DOT/FAA/TC-15/46

Federal Aviation Administration
William J. Hughes Technical Center
Aviation Research Division
Atlantic City International Airport
New Jersey 08405

Accuracy Assessment of Health and Usage Monitoring System Regime Recognition Algorithms

February 2016

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/TC-15/46		

4. Title and Subtitle	5. Report Date
ACCURACY ASSESSMENT OF HEALTH AND USAGE MONITORING SYSTEM REGIME RECOGNITION ALGORITHMS	February 2016
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Chang-Hee Hong	

9. Performing Organization Name and Address		10. Work Unit No. (TRAIS)
Helicopter Association International 1635 Prince Street Alexandria, Virginia 22314-2818	Technical Data Analysis, Inc. 3190 Fairview Park Drive, Suite 650 Falls Church, VA 22191	11. Contract or Grant No. 10-G-020

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
U.S. Department of Transportation Federal Aviation Administration Southwest Regional Office—Aircraft Certification Service, Rotorcraft Directorate 10101 Hillwood Pkwy Fort Worth, TX 76177	Final Report
	14. Sponsoring Agency Code ASW-112

15. Supplementary Notes

The Federal Aviation Administration William J. Hughes Technical Center Aviation Research Division Technical Monitor was Traci Stadtmueller.

16. Abstract

This report details the research conducted under a Federal Aviation Administration (FAA) grant to study the processes by which a rotorcraft operator may implement usage monitoring and advanced usage-based fatigue lifing techniques within the framework of the guidance provided in FAA Advisory Circular 29-2C MG-15. The research focused specifically on the assessment of accuracies of regime recognition algorithms (RRAs) implemented in the Health and Usage Monitoring System (HUMS) and the impact of regime accuracies on the calculated fatigue lives of rotorcraft dynamic components.

The end-to-end paradigm of structural life tracking of rotorcraft using HUMS relies on an accurate characterization of the operational flight regime by its RRA; therefore, the algorithms should be validated through a rigorous verification and validation process. Attaining the high level of accuracy requirement, such as the 97% described in the ADS-79D appendix B, is a challenge because of the limitations of the RRAs currently available. In addition, most algorithms are currently validated manually; the results are therefore prone to being subjective. There are no known validation methodologies that will assess the accuracy of the developed RRA objectively and produce a quantitative measure of accuracies.

This report describes a proposed methodology that assesses the accuracy of RRAs objectively and quantitatively using confusion matrix and correlation factors. An example RRA was validated by assessing its accuracy of more than 95% using extensive flight-load survey data and scripted HUMS test data to demonstrate a methodology for how the accuracy requirements are to be substantiated and validated.

To determine the appropriate level of accuracy requirements as a regulatory guidance, the report also discusses current accuracy requirements available in the rotorcraft community, pilot variability of regime recognition, and the impact of regime accuracy on the reliability of structural fatigue damage assessment.

17. Key Words	18. Distribution Statement
Health and Usage Monitoring System, HUMS, Regime Recognition Algorithms, Accuracy Assessment, Confusion Matrix, Pattern Recognition, Component Retirement Time	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	33	

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AC	Advisory Circular
ARL	Army Research Laboratory
CF	Correlation factor
CWC	Composite Worst Case
FAA	Federal Aviation Administration
GW	Gross weight
HUMS	Health and Usage Monitoring System
ID	Identification
MVLR	Multi-Variable Linear Regression
RRA	Regime recognition algorithm
SPO	Symmetric pull out
TDA	Technical Data Analysis, Inc.
V&V	Verification and validation
VH	Level flight speeds

EXECUTIVE SUMMARY

This report details the research conducted under a Federal Aviation Administration (FAA) grant to study the processes by which a rotorcraft operator may implement usage monitoring and advanced usage-based fatigue lifing techniques within the framework of the guidance provided in FAA Advisory Circular 29-2C MG-15. The research focused specifically on the assessment of accuracies of regime recognition algorithms (RRAs) implemented in the Health and Usage Monitoring System (HUMS) and the impact of accuracies on the calculated fatigue lives of rotorcraft dynamic components.

The end-to-end paradigm of structural life tracking of rotorcraft using HUMS relies on an accurate characterization of the operational flight regime by its RRA; therefore, the algorithms should be validated through a rigorous verification and validation process. Attaining the high level of accuracy requirement, such as the 97% described in the ADS-79D appendix B, is a challenge because of the limitations of the RRAs that are currently available. In addition, most algorithms are currently validated manually; therefore, the results are prone to being subjective. There are no known validation methodologies that will assess the accuracy of developed RRAs objectively and produce a quantitative measure of accuracies.

This report describes a proposed methodology that assesses the accuracy of RRAs objectively and quantitatively using confusion matrix and correlation factors. An example RRA was validated by assessing its accuracy of more than 95% using extensive flight-load survey data and scripted HUMS test data to demonstrate a methodology for how the accuracy requirements are to be substantiated and validated.

The report also discusses current accuracy requirements available in the rotorcraft community, pilot variability of regime recognition, and the impact of regime accuracy on the reliability of structural fatigue damage assessment.

1. INTRODUCTION

Helicopter Association International has contracted Technical Data Analysis, Inc. (TDA) to perform analysis under a Federal Aviation Administration (FAA) grant. The overall program objective is to determine a procedure by which a commercial helicopter operator might use Health and Usage Monitoring System (HUMS) technology to establish maintenance credits or revise maintenance intervals according to individualized usage. General guidance of this application process is provided in FAA Advisory Circular (AC) 29-2C MG 15 [1].

The primary research conducted under this grant was reported in detail by Hong et al. [2], including the review of traditional fatigue lifing methodologies for rotorcraft dynamic components, variations of structural life tracking paradigms using HUMS; onboard and post-flight regime recognition processes; the reliability of the HUMS data process as a whole; and the completeness and accuracy of the guidance presented in AC 29-2C MG-15. The research grant was extended to further study the accuracy assessment of regime recognition algorithms (RRAs), which play a central role in the rotorcraft structural life tracking paradigm using HUMS.

The end-to-end paradigm of structural life tracking of rotorcraft using HUMS relies on an accurate characterization of the operational flight regime by its RRA to establish the basis for retirement time of aircraft components. The algorithms should be validated through a rigorous verification and validation (V&V) process using dedicated flight test data with sufficient flight conditions and time. Attaining the high level of accuracy requirement, such as the 97% described in the ADS-79D appendix B, is a challenge because of the limitations of the RRAs currently available. [3]. Currently, most algorithms are validated manually; therefore, the results are prone to being subjective. There are no known validation methodologies that will assess the accuracy of developed RRAs objectively and produce a quantitative measure of accuracies.

This report describes a proposed methodology that assesses the accuracy of RRAs objectively and quantitatively using confusion matrix and correlation factors (CFs). An example RRA was validated by assessing its accuracy of more than 95% using extensive flight load survey data and scripted HUMS test data to demonstrate a methodology for how the accuracy requirements are to be substantiated and validated.

The report also discusses current accuracy requirements available in the rotorcraft community, needs of standardization of regime definitions for commercial helicopters, pilot dependency of regime recognition, and impact of regime accuracy on the reliability of structural fatigue damage assessment.

2. APPROACH FOR ACCURACY ASSESSMENT OF REGIME RECOGNITION

HUMS was introduced in the late 1980s in an effort to improve the reliability of structural life assessments for rotorcraft by accurately monitoring the operational condition of an aircraft or fleet of aircraft [4]. Various RRAs have been developed to identify flight regime profiles as a part of the HUMS usage monitoring process, and regime recognition is commonly carried out by a hierarchical process to identify flight profiles based on the recorded multiple flight parameters [5–7]. To establish RRA as the basis for the retirement time of aircraft components, algorithms should be validated through a rigorous V&V process using dedicated flight test data with

sufficient flight conditions and time. This study is focused on the methodology development to quantitatively assess the accuracy of existing RRAs in an effort to verify and validate the regime recognition codes. A new approach of regime accuracy assessment using the confusion matrix with CFs based on the regime load pattern matching scheme is described in this report.

2.1 DIFFICULTIES IN REGIME ACCURACY ASSESSMENT

Currently, there are few references that present formal accuracy requirements of HUMS RRAs. For military aircraft, ADS-79D appendix B provides detailed guidance for validation of RRAs, including their accuracy requirements: "… such as 97% or greater recognition of the actual flight regimes …" and "… any unrecognized regimes would introduce less than 0.5% under-prediction of fatigue damage fraction based on the design usage spectrum" [3]. Attaining this 97% accuracy requirement by an objective V&V process is challenging because of the limitations with the current RRAs. One of the major issues encountered is that multiple regimes are identified by the code for a supposedly single regime time. Toggling of similar regimes back and forth is also one of the commonly encountered problems [7]. It appears intuitive that more than one single regime is recognized during the changing flight conditions. This is particularly true for transient maneuvers, in which a time-sliced regime recognition logic is implemented even with the help of more advanced logic, such as a neural network and a hidden Markov model.

The regime classification logics are generally based on certain criteria focused on the peak points of maneuvers; as a consequence, the entry and recovery phase regimes, which are highly transient maneuvers, are identified as similar regimes but slightly different from the target regimes. This does not mean that the regimes recognized during the entry and recovery phases are incorrect. Table 1 shows the typical multi-short regimes recognized from one symmetric pull up maneuver. Note that there are jumps in time increments in the table to show more regimes.

Table 1. Example of maneuver regime comparison

time (sec)	Regime_Name_Piloted	Regime_Name_Recognized
4153	SymmPullUp-LSDive	Dive
4154	SymmPullUp-LSDive	Dive
4155	SymmPullUp-LSDive	SymmPullUp-LSDive
4155	SymmPullUp-LSDive	SymmPullUp-LSDive
4156	SymmPullUp-LSDive	SymmPullUp-LSDive
4157	SymmPullUp-LSDive	Descent
4158	SymmPullUp-LSDive	LevelFlight128
4167	SymmPullUp-LSDive	LevelFlight112
4169	SymmPullUp-LSDive	MaxContPwrClimb
4174	SymmPullUp-LSDive	MaxContPwrClimb
4175	SymmPullUp-LSDive	SymmPushOver
4176	SymmPullUp-LSDive	SymmPushOver
4177	SymmPullUp-LSDive	SymmPushOver
4178	SymmPullUp-LSDive	SymmPushOver
4179	SymmPullUp-LSDive	Descent
4180	SymmPullUp-LSDive	Descent

As can be seen in this example case, if the recognized regimes from the output of the regime recognition codes are compared against the scripted regimes from the pilot cards strictly based on their regime names, the end result of accuracy will be unacceptably poor. It requires detailed physical explanations as to why the identified short-duration regimes are part of the target maneuver such that the regime recognition results are acceptable. This manual process makes the V&V of RRAs subjective and time-consuming. As a method to alleviate the stringent assessment with “black and white” criteria, the confusion matrix is studied with CFs using regime load pattern matching to establish an acceptable method that results in a more accurate and quantitative assessment of RRA.

2.2 CONFUSION MATRIX FOR REGIME VALIDATION

A confusion matrix, also known as an error matrix, is a specific table layout that allows for visualization of the performance of an algorithm. It is commonly used for the validation of algorithms in the artificial intelligence field. As an example, Goodrich engineers used the confusion matrix for the validation of their neural network-based RRA [8]. Table 2 shows a simple confusion matrix, in which each column represents the recognized regimes as instances in a predicted class and each row represents the scripted (piloted) regimes as instances in an actual class. As shown in table 2, with a few selected regimes, if strict accuracy criteria are applied, the portion of matching cases between the true values and predicted values in the diagonal (shaded) cells will get the validation credit, resulting in poor accuracy. See figure 1 for a 3-D plot of the example confusion matrix.

Table 2. Example confusion matrix of regime times

Regime Name	Row Labels	4	5	9	10	11	44	56	80	94	123	Grand Total	Accuracy (%)
Descent	4	371	170	75	65	13	35	49	0	22	6	1528	24.3
Dive	5	12	32	0	0	0	0	0	0	0	5	132	24.0
HvrIGE	9	0	0	67	0	0	0	0	0	0	0	89	75.2
HvrOGE	10	0	0	7	53	0	9	0	0	0	0	95	56.6
IntPwrClimb	11	108	154	17	148	148	38	7	0	104	1	3534	4.2
LTurn	44	16	0	0	5	0	109	8	0	0	0	302	36.1
LTurnHvr	56	0	0	0	98	0	6	84	0	0	0	204	41.0
RRollingPullup-Dive	80	53	7	0	0	0	62	0	11	109	12	1168	0.9
RTurn	94	28	1	0	0	7	582	0	0	733	0	2546	28.8
SymmPullUp	123	97	2	0	4	0	35	0	0	37	6	925	0.7
	Grand Total	1310	775	1321	2021	271	1444	713	13	1413	78	32675	13.4

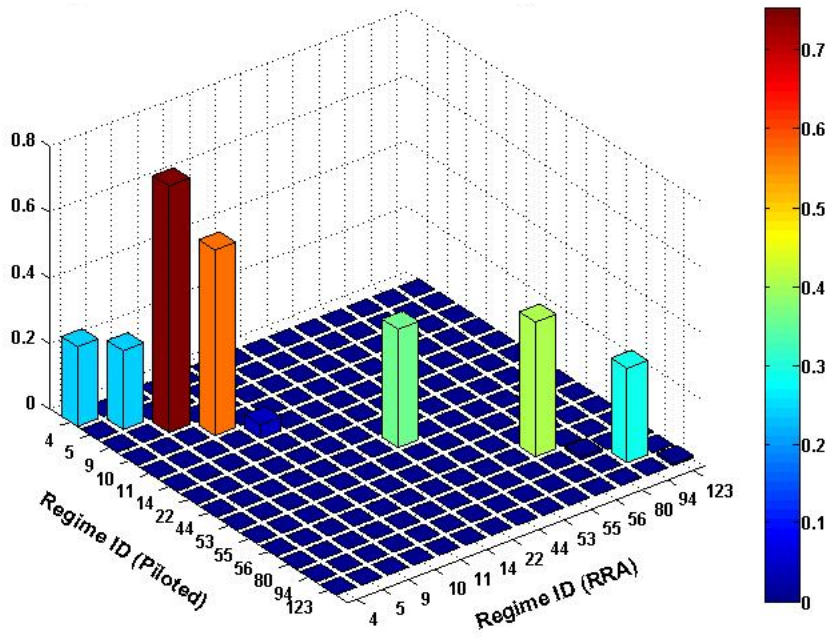

Figure 1. Example 3-D plot of a confusion matrix for RRA accuracy

2.3 REGIME LOAD PATTERNS AND CFS

The off-diagonal terms generally represent short duration regimes identified during the entry and recovery phases of a regime. These off-diagonal terms do not signify incorrectly identified flight conditions. In most instances, they are correctly identified maneuver conditions. The difficulty is determining how to define their correctness based on a particular credit validation method.

To resolve this issue, a CF is conceived to define the validation credit for the off-diagonal terms based on the vibratory load pattern of a regime. Figure 2 shows typical vibratory load patterns on the helicopter pitch link for the symmetric pull out (SPO) maneuvers grouped into two level flight speeds (VH): 0.5 VH and 1.0 VH. In this plot, the vibratory loads are normalized by the maximum vibratory load observed from all the SPO conditions. Note that the three thick blue curves leveled as the same "0.5VH/100% SYM PO 2.0G" are from the same VH and vertical acceleration but different gross weight (GW) and altitude conditions, so their load pattern curves are similar.

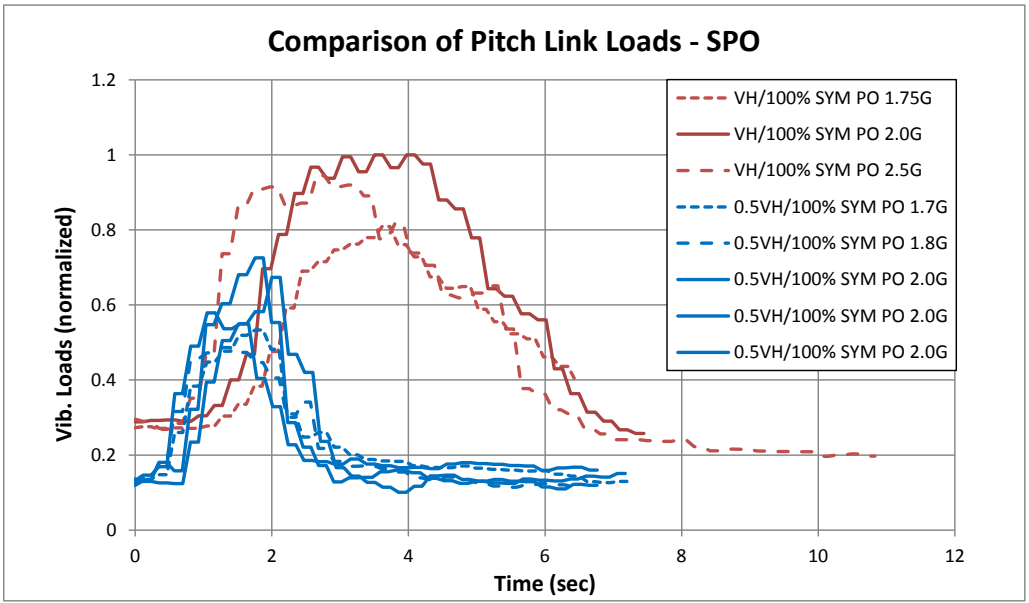

Figure 2. Example of vibratory pitch link load patterns

Multiple regimes will be recognized by an RRA for a given maneuver duration and are known as sub-regimes. Each sub-regime will have its own maximum vibratory load level. It is intended to examine how well maneuver load patterns are matched to the presumed load patterns identified from the flight load survey. The similarity or goodness of load patterns between the flight load survey and the RRA would provide the CF of each recognized sub-regime. Figure 3 shows an example of the regime load patterns normalized by peak values of their SPO groups. The thick curves in the plot represent possible criteria for the pattern-matching boundaries and, in this case, +/-20% of the average values at each time slice are plotted.

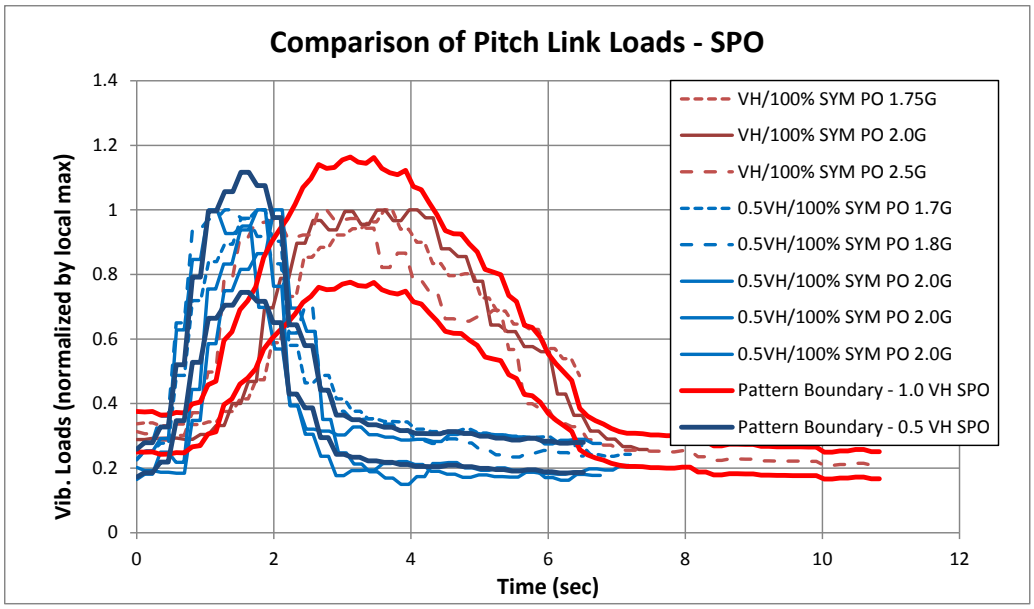

Figure 3. Example of normalized load patterns of SPO with +/-20% pattern boundaries

2.4 MULTIVARIABLE LINEAR REGRESSION ANALYSIS FOR LOAD PREDICTION

The pattern matching to calculate the required CF requires obtaining the maneuver load pattern for the given regime by pilot cards. For this effort, the following five typical HUMS flight parameters are selected as input values for the Multivariable Linear Regression (MVLR) analysis: pitch and roll attitudes, rate of climb, Vertical Acceleration, and Never Exceeding Speed, together with the measured pitch link vibratory loads. The regime load vector is defined by the following equation:

$$\mathrm{Out}(x) = w^T x = w_1 x[1] + w_2 x[2] + \ldots w_m x[D] \tag{1}$$

where $x[i]$ are the regression matrices, and the coefficient vector w at the maximum likelihood of the prediction is defined as [9]:

$$w = (X^T X)^{-1}(X^T Y) \tag{2}$$

Figures 4 and 5 show the comparison of measured and predicted vibratory load patterns of highly transient maneuver regimes using the 5-parameter regression analysis.

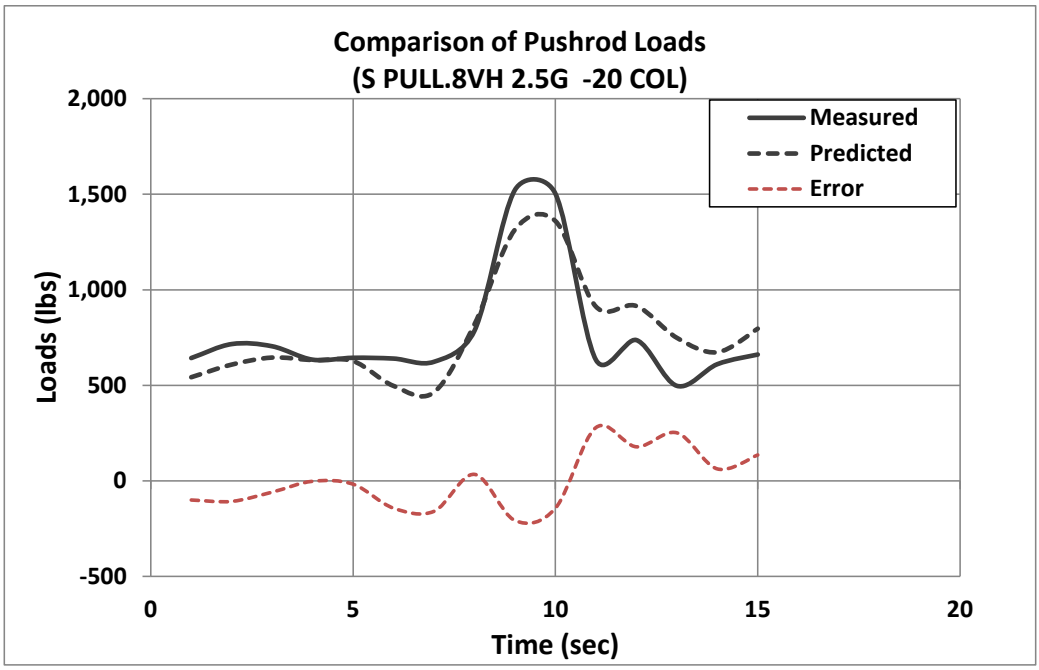

Figure 4. Example comparison of measured and predicted pitch link vibratory loads – SPO

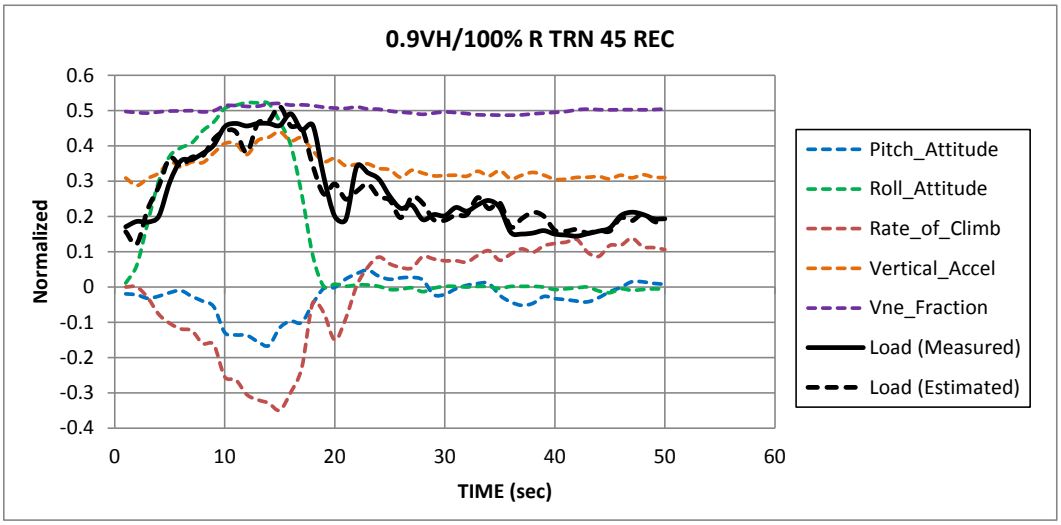

Figure 5. Example comparison of measured and predicted pitch link vibratory loads – R turn

As shown in figures 4 and 5, the vibratory load patterns are predicted with fair accuracy by the MVLR analysis from typical flight parameters. It is also possible to include more parameters in the regression analysis, such as engine torques, rates of pitch, and yaw, if they are needed for particular flight conditions. The prediction accuracy of the analysis depends on how the input data is grouped for similar flight conditions; in this study, the survey data was grouped based on their GW; density altitude; vertical acceleration or load factor; and speed. With this grouping of survey data, the load pattern prediction is good enough to be used for pattern matching in the RRA accuracy assessment.

2.5 REGIME LOAD PATTERN PREDICTION FROM HUMS DATA

Once the required regression coefficient set for the five parameters of each regime is obtained from the flight load survey data, it can be used to predict load patterns from the same five parameters of each HUMS regime provided with the scripted HUMS data. Figure 6 shows the predicted vibratory load pattern plotted with the five flight parameters from the HUMS in different colored and dashed curves.

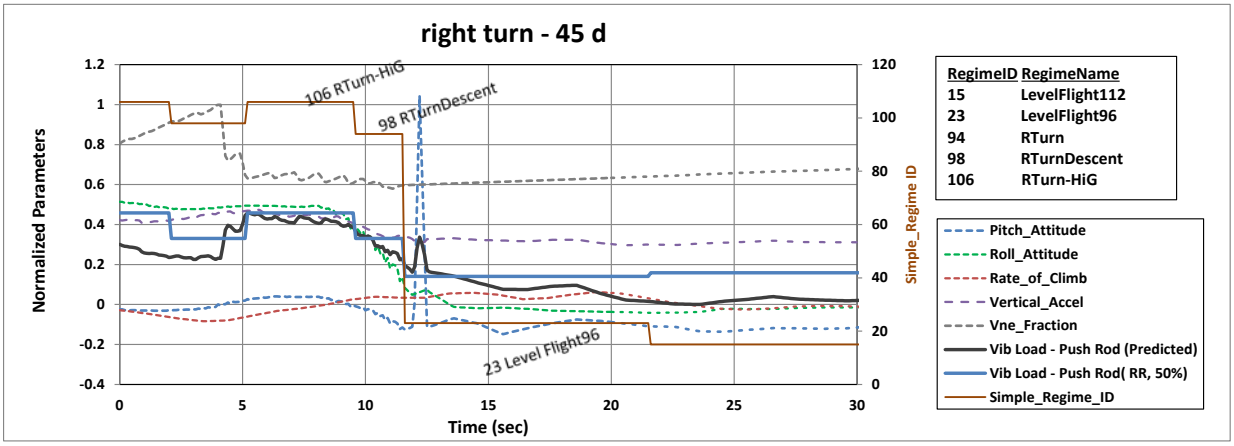

Figure 6. Example of predicted pitch link vibratory loads for a 45-degree R turn

In the figure, the secondary vertical axis on the right side represents the regime identification (ID) numbers, and the short regimes are plotted in thin brown straight lines as a reference. It also shows that the regression coefficient set responds to the signal spike at approximately 12 seconds elapsed time. Note that, in figure 6, the load levels of the sub-regimes identified by the RRA are also plotted in thick blue lines, which are explained in section 2.6.

2.6 REGIME LOAD PATTERNS FOR RRAS

Once multiple regimes are recognized by RRAs, vibratory load levels can be obtained from supplied fatigue damage tables (not presented in this report). To be more accurate, a 50% damage level from the Weibull distribution curve may be used to represent the load level of each regime in place of the load level in the damage table, which is typically selected at the 95% level for conservatism [10].

Figure 7 shows a comparison of the two load patterns: the predicted loads by the regression coefficients in a black curve, and the assigned loads for each sub-regime from the RRA in blue lines. Note that each load pattern is normalized by its peak load to compare patterns. The CF is obtained at each time point by examining the two load patterns and is used as a measure of credit validation for the RRA.

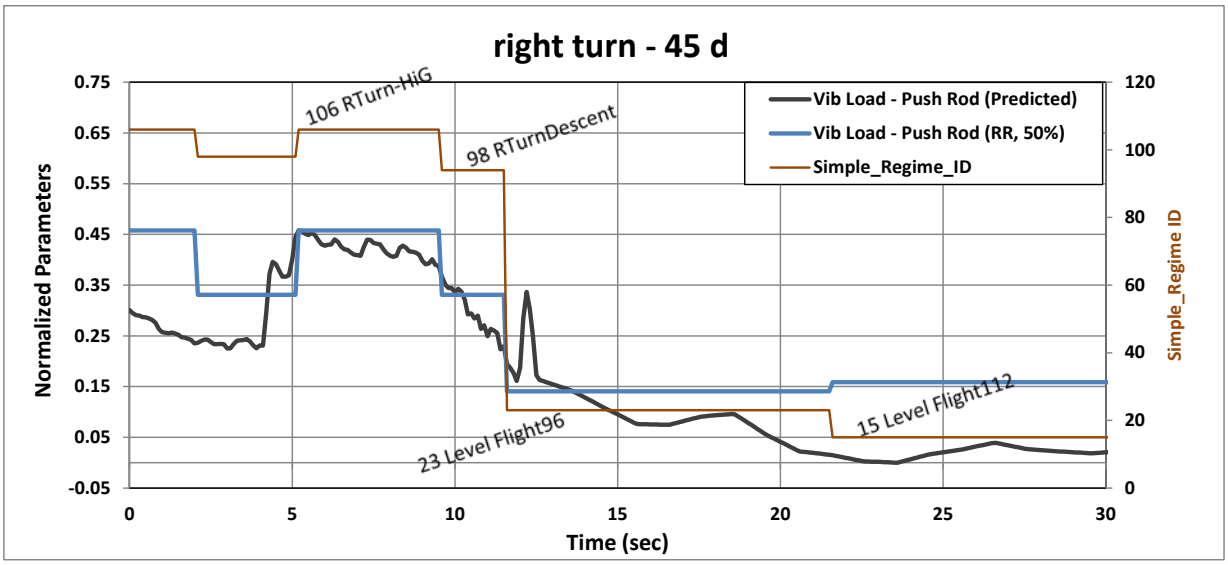

Figure 7. Example of predicted vibratory loads with regime loads from RRA for a 45-degree R turn

2.7 CALCULATION OF CFs

Using comparisons of the two normalized load patterns, the CF at each time point can be calculated from the distance between the two patterns at each time point. Because the sub-regime loads for the RRA are constant during their durations and assigned by the maximum vibratory load levels measured in the flight load survey, it is appropriate to use this local maximum for the predicted load patterns of the same regime durations.

Figure 8 shows the two load patterns with their constant load levels at each sub-regime. Note the upper and lower criteria boundaries for the pattern matching shown in red dashed lines.

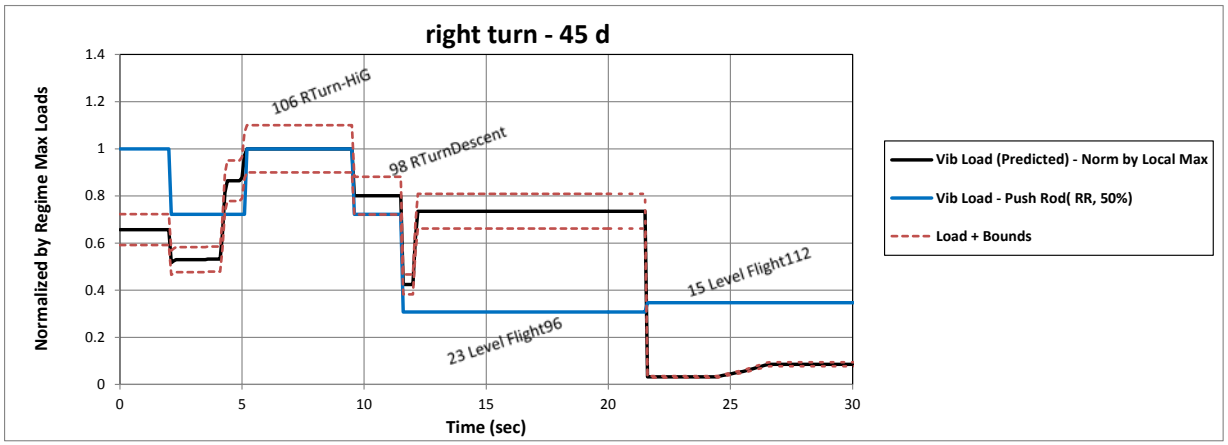

Figure 8. Example of predicted local maximum vibratory loads and regime loads from RRA

The predicted local maximum load levels are unexpectedly high between 12 and 22 seconds for the Level Flight 96 with regime ID 23. This is due to the pitch attitude signal spike, as shown in figures 6 and 7. These types of erroneous load levels will result in an incorrect accuracy assessment of the RRA.

The black line in figure 9 shows the differences in vibratory loads between predicted and the RRA. This allows for the calculation of the CFs at each time point by examining the magnitude of the load differences against the boundaries.

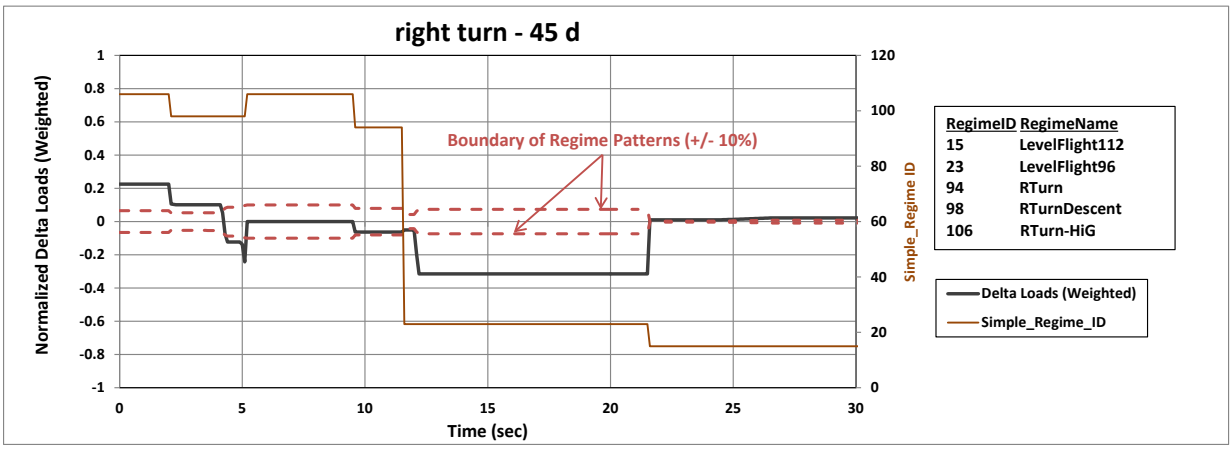

Figure 9. Example of load differences between predicted loads and regime loads from RRA

2.8 LOGIC FOR CFS

Pattern matching is one of the methodologies for pattern recognition. In this study, the following simple logic is used to determine the required CF using the difference of load (ΔLoad, ΔL) between the predicted and RRA load patterns:

$$\text{If } \Delta\text{Load is within the boundaries, CF} = 1.0, \tag{2}$$

$$\text{If } \Delta\text{Load is outside the boundaries, CF} = (1 - \Delta L)/(1 - \Delta L_{bnd}) \tag{3}$$

where ΔL_{bnd} is the load level of pattern boundaries.

3. PRELIMINARY RESULT OF ACCURACY ASSESSMENT

A computer program was developed to implement the accuracy assessment methodology of RRA explained in section 2. Figure 10 shows a block diagram for this accuracy assessment process. By following this process, TDA's regime recognition codes were evaluated to demonstrate the methodology. A substantial amount of flight load survey data and scripted HUMS data were used during the development of the assessment process and for the final evaluation of RRA accuracy.

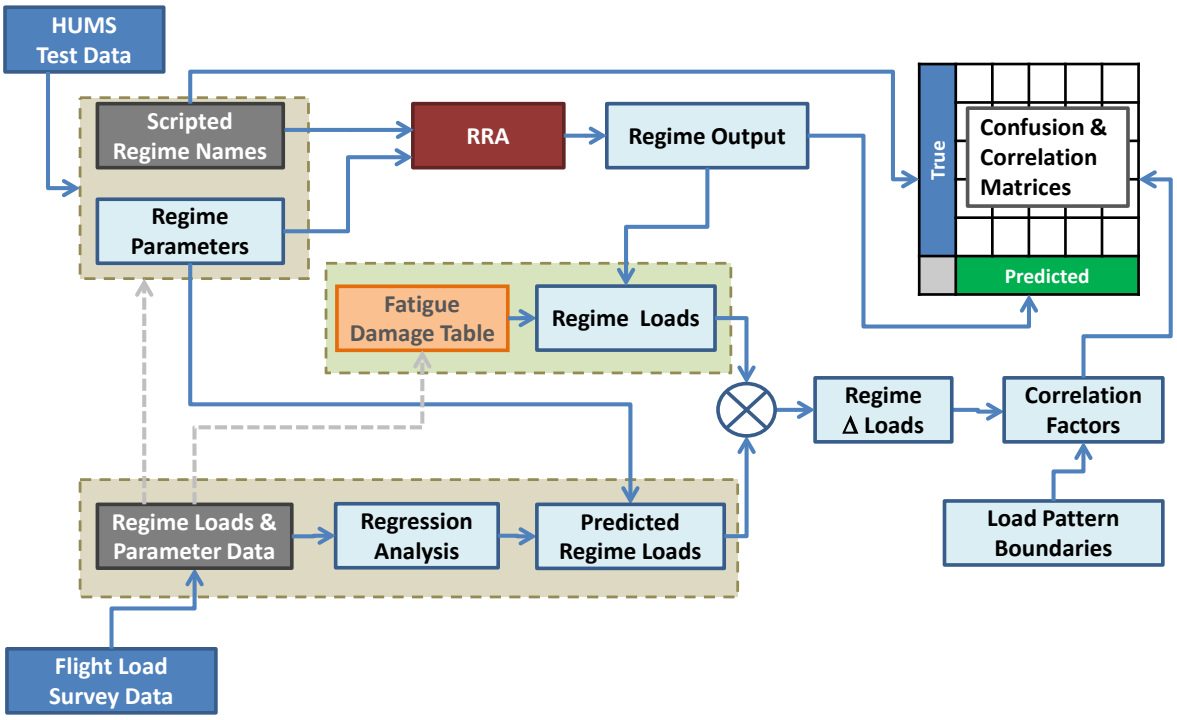

Figure 10. Block diagram for the accuracy assessment of RRAs

In this methodology demonstration study, a +/-10% was discretionally selected as the upper and lower pattern boundaries based on the observation of load pattern scatter for highly transient maneuvers, as shown in figures 2 and 3. Table 3 shows the confusion matrix for a few selected regimes with their adjusted regime recognition accuracies. Figure 11 shows the confusion matrix of regime accuracies in which a uniform distribution of accuracies across the entire matrix area can be observed. Table 3 shows that there are exact matches of regimes on the diagonal (shaded) cells and the off-diagonal cells that have high accuracy validation credits by the assessment methodology described in this report.

Table 3. Example confusion matrix of regime times

Regime Name	Row Labels	4	5	10	11	26	44	94	123	128	Accuracy (%)
Descent	4	371	167	62	9	1	31	23	6	1	84.8
Dive	5	8	32						5	1	82.7
HvrOGE	10			53			7				92.9
IntPwrClimb	11	94	134	141	148	1	35	99	1	4	56.5
LevelFlight	14	23			49	1	28	13		0	94.9
LRollingPullUp	26	45	18			13	122		7	9	69.9
LTurn	44	29		4	7	10	579			0	91.3
RTurn	94	14	0	0	0			1161		2	95.8
SymmPullUp	123	72	2	2		12	32	35	6	5	63.1
SymmPushOver	128	110	33		8		54	9	15	31	85.2

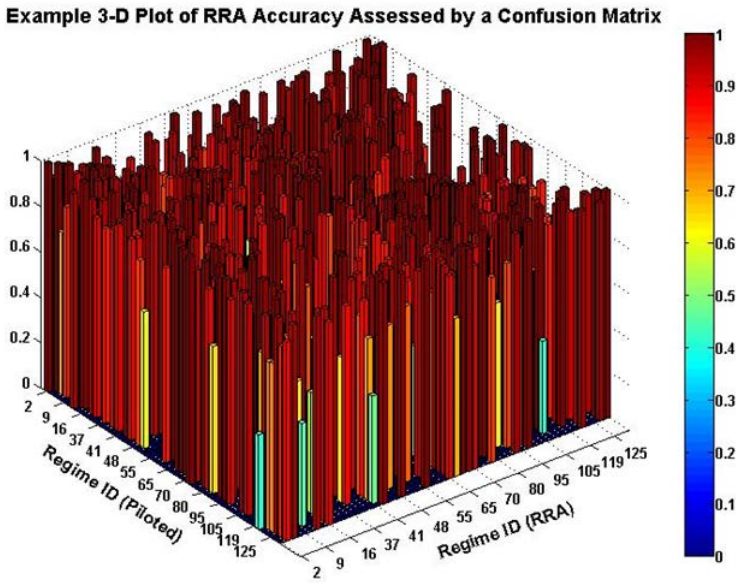

Figure 11. Example 3-D plot of confusion matrix for RRA accuracy

More than 90% accuracy was assessed with a +/-10% credit boundary for the entire scripted regime database of more than 1000 maneuver regimes with 30,000 seconds of accumulated regime times.

Table 4 shows the summary of accuracies for different maneuver groups and credit boundary sizes.

Table 4. Comparison of assessed accuracies

	Accuracy (%)		
	Credit Boundary (+/-)		
Maneuver Group	10%	15%	20%
Steady	91.7	93.3	94.5
Transient	89.9	91.5	93.0
Ground/TO/Lndg	87.3	88.3	90.0
Overall	90.1	91.7	93.0

It was found that the HUMS data and the available flight survey data include various types of errors, such as signal noise, missing parameter data, and incorrect regime names. In addition, many of the regimes have long durations of more than 50 seconds even for transient maneuvers, and sometimes with repeated maneuvers. Regime data with a long duration and a mixture of steady and transient maneuver conditions reduces the accuracy of regression analysis. Eliminating some or all of these errors will improve the overall accuracy. It is also conceivable that a tailored credit boundary scheme for different maneuver groups can be applied for more accurate assessment.

4. IMPROVEMENT OF REGIME ACCURACY ASSESSMENT

As discussed in section 3, it was observed that there are a few areas that require additional measures to obtain improved accuracies. Three measures, among others, are incorporated in the process and discussed in sections 4.1–4.3.

4.1 ERRONEOUS DATA SPIKES

The current HUMS data available for the study included numerous flight records with erroneous data spikes, as shown in figure 7. It was observed that the vertical acceleration is prone to this type of anomaly, which results in the erroneous accuracy assessment due to the inaccurate reference peak load level for the sub-regime time, as shown in figure 8. This type of data spike was removed by manually smoothing out the data. Figure 12 shows the corrected reference load level between 12 and 22 seconds highlighted in the dashed green circle after the data spike was removed.

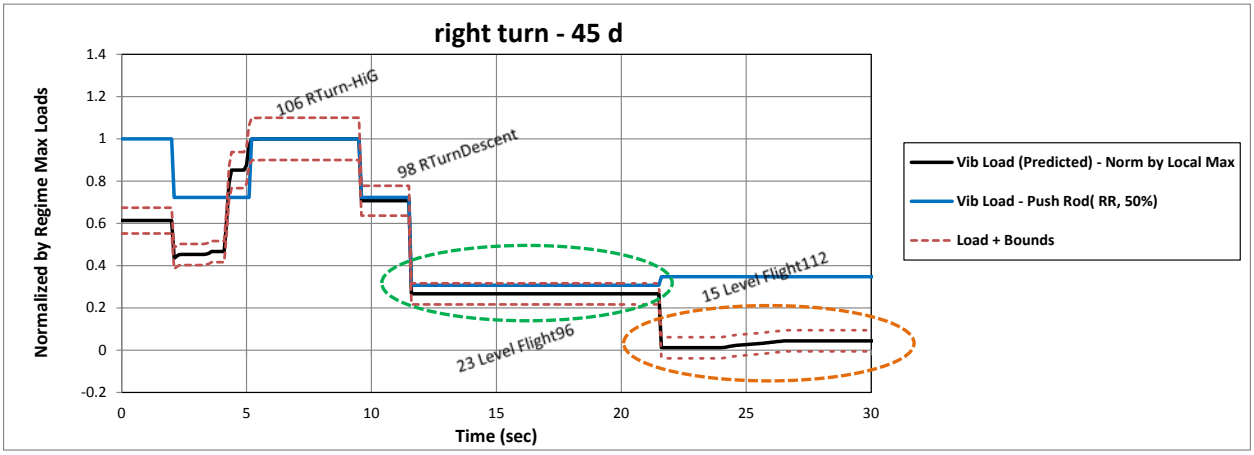

Figure 12. Improved local maximum vibratory loads and accuracy criteria boundaries

4.2 MINIMUM PATTERN CRITERIA BOUNDARIES

During the preliminary accuracy assessment, it was observed that the load pattern criteria boundaries, which are proportional to the local maximum load levels, could become small when the actual load levels become small. This could create a situation of unreasonably tight accuracy criteria. Therefore, +/-5% of maximum predicted regime load was assigned as the minimum level of accuracy criteria boundary during the regime time. Figure 12 shows an example of improved upper- and lower-criteria boundaries highlighted in the dashed orange circle in the 22–30 seconds time interval, and figure 8 shows the tight boundaries in the same elapsed time period before the change was made.

4.3 NEW LOAD PATTERN ALIGNMENT OPTION

To compare the two load patterns (one from the predicted loads using MVLR analysis as the reference and the other from the RRA output), they must be aligned (i.e., normalized by the maximum regime load, as described in section 2.7). It was observed that for the flight regimes of

RRAs in which the regime loads do not change significantly during their duration, there could still be a sudden increase of load levels in short durations due to the recognized regimes that have high load levels assigned. As shown in figure 13, aligning by the maximum load creates a situation in which the loads of the two sources are not matching well in the majority of the regime durations while they are close to each other. To avoid this type of unwanted situation, a second option of load pattern alignment is implemented in which the average regime load, in place of the maximum load, is used to normalize the regime loads. Figure 14 shows that the two regime load patterns match closely when they are normalized by the average regime load.

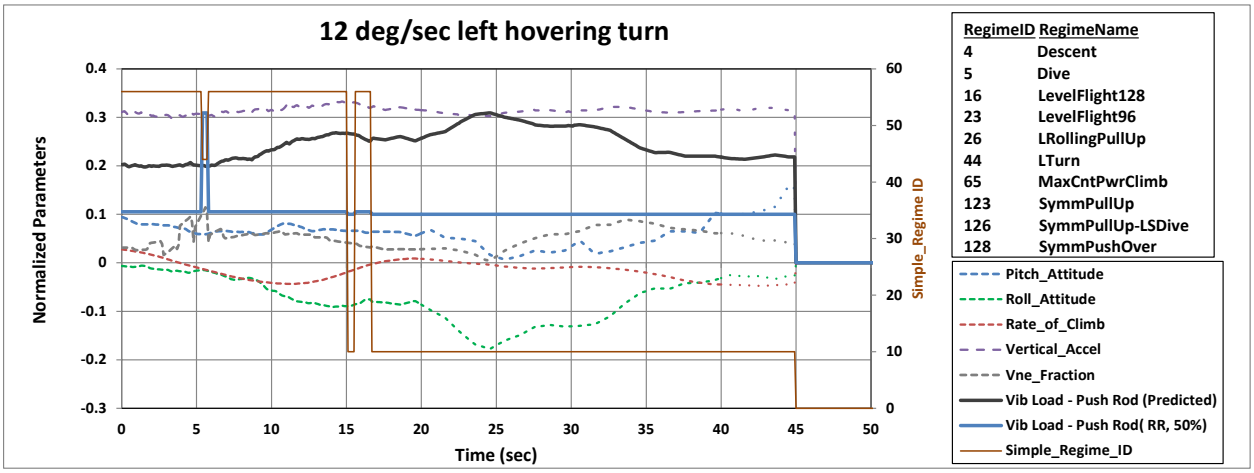

Figure 13. Comparison of two load patterns normalized by the maximum regime load

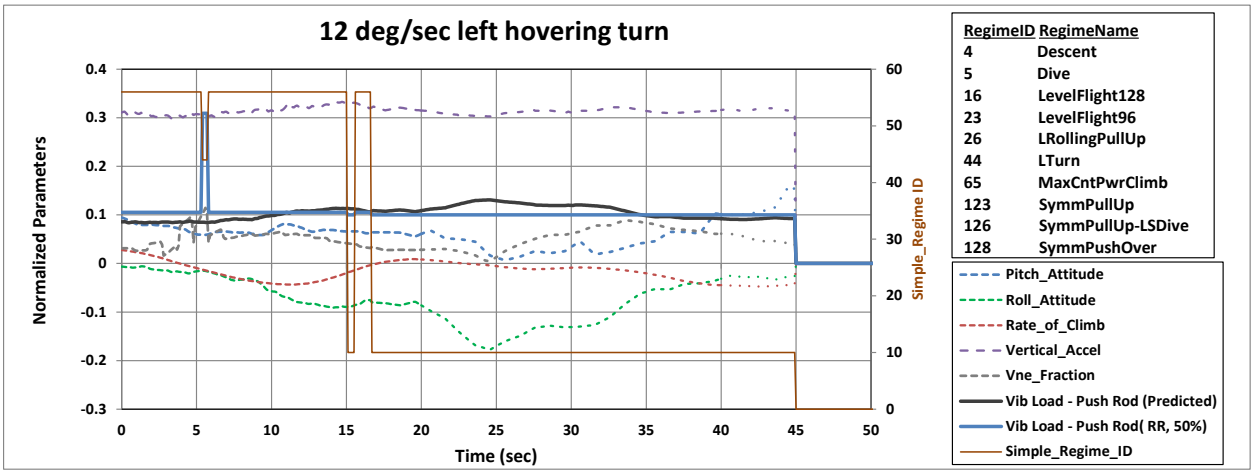

Figure 14. Comparison of two load patterns normalized by the average regime load

4.4 SUMMARY OF REVISED ACCURACY ASSESSMENT RESULT

The three improvement measures described in sections 4.1–4.3 were incorporated in the accuracy assessment process of RRA, and the HUMS output was reevaluated to assess its accuracy. As shown in table 5, the reassessed overall accuracies increased approximately 3% in four different load pattern criteria, or credit boundaries. More detailed accuracies for selected regimes are

shown in table 6 and figure 15, in which more than 95% accuracies were assessed for all regimes except the last two regimes of “Power Dive” and “Steady Auto.”

Table 5. Comparison of reassessed overall accuracies for various credit boundaries

Maneuver Group	Accuracy (%)			
	Credit Boundary (+/-)			
	5%	10%	15%	20%
Steady	92.4	93.8	95.3	96.6
Transient	92.3	93.6	95.0	96.2
Ground/Takeoff/Landing	95.0	95.9	96.9	97.7
Grand Total	92.5	93.8	95.2	96.4

Table 6. Comparison of reassessed regime accuracies for various credit boundaries

Maneuver Group	Accuracy (%)			
	Credit Boundary (+/-)			
	5%	10%	15%	20%
Hover	95.9	97.3	98.4	98.8
Level Flight 90% V MAX	97.4	97.8	98.3	98.6
Max. Cont. Power Climb	92.7	94.0	95.3	96.5
30° AOB Left Turn	94.2	95.1	96.2	97.2
30° AOB Right Turn	94.9	95.6	96.4	97.1
45° AOB Left Turn	94.4	96.2	97.5	97.9
45° AOB Right Turn	94.7	97.2	98.9	99.5
2.0g Pullout	91.9	93.1	94.6	96.0
1.5g Pullout	92.5	93.7	95.3	96.7
Power Dive	82.0	84.1	86.6	89.0
Steady Auto	82.9	85.4	87.8	89.5
Grand Total	92.5	93.8	95.2	96.4

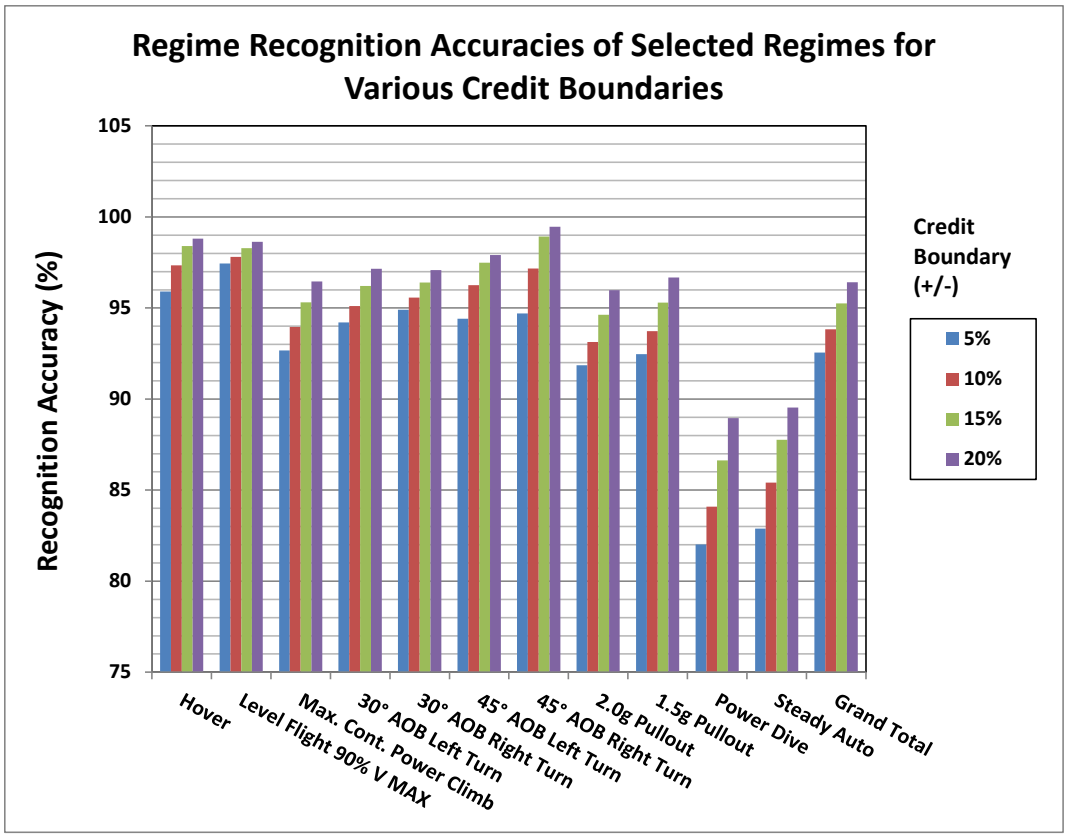

Figure 15. Comparison of reassessed regime accuracies for various credit boundaries

Based on these newly assessed accuracies, it is concluded that a +/-15% credit boundary is a reasonable choice as the load pattern criteria boundary. It is expected that this level of credit boundary will produce a minimum of 95% accuracy from the RRA output.

5. PILOT VARIABILITY OF REGIME RECONGNITION

Pilot variability is an area of concern with respect to using HUMS regime recognitions—specifically, the concern about whether the pilot's individual techniques affect the damage data accumulated. The U.S. Army Research Laboratory (ARL) studied the pilot variability using the Analysis of Variance to analyze the variance that occurs when there are various pilots following the same flight regimes and found that the pilot variability has no impact on the effectiveness of the RRA at the 95% confidence level [11]. A limited amount of the same flight test data used in the ARL study was made available to TDA to investigate the pilot variability. The results are summarized in section 5.1.

Considering that the previous ARL study already concluded that there is no visible variability between pilots, the current study was focused on the qualitative comparison of two pilots: the pilot (pilot #1) and the copilot (pilot #2). The study was also limited to the left turn maneuver with its meaningful number of data points available from the given data files.

5.1 FLIGHT DURATIONS OF LEFT TURN MANEUVERS

The flight test data shows that there are some differences in regime flight times between the two pilots: the average regime duration to complete the left turn maneuver is 18.6 seconds for pilot #1 and 22.4 seconds for pilot #2. Figure 16 shows the comparison of regime durations to complete the targeted left turns between the two pilots. In one case, pilot #2 took 40 seconds to complete the maximum 37-degree left turn maneuver. The figure also shows that the achieved bank angles for pilot #1 are concentrated at the targeted 30-degree and 45-degree turn maneuvers, whereas those of pilot #2 are spread out.

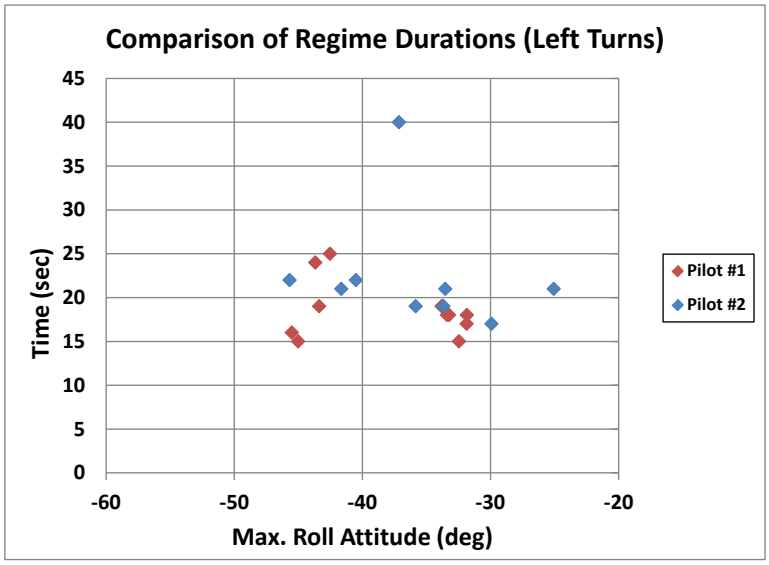

Figure 16. Comparison of left turn durations between the two pilots

From these two observations (longer regime durations and more spread out bank angles), it would appear that pilot #2 might have less experiences in handling these high turn maneuvers. Further detailed investigations of the flight test data reinforces this speculation; figures 17 and 18 show a comparison of the time histories of major flight parameters between the two pilots. High degrees of fluctuation in pitch, roll, and yaw rates can be seen from the curves for pilot #2 in figures 17 and 18, indicating rougher maneuvers.

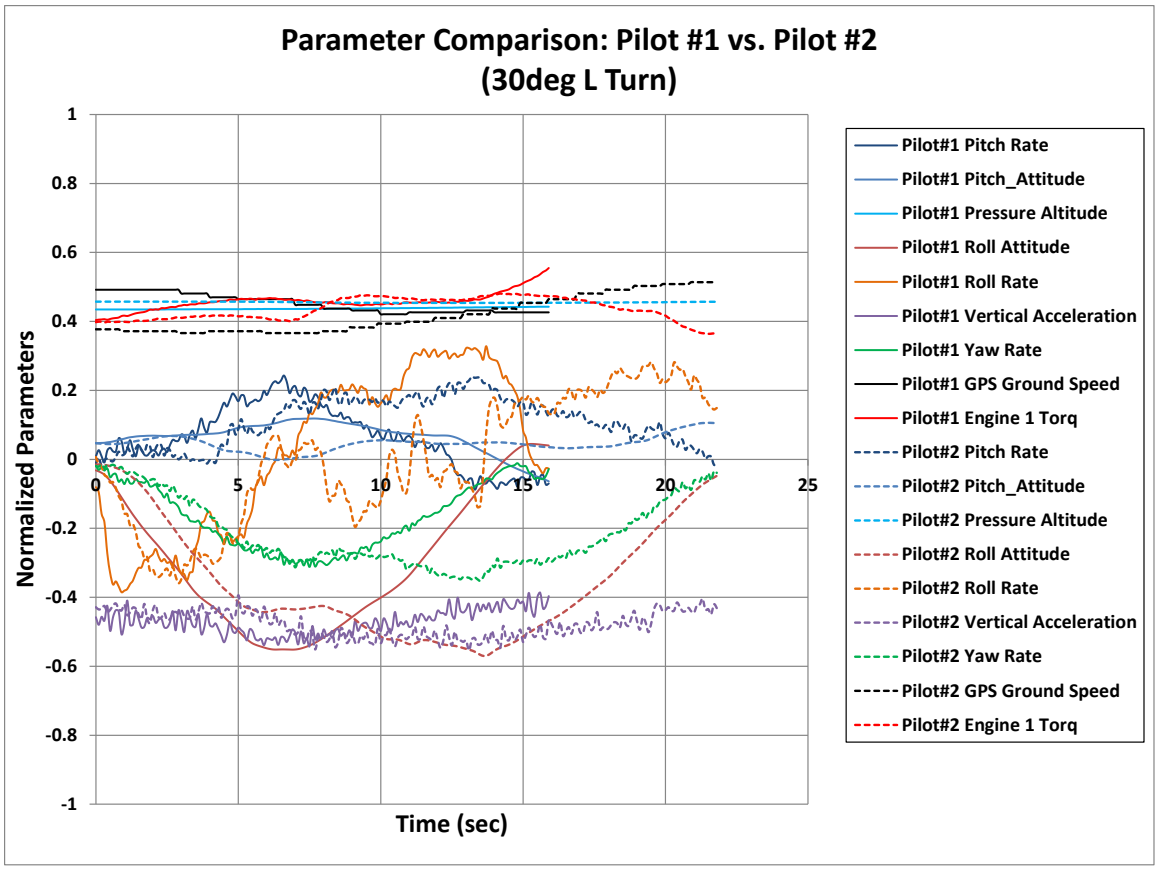

Figure 17. Flight parameter comparison of 30-degree left turns between the two pilots

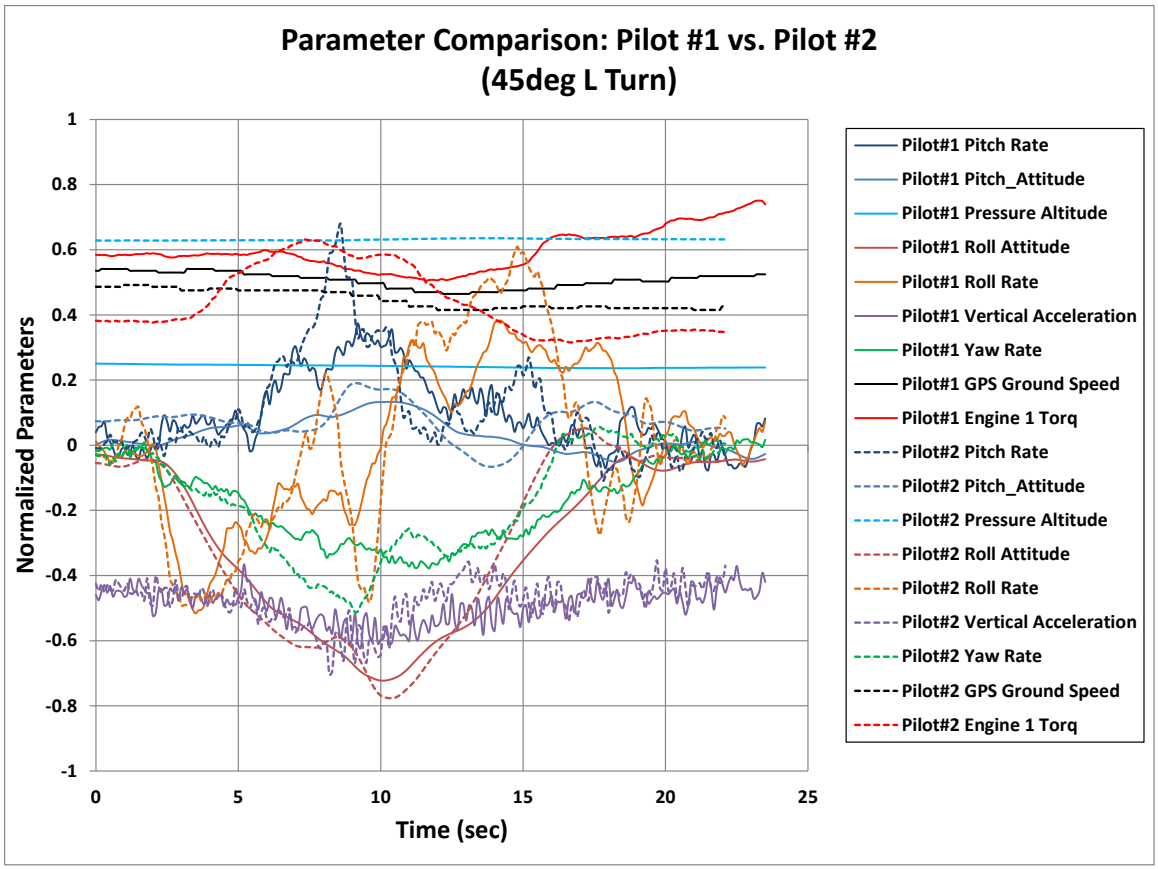

Figure 18. Flight parameter comparison of 45-degree left turns between the two pilots

5.2 COMPARISON OF MIN/MAX FLIGHT MANEUVER PARAMETERS

Pilot #1 may be a more experienced pilot such that he could control the turn maneuvers more smoothly compared to pilot #2 (the copilot). The higher fluctuations shown in figures 17 and 18 for pilot #2 could result in more fragmented regime recognitions and therefore could impact the results of the regime recognition. Such an impact of pilot variability due to the fluctuation of flight parameters may not be clearly identified by statistically analyzing the average or min/max values of flight parameters as described in the ARL report [11]. As shown in Figure 19, the min/max values of major flight maneuver parameters do not exhibit visible differences between the two pilots. Unfortunately, this observation could not be confirmed because the RRA that ran with this flight test data was not available to TDA.

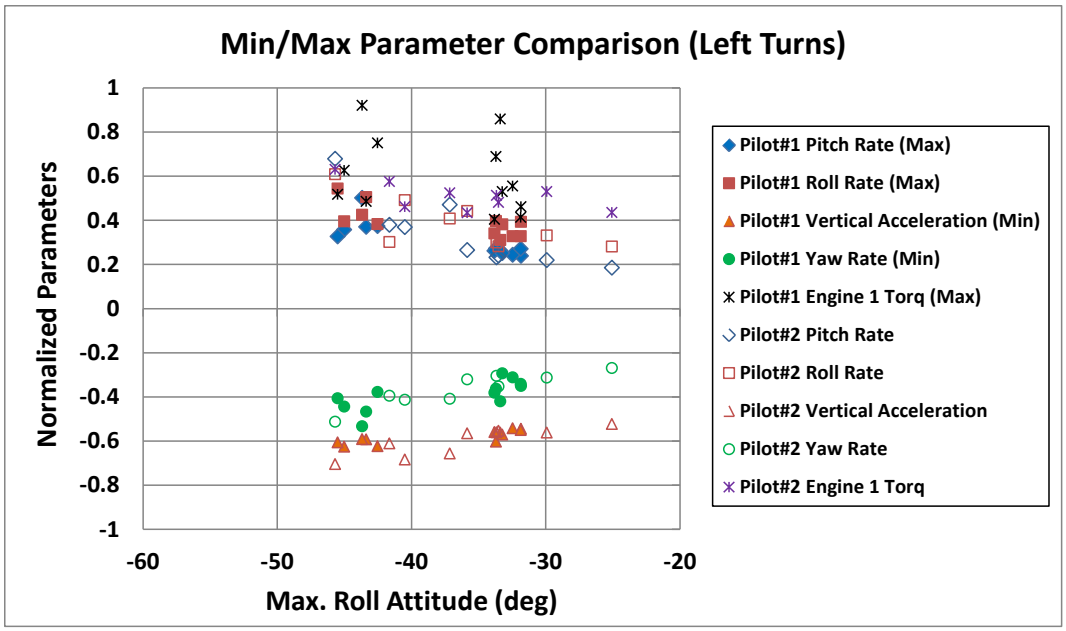

Figure 19. Comparison of min/max flight parameters between the two pilots

6. IMPACT OF REGIME RECOGNITION ACCURACY ON THE COMPONENT FATIGUE LIFE

One of the primary goals of using HUMS is to achieve more accurate structural life tracking of critical safety items of rotorcraft. Therefore, it is important to understand how the accuracy level of RRAs in the end-to-end paradigm of structural life tracking using HUMS will affect the assessed fatigue lives of rotorcraft structural components. In this section, the impact of regime recognition accuracy on the fatigue lives of dynamic components is described.

6.1 USE OF COMPOSITE WORST CASE SPECTRUM AS A REFERENCE USAGE SPECTRUM

It is common that scripted HUMS flight test data are highly skewed to the transient maneuver regimes because they get more attention during the flight testing for their higher-damaging flight conditions. Figure 20 shows the comparison of regime times and occurrences between the usage spectrum of the example scripted HUMS data and the Composite Worst Case (CWC) usage spectrum. It shows that the scripted HUMS data includes relatively small amounts of times for steady flight conditions, such as hover and level flights, and much larger times for high-maneuver conditions, such as turns and pull outs. In this example case, transient maneuver regimes take approximately 60.7% of the total HUMS flight time compared to the 16.5% of the flight time for the CWC usage spectrum.

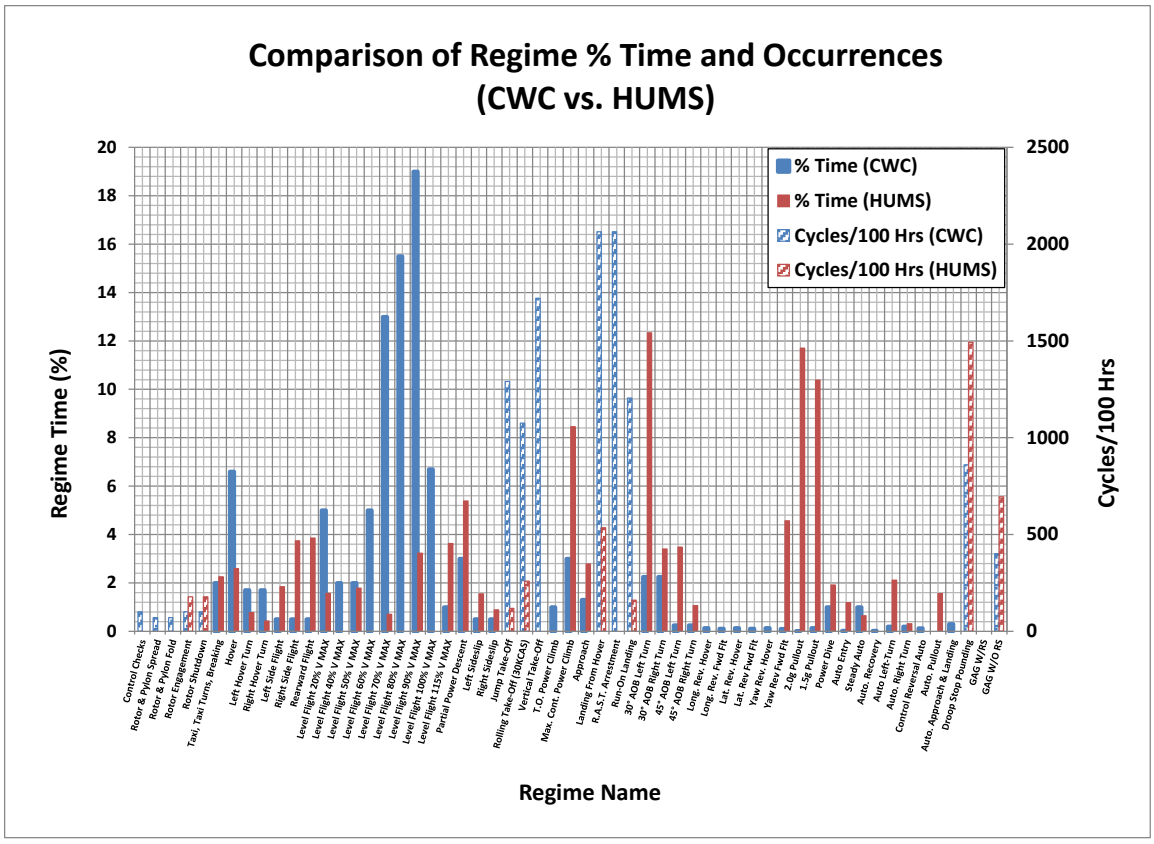

Figure 20. Comparison of regime times and occurrences

To study the impact of regime recognition accuracy on the fatigue lives of components, it is important to establish a standard reference usage spectrum such that the component lives can be assessed and compared for different levels of regime accuracies. Considering that the CWC usage spectrum is commonly used in establishing the baseline fatigue life of a component during the aircraft development, the CWC usage spectrum is chosen to study the impact of regime recognition accuracy levels. By doing this, it is assumed that the accuracy of each regime calculated earlier remains unchanged for different regime times.

6.2 RECOMMENDED ACCURACY REQUIREMENT FOR REGIME RECOGNITION

As explained in section 6.1, the fatigue life of a component can be calculated using the CWC spectrum as a reference spectrum at different levels of the regime accuracies. Figure 21 shows an example trend of fatigue life for a main rotor rotating swashplate for various levels of regime accuracies. The lives along the upper boundary curve are calculated by subtracting the margin of regime error from the reference usage (best case scenario), whereas the lives on the lower boundary are calculated by adding the regime error to the mean usage (worst case scenario). Figure 22 shows the component lives normalized by the CWC life.

Based on figures 21 and 22, it is conceivable that the 95% regime accuracy would provide a minimum accuracy of approximately 90% for the component life calculated, as denoted by the dashed orange line (see figure 22).

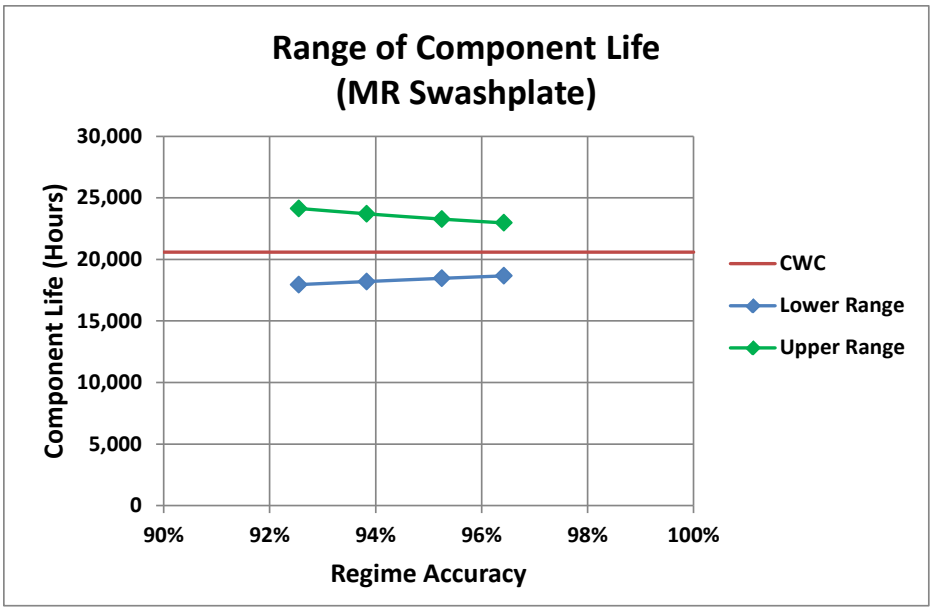

Figure 21. Example range of main rotor rotating swashplate life vs. regime accuracy

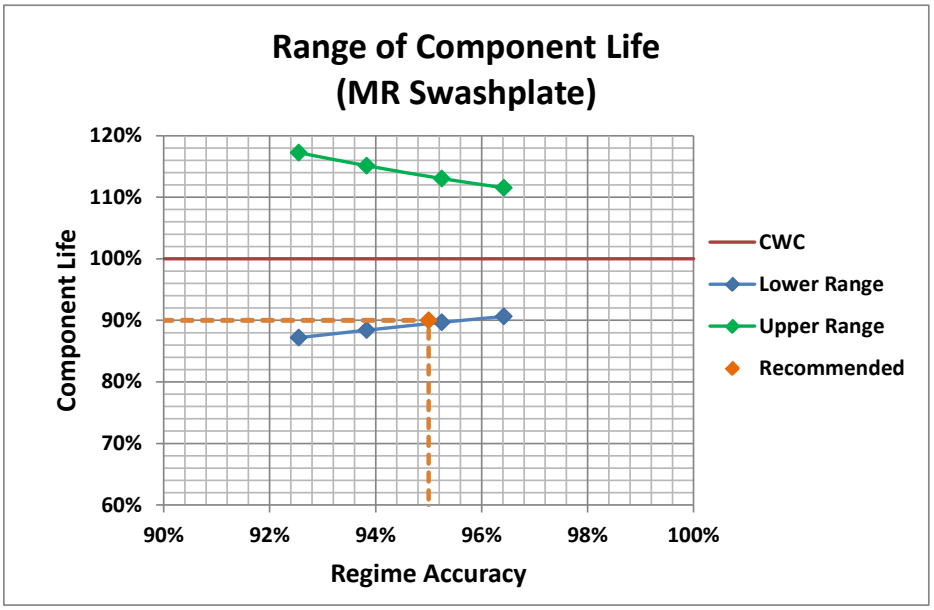

Figure 22. Example % range of main rotor rotating swashplate life vs. regime accuracy

Figures 23 and 24 also show similar trends of fatigue life for a main rotor pitch control horn, in which the 95% regime accuracy would provide a minimum accuracy of 88% for the component life. Note that the pitch control horn has a higher damage rate due to the large scatter in its fatigue test data, causing the higher uncertainty [10].

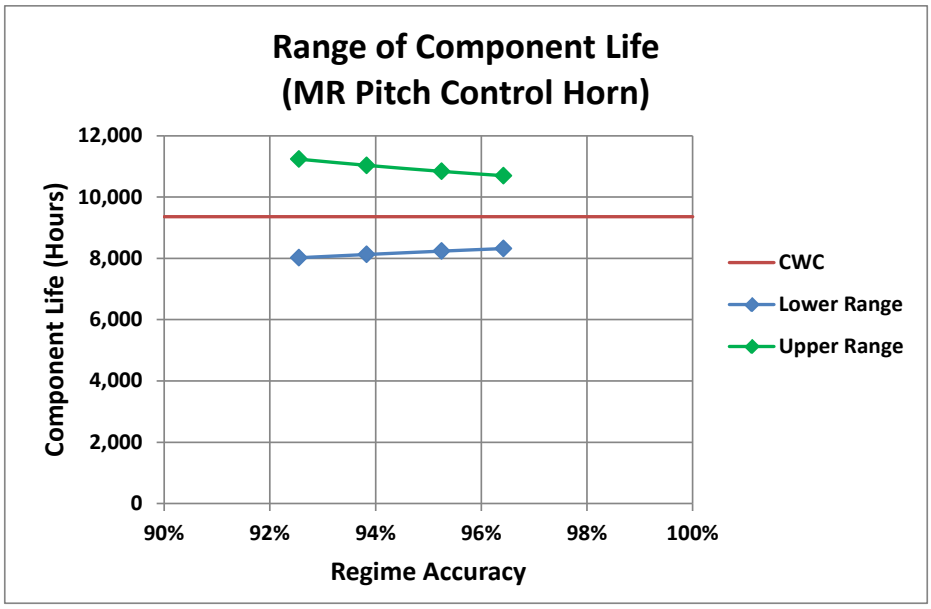

Figure 23. Example range of main rotor pitch control horn life vs. regime accuracy

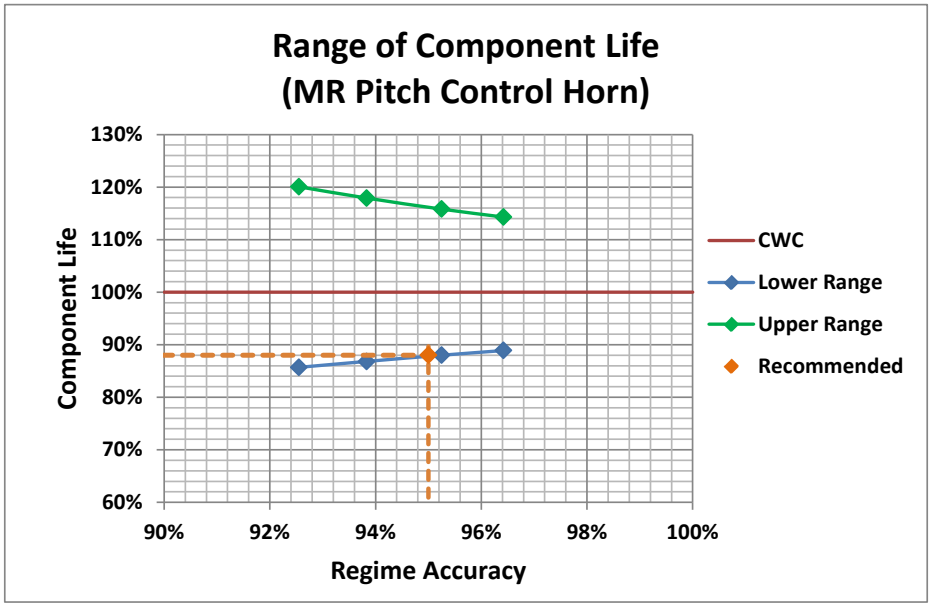

Figure 24. Example % range of main rotor pitch control horn life vs. regime accuracy

Considering that the component lives with these minimum accuracies are conservative ones, it is recommended to use the 95% regime recognition accuracy as a requirement in practice with the expected impact of regime accuracies of approximately 10% on the component life.

7. CONCLUSIONS

An accuracy assessment methodology for the verification and validation of Health and Usage Monitoring System (HUMS) regime recognition algorithms (RRAs) is developed using a confusion matrix with control factors (CFs) based on the comparison results of regime vibratory

load patterns. More than 90% of regime recognition accuracies were assessed for the evaluated regime recognition code. Pilot variability and the impact of regime accuracy on the component fatigue life were studied. The following are highlights and conclusions of the study:

- The scripted regimes by pilot cards and the recognized regimes by an RRA were compared using the confusion matrix to assess the accuracy of the RRA
- CFs were calculated by using a pattern matching of regime vibratory loads
- A multivariable regression analysis was used with five typical flight parameters to predict regime vibratory load patterns
- More than 90% accuracy was observed for the regime recognition code evaluated with various pattern-matching criteria
- More than 95% accuracy was observed using a +/-15% pattern-matching criterion, and this criterion is recommended in practice
- 95% is recommended as a requirement of regime accuracy for a minimum accuracy of 90% for a component life using HUMS

8. REFERENCES

1. "Airworthiness Approval of Rotorcraft Health Usage Monitoring Systems (HUMS)," FAA AC 29-2C MG 15, February 12, 2003.
2. Hong, C., Algera, D., Green, D., DiCampli, E., and Chin, H., "Validation of FAA AC-29-2C MG-15 for Usage Monitoring Credits," FAA report DOT/FAA/TC-15/11, to be published in 2015.
3. "Aeronautical Design Standard Handbook for Condition Based Maintenance Systems for US Army Aircraft," U.S. Army ADS-79-HDBK, Appendix F, "Fatigue Life Management," February 17, 2009.
4. Everett, Jr., R.A., Bartlett, Jr., F.D., and Elber, W., " Probabilistic Fatigue Methodology for Six Nines Reliability," NASA TM 102757, AVSCOM TR 90-B-009, December 1990.
5. Shenliang, W., Bechhoefer, E., and David, H., "A Practical Regime Prediction Approach for HUMS Application," *63rd AHS Forum*, Virginia Beach, Virginia, May 2, 2007.
6. Bechhoefer, E., "The Best Regime Recognition Algorithm for HUMS," *AHS International Condition Based Maintenance Specialists' Meeting*, February 12–13, 2008.
7. Bates, P., Davis, M., and Sadegh, P., "Rotorcraft Dynamic Component Usage Based Maintenance Process," *66th AHS Annual Forum*, Phoenix, Arizona, May 11–13, 2010.

8. He, D., Shenliang, W., and Bechhoefer, E., “A Regime Recognition Algorithm for Helicopter Usage Monitoring,” *Aerospace Technologies Advancements*, January 2010, pp. 492.

9. Moore, A., “Regression and Classification with Neural Networks,” Lecture Notes, School of Computer Science, Carnegie Mellon University, 2001.

10. Hong, C-H, Algera, D., and Chin, H., “Probabilistic Fatigue Damage Modeling for the Reliability Based Rotorcraft Component Life Assessment Using HUMS,” Airworthiness, CBM and HUMS Specialists’ Meeting, Huntsville, Alabama, February 9–11, 2013.

11. Bradley, N., “Federal Aviation Administration Health and Usage Maintenance Mock Certification Pilot Variability Written Documentation,” U.S. Army Research Laboratory Report ARL-TR-6922, May 2014.