Defence Science Journal, Vol. 66, No. 1, January 2016, pp. 03-10, DOI : 10.14429/dsj.66.9196

Application of Quantitative Measures for Analysing Aircraft Handling Qualities

Archana Hebbar[#,*] and Abhay A. Pashilkar[#]

[#]*CSIR-National Aerospace Laboratories, Bangalore - 560 017, India*
[*]*E-mail: p_archana@nal.res.in*

ABSTRACT

The ease and precision with which pilot is able to handle the designated task determines the aircraft's handling qualities. Accordingly, the most common methodology for determining aircraft's handling qualities is through pilot opinions or through questionnaires. These subjective means of analysis is not reliable as the sole source of judgments. Quantitative metrics to analyse the task difficulty based on pilot's performance, supplemented with subjective decision, can provide better insight into pilot workload levels and in turn the aircraft's handling qualities. Application of few objective performance measurement techniques to flight data of a high performance fighter aircraft is discussed in this paper. Pilot/aircraft's performance under different configurations is analysed. Analysis results show that pilots usually tend to give more priority to pitch axis in case of dual axis tracking task. And pilots are therefore more aggressive in accomplishing pitch axis tracking task than in roll. Workload assessments were also performed by comparing the results of single axis tracking experiments conducted using a high fidelity flight simulator with the flight data. It is seen that when roll axis control task is exercised as the primary task, pilot's aggressiveness levels in controlling the roll control inceptor is significantly less, along with improved tracking accuracy levels.

Keywords: Head up displays, root mean square error, power spectral density, pilot inceptor workload, induced oscillations, handling qualities

NOMENCLATURE

δ_i	Stick deflection
t_n	Flight time
Y_j	Fourier coefficients of stick data
Y_k	Time series of stick data
Y_c	Time series of tracking command
ρ	Level of significance
HQ	Handling qualities
RMSE	Root mean square error
HUD	Head up display
PIW	Pilot inceptor workload
PIO	Pilot induced oscillations
HTF	Highest task frequency
PSD	Power spectral density
ANOVA	Analysis of variance
CG	Centre of gravity

1. INTRODUCTION

Handling qualities (HQ) of an aircraft is principally determined by the performance of pilot and vehicle acting together as a system. HQ is most commonly analysed by means of pilot's opinion or by his/her judgements through questionnaires such as Cooper Harper rating. Such subjective assessment methods are proven to be prone to pilot's personal factors like pilot's mood, behaviour, background, surrounding environment and so on[1]. Hence for better understanding of the pilot-vehicle system, there is a need to quantify the same. In this paper, some quantitative tools are applied to real flight data from a high performance fighter aircraft and pilot/aircraft's performance under different configurations is analysed.

Previous human factor studies show the advantage of using quantitative assessment tools and techniques in addition to subjective judgements. A study conducted by Rantanen and his team in 2004 for the Federal Aviation Administration[2] describes the development of 9 metrics for pilot performance that were derived from time series of various flight parameters. Takallu[3], *et al*. examined the impact of different types of synthetic vision displays on pilot performance in a simulator environment using simple time domain metrics. In another study, Edmund[4], *et al*. observed the pilot's control activity using power spectral density (PSD) in three different simulator configurations. The work reported in this paper is different from the above with regards to the way in which available tools are combined together and applied to a practical aircraft HQ task.

Analysis is carried out with two different approaches of objective measurement metrics. One is to analyze the deviations from the desired flight path and the other is to analyse the perceived difficulty to perform the task. Tracking errors are analysed by standard statistical tools such as root mean square error (RMSE) and response delay. Task difficulty is indicated through inceptor control strategy measurement metrics. Both metrics put together can provide the workload as perceived by the pilot.

Data collated from head up display (HUD) tracking tasks conducted in flight using a high performance fighter aircraft is analysed for comparisons regarding tracking error and pilot's control strategy. Also, data from similar tracking tasks

Received 03 September 2015, revised 05 December 2015, online published 27 January 2016

conducted in simulated environment for the same aircraft is compared against real flight results. Points discussed in the present study

- Difference in pilot's control strategy with respect to pitch and roll control columns.
- Effect of flight configuration on tracking performance.
- Significance of task repetitions on the tracking errors and control behaviour.
- Comparison with respect to real flight data and pilot in the loop simulations.

2. FLIGHT TEST DATA AND THE ANALYSIS TOOLS

2.1 Flight Test Data

According to David[5], *et al*., the most important design tool for finding out and correcting PIO is to command unexpected trigger events and check pilot-vehicle interactions to such events. Step and ramp HUD tracking is one such HQ task which is used in this study. This task comprises of step and ramp commands for both pitch and roll axis simultaneously. The target symbol is programmed to move in both axes on the HUD as per the command. Pilot needs to maintain minimum error with respect to the target symbol and another HUD symbol representing the aircraft nose.

The flight data used herein is a product of the various flight tests of a high performance fighter aircraft conducted as part of analysing the aircraft's handing qualities. Data from six flight sorties in two different aircraft configurations (Table 1) are analysed. The tracking command segments are repeated more than once in most of the cases. Three different pilots have flown these configurations. Trends of discrete tracking command in pitch and roll angles are identified for the complete flight duration. Test flights were conducted for varying flight conditions (speed ranging from 165 knots to 470 knots; altitudes up to 32000 ft; mean angle of attack ranging from 4° to 9°).

Table 1. Flight configurations

S. No.	Pilot	No of tracking segments	Configuration	
P1	Pilot 1	2	Heavy stores configuration	Config1
P2	Pilot 2	4		
P3	Pilot 2	3		
P4	Pilot 3	6		
P5	Pilot 3	3	Normal stores configuration	Config2
P6	Pilot 3	1		
Total		**19 sorties**		

First 5 seconds data is not considered in order to avoid large deviations at start. The pitch and roll tracking commands and respective pilot responses for all the flight sorties put together are shown in Fig. 1.

Data from pilot in the loop simulations conducted on a high fidelity fixed base simulator for single axis tracking task is analysed and compared with the flight test results. The mathematical model used in the simulator has been extensively validated against flight test data of the aircraft. Data for 3 segments of pitch tracking and 2 segments of roll tracking is analysed.

2.2 Analysis Tools and Techniques

Following are the quantitative assessment tools used in this study.

2.2.1 *Pilot Control Strategy Measurement*

A key aspect of HQ design is to understand the way pilot acts on the control columns during flight. In this regard, two metrics, one using time series data and the other using frequency domain measures are discussed in this paper.

a. Pilot Inceptor Workload Metric

Pilot inceptor workload (PIW) is a time domain metric to quantify pilot gain. It is a two dimensional plot of pilot's aggressiveness versus duty cycle[6]. Pilot's aggressiveness is the pilot's stick speed. It describes the way pilot acts on the inceptor during flight. Duty cycle is the percentage of time pilot changes his input on the stick. It is a measure of pilot's effort. Aggressiveness is computed as the root mean squared average of rate of change of control stick position

$$Aggressiveness = \sqrt{\frac{1}{n-1}\sum_{i=2}^{n}\left(\frac{\delta_i - \delta_{i-1}}{t_i - t_{i-1}}\right)^2}$$

Duty cycle is represented as

$$Dutycycle = 100\% * \frac{1}{t_n - t_2}\sum_{i=2}^{n} x_i$$

$$\text{where } x_i = \begin{cases} 0;\ for\ \dfrac{\delta_i - \delta_{i-1}}{t_i - t_{i-1}} < noisethreshold; |\delta_i| < \delta_{max} \\ 1; otherwise \end{cases}$$

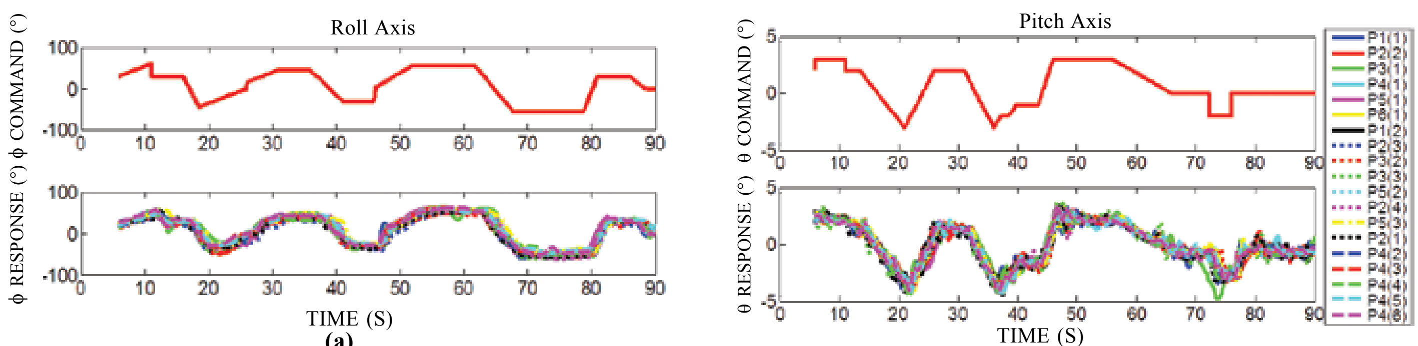


Figure 1. Command and response plots.

Here t_2 is the start time +1 and t_n is the end time of the data set; n is the number of data points; δ_i are the discrete values of the stick deflections in mm and δ_{max} is the maximum stick deflection. Noise threshold is taken as 0.5% of the inceptor's total displacement range per time increment.

A typical PIW plot is shown in Fig. 2. Duty cycle and aggressiveness tend to increase as pilot works for better performance. High aggressiveness and low duty cycle represents occasional fast inputs. This can be a case wherein pilot anticipates a change based on his pre-existing knowledge, applies input and waits for the aircraft to settle. However, in a target tracking task, pilot needs to track a target, moving with a predefined frequency. Hence, occasional fast stick inputs are not expected. Low aggressiveness and high duty cycle represents a constant, but slow stick movement. Low aggressiveness and low duty cycle can mean that the pilot is not in the loop, that is, he is not tracking the command signals well.

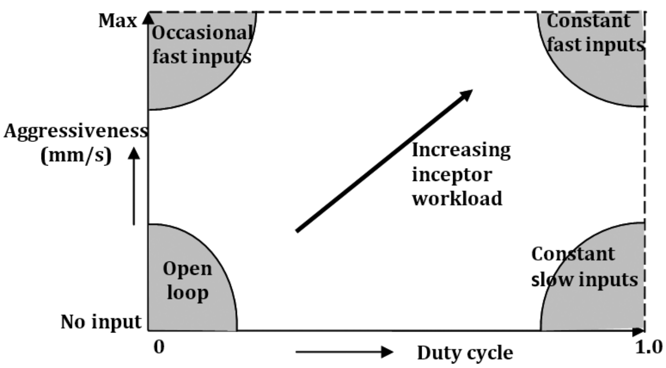


Figure 2. PIW plot.

b. Frequency Domain Measures

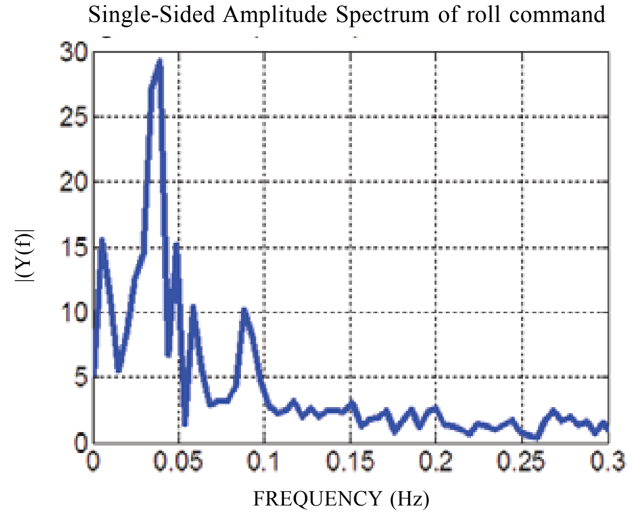
PSD plot of pilot's stick inputs is an important metric to measure pilot's control activity[7]. These plots determine the frequencies that the pilot uses during the task. Taking the Fourier transform of the time series data Y_k, we get the spectral decomposition

$$Y_k = \frac{1}{N}\sum_{j=1}^{N} \tilde{Y}_J \, e^{(\frac{2\pi i}{N})(k-1)(j-1)}$$

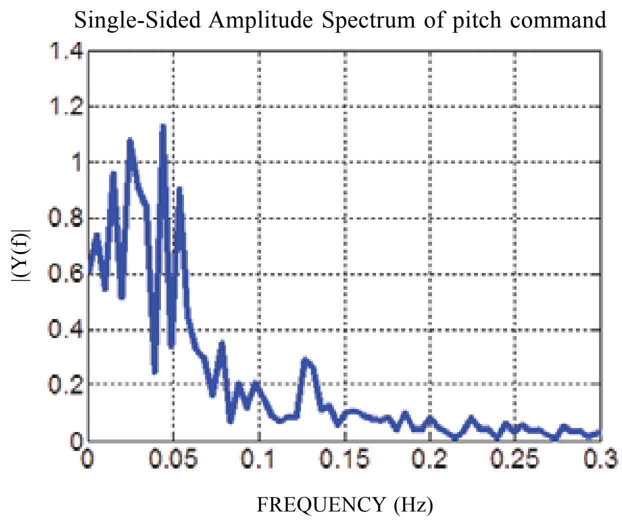
where Fourier coefficients $\tilde{Y}_J = \sum_{k=1}^{N} Y_k e^{(\frac{-2\pi i}{N})(k-1)(j-1)}$ and $i = \sqrt{-1}$

PSD is then given by $\frac{\left|\tilde{Y}_J\right|^2}{N}$

The tracking task discussed here is a discrete tracking task with a set of step and ramp inputs, at frequency around 0.04Hz (Fig 3). This is called the highest tracking frequency (HTF). PSD plot is divided into two areas of interest: Below HTF and above HTF. PSD peaks above HTF are scattered and are distributed over a range of frequencies. They are the result of aggressive pilot control movements (that is pilot gives high amplitude commands and later corrects in both the directions several times) or due to degradation in aircraft's HQ due to aircraft dynamics or both. Further the complete frequencies of interest are divided into five bands based on the cut off frequency (0-0.01Hz, 0.01-0.05Hz, 0.05-0.1Hz, 0.01-0.15Hz and 0.15-0.25Hz).



(a)



(b)

Figure 3. Tracking command frequencies.

Correlation between successive pilot commands can be represented by autocorrelation coefficient (r_h). r_h presents a measure of the correlation between successive data points (Y_k and Y_{k+h}) of time series $Y = \{Y_1, Y_2, ...Y_N\}$.

$$r_h = \frac{\sum_{k=1}^{N-h}(Y_k - \bar{Y})(Y_{k+h} - \bar{Y})}{\sum_{k=1}^{N}(Y_k - \bar{Y})^2}$$

where $\bar{Y} = \frac{1}{N}\sum_{k=1}^{N} Y_k$

A plot of r_h versus lag h gives a measure of how well a subsequent measurement can be predicted from a previous value. r_h varies from 0 to 1 where values close to zero indicate little correlation.

2.2.2 Tracking Accuracy Metrics

These metrics are basically used for event analysis and are fixed based on the test scenario. The time domain metrics used here are, RMSE, percentage of time spent outside tolerance interval (TD), number of deviations outside tolerance interval (ND) and the response delay[2].

RMSE summarises the overall position error. For a command of Y_c and inceptor deflections of Y_k (for n number of samples), RMSE is computed as

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n}(Y_{k,i} - Y_{c,i})^2}}{n}$$

ND counts the occurrences of the aircraft outside the predetermined tolerance interval. While RMSE collects the error magnitude information, ND measures the velocity errors. TD indicates the total time pilot spends outside tolerance.

$$\text{TD} = \frac{\text{cumulative time spent outside a given tolerance} * 100}{\text{Total time of segment}}$$

Response delay signifies the time taken to respond to changes in the command angle.

2.2.3 Analysis of Variance[7]

Analysis of variance (ANOVA) with a level of significance of 0.05 is used to analyse the significance of difference in performance.

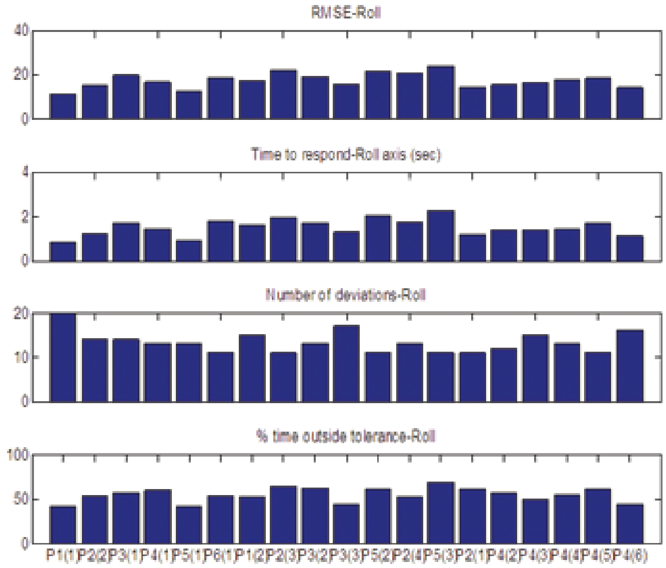
3. COMPARISON BETWEEN PITCH AND ROLL AXIS TRACKING IN FLIGHT DATA

Figure 4 shows the errors in roll and pitch attitude for all the sorties. 10% of the maximum command range is taken as the tolerance interval. In general, the RMSE in roll attitude are significantly higher than the errors in pitch [$F(1,36)=465.89$; $\rho<3.456e^{-22}$]. This is primarily due to the fact the task difficulty for roll axis tracking (roll command max range: $\pm 60°$) is higher than for pitch axis (pitch command max range: $\pm 3°$).

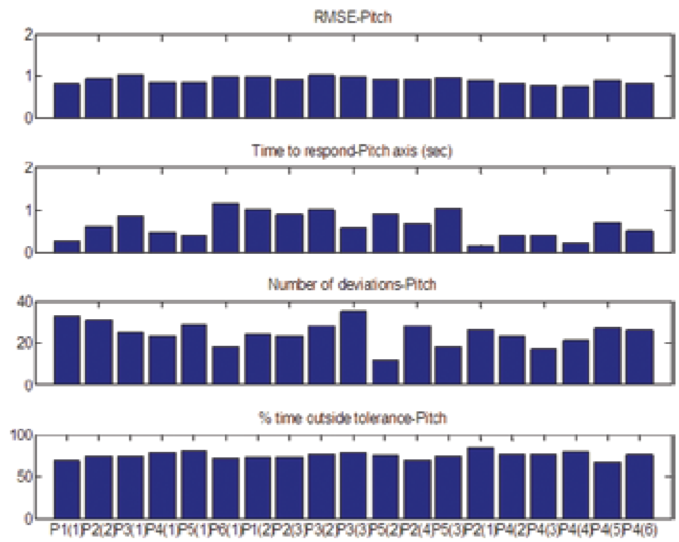
When comparing pitch and roll tracking accuracies, it is seen that pilot responds faster to changes in pitch command. Also on an average, ND and TD across all sorties is more in pitch axis. It means that although the magnitude of errors is less, pilots have attempted for more number of corrections in pitch axis to reduce error.

When analysing PIW metric, it can be seen that all pilots show high levels of aggressiveness with fast control stick movements. When comparing between pitch and roll control column deflections (Fig. 5), it is seen that pilots are slightly more aggressive during pitch than roll tracking. Percentage of time the control stick is used is also slightly higher in pitch.

Similar conclusion can be made by analysing PSD metric. The highest task frequency for both pitch and roll commands is ~0.04Hz. It is seen from Figs. 6(a) and (b) that the amplitude of PSD peaks of both pitch and roll control column deflections are larger near these frequencies. In general, pilots give large amplitude deflections at lower frequencies and at frequencies near HTF. On comparing the control strategy between pitch and roll stick deflections; it is found that amplitude of PSD peaks is more uniform across all frequencies other than HTF in case



(a)



(b)

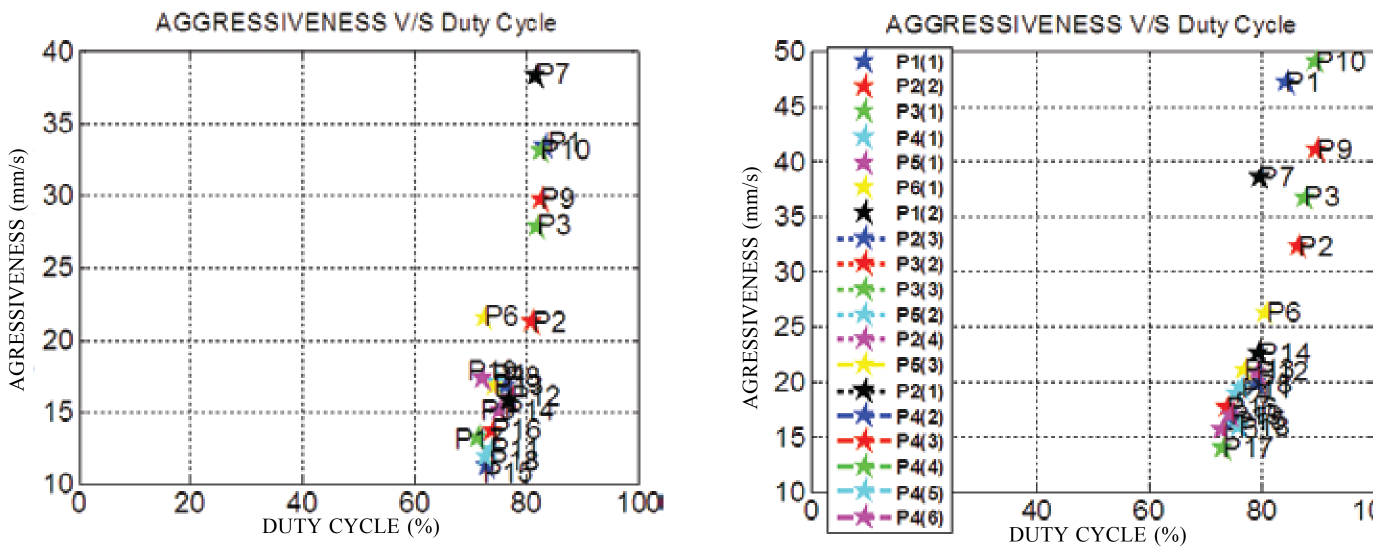
Figure 4. Error in roll and pitch axes: (a) Roll axis and (b) Pitch axis.

of roll axis control. On the contrary, high amplitude PSD peaks are observed at different frequencies in pitch control column (Table 2). This means that pilots have exerted more effort to accomplish pitch axis tracking than in roll axis.

The inference is that higher levels of stress are felt when flying a pitch axis tracking task than the roll axis tracking. In effect, pilot's tracking accuracy and hence his performance is better in case of in pitch axis.

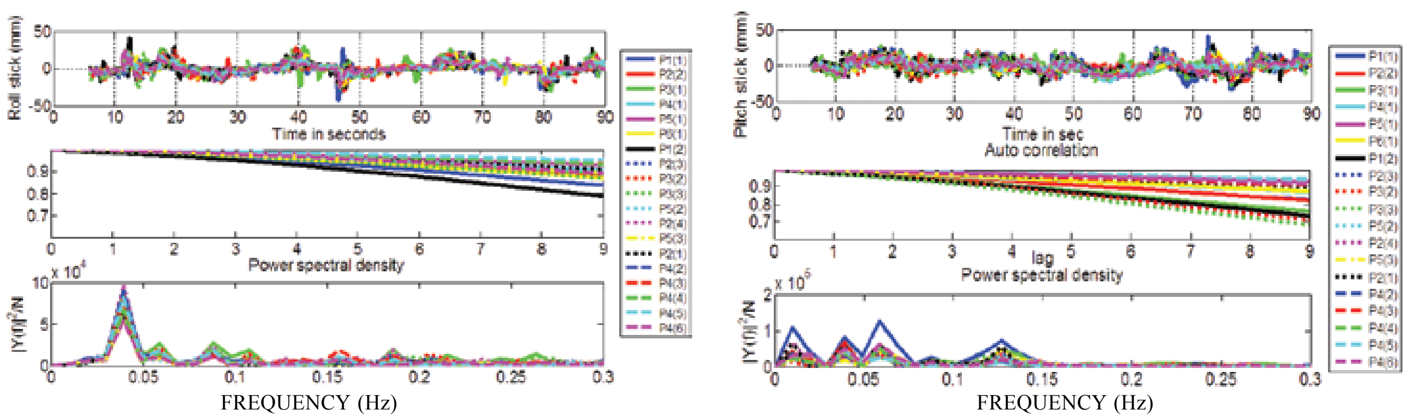
4. COMPARISON OF FLIGHT CONFIGURATIONS

Flight sorties conducted by the same pilot in both configurations described in Table 1 are compared first for tracking errors. It is found that error in pitch axis is dependent on the flight configuration [$F(1,6) = 11.44$, $\rho<0.0148$]. Error is more in config2. When comparing the pilot control strategy, it is found that mean power used for pitch control column is significantly higher at frequencies greater than HTF (0.1-0.25Hz) in config2. It means that pilot tends to give large inputs



(a) (b)

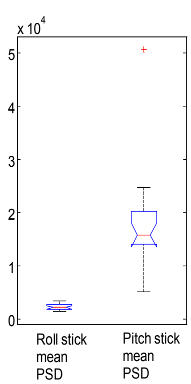
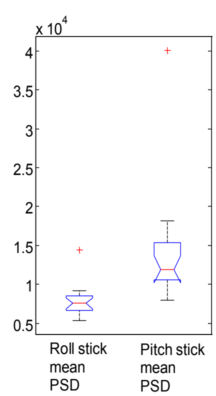
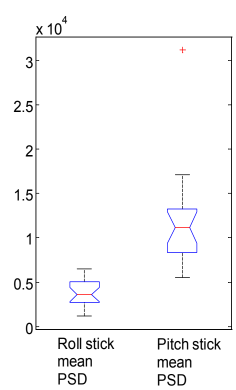
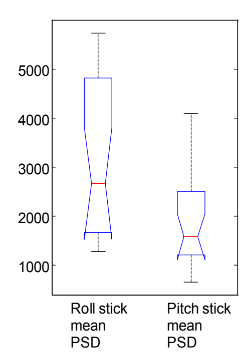
Figure 5. PIW plot of pitch and roll control axes: (a) Roll axis and (b) Pitch axis.

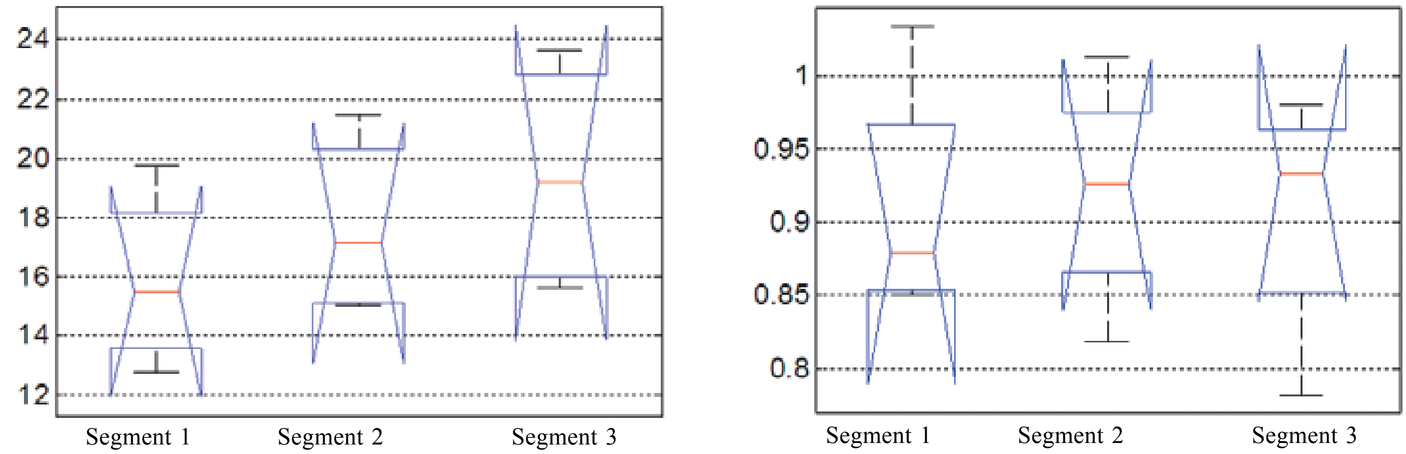


(a) (b)

Figure 6. Frequency domain measure for control strategy: (a) Roll stick and (b) Pitch stick.

Table 2. Comparison of power used for control column deflection at different frequencies between pitch and roll control columns using ANOVA

0-0.01Hz	0.01-0.05Hz	0.05-0.1Hz	0.1-0.15Hz	0.15-0.25Hz
F(1,36)=52.33 ρ=1.62Xe-08	F(1,36)=0 ρ=0.9623	F(1,36)=12.7, ρ=0.001	F(1,36)=35.63 ρ=7.63Xe6-07	F(1,36)=10.36, ρ=0.0027
 Higher powers for pitch stick	No significant difference	 Higher powers for pitch stick	 Higher powers for pitch stick	 Higher PSD for roll stick



(a) (b)

Figure 7. Comparison of pitch and roll errors in repetitive segments: (a) Pitch angle tracking errors and (b) Roll angle tracking errors.

at a faster pace that is, fast and jerky control. In other words, in config2, it is necessary for the pilot to dedicate more effort, give more frequent inputs to the elevator control and make larger inputs, in an attempt to maintain less tracking error. However, no significant difference is found in roll error with respect to flight configuration at ρ =0.05 [F(1,6)=1.23, ρ<0.3099]. Table 3 shows the results of comparison of power used for control stick deflections between the two configurations using ANOVA for all the frequencies of interest.

To understand the reason behind the difference in pilot control behaviour with respect to the flight configurations, the CG locations for both configurations are computed. It is seen that both have aft CG in longitudinal plane, but the CG of configuration 2 is ~0.64% MAC ahead of that of config1. Hence, probably config1 has slightly better manoeuvrability than config2.

5. SIGNIFICANCE OF TASK REPETITIONS ON PILOT'S PERFORMANCE

Within a single flight sortie, tracking tasks of short durations were repeated many number of times. These tracking segments were identified and are compared to analyse whether there is any degradation in performance during repetitive segments. These results show no significant variation in both pitch and roll errors with respect to 3 repetitive flight segments in a single flight sortie at ρ =0.05 [roll angle: F(2,9)=1.07,ρ<0.3834; pitch angle: F(2,9)=0.03, ρ<0.9726]. The control strategy in pitch and roll control columns also does not change with repetitions.

6. COMPARISON OF PILOT CONTROL STRATEGY IN REAL FLIGHT AND IN SIMULATOR

Single axis tracking tasks (pitch and roll separately) conducted using the flight simulator are compared with flight sorties with minimum tracking errors. PSD plots indicate close representation of the simulator HQ with the actual aircraft (Fig. 8(a), (b)).

The effect of workload with respect to single axis and dual axis tracking can be well understood in these test cases. Roll axis tracking accuracy improved significantly (F(1,2)=173.62, ρ <0.0057) when exercised as the primary task. Also lower levels of power is used for roll control (F(1,2)=579.13, ρ <0.0017).

Figure 9 shows that pilot's pitch commands are more correlated in case of simulator data. Hence predictions of subsequent stick movements are more predominant in a simulated environment than in real flight, may be due to higher levels of workload during flight. Such a correlation is not found in roll axis because pitch axis is given more priority in case of the dual axis task.

7. CONCLUSIONS

The quantitative performance metrics for measuring pilot/aircraft's HQ is applied for real flight data as well as for the data from flight simulator. Tracking errors and pilot's control strategy for all the tracking segments are analysed. Different aspects of performance such as significance of task

Table 3. Significance of pilot control strategy with respect to flight configuration using ANOVA

		Roll stick (mm)		
0-0.01Hz	**0.01-0.05Hz**	**0.05-0.1Hz**	**0.1-0.15Hz**	**0.15-0.25Hz**
F(1,6)=3.17, ρ<0.1251	F(1,6)=8.89, ρ<0.0246 (Lower control column power in config2)	F(1,6)=0.39, ρ<0.3928	F(1,6)=1.45Xe-5, ρ<0.9971	F(1,6)=0.17, ρ<0.6935
		Pitch stick (mm)		
0-0.01Hz	**0.01-0.05Hz**	**0.05-0.1Hz**	**0.1-0.15Hz**	**0.15-0.25Hz**
F(1,6)=0.73, ρ<0.427	F(1,6)=3.65, ρ<0.1046	F(1,6)=0.04, ρ<0.8469	F(1,6)=10.24, ρ<0.0186 (Higher control column power in config2)	F(1,6)=7.18, ρ<0.036 (Higher control column power in config2)

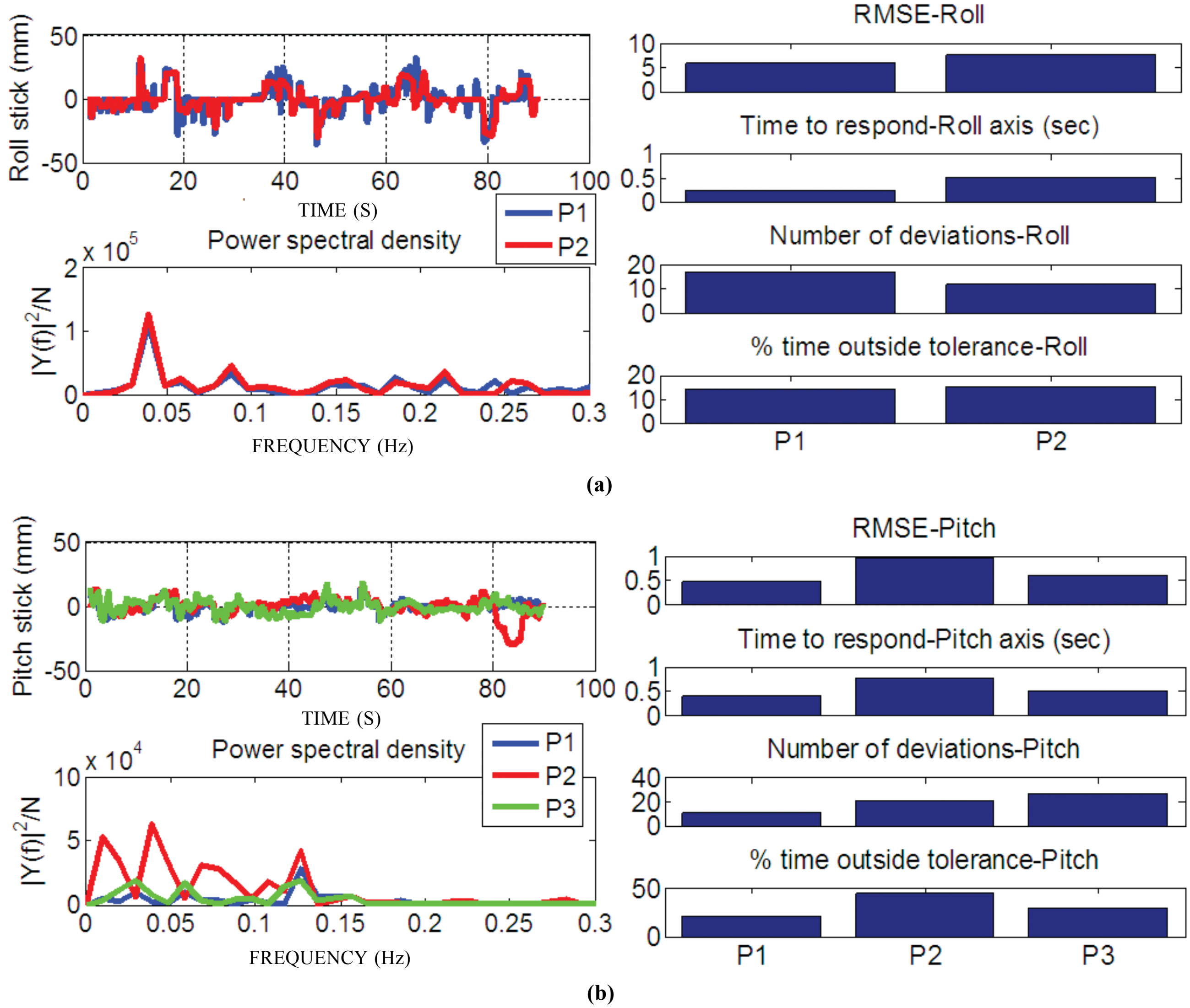


Figure 8. Roll and pitch axis tracking performance in simulator: (a) Roll axis and (b)Pitch axis.

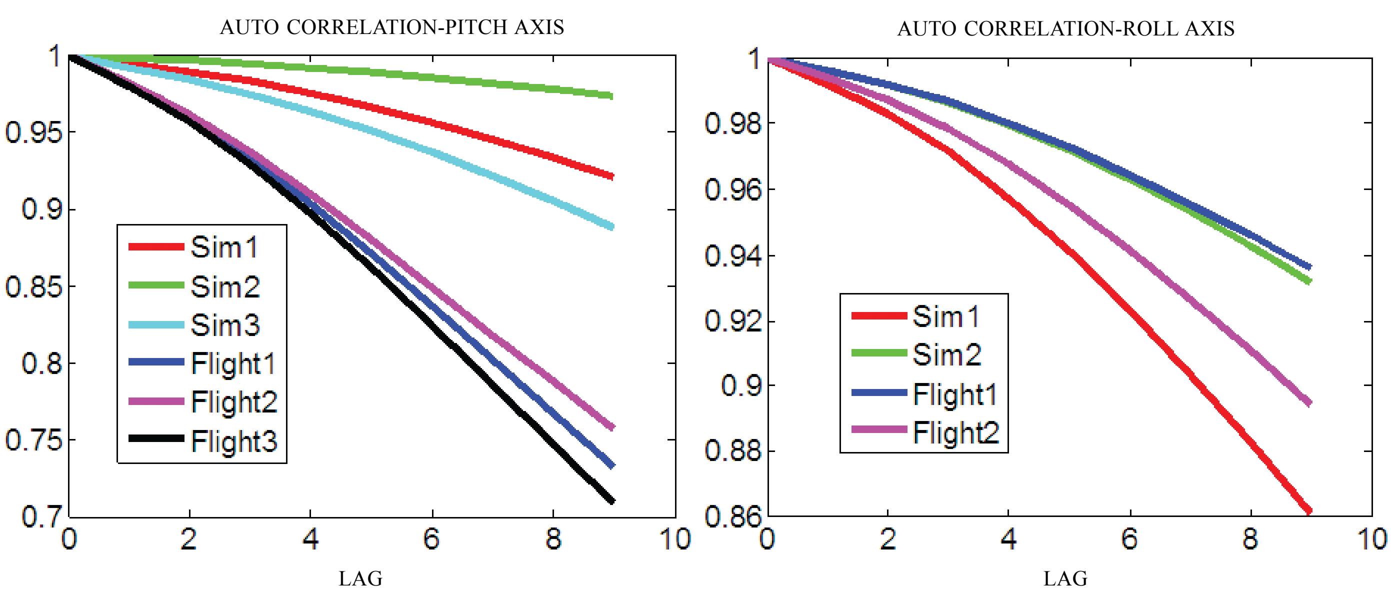


Figure 9. Comparison between results of flight and simulator data.

repetitions on tracking errors/control strategy and difference in pilot's control strategy between pitch/roll control columns are discussed. The change in pilot's performance with change in aircraft's HQ is evident in section 4.

One of the other major observations that are proven quantitatively from this analysis is that pilots are in general more responsive to pitch axis commands than roll axis when performing dual axis tasks. All pilots have used higher powers to accomplish pitch tracking task than roll, even though magnitude of pitch command is lesser than the other. Pilot's performance improved significantly in case of single axis tasks wherein the demands on mental resources are greatly reduced.

Another observation from the analysis is that correlation between successive pitch control commands is better in the simulator experiments than in real flight. It means that anticipation of control is better in the simulator data. However, as the simulations are restricted to single axis alone, it is difficult to arrive at a conclusion. The results shall improve if similar tracking tasks are conducted in both axes simultaneously in future.

REFERENCES

1. Endsley, Mica R. & Jones, Debra G. Designing for situation awareness-An approach for user centered design. Ed 2., CRC Press, 2004, pp. 266-267.
2. Rantanen, E.M.; Johnson, N.R. & Talleur, D.A. The effectiveness of a personal computer aviation training device, a flight training device, and an airplane in conducting instrument proficiency checks: Objective pilot performance measures. AHFD-04-16/FAA-04-6. November 2004, **2**.
3. Takallu M.A.; Wong, D.T. & Uenking, M.D. Synthetic vision systems in GA cockpit-evaluation of basic manoeuvres performed by low time GA pilots during transition from VMC to IMC. *In* International Advanced Aviation Technology Conference. Anchorage, AK, 2002.
4. Field, E.J. & Giese, Sean E.D. Appraisal of several pilot control activity measures. *In* AIAA Atmospheric Flight mechanics and conference and Exhibit, San Francisco, California, AIAA 2005-6032, 2005, pp 1-18.
   doi: 10.2514/6.2005-6032
5. Mitchell, David G.; Kish Brian A. & Seo, John S. A flight investigation of pilot-induced oscillation due to rate limiting. *In* IEEE Aerospace Conference, 1998.
   doi: 10.1109/AERO.1998.685777
6. Shepherd, Michael J.; Adam, MacDonald; William, R. Gray III & Richard, G. Cobb. Limited simulator aircraft handling qualities evaluation of an adaptive controller. *In* IEEE Aerospace Conference Paper 1292, March 7-14 2009.
   doi: 10.1109/AERO.2009.4839613
7. Ebbatson, M.; Huddlestone, J.; Harris, D. & Sears, R. The application of frequency analysis based performance measures as an adjunct to flight path derived measures of pilot performance. *Human Factors Aerospace Safety*, 2007, **6**(4), 383-394.
8. Ostertagov, Eva & Ostertag, Oskar. Methodology and application of one way ANOVA. *J. Mech. Eng.*, 2013, 1.7, 256-261.

ACKNOWLEDGMENTS

The work is carried out under a project sanctioned by Aeronautical Research and Development Board (ARDB). The authors sincerely thank, Dr Vijay V Patel, Scientist G, ADA, Dr Girish Deodhare, TD IFCS, ADA for their valuable comments and encouragement.

CONTRIBUTORS

Mrs P. Archana Hebbar is currently pursuing her Masters (Electronics) from VTU, Belgaum. She is presently working as Scientist in Flight Mechanics and Control Division, NAL. She has experience in design and development of flight simulator facilities and is currently working in the field of Human Factors in aviation.
She has contributed in the formulation of the test cases, application of the methodology and analysing the results.

Dr Abhay A. Pashilkar received his BTech(Hons) in Aerospace Engineering from IIT Kharagpur, ME from the Indian Institute of Science, Bangalore, in 1993 and PhD from the Indian Institute of Science, in 2002. He is a member of the National Control Law Team for LCA. He is Group Head, Flight Simulation in Flight Mechanics and Control Division, NAL. His areas of interest are: Modelling, simulation and control of aerospace vehicles.
His main contribution to this paper was in the discussion of the results, reasoning and conclusions.