European Rotorcraft Forum

THE DEVELOPMENT OF A PILOT CONTROL ADAPTATION METRIC FOR SIMULATION PERCEPTUAL FIDELITY ASSESSMENT

Wajih A. Memon, Neil Cameron, Mark D. White, Gareth D. Padfield
The University of Liverpool, UK

Linghai Lu
Cranfield University, UK

Abstract

This paper reports the use of a control compensation metric to examine pilot adaptation in the objective assessment of simulation perceptual fidelity. The utility of the proposed metric to quantify different levels of pilot control compensation, hence adaptation, whilst flying low and high aggression tasks is explored. The tasks were conducted by different test pilots using the Heliflight-R simulator to examine the effect of additional transport delays on overall simulation perceptual fidelity. A weighted adaptive control compensation metric shows strong correlation with (Cooper-Harper) Handling Qualities and Simulation Fidelity Ratings awarded for each of the tasks. Moreover, in combination with a time-varying frequency-domain exposure, the metric is shown to be insightful for understanding variations in the pilots' assessment of simulation perceptual fidelity.

1. INTRODUCTION

During flight trials, assessment of the Handling Qualities (HQs) of an aircraft is traditionally undertaken with Test Pilots (TPs) awarding Handling Qualities Ratings (HQRs) using the scale developed by Cooper and Harper in 1969 [1]. While the scale was developed for the evaluation of an aircraft's HQs, it has also been used as a metric for supporting the assessment of flight simulation fidelity [2–4]. A fundamental element of the HQR process is for the TP to report the control compensation required to overcome any vehicle system deficiencies that could inhibit them from flying a task to operationally relevant performance and safety standards. Cooper and Harper defined pilot compensation as [1]:

"*The measure of additional pilot effort and attention required to maintain a given level of performance in the face of less favourable or deficient vehicle characteristics*"

The HQR scale incorporates a decision tree structure (Appendix A: Figure A1) which ensures that the HQR couples:

a) the extent of the HQ deficiencies (no worse than mildly unpleasant (Level 1), minor to very objectionable (Level 2), major (Level 3) or so severe that there is a high risk of loss of control (HQR 10)),

with,

b) the achieved performance (desired, adequate, inadequate),

and,

c) the required compensation (not a factor, minimal, moderate, considerable, extensive, maximum tolerable, controllability in question) [1].

Drawing on the experience of the HQ community in capturing pilot subjective assessment, and following the methodology utilised within the HQR scale, a Simulation Fidelity Rating (SFR) scale for capturing pilot's assessment of overall simulation fidelity was developed at The University of Liverpool (UoL) [5]. The scale employs some of the key concepts that are critical to the effectiveness of the simulation devices. These include the transfer of training, comparative task performance and task strategy adaptation.

As with the structure of the HQR scale, the SFR scale also incorporates a decision tree format (Appendix A: Figure A2) which couples together:

a) the simulation fidelity characteristics/deficiencies (fit for purpose (Level 1), fidelity warrants improvement (Level 2), not fit for purpose-improvement mandatory (Level 3) or so severe that the task cannot be performed (Level 4)),

with,

b) the comparative performance (equivalent, similar, not similar),

and,

c) the task strategy adaptation (negligible, minimal, moderate, considerable, excessive) [5].

For simulation fidelity assessment using the SFR

scale, the relationship between task performance and task strategy adaptation is similar to that between task performance and compensation in the HQs evaluation. The task strategy adaptation refers to various aspects of a pilot's behaviours related to the overall task; this includes the pilot's control activity, visual scan patterns, mental workload, eye/head movement and utilisation of cues [5]. Pilot control adaptation is a key element of the SFR scale and is considered in this paper to be reflected in the changes in compensation.

Over the years, these rating scales have gained significant international acceptance and worldwide adoption in various research and development activities [6–8]. The ratings are likely to be 'affected' by individual pilot biases, the subjective element of making the award. This includes pilot training and operational background, situational awareness, fatigue level and environmental factors [9]. With these variable influences on subjective assessment, coupled with a desire to better understand the pilot-system interactions, the development of a robust metric for the quantification of the pilot control compensation when flying a task has been a subject of interest for the rotorcraft modelling and simulation community for many years [2,9–14]. The development of such a metric can support the pilot's subjective assessment of the differences in compensation that might be experienced in flight and simulation. Here we refer to the differences in pilot control compensation, flight versus simulation, as control adaptation.

The research reported in this paper builds on the authors' published work [15] where they introduced a new metric that captures control compensation based on time-domain analysis.

2. BACKGROUND

The Flight Science and Technology group at UoL is engaged in research to quantify the fidelity of rotorcraft simulators [2,8,16-17], for its use in design, development, crew training, and qualification. The research has addressed both the 'accuracy' of the flight model (predictive fidelity) and the 'realism' of the integrated simulation as experienced by a pilot (perceptual fidelity) [18]. The present paper focuses on the perceptual element of simulation fidelity, particularly on the quantification of pilot compensation and what we describe as adaptation.

Current simulator certification standards, e.g. EASA's CS-FSTD(H) [19] and FAA's CFR Part 60 [20], provide criteria for the assessment of the component fidelity of civil flight simulators that define the acceptable match of flight test data between aircraft and simulation models. The criteria provide qualification performance standards for objective evaluation based on tolerances, i.e., acceptable differences between flight and simulation. However, it is understood that meeting these standards will not guarantee good overall perceptual fidelity [21]. Whilst an assessment of the simulator's perceptual fidelity is undertaken by an evaluation pilot during certification, robust objective metrics for this type of assessment have still to be generally agreed on [5].

Previous research conducted by the authors examined the utility of a particular form of 'Control Compensation Metric' (CCM) in examining pilot workload [15]. The study was aimed at the development of a metric to quantify and predict the extent of pilot control compensation required to fly a wide range of mission task elements in flight and simulation trials. This paper reports further development of the CCM that helps in the objective assessment of the simulation perceptual fidelity to understand pilot adaptation; the ways in which pilots change their control strategy when transferring from simulator to flight and vice-versa. A weighted-adaptive CCM is used to quantify the extent of pilot control adaptation to examine, in this paper, simulator 'deficiencies' produced by different simulator time (transport) delays.

In previous time-delay research, measures of the pilot control activity were used in attempts to quantify pilot adaptation and understand its correlation with the subjective assessment [16,22]. However, a suitable standardised metric to quantify and predict different levels of pilot control compensation has remained a challenge and a goal in the UoL research. The research presented in this paper addresses this requirement using the newly developed CCM, sensitive to the changes in the pilot control compensation, to complement the subjective assessments [15].

3. RESEARCH METHODOLOGY

This research has used test data gathered in a previous research project: "Flight Simulation Fidelity for Rotorcraft Design, Certification and Pilot Training" [16]. As part of this work, a series of experiments were conducted using the HELIFLIGHT-R simulator [23] to examine the effect of Additional Transport Delay (ATD) (+100, +200 and +300ms) on simulation perceptual fidelity [16,22]. The nominal transport delay associated with the HELIFLIGHT-R simulator at the time of testing was approximately +100ms [16]. Therefore, the transport delay to be investigated is the ATD.

The essence of the initial study was to use the SFR scale [5] to examine the effect of changes in the simulator time delay, (as well as additional cross-coupling and stick gearing/force parameters) on perceptual fidelity. As described in the Introduction, the SFR scale couples together the pilot's task strategy adaptation with comparative task

performance; either between flight and simulation or between different simulation configurations. Pilot control adaptation is a key element of the SFR scale and is considered in this paper to be reflected in the changes in compensation. A baseline simulation model of the National Research Council's Bell 412 Advanced Systems Research Aircraft [24] was used in this study. The ATD modifications were introduced to assess the impact on perceptual fidelity on a low task-aggression Precision Hover (PH) and a higher task-aggression Acceleration-Deceleration (AD) Mission Task Element (MTE) [25]. The MTE descriptions are provided in Appendix B.

3.1. Objective Metrics Development

3.1.1. Control Compensation Metric (CCM)

The CCM has been derived using the control attack concept formulated by Padfield et al. [26]. The control (η) attack metric, A_η, is based on the rate and magnitude of the pilot's control input, along the lines of the attitude quickness parameter to quantify agility [25]. The A_η parameter (Eqn.1) characterises each discrete control input and is defined as the ratio of the peak rate of control displacement, $\dot{\eta}_{pk}$, to the magnitude of the change in the control displacement, $\Delta\eta$ [26].

(1) $$A_\eta = \frac{\dot{\eta}_{pk}}{\Delta\eta}$$

In broad terms, higher attack values indicate rapid-small control deflections while a lower attack value indicates slower-large control deflections. As with the ADS-33 attitude quickness parameter [25], the attack approximates the inverse of the time to change for a simple ramp. An example of a high attack and low attack pilot control input is shown in Figure 1.

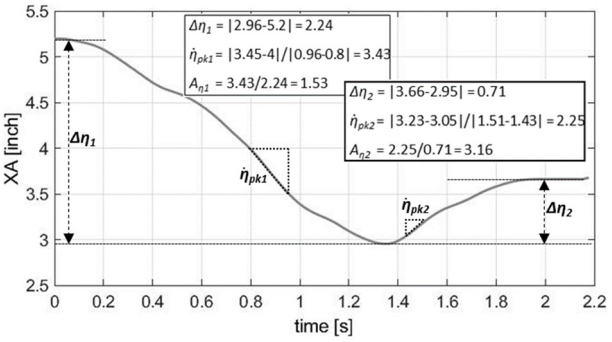


Figure 1: Lateral control deflection time history showing high ($A_{\eta 2}$) and low ($A_{\eta 1}$) values of attack

Using the control attack concept, a number of metrics have previously been explored to characterise the pilot control activity in a particular control of interest [2]. Two of those metrics which are utilised in the development of the CCM are the Attack Number ($A_{\eta N}$) and Attack Activity Rate ($A_{\eta R}$) metrics:

a) Attack Number ($A_{\eta N}$): the total number of times that the pilot displaces a particular control having a magnitude higher than a specified threshold value; selected as 2.5% of the full control range. The appropriate threshold has been determined to capture 'productive' control inputs representing guidance and stabilisation control activity [15].

b) Attack Activity Rate ($A_{\eta R}$): the ratio of the total number of A_η points to a defined time; this might be the duration of the MTE to give an 'average' value or a particular time slice to give a 'local' value. It can be defined as the 'busyness' metric. The $A_{\eta R}$ has the same units as the A_η parameter (i.e., 1/s).

In the development of the CCM, the $A_{\eta R}$ metric was extended in the form of a weighted-adaptive control compensation metric, used to correlate subjective pilot assessments with pilot control activity. A weighted-adaptive metric combines all control's $A_{\eta i}$ weighted with the corresponding fraction of the total control attacks, $A_{\eta Tot}$, applied in that axis (Eqn. 2).

(2) $$\frac{A_{\eta Ni}}{A\eta_{NTot}}$$

where *i* corresponds to the four control axes (lateral '*XA',* longitudinal *'XB',* collective *'XC' and* pedal *'XP'*).

The values for each control axes are then summed to obtain an accumulated weighted adaptive metric '$A_{\eta RC}$' (Eqn. 3).

(3) $$A_{\eta RC} = \sum_{i=XA}^{XP} A_{\eta R_i} \cdot \frac{A_{\eta Ni}}{A\eta_{NTot}}$$

Alongside the $A_{\eta RC}$, the combined $A_{\eta R}$ for the primary $A_{\eta R\text{-}Pr}$ and secondary $A_{\eta R\text{-}Sec}$ controls separately are computed in the adaptation investigations. To calculate these, the same formulation as defined in (Eqn. 3) is used by accumulating the specific controls only. The primary and secondary controls are different for each of the two MTEs examined, PH and AD, based on their task performance requirements. For the PH MTE, lateral and longitudinal are primary, with collective and pedal control axes as secondary. For the AD MTE, the longitudinal control axis is primary; lateral, collective and pedal control axes are secondary [15].

In this paper, the utility of the combined adaptive CCM in quantifying different levels of pilot control compensation (i.e., adaptation) will be demonstrated by examining its correlation with the HQRs and SFRs awarded by different TPs.

3.1.2. Time-localised Attack Metric

In a previous study [2], the average attack activity rate for a complete MTE was used, but this averaging masks local peaks in pilot control compensation

within the MTE. In the HQ assessment process adopted at Liverpool, pilots are required to assign the HQR based on the most challenging phase of an MTE, rather than an average. To support the development of the CCM [15], a time-varying attack activity rate metric $A_{\eta R}^{loc}$, was formulated, using a 5sec time window (with 2.5secs overlap), to capture local control attack activity rate peaks ($A_{\eta R}^{pk}$) during an MTE and to identify where and how the pilot is 'working the hardest' [15].

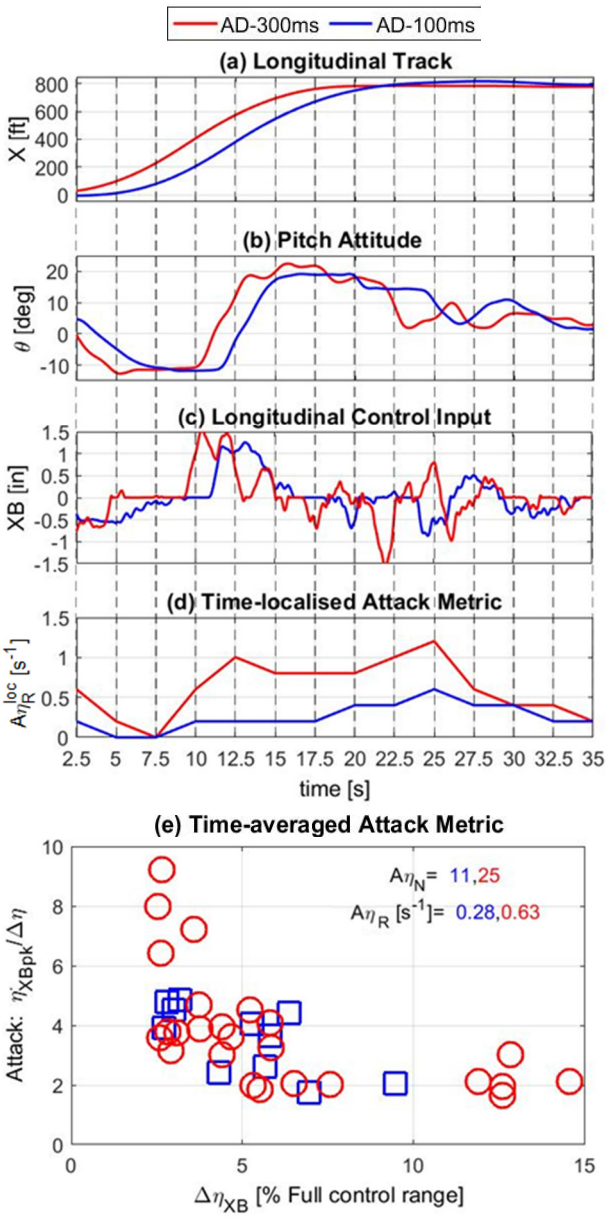


Figure 2: Longitudinal track, Pitch, Longitudinal cyclic, Time-averaged and Time-localised $A_{\eta R}$ metrics for AD MTE (+100ms vs +300ms ATD)

Figure 2 shows a longitudinal axis example comparison of the time-localised (d) and time-averaged (e) attack metrics performed in the AD MTE, for two different flight model configurations, ATD +100ms and +300ms. The time-localised attack metric $A_{\eta R}^{loc}$ captures local control peaks during the evolution of an MTE and allows differences in pilot control behaviour i.e., adaptation, due, for example, to a change in the simulation environment.

Comparing the two simulation test cases, the time-averaged attack metric in Figure 2 (e) shows 11 $A_{\eta N}$ with 0.28/s $A_{\eta R}$ for the +100ms case and 25 $A_{\eta N}$ with 0.63/s $A_{\eta R}$ for the +300ms case. This gives an overall picture of control compensation, but does not provide any information about the phase(s) within the MTE where the pilot might be struggling or compensating the most (e.g., peak compensation phase(s)). On the other hand, the time-localised $A_{\eta R}^{loc}$ plot in Figure 1(d) provides key information of the $A_{\eta R}$ during different phases through the MTE; thus identifying where the pilot is working the hardest.

From Figure 1(d) it can be seen that, the $A_{\eta R}$ for +100ms case remains lower than that for +300ms case throughout the MTE. In +300ms case, two obvious $A_{\eta R}^{pk}$ ≥1/s, at 12.5 and 25secs, correspond to the pitch-reversal (deceleration) and level-off (hover capture) phases of the MTE, respectively. However, in +100ms case, the $A_{\eta R}$ between t=10-17.5secs remains constant at 0.2/s and only one obvious $A_{\eta R}^{pk}$ of 0.6/s is observed at t=25secs. Overall, for the +300ms case, larger variations can be seen in $A_{\eta R}^{loc}$.

With such large variations in the $A_{\eta R}$ through the MTE, using the 'time-averaged attack can be misleading when compared with the subjective HQ ratings, and will not help in the identification of the dominant compensation/adaptation phase(s).

3.1.3. Spectackogram: Spectrogram and Attack Metric Composite

To further demonstrate the applicability of the $A_{\eta R}$ metric in assessing pilot control adaptation, the Fourier transform-based spectrograms [27] were combined with the $A_{\eta R}^{loc}$ metric to develop a composite time-varying frequency-domain exposure, called the 'Spectackogram'.

Figure 3 shows an example of a spectackogram plot. The central spectrogram contour-plot is formulated using a moving 4 secs window, selected to provide a suitable resolution in both time and frequency, thus highlighting key features of the control signal. The spectrogram shows a three-dimensional visual representation of the spectrum of frequencies within the control input as a function of time, with the colour-bar indicating the magnitude of the Short-Time Fourier Transform (STFT) of the control signal; the outer plots show the time history (upper), Fast Fourier Transform (FFT, left) and $A_{\eta R}^{loc}$ (lower).

Using this presentation format, details in different phases of the MTE can be observed, along with information for the frequency content as a function of time. In Figure 4, the $A_{\eta R}^{loc}$ plot shows the first $A_{\eta R}^{pk}$ at t=12.5secs, corresponding to the pitch-reversal (deceleration) phase. At the same location in time, the spectrogram plot shows a high-amplitude/low-frequency 'hot-spot' <0.3Hz. Another medium-amplitude/high-frequency hot-spot, with frequency content up to 0.6Hz, is observed between t=20-25secs window, where the $A_{\eta R}^{pk}$ increases up to 1.2/s. This corresponds to the level-off (hover capture) phase. The $A_{\eta R}^{loc}$, in combination with the Spectrogram, provides complementary insight into the compensation required to complete the task.

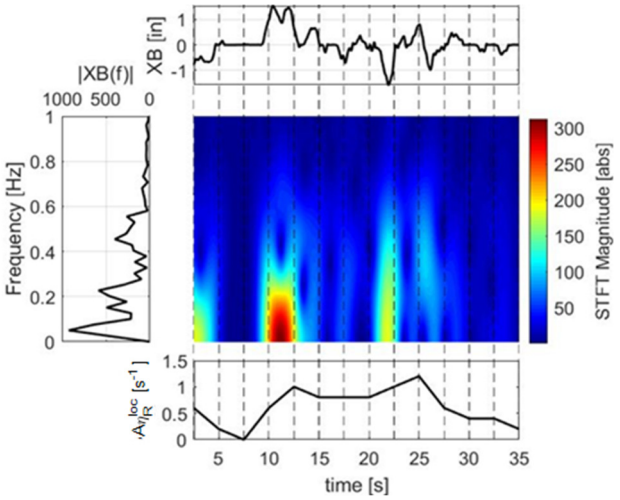


Figure 3: Spectackogram composite for +300ms ATD case (AD MTE)

In this section, the utility of the proposed metric has been demonstrated for an example AD MTE case only. In the following section, its application is explored and discussed in detail by examining the correlation with the HQRs and SFRs awarded by the pilots for the two MTEs with different task performance requirements and dominant control axes. The examination consists of two case studies, one from HQR, the other from SFR, assessments to demonstrate the application of the developed metrics and the spectackogram charts.

4. CORRELATION OF ADAPTATION METRIC WITH PILOT SUBJECTIVE ASSESSMENT

4.1. Case Study I: HQRs vs CCM

Simulator test data gathered from two MTEs (PH and AD) were analysed to quantify pilot compensation and compared with the HQRs that provide the subjective assessment of compensation. Comparisons of the peak adaptive CCM with the HQRs are presented in Figure 4, showing results for all four controls combined $A_{\eta RC}^{pk}$ (a), and separated primary $A_{\eta R\text{-}Pr}^{pk}$ (b) and secondary $A_{\eta R\text{-}Sec}^{pk}$ (c) controls, for a single test pilot flying both MTEs. The figure also shows the best fit straight-line (dashed) between the HQRs and CCM, along with the coefficient of determination (r^2), and 85 % prediction boundary calculated using the T-distribution [28]. The coloured regions on the plots differentiate the HQ levels from the HQR scale, with the associated compensation descriptors on the secondary y-axis.

It can be seen from Figure 4(a) that three of the AD MTE configurations are located in the top-right region of the band suggesting higher levels of control compensation. The low-ATD PH MTE configurations are located in the bottom-left region, with correspondingly lower levels of pilot compensation. Moreover, a strong positive correlation is obtained between the HQRs and $A_{\eta RC}^{pk}$, with the r^2 value of 96 %. As expected, as the ATD increases, the HQR worsens as the pilot has more difficulty achieving the precision requirements.

The correlations for separated primary and secondary controls are lower (see Figure 4(b) and (c)). This can be explained by the different sources of deficiency and consequent compensation in the different cases for the same HQR. For example, in +0ms PH case, the dominant compensation was from the primary controls (*XA, XB)*, whereas for the ATD+100ms PH case, the dominant compensation was from the secondary controls (*XC, XP).* The $A_{\eta RC}^{pk}$ for both the cases is similar (0.275-0.325) so reflecting the different contributions to overall compensation. The separation of the controls into primary and secondary is found to be important for understanding the respective contribution of each to the corresponding $A_{\eta RC}$ values since, within each of these cases, different controls will be dominant during different phases of the MTEs. To illustrate this, two extreme HQR cases; (HQR 2 (+0ms PH) and HQR 7 (+300ms AD)) have been selected for comparison, to represent '*not a factor*' and '*maximum tolerable'* levels of control compensation.

For the ATD+300ms AD case, the $A_{\eta RC}^{pk}$ is more than 300% of the 0ms PH case (see Figure 4(a)). Correlating this observation with the Spectackogram plots in Figure 5, the magnitude, as well as the frequency content for the ATD+300ms AD case is significantly larger than the ATD+0ms PH case. In terms of the primary and secondary controls separately, the secondary controls for the ATD+300ms AD case dominate with $A_{\eta R\text{-}Sec}^{pk}$ being 150% of the $A_{\eta R\text{-}Pr}^{pk}$. Whereas, for the ATD+0ms PH case, the primary controls dominate with $A_{\eta R\text{-}Pr}^{pk}$ being 200% of the $A_{\eta R\text{-}Sec}^{pk}$.

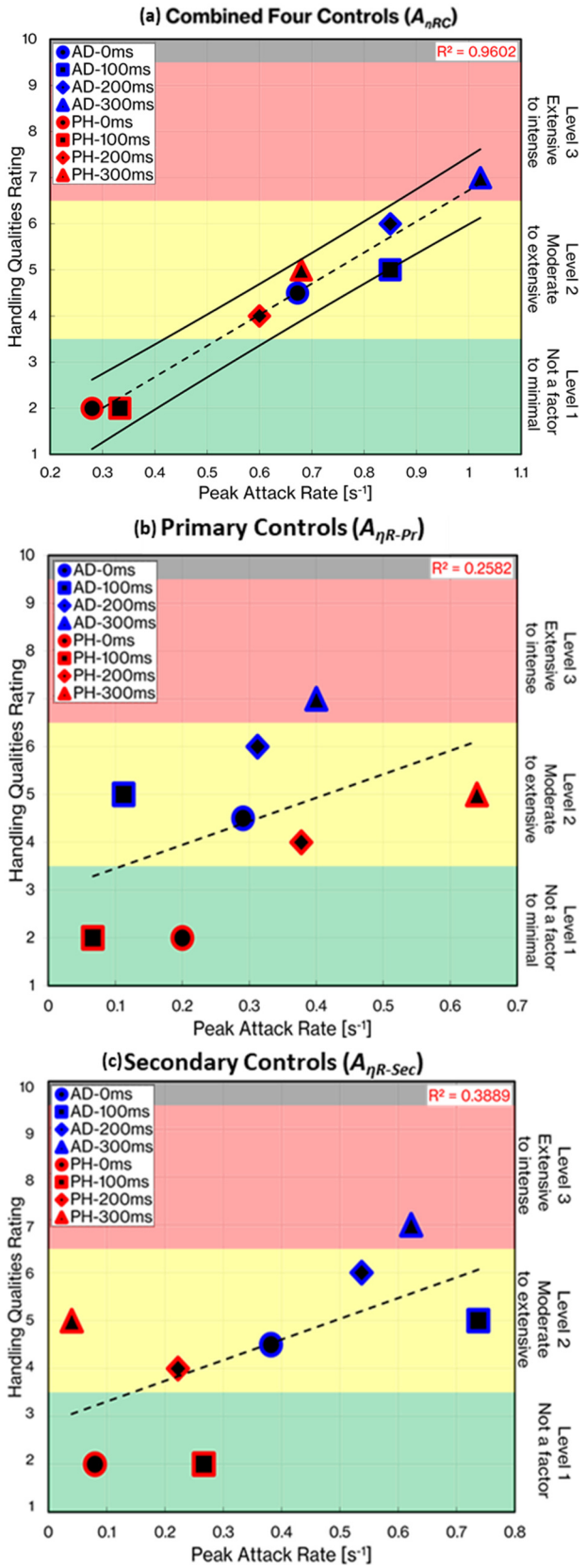


Figure 4: HQR vs peak-combined $A_{\eta R}$ for (a) combined, (b) primary and (c) secondary control axes accumulated

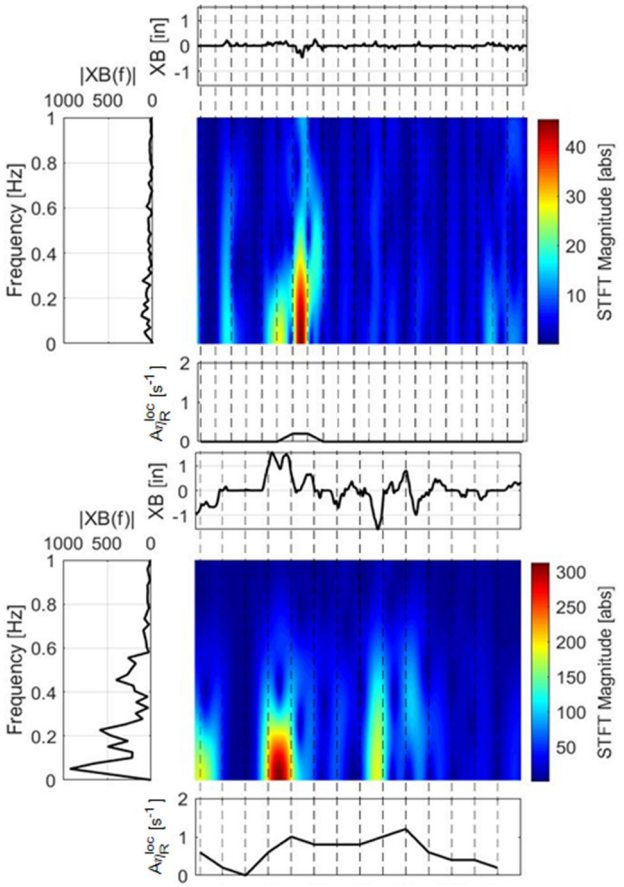


Figure 5: Primary control spectackograms for HQR-2 'PH +0ms' (top) and HQR-7 'AD +300ms' (bottom)

4.2. Case Study II: SFRs vs CCM

In this case study, the CCM metric has been applied to understand the spread in the subjective assessment of the simulation perceptual fidelity across TPs, having different piloting styles and strategies. This is achieved by correlating the combined $A_{\eta RC}{}^{pk}$ CCM with SFRs awarded for the same PH MTE, ATD+100 ms case, extending across all three adaptation levels (see Figure 6). The pilots awarded their SFRs based on adaptation and performance compared with the baseline configuration. The coloured regions on the plots differentiate the simulation fidelity levels from the SFR scale, with the associated adaptation descriptors on the secondary y-axis.

As within the HQR case study (Figure 4(a)), a strong positive correlation is obtained between the SFRs and $A_{\eta RC}{}^{pk}$, with the r^2 value of 97 %. This suggests that the pilot who awarded higher SFR applied higher control attack for the same test condition. A striking result is that, for the same ATD configuration and MTE, significant variance in the SFRs (1-7) between pilots is seen. This may be due to several causal factors, such as different control strategies, task interpretation, pilot training/experience, etc. Using the CCM (Figure 6), it can be seen that the $A_{\eta RC}{}^{pk}$ for

Pilot D is more than 350% of Pilot A and around 120-140% of Pilot B and C.

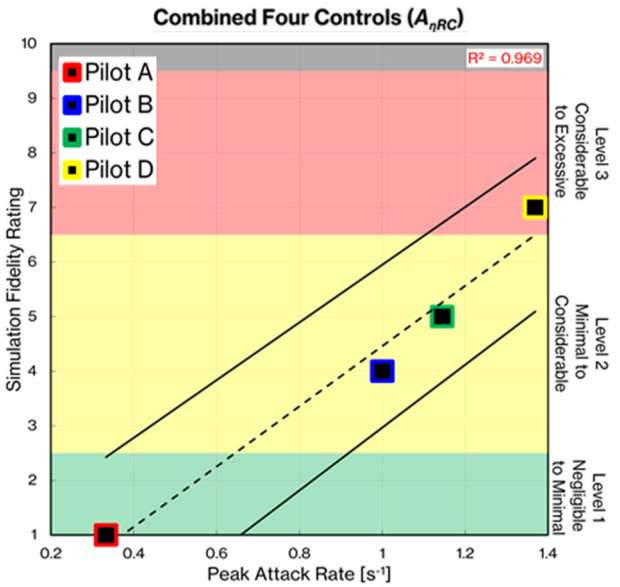


Figure 6: SFR vs $A_{\eta RC}^{pk}$ for four test pilots (PH MTE)

To understand different levels of control compensation applied by different pilots, Figure 7 shows a comparison of the longitudinal 'primary' axis spectackogram charts for the two extreme cases (Pilot A and D), to represent the '*negligible'* and '*excessive'* levels of the task strategy adaptation, respectively. The spectackogram plots show significant differences in the level of control compensation between the two pilots. The magnitude, frequency content and the $A_{\eta RC}^{loc}$ clearly show how control compensation increases with SFR.

For Pilot A, the time-localised attack plot show $A_{\eta R}^{pk}$ of 0.2/s, corresponding to the start of the MTE. After this point, the control activity magnitude remained below the attack threshold. For this case, the pilot commented that he "*perceived no differences"* in comparison to the baseline (ATD+0ms) case. In contrast, for Pilot D, the plots show higher control compensation with multiple $A_{\eta R}^{pk}$ of 1.6/s, which correspond to the hover stabilisation phase of the MTE. For this case, the pilot commented that the "*lateral cyclic was easily excited"* and "*adaptation was required in multiple axes*". The level of control compensation in other axes can be seen from the spectackogram charts provided in Appendix C. For Pilot D case, the plots clearly show multiple hot-spots and $A_{\eta R}$ peaks >1/s. However, for Pilot A, the overall control compensation is significantly lower with the lateral and collective control activity below the attack threshold throughout the MTE.

The spectackogram plots, along with the CCM, provide detailed insight into different levels of control compensation applied by different pilots, which correlate with their subjective assessments.

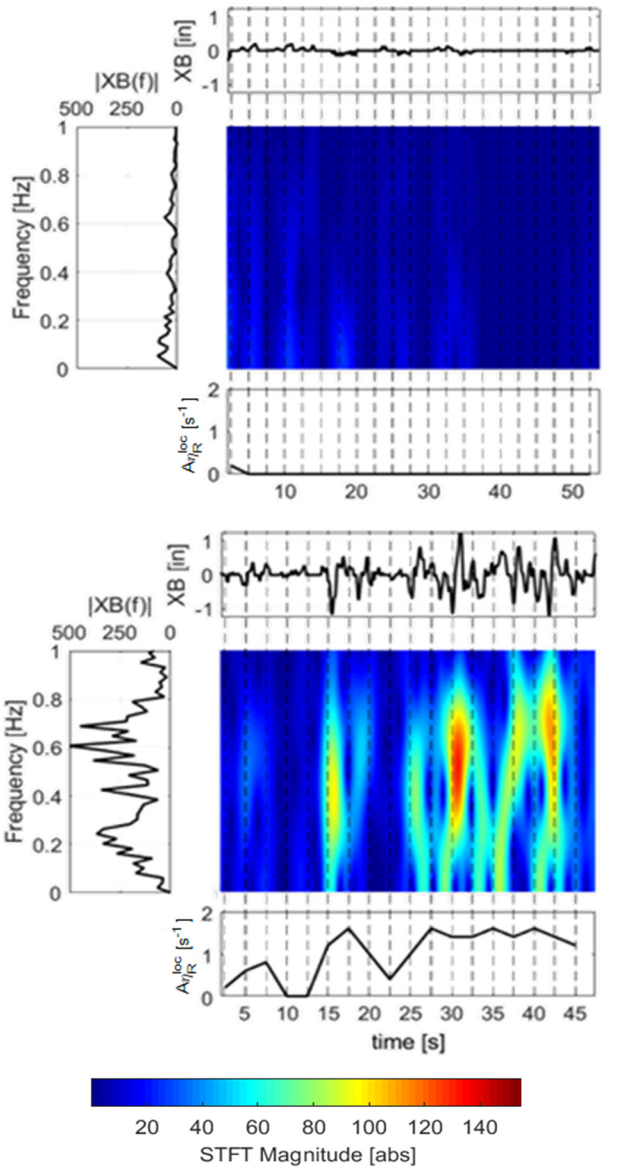


Figure 7: Spectackogram plots for "Pilot-A SFR-1" (top) and "Pilot-D SFR-7" (bottom) for the longitudinal 'primary' axis

5. CONCLUDING REMARKS

The paper has described further development of a control compensation metric, suitable for quantifying and explaining the variations in the perception of handling qualities and simulation fidelity across pilots. One such problem was to understand the spread between pilots in subjective assessment of adaptation when flying the PH MTE with the same configuration. The weighted adaptive metric showed a strong correlation with the handling qualities and simulation fidelity ratings awarded by pilots for two different mission task elements. The following are the key conclusions drawn from this work:

- The proposed CCM has shown a strong correlation with the HQRs and SFRs returned from a range of TPs. Results for separated primary and secondary controls show much weaker correlation with HQRs. However, separation is useful in understanding the different contribution of each control.
- By analysing the correlation between the CCM and pilot subjective assessments across TPs for the same ATD/MTE case, the results have helped in understanding the spread of the SFRs. It was determined that the pilot who awarded higher SFR applied higher control compensation for the same test condition. The use of the CCM confirms that the control compensation plays a significant role in characterising adaptation within the SFRs.
- It has been observed that the pilots who apply higher frequency/magnitude control activity appear to be more sensitive to transport delays. The CCM metric and spectackogram plots have reflected this and have shown strong correlation with the SFRs and HQRs.

The results presented in this paper are based on a relatively limited number of tests and further investigations are required to determine the adaptation level boundaries, in relation to the HQR and SFR levels, which could help in the prediction and verification of subjective assessments.

In the ongoing research, the approach will be utilised to investigate the spread in the simulation fidelity ratings for higher task-aggression mission task elements, e.g., Acceleration-Deceleration, Roll-step.

ACKNOWLEDGEMENTS

The UK's Engineering and Physical Sciences Research Council (EP/P031277/1 and EP/P030009/1) funded the research reported in this paper. The work was carried out using the flight and simulator test data gathered during the Lifting Standards (EP/G002932/1) project. The use of the NRC's Bell 412 ASRA facility is gratefully acknowledged and special thanks to the test pilots who participated in the trails.

REFERENCES

[1] Cooper, G. E., and Harper, R. P., "The Use of Pilot Ratings in the Evaluation of Aircraft Handling Qualities," National Aeronautics and Space Administration, NASA-TND-5153, 1969.

[2] Perfect, P., White, M. D., Padfield, G. D., and Gubbels, A. W., "Rotorcraft Simulation Fidelity: New Methods for Quantification and Assessment," *The Aeronautical Journal*, Vol. 117, (1189), 2013, doi: 10.1017/s0001924000007983.

[3] Atencio, A. Jr., "Fidelity Assessment of a UH-60A Simulation on the NASA Ames Vertical Motion Simulator," National Aeronautics and Space Administration, NASA-TM-104016, 1993.

[4] Hess, R. A., and Malsbury, T., "Closed-loop Assessment of Flight Simulator Fidelity," *Journal of Guidance, Control and Dynamics*, Vol. 14, (1), 1991, pp. 191–197, doi: 10.2514/3.20621.

[5] Perfect, P., Timson, E., White, M. D., Padfield, G. D., Erdos, R., and Gubbels, A. W., "A Rating Scale for the Subjective Assessment of Simulation Fidelity," *The Aeronautical Journal*, Vol. 118, (1206), 2014, pp. 953–974, doi: 10.1017/S0001924000009635.

[6] Mitchell, D. G., "Fifty Years of the Cooper-Harper Scale," *AIAA Scitech 2019 Forum*, San Diego, CL, Jan. 7-11 2019, doi: 10.2514/6.2019-0563.

[7] Anon, "Qualification, Approval and Use of Flight Simulator Training Devices," MAA Regulatory Publications Additional Supporting Material for Regulatory Article (RA) 2375, 2020.

[8] NATO, "Rotorcraft Flight Simulation Model Fidelity Improvement and Assessment," TR-AVT-296-UU, 2021.

[9] Paul, R., and Rhinehart, M., "Exploring Pilot Workload Using Inceptor Time Histories," *Vertical Flight Society's 76th Annual Forum Proceedings*, Virtual, Oct. 5-8, 2020.

[10] Jones, J. G., Padfield, G. D., and Charlton, M. T., "Wavelet Analysis of Pilot Workload in Helicopter Low-level Flying Tasks," *The Aeronautical Journal*, Vol. 103, (1019), 1999, pp. 55–63, doi: 10.1017/S0001924000065106.

[11] Charlton, M. T., Howell, S. E., Padfield, G. D., Jones, J. G., Bradley, R., Thomson, D., and Leacock, G. "A Methodology for the Prediction of Pilot Workload and the Influence on Effectiveness in Rotorcraft Mission Tasks," *24th European Rotorcraft Forum*, Marseilles, France, Sep. 15-17, 1998.

[12] Jennings, S., Craig, G., Carignan, S., Ellis, D. K., and Thorndycraft, D., "Evaluating Control activity as a Measure of Workload in Flight Test," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 49 (1), 2005, doi:

10.1177/154193120504900115.

[13] Bachelder, E. N., Lusardi, J. A., Aponso, B., and Godfroy-Cooper, M., "Estimating Handling Qualities Ratings from Slalom Slight Aata: A Psychophysical Perspective," *Vertical Flight Society's 76th Annual Forum* Proceedings, Virtual, Oct. 5-8, 2020.

[14] Lampton, A. K., and Klyde, D. H., "Power Frequency: A Metric for Analyzing Pilot-in-the-loop flying tasks," *Journal of Guidance, Control and Dynamics*, Vol. 35, (5), 2012, pp. 1526–1537, doi: 10.2514/1.55549.

[15] Memon, W. A., White, M. D., Padfield, G. D., Cameron, N., and Lu, L. "Helicopter Handling Qualities : A Study in Pilot Control Compensation," *The Aeronautical Journal*, 2021. (accepted with minor modifications)

[16] Timson, E., "Flight Simulation Fidelity for Rotorcraft Design, Certification and Pilot Training," Doctoral thesis, The University of Liverpool, 2013.

[17] Cameron, N., White, M. D., Padfield, G. D., Lu, L., Agarwal, D., and Gubbels, A. W. "Rotorcraft Modeling Renovation for Improved Fidelity," *75th Annual Forum & Technology Display*, Philadelphia, PA, May 13-16, 2019.

[18] Cameron, N., Memon, W. A., White, M. D., Padfield, G. D., Lu, L, and Agarwal D., "Appraisal of Handling Qualities Standards for Rotorcraft Lateral-directional Dynamics," *AIAA Scitech 2021 Forum*, Virtual, Jan. 11-15, 19-21, 2021, doi: 10.2514/6.2018-1017.

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

[20] FAA, "Flight Simulation Training Device Initial and Continuing Qualification and Use," 14 CFR Part 60, 2006.

[21] Padfield, G. D., Casolaro, D., Hamers, M., Pavel, M., Roth, G., and Taghizad, A. "Validation Criteria for Helicopter Real-Time Simulation Models: Sketches from the Work of GARTEUR HC-AG12," *30th European Rotorcraft Forum*, Marseilles, France, Sep. 14-16, 2004.

[22] Timson, E., Perfect, P., White, M., Padfield, G. D., Erdos, R., and Gubbels, A. W. "Subjective Fidelity Assessment of Rotorcraft Training Simulators," *American Helicopter Society 68th Annual Forum*, Fort Worth, TX, 2012.

[23] White, M. D., Perfect, P., Padfield, G. D., Gubbels, A. W., and Berryman, A. C., "Acceptance Testing and Commissioning of a Flight Simulator for Rotorcraft Simulation Fidelity Research," *Proceedings of the Institute of Mechanical Engineering*, Part G: Journal of Aerospace Engineering, Vol. 227, (4), 2012, pp. 663–686, doi: 10.1177/0954410012439816.

[24] Manimala, B. J., Walker, D. J., Padfield, G. D., Voskuijl, M., and Gubbels, A. W. "Rotorcraft Simulation Modelling and Validation for Control Law Design," *The Aeronautical Journal*, Vol. 111, (1116), 2007, pp. 77–88.

[25] Anon, "Aeronautical Design Standard Performance Specification Handling Qualities Requirements for Military Rotorcraft," ADS-33E-PRF, 2000.

[26] Padfield, G. D., Jones, J. P., Charlton, M. T., Howell, S. E., and Bradley, R., "Where Does the Workload go when Pilots Attack Manoeuvres? An Analysis of Results from Flying Qualities Theory and Experiment," *20th European Rotorcraft Forum*, Amsterdam, The Netherlands, Oct. 4-7, 1994.

[27] Tritschler, J. K. O'Connor, J. C., Holder, J. M., Klyde, D. H., and Lampton, A. K., "Interpreting Time-frequency Analyses of Pilot Control Activity in ADS-33E Mission Task Elements," *American Helicopter Society 73rd Annual Forum*, Fort Worth, TX, May 9-11, 2017.

[28] Navidi, W. C., "Statistics for Engineers and Scientists," McGraw Hill, New York, NY, 2014.

Appendix A: Pilot Rating Scales

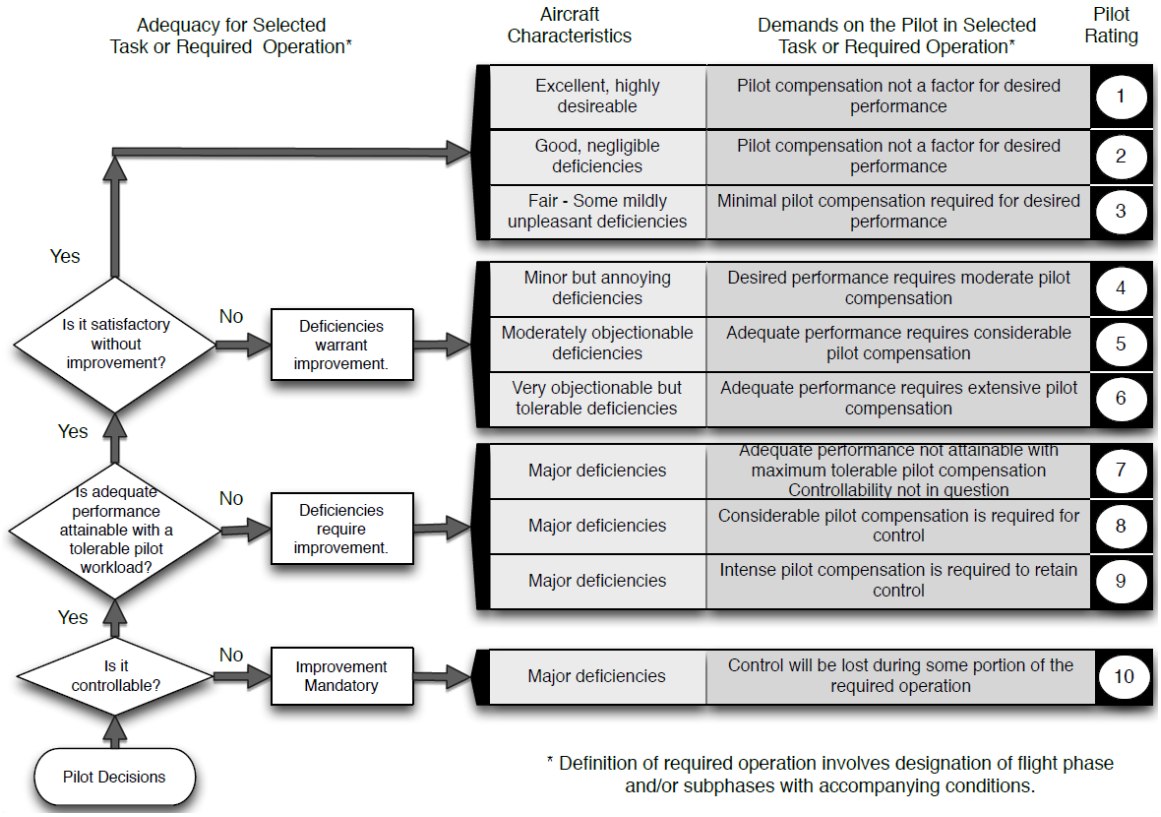


Figure A1: Cooper-Harper Handling Qualities Rating Scale [1]

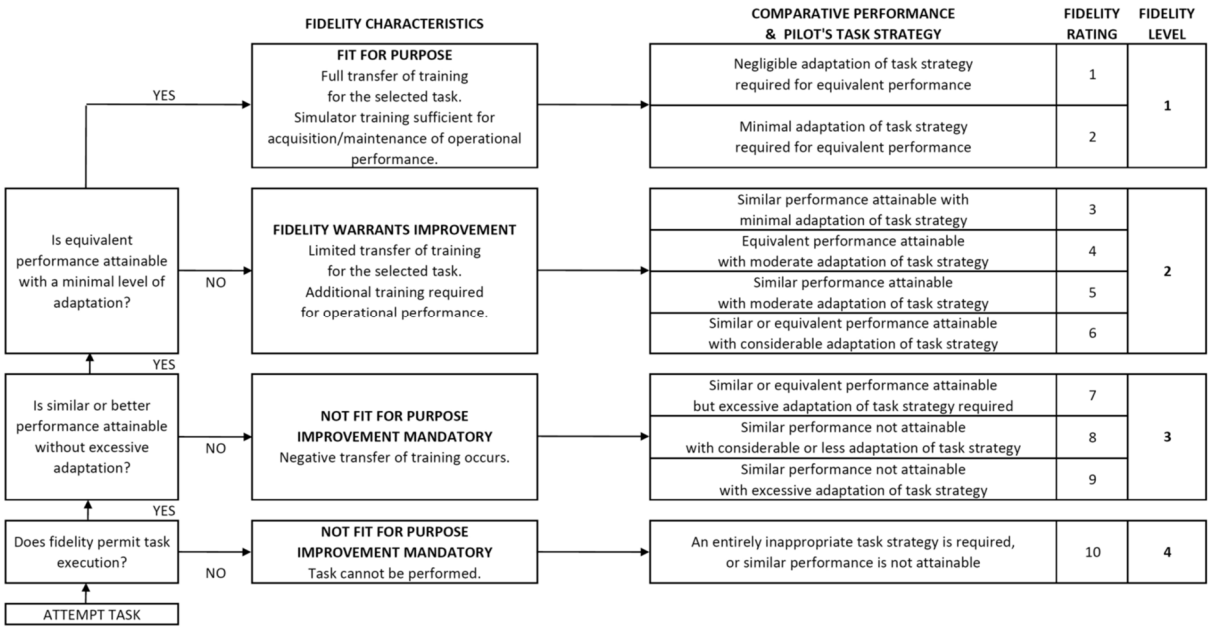


Figure A2: Simulation Fidelity Rating Scale [5]

Appendix B: Mission Task Elements (Cargo/Utility, Good Visual Environment)

Table B1: Precision Hover (PH) MTE definition

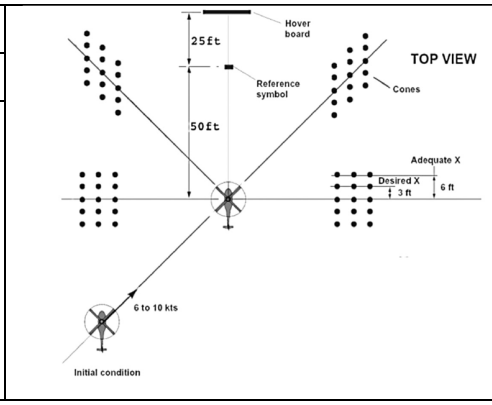
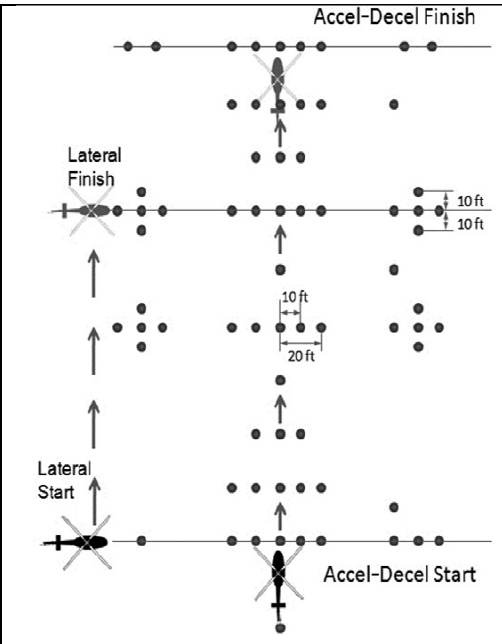
Manoeuvre Description	
Initiate the MTE at a ground speed of 6-10kts, at an altitude <20ft. The target hover point shall be oriented approximately 45° relative to the heading of the aircraft. The target hover point is a repeatable, ground-referenced point from which rotorcraft deviations are measured. The ground track should be such that the aircraft will arrive over the target point.	
Performance Standards	
Desired	**Adequate**
• Attain a stabilised hover within: 5secs • Maintain a stabilised hover for: 30secs • Maintain *lateral* and *longitudinal* within: ±3ft • Maintain altitude within: ±2ft • Maintain heading within: ±5° • There shall be no objectionable oscillations in any axis either during the transition to hover or the stabilised hover	• 8secs • 30secs • ±6ft • ±4ft • ±10°
	

Table B2: Acceleration Deceleration (AD) MTE definition

Manoeuvre Description	
The MTE starts with the aircraft in a stabilised hover. Power is rapidly increased to approximately maximum, maintaining constant altitude using pitch attitude, and holding collective constant during the acceleration to an airspeed of 40kts. Upon reaching the target airspeed, a deceleration is initiated by aggressively reducing the power and holding altitude constant with pitch attitude. The peak nose-up attitude should occur just before reaching the final stabilised hover. The MTE is completed with a stabilised hover for 5secs over reference point at the end. The distance from the starting point to the final hover position is a function of the performance of the rotorcraft, and is determined based on trial runs consisting of acceleration to the target airspeed, and decelerations to hover as described above.	
Performance Standards	
Desired	**Adequate**
• Maintain altitude below: 70ft • Maintain lateral track within: ±10ft • Maintain heading within: ±10°	• 100ft • ±20ft • ±20°
	

Appendix C: Spectackogram plots from Case Study: II (SFRs vs CCM)

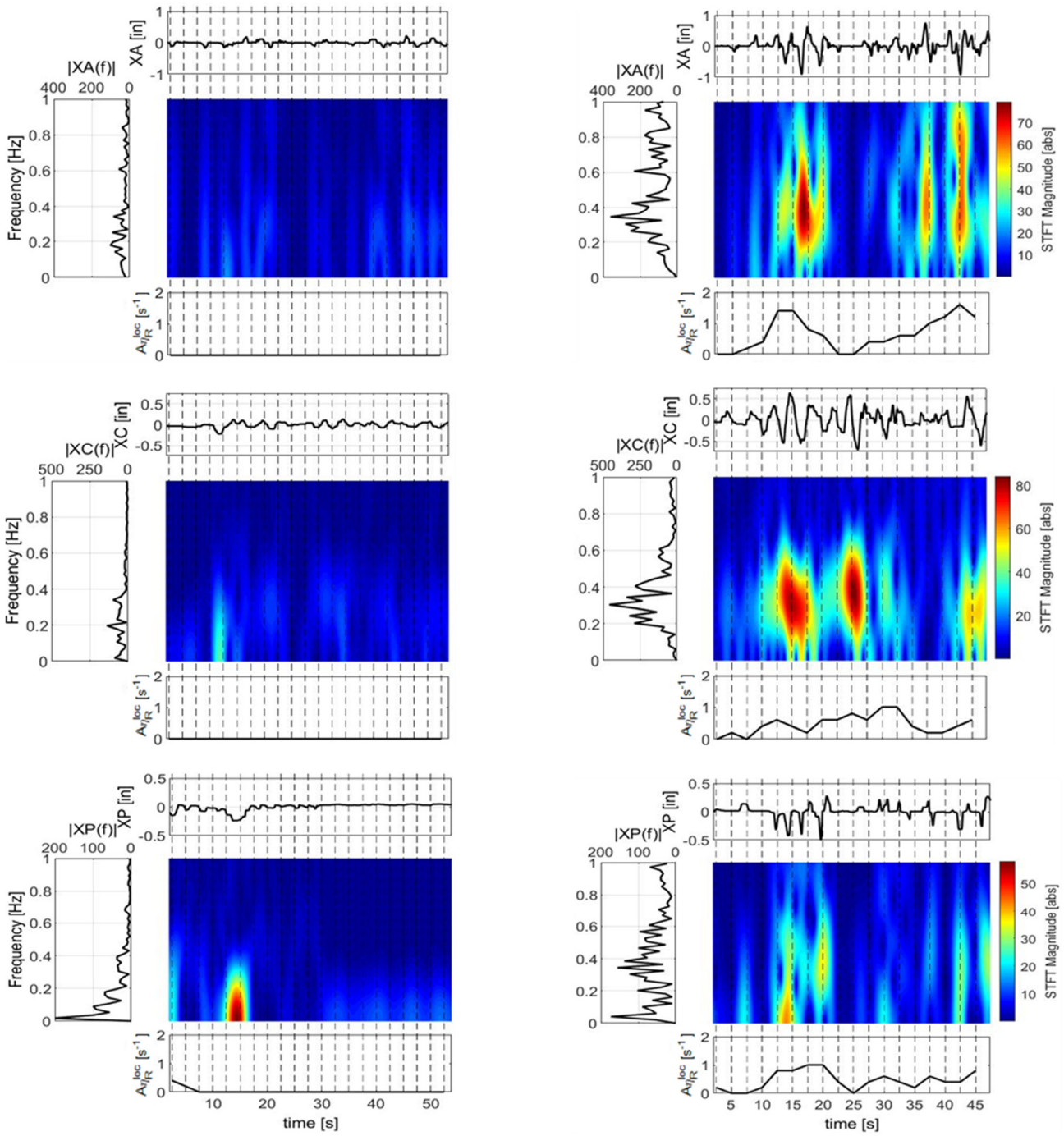


Figure C1: Spectackogram plots for "Pilot-A SFR-1" (left) and "Pilot-D SFR-7" (right) for the lateral 'primary', and collective and pedal 'secondary' axes (PH MTE)