Accuracy Assessment of Regime Recognition Algorithms Using Confusion Matrix

Chang-Hee Hong, PhD

Sr. Engineer, TDA

AHS/FAA CBM HUMS Meeting Feb 9-12, 2015

Program Overview

Objective

- **Develop an accuracy V&V methodology of HUMS RRA , and investigate the impacts of accuracy requirements and mitigations**
- **Define a quantitative accuracy requirement for RRA**

– **Part of *the End-to-End Validation of Structural Usage Tracking Paradigm for Civil Helicopter Operators (AC 29-2C MG-15: "Airworthiness Approval of Rotorcraft HUMS")***

– **FAA Research Grant 10-G-020**

Team

- **Helicopter Association International**
- **Technical Data Analysis, Inc.**
- **Consultants: Dr. H. Chin and D. Green**

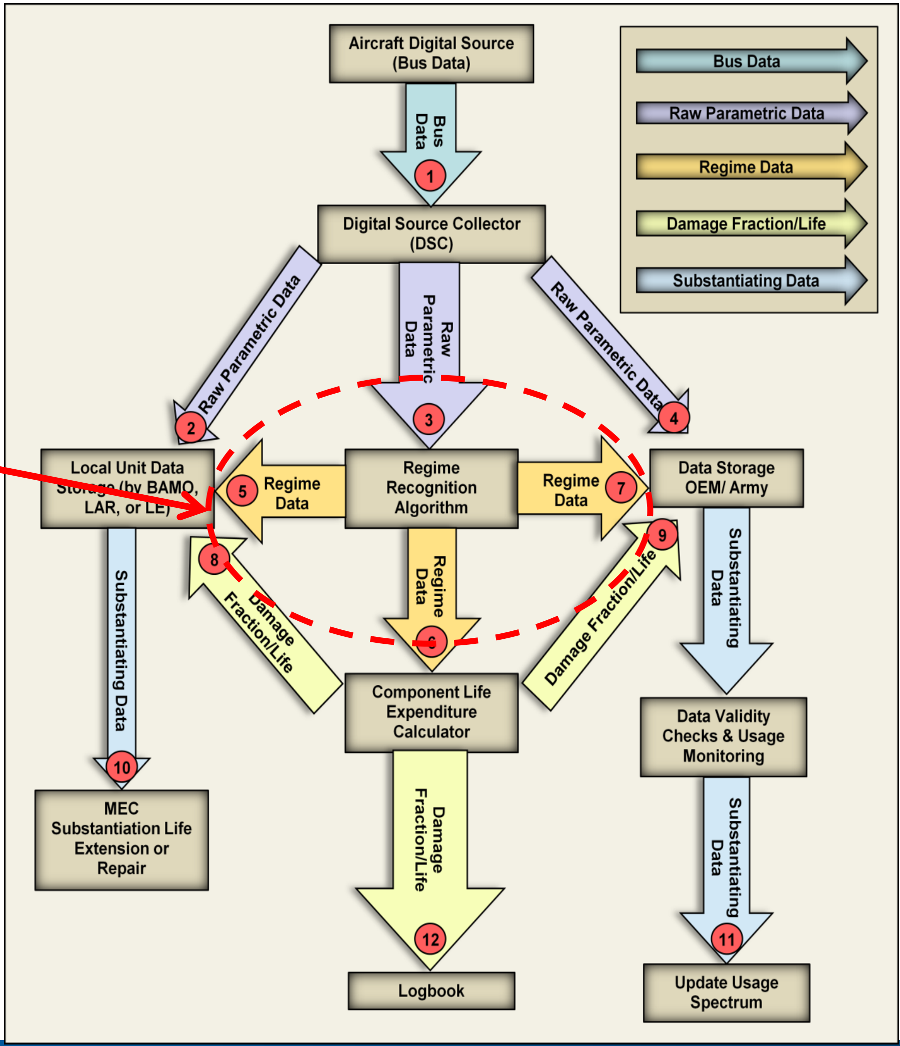
Note: End-to-End: From the starting point of airborne data acquisition to the defined credit without further significant processing

Objective of Usage Monitoring

To improve Fatigue Damage Assessment ***by updating Service Usage Spectrum***

- **Regime Recognition**

  *Accurate characterization of operational flight regimes is **a key characteristic** of the CBM system (ADS-79D).*



For Regime Recognition Algorithms (RRA):

- **What is the adequate level of accuracy requirement?**

- ***How to verify and validate (V&V) the algorithms?***
- ***What is the impact of accuracy levels on the component fatigue damage?***

Introduction

<u>RR Accuracy Requirements for Military Rotorcraft</u> (ADS-79D)

I. 97% Accuracy of Regime Recognition

“Must identify the maneuvers flown, their severity and duration, such that 97% of the entire flight time is properly identified.”

➢ **Shown to be challenging to meet this requirement.**

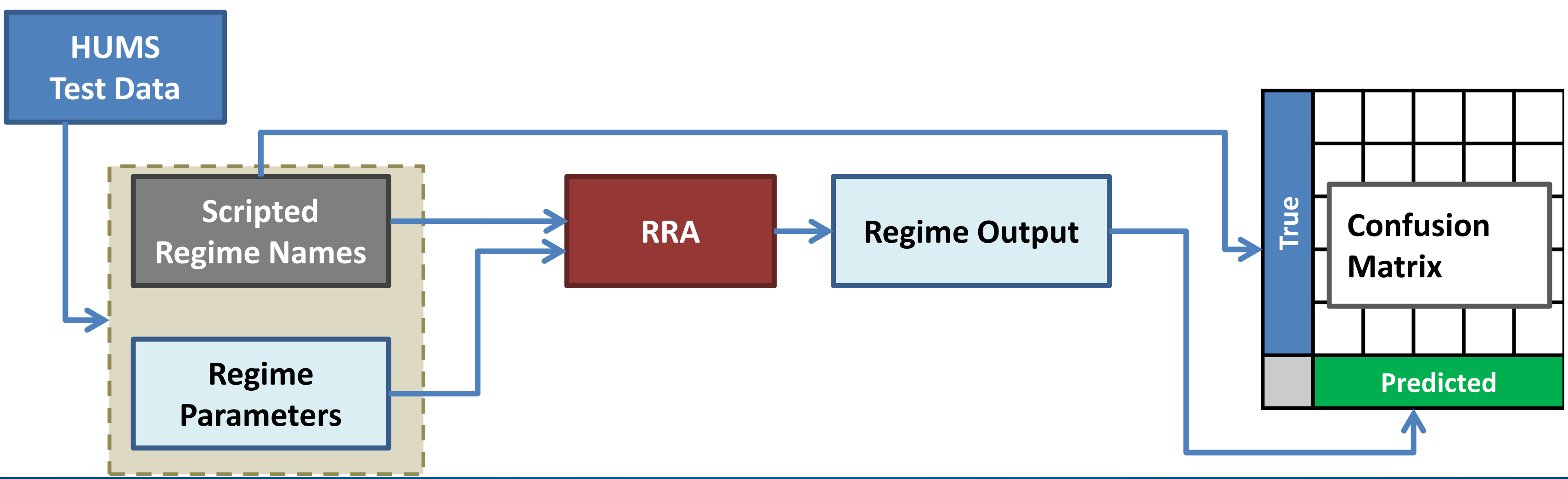
II. Less than 0.5% under-prediction of fatigue damage by any unrecognized regimes

- **Revision D of the handbook *allows for relief from 97%* if one can establish *"sufficient regimes"* based on the damage fraction.**

Validation of RRA

- **Not a simple *black & white* comparison of identified regimes against the ones in the flight test pilot cards.**
  - Variation in regime names, definitions and criteria
  - RR codes are heavily depended on the flight survey data
  - RR Code V&V are mostly done by a laborious manual process

➢ **Very difficult to define and develop an *objective* process for a *quantitative* accuracy measure.**



Difficulties in Regime Recognition

Major issues: Identifying multiple *"short"* regimes in a single maneuver, including toggling

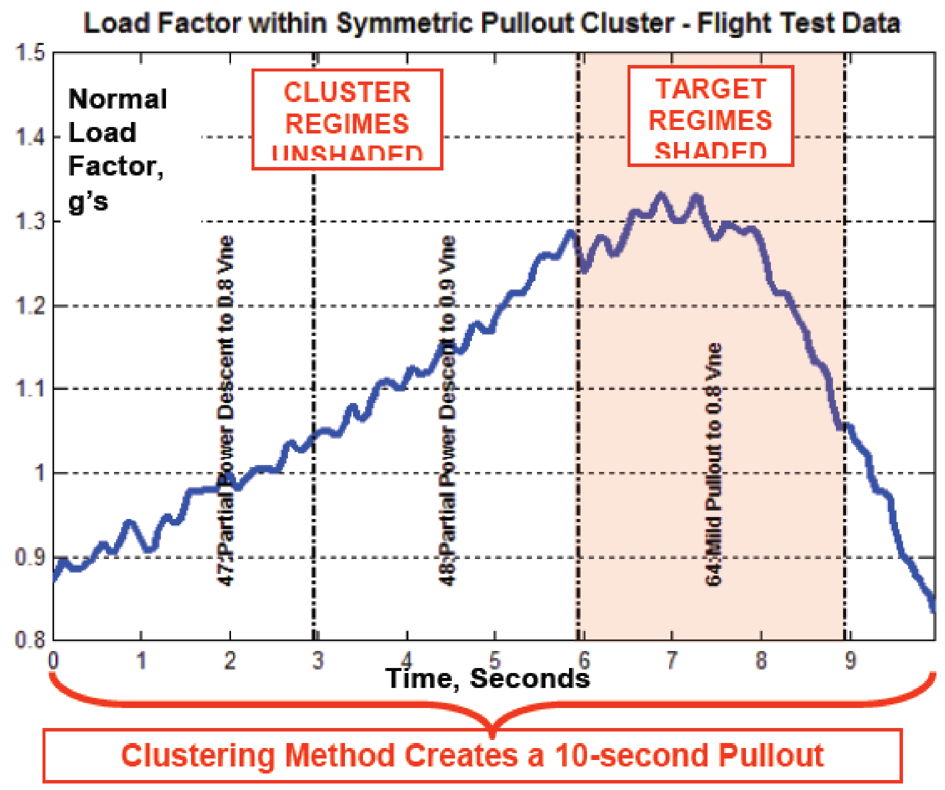
Example Output of RR codes using Scripted Flight Maneuver Data

time	Angle_of_Bank	Pitch_Attitude	Rate_of_Climb	Vertical_Accel	Vne_Fraction	Regime_Piloted	Regime_Recognized
4238	1.89	-5.47	-2224.08	1.69	79.05	SymmPullUp-LSDive	SymmPullUp-LSDive
4239	1.78	14.70	-1509.75	1.74	77.88	SymmPullUp-LSDive	SymmPullUp-LSDive
4240	0.46	20.01	-356.74	1.54	72.74	SymmPullUp-LSDive	SymmPullUp
4241	0.53	13.68	625.93	1.11	69.21	SymmPullUp-LSDive	MaxContPwrClimb
4242	1.24	11.16	1119.25	1.01	70.82	SymmPullUp-LSDive	MaxContPwrClimb
4243	0.87	6.70	1084.39	0.88	70.97	SymmPullUp-LSDive	SymmPushOver
4244	0.38	4.06	807.51	0.87	71.13	SymmPullUp-LSDive	MaxContPwrClimb
4245	0.06	2.45	543.65	0.90	71.28	SymmPullUp-LSDive	MaxContPwrClimb
4246	0.06	1.28	271.37	0.93	71.43	SymmPullUp-LSDive	LevelFlight112
4247	0.11	0.66	28.91	0.95	71.59	SymmPullUp-LSDive	LevelFlight112
4268	1.04	4.02	470.09	1.07	74.80	SymmPullUp-LSDive	LevelFlight128
4269	0.95	4.58	510.28	1.05	74.95	SymmPullUp-LSDive	MaxContPwrClimb
4270	0.62	4.81	516.24	1.02	75.10	SymmPullUp-LSDive	MaxContPwrClimb
4271	0.44	4.21	402.82	0.98	75.26	SymmPullUp-LSDive	LevelFlight128
4275	0.10	2.09	334.18	1.06	75.87	SymmPullUp-LSDive	LevelFlight128
4276	0.14	1.67	485.56	1.10	76.02	SymmPullUp-LSDive	LevelFlight128
4277	0.64	2.83	594.73	1.12	76.17	SymmPullUp-LSDive	MaxContPwrClimb
4278	1.34	5.35	670.15	1.16	76.33	SymmPullUp-LSDive	MaxContPwrClimb
4279	1.29	9.87	856.35	1.28	76.26	SymmPullUp-LSDive	MaxContPwrClimb
4280	1.95	12.39	1255.89	1.25	73.41	SymmPullUp-LSDive	MaxContPwrClimb

Difficulties in Regime Recognition

- It is common to recognize multiple regimes in a transient maneuver time by "rule-based" RRA.

  - ➢ How to handle (treat) these multiple sub-regimes?



(AHS 2010, Sikorsky)

Difficulties in Regime Recognition

- **Simple *black & white* comparison results in a poor accuracy**
- **To deal with multiple regimes identified --> Use Confusion Matrix (Error Matrix):**
  - Visualize algorithm performance
  - Provide a *quantitative* system accuracy (by definition)

Example of Confusion Matrix and RRA Accuracy

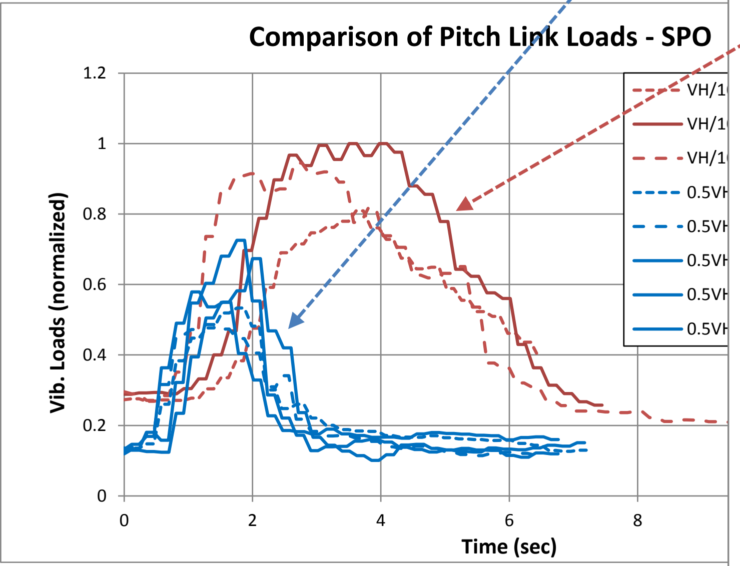
Regime Name	Row Labels	4	5	9	11	44	80	94	123	Grand Total	Accuracy (%)
Descent	4	371	170	75	13	35	0	22	6	1528	24.3
Dive	5	12	32	0	0	0	0	0	5	132	24.0
HvrIGE	9	0	0	67	0	0	0	0	0	89	75.2
IntPwrClimb	11	108	154	17	148	38	0	104	1	3534	4.2
LTurn	44	16	0	0	0	109	0	0	0	302	36.1
RRollingPullUp-Dive	80	53	7	0	0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

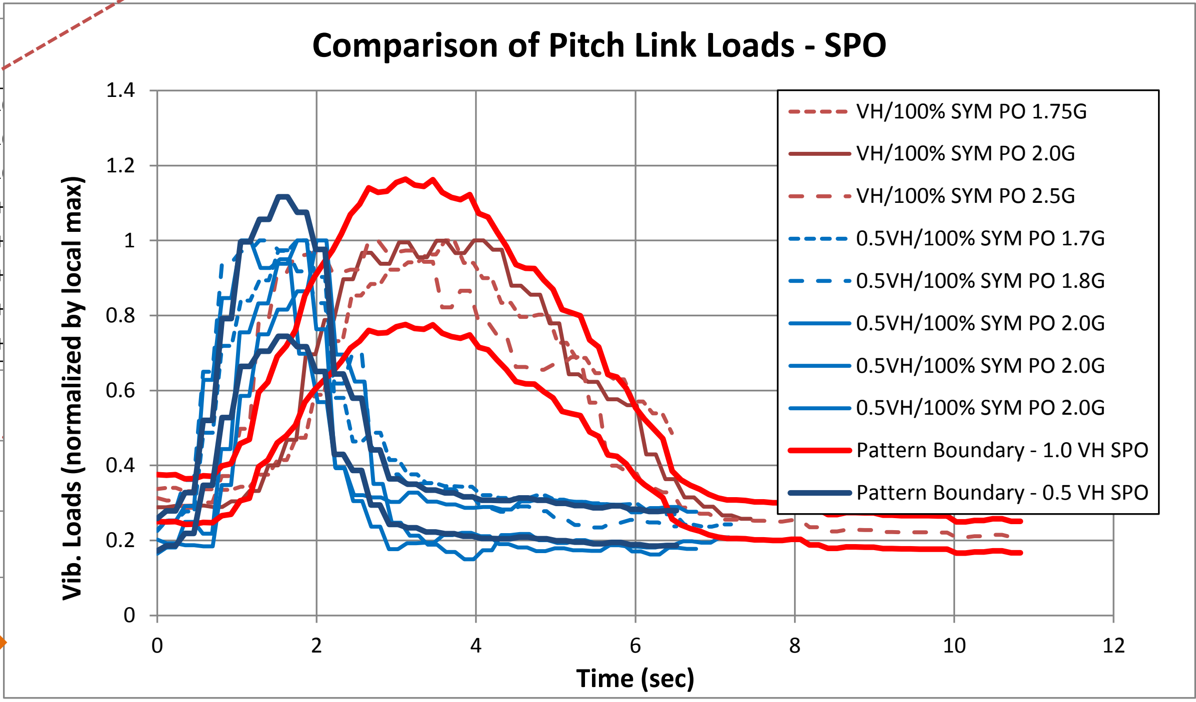
Use Correlation Factors for Improved Accuracy

- Adjust accuracy based on the correlation factors
  - Use regime vibratory load patterns

Example of Vib. Load Patterns – Transient Maneuvers

– Sym. Pullouts: 0.5 VH vs. 1.0 VH



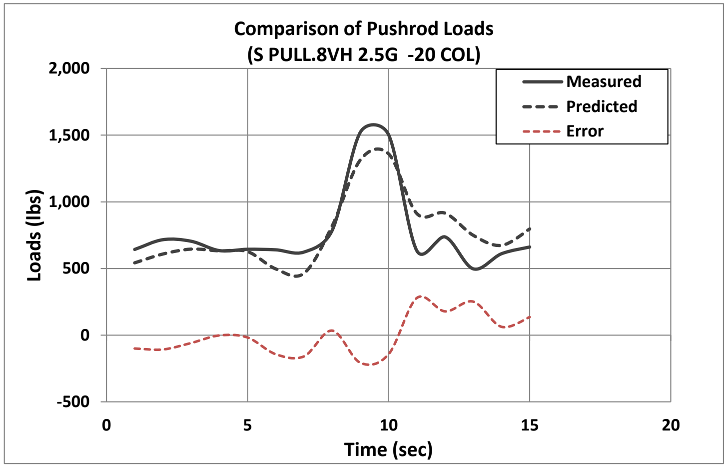


How to *predict* regime vibratory loads accurately?

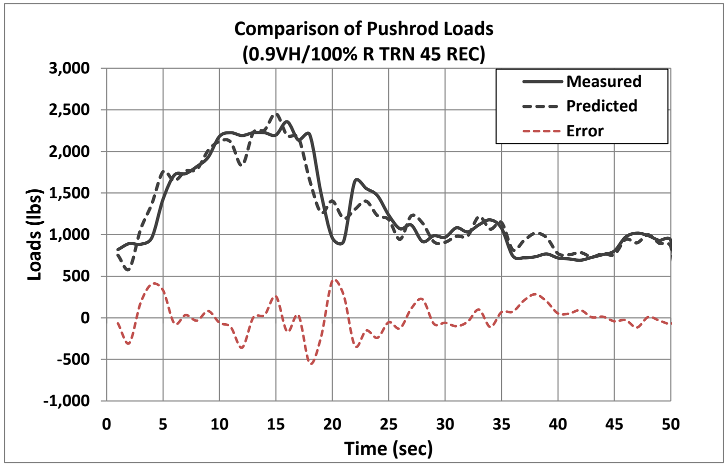
- HUMS does not record component load data
- Use Multi-variable Linear Regression Analysis (MVLR)
  - 5 maneuver parameters with known vib. loads (flight load survey)
    - Pitch
    - Roll
    - Rate of climb
    - Vertical Acceleration
    - Vne Fraction

Examples of maneuver loads prediction

2.5G Sym. Pullout



R Turns – 45 D



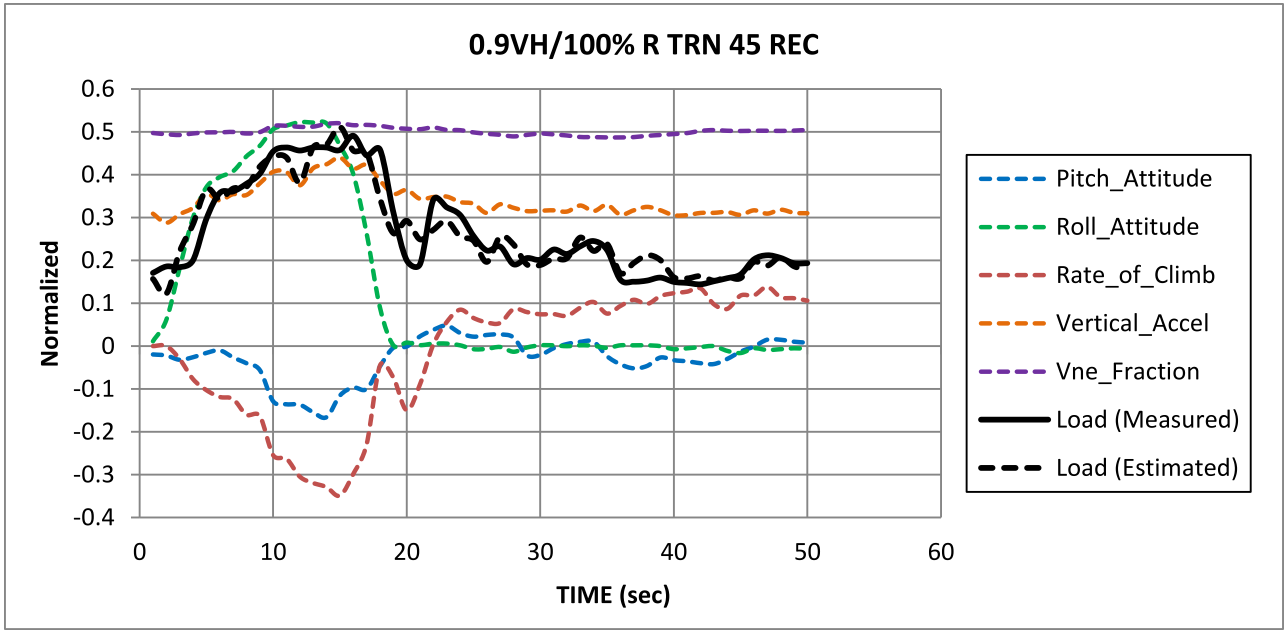
How to *predict* regime vibratory loads accurately?

MVLR Eq.: $\text{Out}(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{x} = w_1x[1] + w_2x[2] + \ldots . w_mx[D]$

where, max. likelihood *w* is: $\boldsymbol{w} = (X^TX)^{-1}(X^TY)$

- ***Prediction is good enough for the regime load pattern recognition.***

Example of measured and predicted regime loads

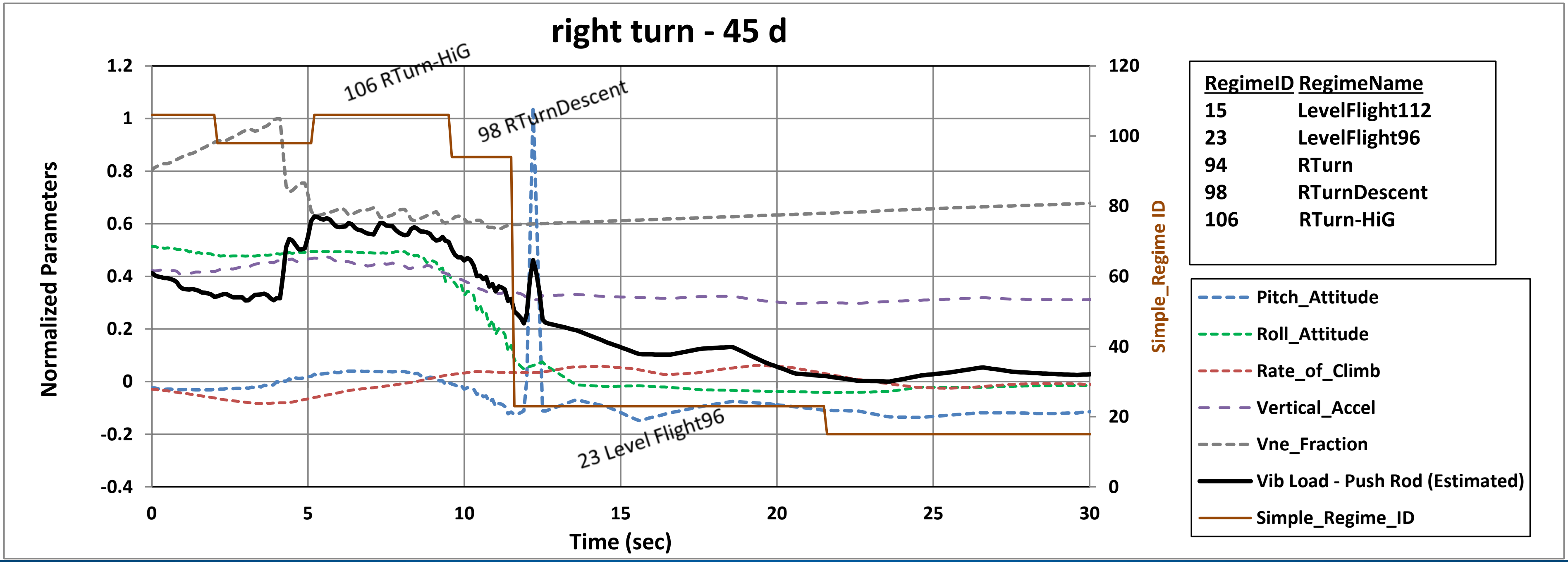


How to *predict* regime vibratory loads accurately?

- Survey data is grouped by GW, HD, Ng and Vne
- Regression coefficients are calculated for each group:

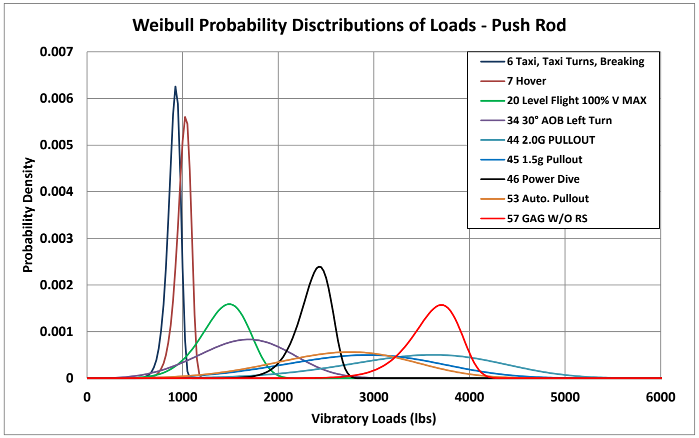
➢ Using the calculated regression coeff. sets, regime load can be predicted from the HUMS parameter data

Example of regime vib. loads predicted from the HUMS flight parameters

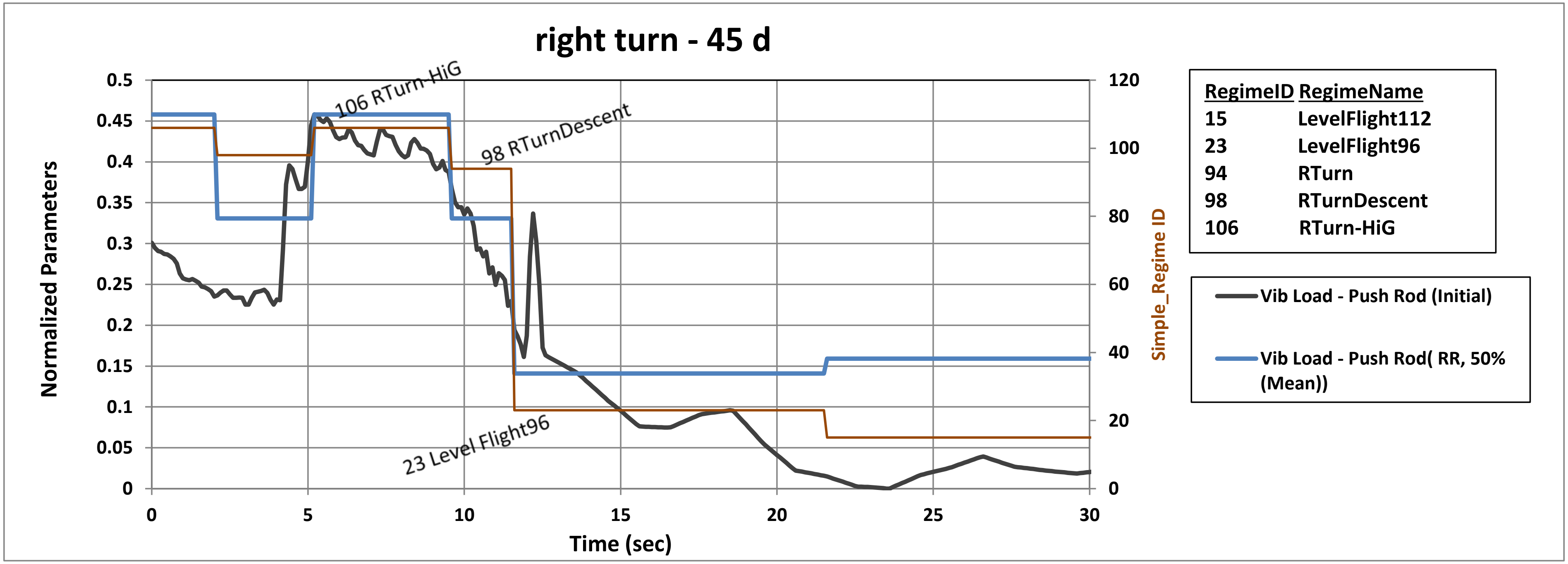


Use Load Patterns for Improved Assessment

- For RRA, use mean load levels (at 50%) from the regime load distribution curve
- Use predicted value (from regression analysis) –> as a *true value*



Comparison of pitch link vib. loads: Predicted vs. RRA



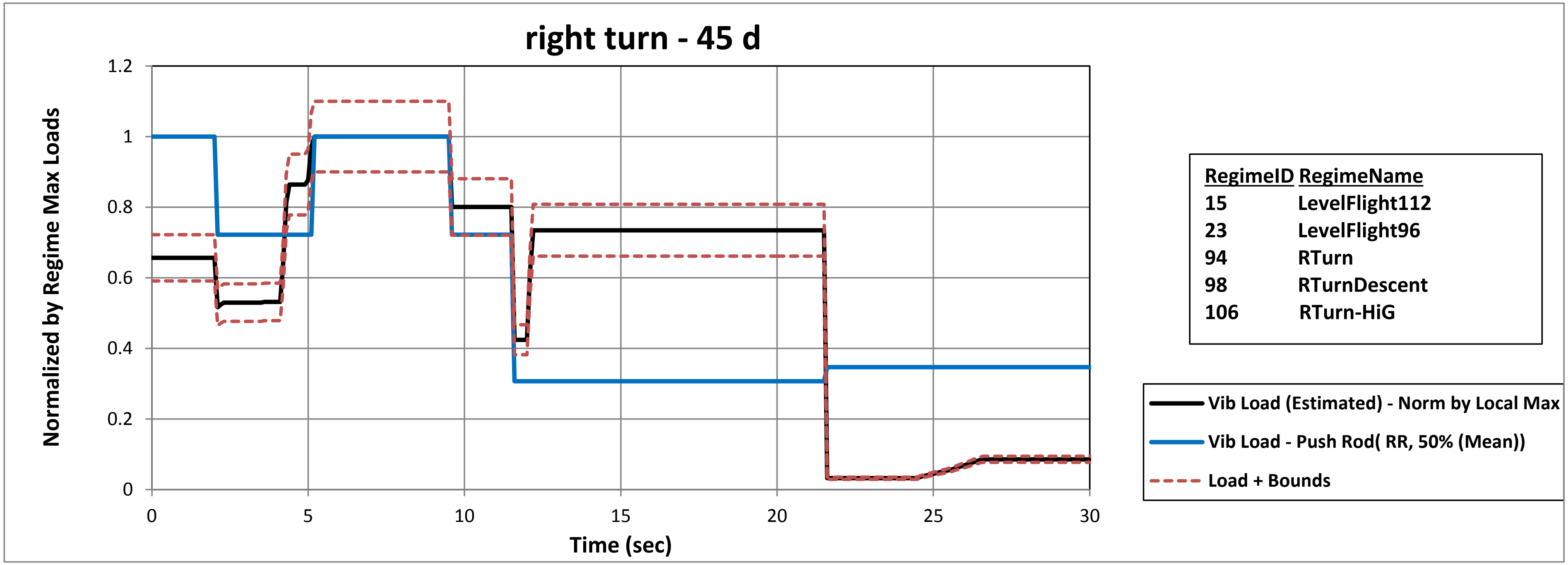
Use of Load Patterns for Improved Assessment

- **Calculate the differences (Δ Loads)**

$$\Delta \text{ Load} = \text{RRA Load} - \text{True Value}$$

  - Normalized by local regime max. loads

Example of load patterns with the criteria boundaries (+/- 10%)

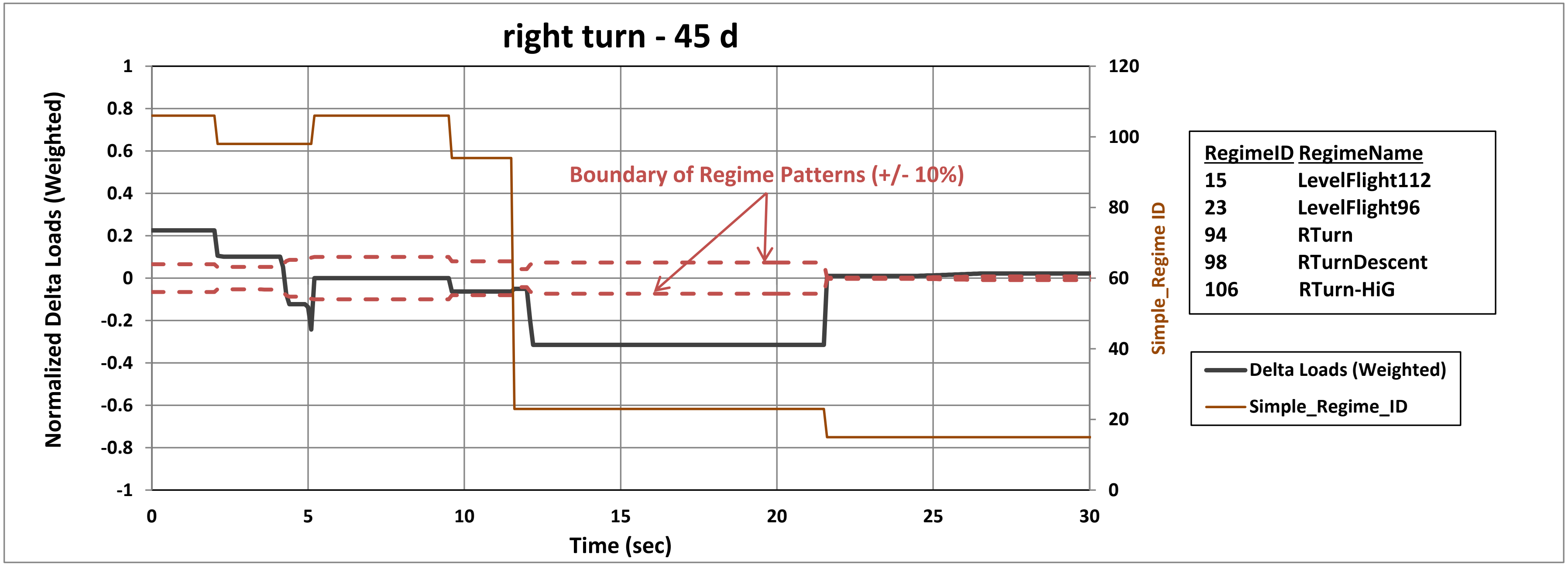


Logic to determine Correlation Factors

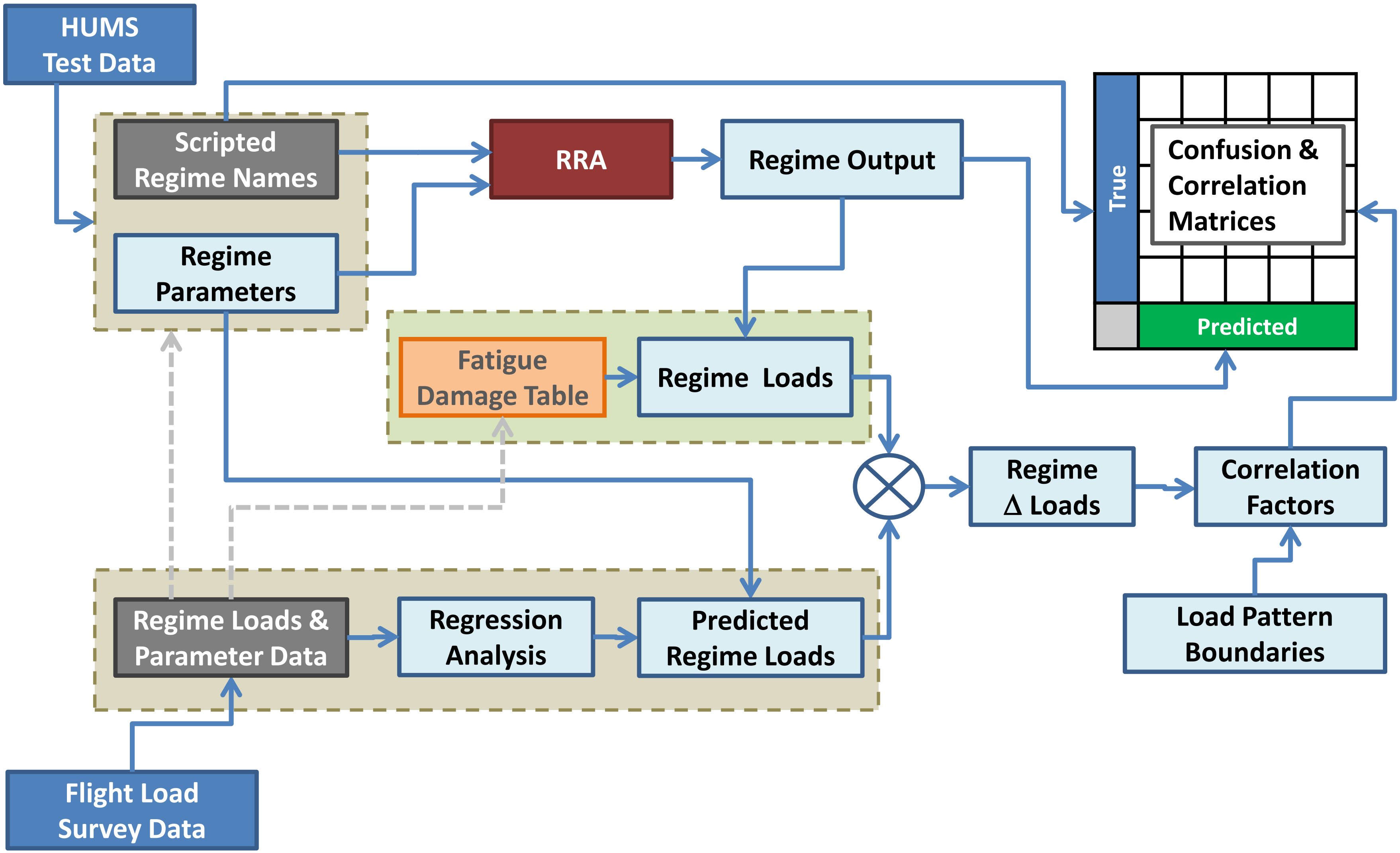
1) If ΔLoad is *within the boundary*: CF = 1.0
2) If ΔLoad is *outside the boundary*: CF = $(\Delta L - \Delta L_{bnd.})/(1 - \Delta L_{bnd.}) < 1.0$

➢ *More accurate v&v: based on both magnitudes and patterns*

Example ΔLoads with the load pattern boundaries (+/- 10%), (V&V Credit 90.3%)



Accuracy Assessment of HUMS Regime Recognition Algorithms



Summary of Assessed RRA Accuracy

- Accuracy increases with bigger credit boundaries.
- ➢ Tailored boundaries with data cleanup will provide higher accuracy.

Maneuver Group	Accuracy (%)		
	Credit Boundary (+/-)		
	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

Example of Regime Time Confusion Matrix with Assessed Regime Accuracies

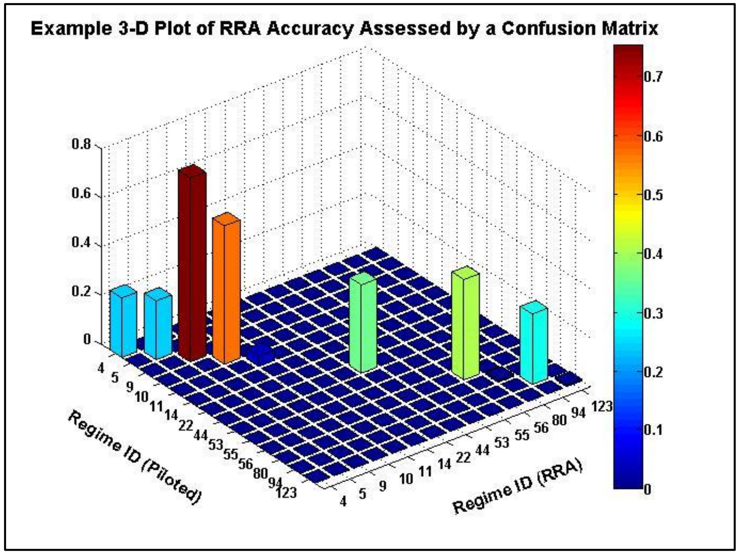
(Note: Accuracy is calculated before cleaning up the data with +/- 10% boundary.)

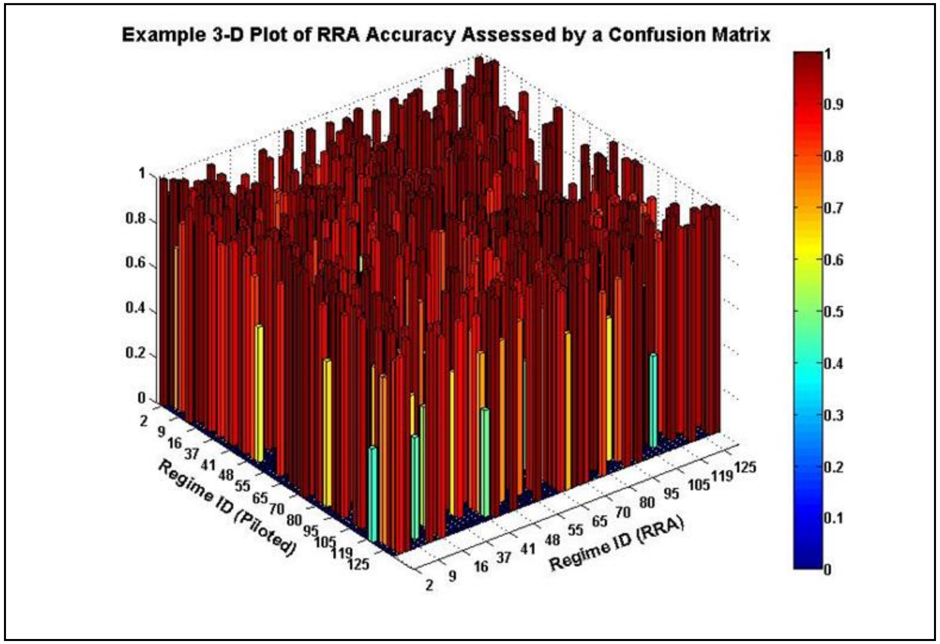
Regime Name	Regime ID	4	5	10	11	26	44	94	123	128	Accuracy (%)
Descent	4	371	167	62	9	1	31	23	6	1	84.8
Dive	5	8	32						5	1	82.7
HvrOGE	10			53			7				92.9
IntPwrClimb	11	94	134	141	148	1	35	99	1	4	56.5
LevelFlight	14	23			49	1	28	13		0	94.9
LRollingPullUp	26	45	18			13	122		7	9	69.9
LTurn	44	29		4	7	10	579	581		0	93.6
RTurn	94	14	0	0	0			264		2	91.6
SymmPullUp	123	72	2	2		12	32	35	6	5	63.1
SymmPushOver	128	110	33		8		54	9	15	31	85.2

Summary of Assessed RRA Accuracy

- Found numerous abnormalities in the data
- Data cleanup will increase the accuracy
- A Confusion Matrix bar chart provides a good visual distribution of assessed regime accuracies.

Bar Charts for the Example CM's (before and after adjusted)





Summary

- Confusion Matrix with correlation factors is used for an accuracy assessment of HUMS RRA.
- Vibratory load patterns are used to obtain correlation factors to improve the accuracy measure.
  - The mean load levels (50 percentile) of the regime loads used for the fatigue damage calculation are compared against the predicted loads by regression analysis.
- Study result shows that the assessed regime accuracy is over 90% for the RR codes examined.
- A tailored credit boundary scheme and data cleanup will increase the accuracy.

Future Work

- **Assess the impact of RRA accuracy on the Component Fatigue Life**
  - **Generate the standard usage spectrum (CWC) from the scripted HUMS data**
    - **Current scripted HUMS data is highly skewed to the transient maneuvers**
  - **Study the sensitivity of component fatigue lives on the RRA accuracy.**

Discussion

FAA POC: Paul Swindell
Rotorcraft Program Manager
FAA William J. Hughes Technical Center
(609) 485-8973
paul.swindell@faa.gov

TDA POC: Chang-Hee Hong
Sr. Engineer
(703) 226-4067
chong@tda-i.com

HAI POC: Ed DiCampli
Executive Vice President
(703) 683-4646
ed.dicampli@rotor.com