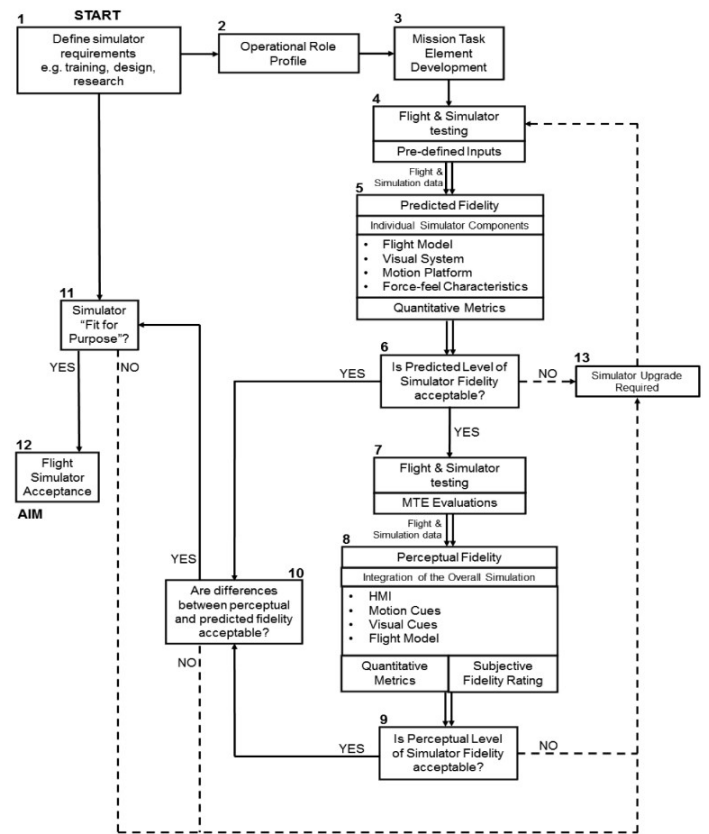


Figure 10. Methodology for integrated predicted and perceptual simulator fidelity assessment [Perfect et al. (2013)].

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 10), which will set the required level of fidelity. Once the purpose of the simulator has been defined, testing on the simulator and the simulated aircraft can be conducted (Block 4). This leads to the computation 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 analyzed 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 (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 HQ, 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 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 recognized that a simulator may be fit for some purposes but not others and thus have limited fidelity.

8. CONCLUDING REMARKS

The credibility of any simulation system depends on its fidelity. Simulation fidelity, in particular simulator fidelity, has an inherent multidimensional and multifaceted nature. Its assessment and specification can be a complex and time consuming task. The present paper has brought together different aspects of simulation fidelity evaluation methods and metrics for rotorcraft as applied to flight dynamics modelling, handling qualities evaluations and ground-based simulator developments. The following conclusions can be drawn from the presented review:

- Frequency-domain and Time-domain Integrated Cost Functions are useful metrics for assessing the predictive accuracy of any rotorcraft simulation model. In addition to these metrics, ValCrit-F and ValCrit-T metrics represent equivalent metrics based on Chi-square statistics.
- The development of the Maximum Unnoticeable Added Dynamics (MUAD) stems from the concept of equivalent system approach. Their assigned boundaries can be used to evaluate the mismatch between an actual aircraft response and a lower-order equivalent system (LOES) model. The extension of MUAD to the “Allowable Error Envelopes” AEE for validation of rotorcraft simulation model in a ground-based simulator should be further developed and tested separately for fixed-base piloted simulations, moving-base piloted simulations and for in-flight piloted simulations.
- Fidelity is a fundamental concept in development of ground-based simulators and should be seen as comprised of at least two elements: objective fidelity (also called predictive fidelity or physical fidelity) involving a series of ‘predictive’ metrics for flight model fidelity and also for simulator’s components including motion, visual, control loading and sound replication

and 2) subjective fidelity (also called perceptual or behavioral fidelity) involving the subjective opinion of the pilot, which is always a combination of how the complete system is able to replicate the cognitive skills required on the flight deck. The “Qualification Test Guide” (QTG) is used by FAA and EASA as the functional performance standard to assess simulator performance.

- Given the predictive and perceptual fidelity relationships discovered, perceptual and predictive fidelity should not be dealt with in isolation. Herein, the Simulation Fidelity Rating (SFR) scale aims to create the foundation on which predictive and perceptual fidelity can be tackled together in order to enhance the realism of the integrated simulated pilot experience.

Authors contact:

Marilena D. Pavel m.d.pavel@tudelft.nl

Mark Tischler mrtischler@aol.com

Mark White mdw@liverpool.ac.uk

Olaf Stroosma o.stroosma@tudelft.nl

Michael Jones michael.Jones@dlr.de

David Miller david.g.miller@boeing.com

Vincent Myrand-Lapierre vincent.myrandlapierre@cae.com

Michel Nadeau-Beaulieu michel.nadeaubeaulieu@cae.com

Armin Taghizad armin.taghizad@onera.fr

ACKNOWLEDGMENTS

This work has been performed as part of The Applied Vehicle Technology (AVT) Panel of the NATO Science and Technology Organization (STO), NATO AVT-296.

REFERENCES

Ad Hoc Advisory Subcommittee on Avionics, Controls, and Human Factors (1979), “Proposed Study of Simulation Validation/Fidelity for NASA Simulators”, November, Washington, D.C., National Aeronautics and Space Administration.

ADS-33E (2000), “Aeronautical Design Standard, Performance Specification, Handling Qualities Requirements for Military Rotorcraft”, US Army, ADS-33E-PRF.

Advani, S. K., and Hosman, R. J. A. W. (2006), “Revising Civil Simulator Standards – An Opportunity for Technological Pull”, AIAA Modeling and Simulation Technologies Conference and Exhibit, Paper No. AIAA-2006-6248.

Advani, S. K., Hosman, R. J. A. W., and Potter, M. (2007), “Objective motion fidelity qualification in flight training simulators”, AIAA Modeling and Simulation Technologies Conference and Exhibit, Paper No. AIAA-2007-6802.

Advani, S. K., and Wilkinson, C. H. (2001), “Dynamic Interface Modelling and simulation – a Unique Challenge”,

- Royal Aeronautical Society Conference on Helicopter Flight Simulation, London, UK.
- AGARD (1980). Fidelity of Simulation for Pilot Training (Report No. AGARD-AR-159). Paris: Advisory Group for Aerospace Research and Development
- Buchholz, J. J., Bauschat, J. M., Hahn, K. U., and Pausder, H. J. (1996), "ATTAS & ATTheS In-Flight Simulators: Recent Application Experiences and Future Programs", AGARD-CP577, Paper no. 31, April.
- Cooper, G. E., and Harper, R. P. Jr. (1969), "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities", NASA TN D-5133.
- EASA CS-FSTD(H) (2012), "Certification Specifications for Helicopter Flight Simulation Training Devices", European Aviation Safety Agency EASA CS-FSTD(H), June, <https://www.easa.europa.eu/certification-specifications/cs-fstdh-helicopter-flight-simulation-training-devices>, Retrieved on June 4th, 2020.
- FAA 14 Part 60 (2016), "Flight Simulation Training Device Initial and Continuing Qualification and Use", FAA 14 Part 60 Appendix C to Part 60-Qualification Performance Standards for Helicopter Full Flight Simulators.
- Gross D.C., and Freeman, R.H. (1997), Measuring Fidelity Differential in HLA (high level architecture) Simulations, '97 Fall Simulation Interoperability Workshop, September, paper no. 125.
- Hamel, P. G., and Jategaonkar, R. V. (1996), "The Evolution of Flight Vehicle System Identification", *Journal of Aircraft*, 33(1), pp. 10-28, Jan.
- Haverdings, H., Dequin, AM, Basset, P. M., von Grünhagen, W., Kampa, K, Massey, C., McCallum, A., and Meerwijk, E. L. (2000), "Mathematical modeling for the prediction of helicopter flying qualities within GARTEUR -phase 3", European Rotorcraft Forum, The Hague, Netherlands, Sept. 26–29.
- Havighurst, L.C., Fields, L.E., and Fields, C.L. (2003). High versus low fidelity simulations: does the type of format affect candidate's performance or perceptions. In proceedings from the 27th annual IPMAAC conference on personnel assessment.
- Hays, R.T., and Singer, M.J., (1989), "Simulation Fidelity in Training System Design: Bridging the Gap between Reality and Training", Springer-Verlag, London.
- Heffley, R. K., Clement, W.F., Ringland, R.F., and Jewell, W.F. (1981), "Determination of motion and Visual System Requirements for Flight Training Simulators," Systems Technology Inc., Technical Report No. 546, August.
- Hess, R. A., and Malsbury, T. (1991), "Closed loop assessment of flight simulator fidelity", *Journal of Guidance, Control and Dynamics*, 14(1), pp 191–197.
- Hess, R. A., and Siwakosit, W. (2001), "Assessment of flight simulator fidelity in multi-axis tasks including visual cue quality", *Journal of Aircraft*, 38(4), pp 607–614.
- Hodge, S. J., Perfect, P., Padfield, G. D. ,and White, M. D. (2015a), "Optimising the Yaw Motion Cues Available from a Short Stroke Hexapod Motion Platform", *The Aeronautical Journal*, January, 119(1211), pp. 1-21.
- Hodge, S. J., Perfect, P., Padfield, G. D., and White, M. D. (2015b), "Optimising the Roll-Sway Motion Cues Available from a Short Stroke Hexapod Motion Platform", *The Aeronautical Journal*, January, 119(1211), pp. 23-44.
- Hodgkinson, J. (1998), *Aircraft Handling Qualities*, AIAA Educational Series, AIAA, Reston, VA.
- Hodgkinson, J. (2005), "History of Low-Order Equivalent Systems for Aircraft Flying Qualities", *Journal of Guidance, Control, and Dynamics*, July-August, 28(4), pp. 577-583.
- Hosman, R. J. A. W., and Advani, S. K. (2012), "Status of the ICAO Objective Motion Cueing Test," Flight Simulation Research, New Frontiers Conference Proceedings, Royal Aeronautical Society, London.
- Jones, M., White, M., Fell, T., and Barnett, M. (2017), "Analysis of Motion Parameter Variations for Rotorcraft Flight Simulators", Proceedings of the 73rd AHS Annual Forum & Technology Display, Fort Worth, Texas, USA, May 9-11.
- Jones, M. (2018), Enhancing motion cueing using an optimisation technique. *The Aeronautical Journal*, 122(1249), 487-518, doi:10.1017/aer.2017.141.
- Lane N.E., and Alluisi E.A. (1992), "Fidelity and Validity in Distributed Interactive Simulations: Questions and Answers", Document D-1066, Institute for Defense Analyses, Alexandria VA, November.
- McCallum, A. T., and Charlton, M. T. (2001), "Structured Approach to Helicopter Simulator Acceptance", The Challenge of Realistic Rotorcraft Simulation, November Royal Aeronautical Society Conference, London, UK.
- McDonald B. (1998), "A proposed process for defining required fidelity of simulations", Spring Simulation Interoperability Workshop, March, Paper 234.
- Meyer R.J. (1998), "You can say Resolution, I say Fidelity, You say accuracy, I say detail, ...Let's not call the whole thing off!", Spring Simulation Interoperability Workshop, March, paper 245.

- Mitchell, D. G., Doman, D. B., Key, D. L., Klyde, D. H., Leggett, D. B., Moorehouse, D. J., Mason, D. H., Raney, D. L., and Schmidt, D. K., (2004) Evolution, Revolution, and Challenges of Handling Qualities, *Journal of Guidance, Control and Dynamics*, 27(1), January–February 2004, pp.12-28
- Mitchell, D. G., Hoh, R. H., He, C., and Strobe, K. (2006b), “Development of an aeronautical design standard for validation of military helicopter simulators”, 62nd Annual Forum of the American Helicopter Society, Phoenix, AZ.
- Mitchell, D. G., Nicoll, T. K., Klyde, D. H., and Schulze, C. (2009), “Effects of Time Varying Rotorcraft Dynamics on Pilot Control”, AIAA Atmospheric Flight Mechanics Conference, AIAA 2009-6055, Aug.
- MIL-STD-1797B (2006), “Flying Qualities of Piloted Aircraft”, Department of Defense Interface Standard.
- Miletović, I. (2020), “Motion Cueing Fidelity in Rotorcraft Flight Simulation: A New Perspective using Modal Analysis”, Delft University of Technology, PhD Thesis
- Miletovic I., Pavel, M.D., Pool, D.M, Stroosma, O., van Paassen, M.M., Mulder, M., Wentink, M. (2021) “Eigenmode Distortion: A New Perspective on Motion Cueing Fidelity in Flight Simulation”, *Journal of Guidance, Control and Dynamics*, April.
- Miller, D., Kocher, E., Hughes, D. F., Grugan, R. A., and Taylor, J.E. (2009), “Complementary Moving Base and Dynamic Motion Seat Cueing for Rotorcraft Simulation”, AIAA Modeling and Simulation Technologies Conference, Chicago.
- Mihram, Arthur G. (1972), “Some Practical Aspects of the Verification and Validation of Simulation Models”, *Journal of the Operational Research Society*, 23, pp17–29.
- Morgan, J. M., and Baillie, S. W. (1993), “ADS-33C Related Handling Qualities Research Performed using the NRC Bell 205 Airborne Simulator,” *Piloting Vertical Flight Aircraft: A Conference on Flying Qualities and Human Factors*, NASA CP 3220, January 20–22, 1993.
- Myers III, Paul L., Starr, Arnold, W., and Mullins, Kadie (2018), *Flight Simulator Fidelity, Training Transfer, and the Role of Instructors in Optimizing Learning*, *International Journal of Aviation, Aeronautics, and Aerospace*, 5(1). <https://doi.org/10.15394/ijaaa.2018.1203>
- Myrand-Lapierre V., Tischler M.B., Pavel M.D., Nadeau-Beaulieu M., Stroosma O., and Gubbels, A. W. (2020), “Bell 412 Modeling and Model Fidelity Assessment for Level-D Training Simulators”, *Vertical Flight Society Forum 76th*, Oct. 6-8, Virginia Beach.
- Pace, D. K., (2015), “Fidelity, Resolution, Accuracy, and Uncertainty”, Chapter 3 in Springer-Verlag London 2015 M. L. Loper (ed.), *Modeling and Simulation in the Systems Engineering Life Cycle*, Simulation Foundations, Methods and Applications, DOI 10.1007/978-1-4471-5634-5_3.
- Padfield, G. D., Charlton, M. T., and McCallum, A. T. (1996), “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.
- Padfield G.D., (2013) “Rotorcraft Handling Qualities Engineering: Managing the Tension between Safety and Performance”, 32nd Alexander A. Nikolsky Honorary Lecture, *J. of The American Helicopter Society* 58, 011001, pp.1-27, DOI: 10.4050/JAHS.58.011001
- Padfield, G. D., McCallum, A. T., Dequin Haddon, D., Kampa, K. Basset, PM., and von Grünhagen, W., (1996), “Predicting Rotorcraft Flying Qualities through Simulation Modelling. A Review of Key Results from GARTEUR HC/AG-06”, 22th European Rotorcraft Forum, Brighton, United Kingdom, 16 - 19 September.
- Padfield, G. D., Pavel, M. D., Casolaro, D., Roth, G., Hamers, M., and Taghizad, A. (2004), “Validation criteria for helicopter real-time simulation models, Sketches from the work of GARTEUR HC/AG-12”, 30th European Rotorcraft Forum, Marseilles, France.
- Pavel, M. D., White, M. D., Padfield, M.D., Roth, G., Hamers, M., and Taghizad, A. (2013), “Validation of mathematical models for helicopter flight simulators current and future challenges”, *The Aeronautical Journal*, 117(1190), pp. 343 – 388, April.
- Penn, H. (2013), “Investigating the Feasibility of applying MUAD to Helicopter Simulator Model Validation: An analysis of the effect of multiple simulators on Allowable Error Envelopes using multiple pilots”, Unpublished M.Sc. thesis, Delft University of Technology, July.
- Perfect, P., White, M. D., Padfield, G. D., and Gubbels, A.W. (2013), “Rotorcraft Simulation Fidelity: New Methods for Quantification and Assessment”, *The Aeronautical Journal*, March, 117(1189), pp. 235-282.
- Perfect, P., Timson E., White, M. D., Padfield, G. D., Erdos, R., and Gubbels, A. W. (2014), “A Rating Scale for the Subjective Assessment of Simulation Fidelity”, *The Aeronautical Journal*, August, 118(1206), pp. 953 – 974.
- Reardon, S. E., Beard, S. D., and Aponso, B.L. (2014) “Effects of Motion Filter Parameters on Simulation Fidelity Ratings”, 70th Annual Forum of the American Helicopter Society, Montreal, Canada May 20–22.

- Rehmann, Albert J. (1995), "A Handbook of Flight Simulation Fidelity Requirements for Human Factors Research", DOT/FAA/CT-TN95/46.
- Reid, L. D., and Nahon, M. A. (1985), "Flight simulation motion-base drive algorithms: Part 1 – Development and testing of the equations", University of Toronto, Institute of Aerospace Studies, UTIAS Report No. 296.
- Roscoe, M. F., and Thompson, J. H. (2003), "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, May.
- Roza, Z.C. (2005), "Simulation Fidelity Theory and Practice, A Unified Approach to Defining, Specifying and Measuring the Realism of Simulations", Ph.D. Dissertation, Delft University of Technology, January.
- Schroeder, J. A. (1999), "Helicopter Flight Simulation Motion Platform Requirements", NASA-TP-1999-208766.
- Schroeder, J. A., Chung, W. W. Y., and Hess, R. A. (2000), "Evaluation of motion fidelity criterion with visual scene changes", *Journal of Aircraft*, 37(4), pp 580–587.
- Sinacori, J. B., (1977), "The Determination of Some Requirements for a Helicopter Research Simulation Facility", NASA-CR-152066.
- Smith, R. E., Hodgkinson, J., and Snyder, R. C. (1981), "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, OH.
- Snedecor G. and Cochran W., (1989), *Statistical Methods*, 8th Edition, ISBN-13: 978-0813815619, Blackwell Publishing
- Tischler, M. B., and Remple, R. K. (2012), *Aircraft and Rotorcraft System Identification: Engineering Methods With Flight-test Examples*, 2nd Edition, AIAA Educational Series, November.
- Tischler, M.B., D'Agosto, S., Cameron, N., Greiser, S., Gubbels, A., Guner, F., He, C., Horn, J., Hui, K., Jones, M., Juhasz, O., Lee, O., Lehmann, R., Miller, D., Myrand-Lapierre, V., Nadeau-Beaulieu, M., Nadell, S., Padfield, G.D., Pavel, M.D., Prasad, J., Ragazzi, A., Richard, S., Scepanovic, P., Seher-Weiß, S., Soong, J., Stroosma, O., Taghizad, A., Tobias, E., White, M.D., Xin, H., and Yavrucuk, I. (2021), "Rotorcraft Flight Simulation Model Fidelity Improvement and Assessment", NATO STO AVT-296 Technical Report.
- Tischler, M. B. (1995), "System Identification Methods for Aircraft Flight Control Development and Validation", NASA Technical Memorandum 110369.
- Wang Y., White M.D., Owen I., Hodge S. and Barakos G. (2012), "Effects of visual and motion cues in flight simulation of ship borne helicopter operations", Paper 137, 38th European Rotorcraft Forum, Marriott Hotel, Amsterdam, 4-7 September.
- White, M. D., Perfect, P., Padfield, G. D., Gubbels, A. W., and Berryman, A. C. (2010), "Progress in the Development of Unified Fidelity Metrics for Rotorcraft Flight Simulators", 66th Annual Forum of the American Helicopter Society, Phoenix, AZ, May.
- White, M. D., Pavel, M. D., Jones, M., Spira, D., Cuzieux, F., and van den Vorst, J. (2020), "Rotorcraft Simulation Fidelity Assessment Predicted and Perceived Measures of Fidelity", Report for GARTEUR HC/AG-21, GARTEUR TP-189, March.

APPENDIX A

This appendix presents the basic structure of the Simulation Fidelity Rating (SFR) Scale.

		Comparative Performance		
		Equivalent	Similar	Not Similar
Strategy Adaptation	Negligible	LEVEL 1 <i>Full Transfer of Training</i> (SFR 1-2)	LEVEL 2 <i>Limited Transfer of Training</i> (SFR 3-6)	LEVEL 3 <i>Negative Transfer of Training</i> (SFR 7-9)
	Minimal			
	Moderate			
	Considerable			
	Excessive			

Figure A1 Simulation Fidelity Rating Fidelity Matrix [Perfect et al. (2014)]

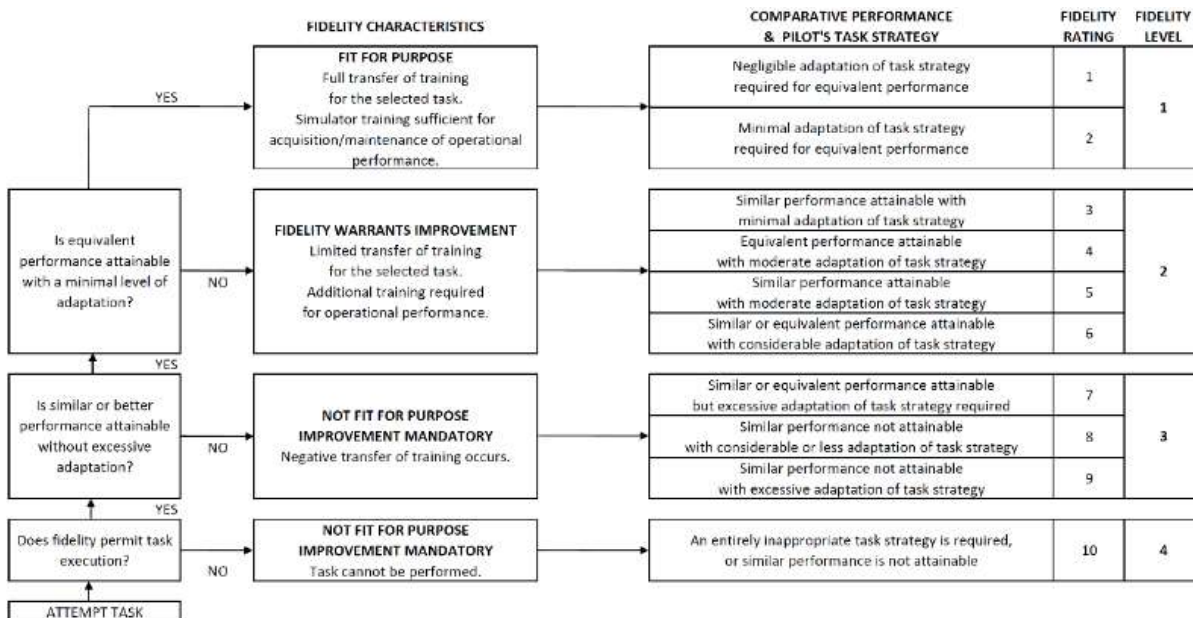


Figure A2 Simulation Fidelity Rating Scale [Perfect et al. (2014)]