U.S. Department of Transportation
Federal Aviation Administration
National Airport Pavement Safety
Facility ANG-E262
William J. Hughes Technical Center
Atlantic City Int'l Airport, NJ 08405

FAA Surface Roughness Study
Final Data Report
Boeing 737-800

August 28, 2015

Final Report

This document is available to the U.S. Public through the National Technical Information Service (NTIS), Springfield, Virginia 22161

U. S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof. The United States Government does not endorse products or manufacturers. Trade or manufacturers names appear herein solely because they are considered essential to the objective of the report. This document does not constitute FAA certification policy. Consult your local FAA aircraft certification office as to its use.

This report is available at the Federal Aviation Administration William J. Hughes Technical Center's Full-Text Technical Reports page: actlibrary.tc.faa.gov in Adobe Acrobat portable document format (PDF).

Final Data Collection Report Documentation Page

1. Report Title	FAA Surface Roughness Study Final Data Report Boeing 737-800
2. Report Date	August 28, 2015
3. Performing Organization Name and Address	Cherokee CRC, LLC 10838 E Marshall St., Ste. 220 Tulsa, OK 74116
4. Author(s)	Skip Hudspeth; David Stapleton; Jeard Ballew John Sparkman, PMP;
5. Sponsoring Agency Name and Address	U.S. Department of Transportation Federal Aviation Administration National Airport Pavement Safety Facility ANG-E262 William J. Hughes Technical Center Atlantic City Int.'l Airport, NJ 08405
6. Acknowledgement	FAA Branch AFS440 for providing the Boeing 737-800 simulator and support, Gordon Hayhoe – FAA Aviation Research & Technology Development Office for providing Profile data and reference materials.
7. Introduction	CCRC was tasked under contract #DTFACT-09-D-00018, Delivery Order 1 to develop a subjective rating scale in the form of a questionnaire and modify the Roughness Model in the Boeing 737-800 simulator located at the FAA's Flight Standard's Flight Operation Simulation Laboratory (AFS-440). This Report Covers the Final Study Data Collection conducted June 4 – 6 and June 11 – 13, 2013.
8. Key Words	Roughness, runway surface profiles, Simulation, Boeing 737-800, subjective rating questionnaire

Table of Contents

Section			Title	Page
1.0			**INTRODUCTION**	10
	1.1		Summary	10
	1.2		Background	10
	1.3		Roughness study objectives	11
2.0			**ROUGHNESS STUDY DEVELOPMENT**	11
	2.1		Enhancements to the flight simulator	12
	2.2		Real world roughness profiles	17
	2.3		Generic profiles	20
	2.4		Test scenarios	20
	2.5		Rating form development	22
	2.6		Pilot briefings	24
	2.7		Pilot rating panel selection	25
3.0			**FINAL STUDY TESTING**	25
	3.1		Roughness rating sessions	25
	3.2		Collected simulator parameters	25
	3.3		Initial data evaluation	26
		3.3.1	Evaluation of accelerometer data for consistency	26
		3.3.2	ISO processing of accelerometer data	27
		3.3.3	Additional indices	29
4.0			**DATA ANALYSIS**	30
	4.1		Analysis overview	31
	4.2		Pilot numerical ratings as a function of four ISO roughness factors	32
	4.3		Calculation of the unacceptable range for each ISO index	46
		4.3.1	The numerical rating at which a ride becomes unacceptable	46
		4.3.2	The unacceptable range for each ISO index	49
	4.4		Comparison of pilot ratings for rides with high crest factor	54
	4.5		Military and repeat pilot bias	54
	4.6		Human vibration limits	58
	4.7		Computation of best fit shifted logarithmic curves to data	60
	4.8		Comparison of the best shifted logarithmic fit to other fits	69
	4.9		Comparison of best fits when data from the preliminary study is included	76
	4.10		Variation in pilot ratings	82

4.11	Comparison of simulator and ProFAA roughness model results	84
5.0	**CONCLUSIONS**	86
6.0	**FUTURE WORK**	90
7.0	**ACKNOWLEDGEMENTS**	90
8.0	**REFERENCES**	91

ATTACHMENTS

APPENDIX A – Additional Statistical Graphs
APPENDIX B – List of Real World Taxiway and Runway Profiles
APPENDIX C - Example of Real World Surface Profile Formatting for Simulator Use
APPENDIX D – Real World Profile Height and Simulator Cockpit Acceleration Graphs
APPENDIX E – Flight Simulator Data Collection Parameter List
APPENDIX F – Accelerometer Signal Processing and ISO Index Computations
APPENDIX G - Cockpit Accelerometer Specifications
APPENDIX H - MATLAB Code to Generate Shifted Logarithmic Fits
APPENDIX I – Pre-Brief Presentation
APPENDIX J – Post-Flight Questionnaire
APPENDIX K – Pilot Rating Form
APPENDIX L – Pilot Background Data

LIST OF FIGURES

Figure		Page
1	Boeing 737-800 Simulator – Cockpit Accelerometer (G) Vs. Time (s)	15
2	ProFAA 727 Simulation – Cockpit Vertical Acceleration (G) Vs. Time (s)	15
3	Example of Boeing 737-800 simulator generic roughness surface height variations	16
4	FAA surface profiling equipment	18
5	Example of flight simulator surface height variations for a real world runway surface profile. Profile height is shown in both simulator units (feet) and original units (inches)	19
6	Boeing 737-800 Simulator cockpit seating positions	21
7	Preliminary Study Rideability Rating Form	23
8	Final Study Rideability Rating Form	24
9	Overlay of Taxiway RMS Cockpit Acceleration for Sessions 2-12	26
10	Overlay of Runway RMS Cockpit Acceleration for Sessions 2-12	26
11	Frequency response of the weighting procedure compared with the ISO specified frequency transfer function for Weighting W_k. Amplitude ratio plotted on a linear scale, left, and a logarithmic scale, right.	28
12	The Distinct Taxiway and Runway Trends for Final Study Real World Pilot Average Numerical Ratings vs. Weighted RMS.	33
13	The Problem Extrapolating a Quadratic Fit Outside of the Data Range.	34
14	Various Curve Fits by Least Squares to Pilot Average Taxiway Ratings: Linear (Blue), Quadratic (Red), Exponential (Yellow), Exponential Through (0,10) (Green), Logarithmic (Black)	39
15	Various Fits of Average Runway Rating Vs. ISO Index by Least Squares Linear (Blue), Quadratic (Red), Exponential (Yellow), Exponential Through (0,10) (Green), Logarithmic (Black)	43
16	Shifted Logarithmic Fits of Taxiway and Runway Rating Vs. Weighted VDV	45
17	Percentage of Taxiways and Runways Rated Unacceptable Vs. Numerical Rating	47
18	Average Pilot Taxiway and Runway 0-10 Rating vs. Weighted RMS	50
19	Average Pilot Taxiway and Runway 0-10 Rating vs. Weighted MTVV	51
20	Average Pilot Taxiway and Runway 0-10 Rating vs. DKup	52
21	Shifted Logarithmic Fits Highlighting Taxiways with High Crest Factor	54
22	Linear Fits of Average Pilot Rating on 37 Taxiways to Weighted VDV for Military and Non-military Pilots.	55
23	Linear Fits of Average Pilot Rating on 37 Runways to Weighted VDV for Military and Non-military Pilots.	56
24	Linear Fits of Average Pilot Rating on 37 Taxiways to Weighted VDV for Repeat Pilots and Non-repeat/Non-military Pilots.	57
25	Linear Fits of Average Pilot Rating on 37 Runways to Weighted VDV for Repeat Pilots and Non-repeat/Non-military Pilots.	58
26	Comparison of the Quadratic and Shifted Logarithmic Fits to Pilot Rating Vs. Weighted RMS	61
27	Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. Weighted RMS	62

28	Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. Weighted RMS	63
29	Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. Weighted MTVV	64
30	Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. Weighted MTVV	65
31	Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. Weighted VDV	66
32	Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. Weighted VDV	67
33	Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. DKup	68
34	Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. DKup	69
35	Linear Fit of Rating vs. Weighted RMS with Confidence & Prediction Intervals	71
36	Exponential Fit of Rating vs. Weighted RMS With Confidence & Prediction Intervals	73
37	Log. Fit of Rating vs. Weighted RMS With Confidence & Prediction Intervals	75
38	Combined Preliminary and Final Runway Average Ratings vs. Weighted RMS	76
39	Confidence & Prediction Intervals for Logarithmic Fits with Weighted RMS	77
40	Confidence & Prediction Intervals for Logarithmic Fits with Weighted MTVV	78
41	Confidence & Prediction Intervals for Logarithmic Fits with Weighted VDV	79
42	Confidence & Prediction Intervals for Logarithmic Fits with DKup	80
43	Comparison of Shifted Logarithmic Trends Using All Pilot Numerical Ratings vs. Pilot Average Ratings from Final Simulator Runs	82
44	95%, 90% and 50% Prediction Intervals for Logarithmic Fits vs. Weighted RMS	83
45	Cockpit Weighted RMS Acceleration – ProFAA Vs. Boeing 737-800 Flight Simulator	85
46	Cockpit Weighted RMS Acceleration - ProFAA Vs. Boeing 737 Sim Vs. Flex Vs. No Flex for a sample of 10 Final Study Taxiways and 8 Final Study Runways	86
47	Shifted Logarithmic Trend Comparison: All Pilots' Runway Ratings Vs. Final Testing Average Runway Ratings	89
48	Quadratic Trend Comparison: All Pilots' Runway Ratings Vs. Final Testing Average Runway Ratings	89

LIST OF TABLES

Table		Page
1	Pilot Taxiway Correlation Coefficients	21
2	Pilot Runway Correlation Coefficients	21
3	Correlation Coefficients Relating Subjective Pilot Avg. Ratings to Objective ISO Roughness Indices	32
4	Pilot Average Numerical Rating of a Taxiway or Runway When 5%, 10%, or 50% of Pilots Rate the Ride as Unacceptable	49
5	ISO Index Values at Which 5%, 10% and 50% of Pilots Are Estimated to Rate a Taxiway or Runway as Unacceptable	53
6	Index Values at Which Taxiways and Runways Become Unacceptable Based Upon the Index Values at Which 5% of Pilots Rated Rides Unacceptable	53
7	ISO Standards for Discomfort	59
8	ISO Index Values for Which Taxiways and Runways Are Unacceptable to 5% of Pilots	88

LIST OF ACRONYMS

FAA	Federal Aviation Administration
MMAC	Mike Monroney Aeronautical Center
CCRC	Cherokee CRC, LLC
NCHRP	National Cooperative highway Research Program
G	G-force
m/s^2	Meters per second squared
Hz	Hertz
fps	feet per second
s	seconds
CG	Center of Gravity
VWGD	Z-body acceleration [ft/s^2]
VQD	Pitch acceleration body axis [rad/s^2]
VXXM(1)	X-body distance of nose gear from CG [ft]
ModePosZAccel(5)	Flex mode vertical acceleration at cockpit [ft/s^2]
ASTM	American Society for Testing and Materials
RMS	Root-mean-square
rad/s^2	radians per second squared
$a_W(t)$	weighted acceleration
$(W_{k)}$	ISO Weighting
ISO	International Organization for Standardization
MTVV	Maximum transient vibration value from a running RMS computation
VDV	Fourth power vibration dose value
DK	Spinal response acceleration dose
MPR	Mean Panel Rating
PI	Profile Index
MATLAB	Matrix Laboratory

1.0 INTRODUCTION

1.1 SUMMARY

In a continuation of the Federal Aviation Administration's (FAA's) research to determine acceptable limits for airport pavement roughness, the FAA Roughness Final Study was conducted on June 4 – 6 and June 11 – 13, 2013 using the FAA's Boeing 737-800 flight simulator at the Mike Monroney Aeronautical Center (MMAC) in Oklahoma City. Thirty six pilots participated in the study over twelve test sessions. Subjective pilot ratings of pavement roughness and objective measures of cockpit accelerations were collected for all sessions. A data analysis was performed to correlate the pilot ratings with cockpit accelerations.

1.2 BACKGROUND

Airport surface roughness is controlled very closely during construction and contractors are held to high standards. These standards include maximum variances along the longitudinal and transverse axes of new runway and taxiway construction. However, once the construction is complete, the FAA does not have a reliable method for determining when airport pavement becomes "too rough" due to deterioration. In support of the goal to develop a method for evaluating in-service pavement roughness, this study developed a rating scale for pilots' subjective response to vertical cockpit vibrations excited by longitudinal pavement surface elevation disturbances.

Cherokee CRC (CCRC) began work on the Airport Pavement Roughness study with the FAA in September 2008 using the FAA's Boeing 737-800 full flight simulator in Oklahoma City. The Boeing 737-800 simulator's roughness model was modified to allow the use of real world airport surface roughness profiles and to increase the fidelity of the ground model response to roughness. A methodology was developed for presenting surface roughness profiles and obtaining pilot roughness evaluations. Test scenarios, roughness rating forms, and pre-brief and post-brief sessions were developed and refined during a series of pilot studies in 2010 and 2011. National Cooperative Highway Research Program (NCHRP) highway rideability studies [Ref. 1] were reviewed and used as models for developing the airport pavement rideability studies.

A preliminary roughness study was conducted on November 8 – 10, 2011 using the FAA's Boeing 737-800 full flight simulator located at the Mike Monroney Aeronautical Center in Oklahoma City. The study consisted of four pavement rideability rating sessions with three pilots per session, for a total of 12 subject pilots. Subject pilots were presented with a series of eighty taxiway and runway pavement roughness test scenarios. Subjective pilot ratings and objective measures of cockpit accelerations were collected for all sessions. A data analysis was performed to correlate the pilot ratings with cockpit accelerations. [Ref. 7].

1.3 ROUGHNESS STUDY OBJECTIVES

1) Evaluate the response of the simulator model(s) when the simulator standard roughness models are replaced by measured runway and taxiway profiles and determine if the simulator models accurately reproduce the roughness of the surfaces during test scenarios.

2) Develop a rating scale for pilot subjective response to vertical cockpit vibrations excited by longitudinal pavement surface elevation disturbances. The scale will range from unacceptably rough to very smooth.

3) Use the rating scale to obtain pilot ratings of simulated surface roughness.

4) Using statistical correlation techniques, relate the subjective ratings to objective measures of the cockpit vertical vibrations or objective measures of the properties of the pavement longitudinal elevation profiles.

5) Identify on the rating scale limits for cockpit vibration resulting in unacceptable, or unsafe, ride quality conditions.

2.0 ROUGHNESS STUDY DEVELOPMENT

The following tasks were accomplished in preparation for the preliminary and final roughness studies:

1. Enhanced the Boeing 737-800 flight simulator software to provide realistic cockpit accelerations in response to real world airport surface roughness profiles.

2. Developed simulator test scenarios for obtaining pilot subjective ratings of surface roughness and objective measures of cockpit acceleration and other simulator parameters.

3. Selected and formatted real world taxiway and runway surface elevation profiles for use in the test scenarios.

4. Created samples of simulator generic (random) roughness for inclusion in the test scenarios.

5. Developed routines for capturing time histories of objective test data such as cockpit accelerometer output, profile elevation and flight model parameters.

6. Developed roughness rating forms and rating panel briefings.

2.1 ENHANCEMENTS TO THE FLIGHT SIMULATOR SOFTWARE

The Boeing 737-800 simulator software was modified to allow integration of real world airport surface profiles into the simulator ground model and to enhance the existing simulator generic roughness model. The integration of real world surface profiles into the flight simulator consisted of the following tasks:

1. Integration of real world profiles into the simulator ground model
2. Alignment of taxiway and runway profiles with simulator visual scenes
3. Initial testing of cockpit acceleration response to real world roughness profiles
4. Development of aircraft fuselage flex models to provide realistic cockpit accelerations
5. Integration of ground model rigid body accelerations and flexible mode accelerations into the simulator motion model
6. Tuning of flexible mode and motion models
7. Testing and validation of the simulator roughness simulation enhancements

This report references two units of measurement for acceleration. G-force (G) is used for the simulator cockpit accelerometer output and for the ProFAA simulated cockpit acceleration depicted in Figures 1 and 2 in this section and Meters per second squared (m/s^2) are used for references to acceleration ISO indices in all other sections.

Overview of Boeing 737-800 flight simulator flight and motion systems

The Boeing 737-800 CAE full flight simulator is an FAA certified Level D flight training device, providing a six-degree of freedom motion system, high resolution visual display, and sound systems.

The Boeing 737-800 simulator flight model runs at an iteration rate of 60 Hertz (Hz) and provides a simulation of the aircraft equations of motion and interaction with the ground and air. The flight model assumes a rigid aircraft and implements buffets due to airframe flexing only in specific instances where needed for realism. The flight model sends aircraft linear and rotational velocities and accelerations to the simulator visual and motion systems.

The motion system provides motion cues for aircraft maneuvers. Because the motion system actuators provide limited travel, sustained low frequency accelerations are not possible and the motion software is optimized using a complex set of filters to provide cueing for the most critical aircraft training maneuvers such as takeoffs, aborted takeoffs, and landings.

Integration of real world surface elevation profiles

The real world profiles consist of airport surface elevation changes along the longitudinal axis of the airport runway or taxiway. The elevation profiles are two-dimensional; height varies only with respect to x-distance along the runway.

The Boeing 737-800 simulator flight model runs at a rate of sixty times per second. The highest ground speed for a Boeing 737-800 is approximately 150 knots or 253 feet per second (fps). At

this speed, the flight model responds to a change in surface elevation every 4.22 feet (253fps / 60 seconds (s)) along the surface profile. A surface profile sample spacing of four feet was initially chosen to match the sample spacing with the simulator response rate at the highest anticipated ground speed.

During development of the preliminary and final studies, ground speeds of 20 knots and 100 knots were chosen for the taxiway and runway scenarios respectively. Testing showed that providing a profile resolution higher than the simulator sample rate resulted in increased cockpit accelerations.

For taxiway scenario movement at 20 knots (34 feet per second) the profile height is sampled by the roughness model software every 0.56 feet (34 feet/60 Hz) along the profile. Runway scenarios at 100 knots (168 feet per second) provide profile height sampling every 2.8 feet (168 feet/60 Hz).

In order to improve the cockpit acceleration response, the profile sample spacing was changed to 0.4 feet for taxiway profiles and 2.0 feet for runway profiles.

These simulator profile sample spacing's were chosen to provide height data at higher resolution than the simulator's sample interval while keeping the profile data file size within simulator input limits.

Integration of surface profiles into the simulator flight model

The Boeing 737-800 simulator software was modified to allow selection and input of surface profiles into the simulator run-time software. A simulator instructor station control page was created to facilitate profile selection and control. Linear interpolation was used to calculate profile elevation values between data points. Routines were created for aligning the surface elevation profiles with the flight simulator reposition runway and a parallel taxiway. The surface profile elevation data was integrated into the flight model with individual gear height modeled as a function of the gear's position along the elevation profile. When the simulated aircraft moves along the profile, the aircraft tire and landing gear strut models react to changes in surface height and generate strut forces. The strut forces are input into the flight model equations of motion generating linear and rotational velocities and accelerations at the aircraft center of gravity.

Initial evaluation of the Boeing 737-800 simulator response to roughness profiles

The simulator response to surface roughness was evaluated both objectively and subjectively. Test scenarios were created to move the aircraft along the surface profile at a fixed speed and record test parameters. An accelerometer [Appendix G] mounted below the pilot's seat provided vertical acceleration data. Subjective assessments of the simulator response were made by engineers and two industry pilots. The Boeing 737-800 cockpit vertical accelerations were compared with the predicted accelerations generated by the ProFAA airport pavement simulation program [4].

Initial testing showed the Boeing 737-800 flight simulator motion and visual both responded to the surface roughness profiles; however, the cockpit accelerations were lower in intensity than those predicted by the ProFAA program. Subjectively, the cockpit accelerations felt somewhat soft and lacking the higher frequency vibrations associated with travel on a rough airport surface. Because the Boeing 737-800 flight model only calculated rigid body accelerations, the higher frequency vibrations resulting from body and wing flexing were missing from the simulated cockpit accelerations.

Additional Enhancements to the Boeing 737-800 roughness model

1. Addition of Aircraft Body and Wing Flexing Model

   In order to provide realistic cockpit vibrations in response to surface roughness, a flexible mode simulation was developed to simulate the cockpit vibrations caused by flexing of the aircraft fuselage due to surface roughness. The flexible mode model provided by the FAA uses strut force as input to excite the bending mode accelerations. The model outputs linear bending mode positions, velocities, and accelerations at five modal positions:

   Nose Gear, Left Main Gear, Right Main Gear, Center of Gravity (CG) and Cockpit

   The flexible mode model was implemented in the simulator flight software. Modeling of up to four bending modes was provided, with the number of active bending modes selectable from the instructor station. Cockpit vertical accelerations were calculated using the following formula:

   GCP = (VWGD - VQD * VXXM(1) + ModePosZAccel(5)) / 32.2

VWGD	=	Z-body acceleration [ft/s^2]
VQD	=	Pitch acceleration body axis [rad/s^2]
VXXM(1)	=	X-body distance of nose gear from CG [ft]
ModePosZAccel(5)	=	Flex mode vertical acceleration at cockpit [ft/s^2]

   The simulator flight model transfers only the CG (not cockpit) accelerations to the motion system. The motion software transforms the CG accelerations into cockpit accelerations. In order to send the cockpit flexible mode data to the motion system, the cockpit vertical accelerations were first translated into pitch accelerations at the CG, and then added to the flight model rigid body pitch velocity and acceleration outputs. An accelerometer mounted beneath the simulator cockpit provided measures of actual cockpit acceleration for this study. The addition of the flexible mode simulation provided realistic levels of cockpit acceleration intensity and frequency as compared with the B727 cockpit accelerations predicted by the FAA's ProFAA simulation as shown in Figures 1 and 2.

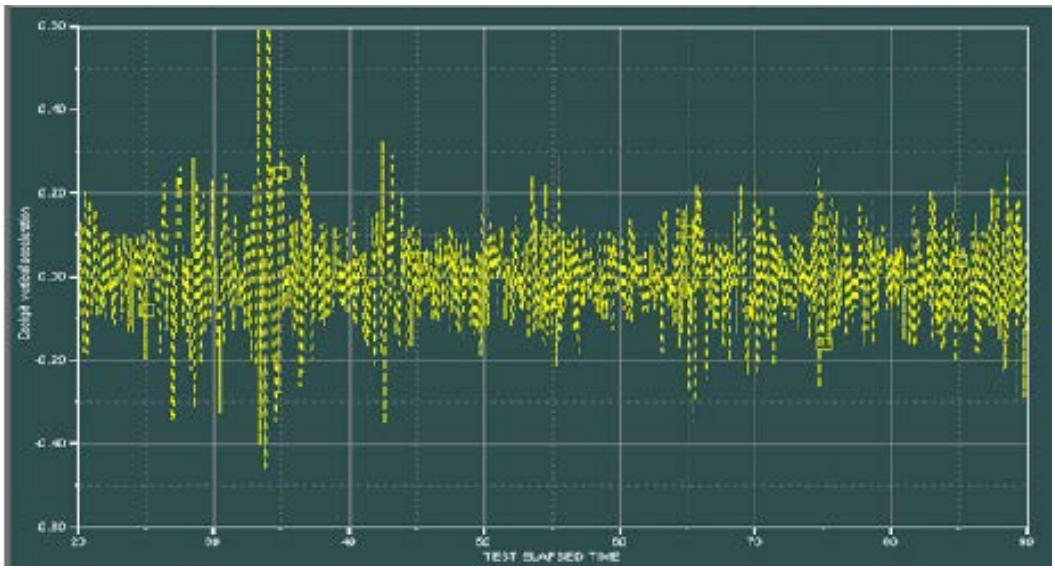


Figure 1. Boeing 737-800 Simulator – Cockpit Accelerometer (G) Vs. Time (s)

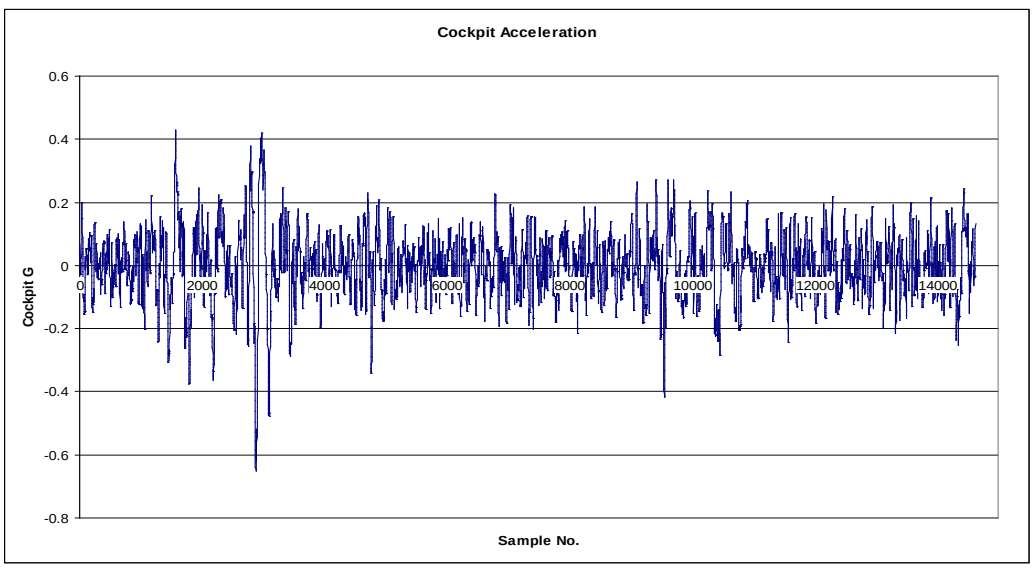


Figure 2. ProFAA 727 Simulation - Cockpit Vertical Acceleration (G) Vs. Time (s)

2. Motion Filter Tuning

The motion system filters the flight model accelerations in order to optimize the motion response within its limited motion range. The filters are tuned to enhance the realism of maneuvers critical for flight training such as takeoff rotation, aborted takeoffs, and deceleration after touchdown. High pass filters are used to limit low frequency motions and maintain the motion actuators near their mid-range. The cockpit vertical response to surface roughness was enhanced by adjusting the vertical high pass filter. The normal settings for the vertical high pass filter are a breakpoint of 2.5 rad/sec (0.4 Hz) with a gain

of 0.7. The breakpoint frequency was decreased to 0.1 rad/sec and the gain increased to 1.0.

Existing generic roughness model

The Boeing 737-800 simulator contains a generic surface roughness model providing five levels of random surface roughness. The generic roughness provides simulation of a random series of surface spalls and mats. The higher order cockpit vibrations associated with the generic roughness are modeled on a spectral analysis of real world aircraft vibrations associated with taxiing at 30 knots on a rough runway. The amplitude of the simulated vibrations is modulated with respect to the simulated aircraft ground speed. Figure 3 provides an example of the surface height variations generated by the generic roughness model.

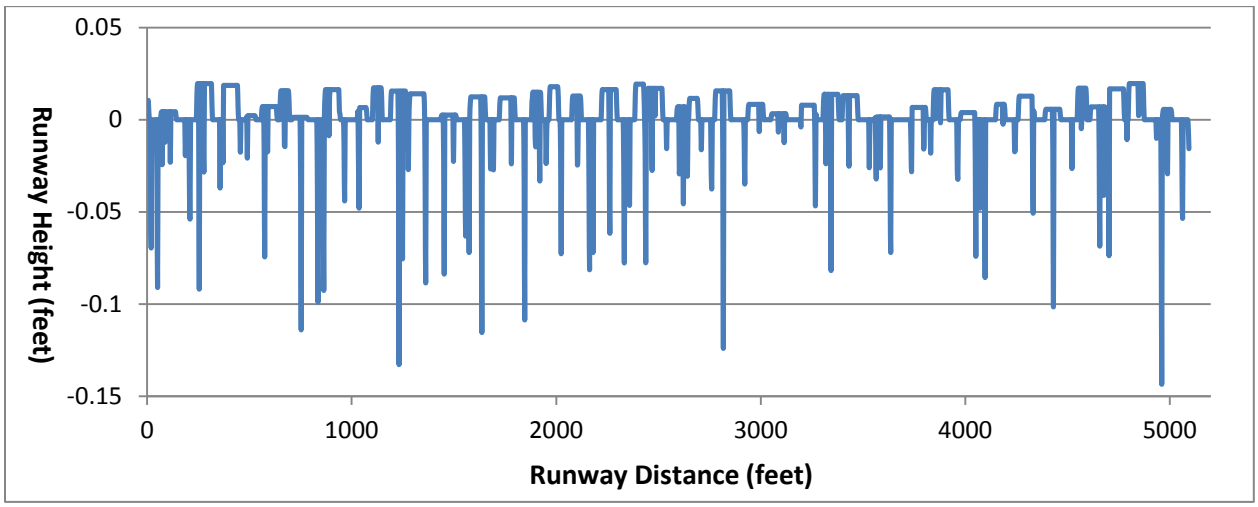


Figure 3. Example of Boeing 737-800 simulator generic roughness surface height variations.

The following are some differences between the simulator generic and real world roughness simulation:

1. The real world roughness profiles provide a range of types of surface roughness (such as localized surface deterioration and crests due to crossing runways/taxiways) and provide examples of isolated discrete jolts. The generic profiles are limited to simulation of spalls and mats with a uniform general intensity of roughness along the length of the profile.

2. Because the real world profile roughness is uniform across the transverse width of the surface, no lateral cockpit accelerations are produced. However, the generic roughness simulation models the random roughness separately for the left and right main landing gear and produces lateral cockpit accelerations.

3. The real world roughness simulation models the cockpit accelerations due to flexing of the aircraft body in response to gear forces from moving along real world roughness profiles. The generic roughness simulates aircraft body flexing effects through the addition of motion buffets with buffet intensity based on aircraft speed and does not simulate vibrations from isolated jolts.

2.2 REAL WORLD ROUGHNESS PROFILES

Real world airport taxiway and runway profile data was obtained from taxiway and runway surface profiles provided by the FAA, Boeing, and Airbus. Appendix B provides a list profiles used for the Boeing 737-800 roughness study. Figure 4 shows an example of surface profiling equipment used by the FAA to capture the airport surface profiles. The three sensors used in the depicted profiling device are:

- Vehicle Elevation; Allied Signal QA700 Accelerometer
- Vehicle-to Pavement Distance; Selcom 2207 Optocator Laser Sensor
- Traveled Distance; Datron DLS-2 Optical Speed and Distance Sensor

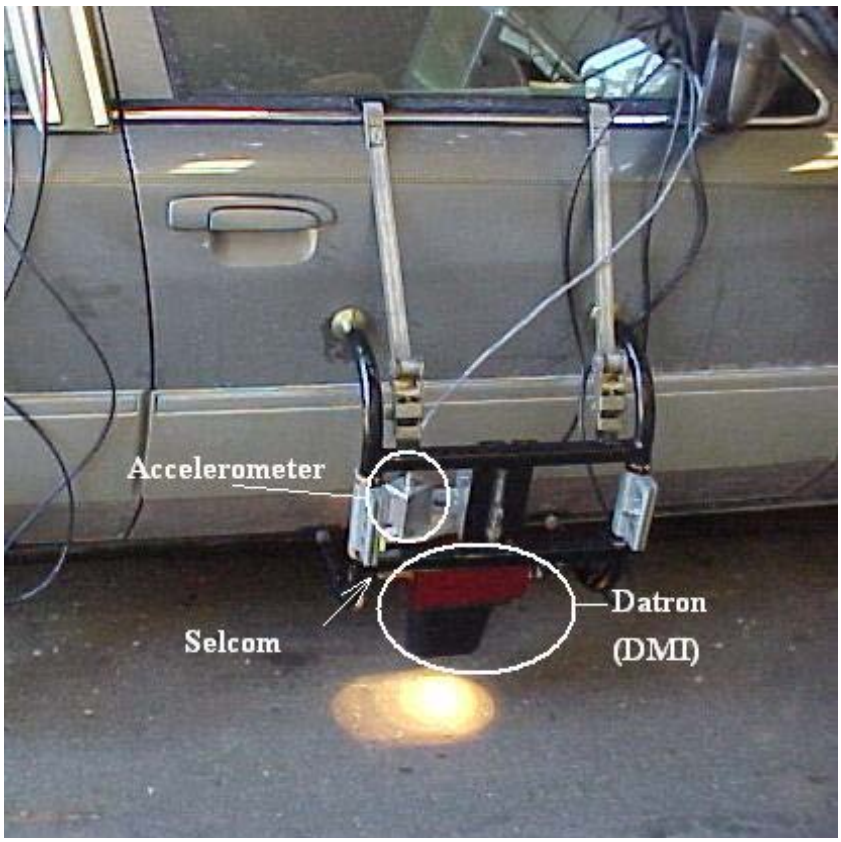


Figure 4. FAA surface profiling equipment

The essential element of an inertial profiling device, which makes the technique feasible, is a high-quality accelerometer. In essence, the accelerometer is the hardware for a single-axis inertial navigation system and is used to measure the absolute vertical position of a point on the test vehicle (the vertical position relative to an inertial reference). The accelerometer is mounted on the test vehicle with its sensitive axis aligned in the vertical direction. Vertical position is computed by double integrating the accelerometer output signal. Long-term drift errors are removed by high-pass filtering, or by other means of compensation. The distance from the accelerometer mounting point to the surface of the pavement is measured with a displacement measurement sensor. The combination of the two measurements then gives the absolute elevation of the pavement surface. Distance traveled by the test vehicle along the pavement is measured with a speed sensor, or a direct-reading distance traveled sensor. A continuous longitudinal elevation profile is therefore measured relative to an inertial reference (with a slowly moving datum if the accelerometer signal has been high-pass filtered).

An FAA designed & manufactured control box is used to integrate the signals from the three sensors mounted on a frame that is temporarily affixed to the data collection vehicle. A standard laptop (notebook) computer is used to collect the raw data from the sensors, to compute the profile from the raw data, and to compute the indices used for the evaluation of pavement condition using FAA software.

For this study, a set of profiles was chosen to provide a wide range of roughness levels and to provide examples of both uniform roughness (consistent roughness intensity along the profile) and non-uniform roughness (variations in roughness intensity along the profile). Appendix B provides a list of the surface profiles selected for this study.

The real world profiles were formatted as follows for integration with the Boeing 737-800 flight simulator ground model:

1. Selected a subsection of each profile with proper length for simulator scenarios:
   - Taxiway scenarios: 1000 foot profile length for 30 second duration at 20 knots
   - Runway scenarios: 5100 foot profile length for 30 second duration at 100 knots
2. Filtered the profiles to remove low frequency variations in height using a 1000 feet. cutoff high-pass filter
3. Changed the profile height units from inches to feet to match the simulator flight model units.
4. Changed the profile height sample spacing to align with the flight simulator's sample rate.
   - Taxiway profiles: 0.4 foot sample spacing
   - Runway profiles: 2.0 foot sample spacing
5. Added a three second smooth lead-in to each simulator profile

Figure 5 provides an example of flight simulator surface height variations generated by a real world roughness profile.

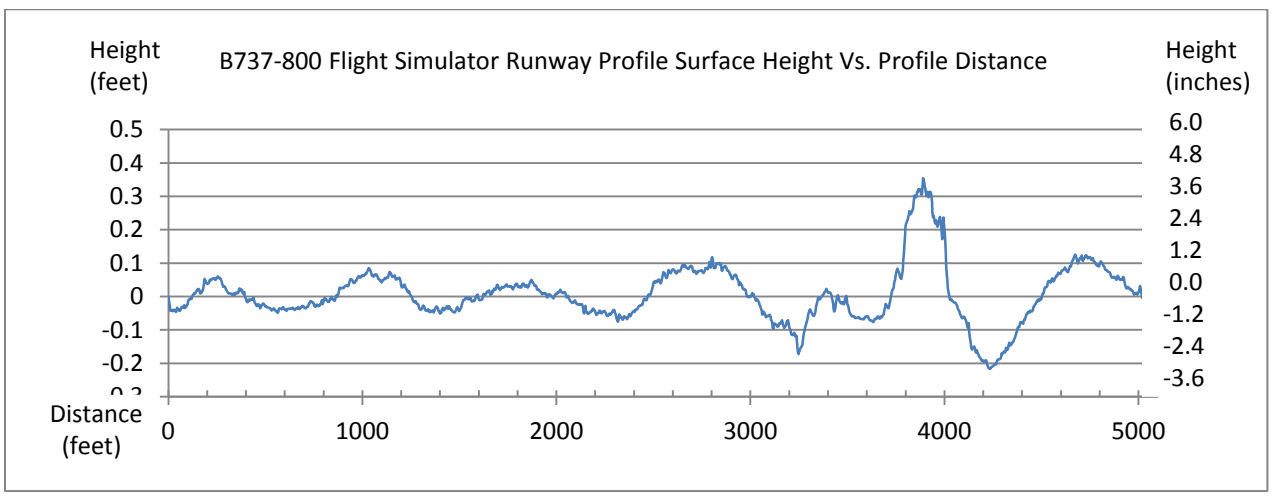


Figure 5. Example of flight simulator surface height variations for a real world runway surface profile. Profile height is shown in both simulator units (feet) and original units (inches)

Additional details for real world profile formatting are provided in Appendix C.

2.3 GENERIC PROFILES

Twenty generic roughness taxiway and twenty generic roughness runway profiles were included in the preliminary roughness study test scenarios for comparison with the real world roughness model. [Ref. 7] The intensity levels of the standard simulator generic roughness were not high enough to obtain unacceptable ratings from pilots; therefore rougher generic roughness profiles were created for the study by increasing the intensity of the generic roughness.

In keeping with the roughness study objectives, the final roughness study focused on pilot ratings of real world surface roughness profiles. For the final study the number of generic profiles was reduced to three taxiways and three runways to allow increasing the number of real world profiles to thirty-seven each for taxiways and runways.

2.4 TEST SCENARIOS

A methodology was developed for presenting surface roughness profiles for evaluation and obtaining pilot roughness ratings and objective simulator data. Test scenarios, roughness rating forms and pilot briefings were developed with input from human factors specialists, and refined during a series of pilot studies in 2010 and 2011. Early scenario designs explored the following variations:

- Length of test scenarios
- Speed of aircraft movement
- On-ground movement only vs. takeoff and landing movement
- Pilot-controlled vs. automated scenarios

National Cooperative Highway Research Program (NCHRP) highway rideability studies [Ref. 1] were reviewed, leading to the following scenario design for the preliminary and final studies:

Pilots were presented with a set of 40 taxiway and 40 runway scenarios providing a range of surface roughness.
Scenarios were designed to provide automated movement along the taxiway or runway for 30 seconds. The takeoff and landing scenarios were rejected because of the length of time required (approximately two minutes for takeoff and three minutes for each landing scenario).

The pilots assumed the role of a non-flying pilot or observer during the scenarios with no input to the flight controls, allowing the pilots to focus their full attention on assessing the ride quality. Using pilots in a ‘non-flying’ role allowed placement of a third rater in the cockpit (seated in the observer’s seat just behind the pilots) as shown in Figure 6, allowing collection of ride quality ratings from three subjects simultaneously.

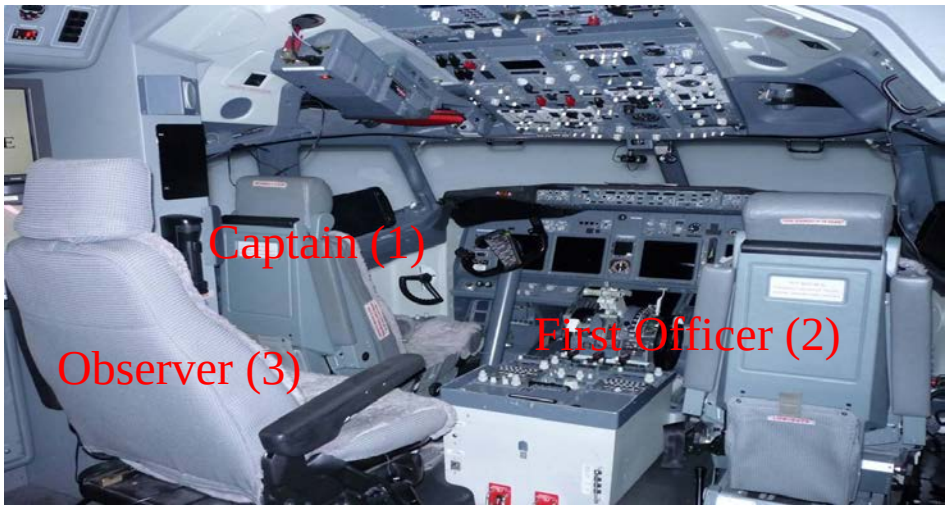


Figure 6. Boeing 737-800 Simulator cockpit seating positions

A statistical analysis of preliminary study ratings showed no significant differences with regards to seat position. The matrices of correlation coefficients for the pilots' average numerical ratings of taxiways and runways were computed from the pilots' responses and found to be

Table 1. Pilot Taxiway Correlation Coefficients

Taxiway Correlation	Captain	1st Officer	Observer
Captain	1.000	0.989	0.991
1st Officer	0.989	1.000	0.990
Observer	0.991	0.990	1.000

and

Table 2. Pilot Runway Correlation Coefficients

Runway Correlation	Captain	1st Officer	Observer
Captain	1.000	0.985	0.990
1st Officer	0.985	1.000	0.987
Observer	0.990	0.987	1.000

respectively. A correlation coefficient near 1.0 (100%) indicates high correlation, with 1.0 indicating perfect correlation.

Since average numerical ratings from pilots in different seats were about 99% correlated it was concluded that each pilot received a similar experience regardless of seat location. (See Figure 6)

Pilot raters were presented with taxiway and runway profiles providing a range of roughness from very smooth to very rough as well as examples of both uniform and non-uniform pavement roughness. A standardized roughness rating form was used for capturing pilot ratings of rideability.

For the preliminary study, a total of eighty scenarios were provided in the following categories:

- 20 Real world taxiway scenarios at 20 knots
- 20 Real world runway scenarios at 100 knots
- 20 Generic taxiway scenarios at 20 knots
- 20 Generic runway scenarios at 100 knots

For the final roughness study, the number of real world scenarios was increased to allow presentation of a wider range of real world roughness. Additional real world profiles were selected to provide examples of specific types of non-uniform roughness such as crests occurring at intersecting runways and taxiways. A total of eighty final study scenarios were provided in the following categories:

- 37 Real world taxiway scenarios at 20 knots
- 37 Real world runway scenarios at 100 knots
- 3 Generic taxiway scenarios at 20 knots
- 3 Generic runway scenarios at 100 knots

2.5 RATING FORM DEVELOPMENT

The roughness rating form was developed for obtaining pilot subjective ratings of airport surface rideability as well as evaluating need for improvement. The roughness rating form was modeled after the American Society for Testing and Materials (ASTM) rating form [Ref 1] and provides the following inputs:

Rideability Level	Numeric value from 0 to 10 indicating the ride roughness with 10 = perfectly smooth and 0 = impassable
Need-for-Improvement	Checkbox selection indicating the need for surface improvement

The Preliminary Study rating form (shown in Figure 7) provided the following Need-for-Improvement selections:

- Acceptable: Ride Quality Does Not Need Improvement
- Uncomfortable: Recommend Ride Quality Improvement
- Unacceptable: Ride Quality Must Be Improved

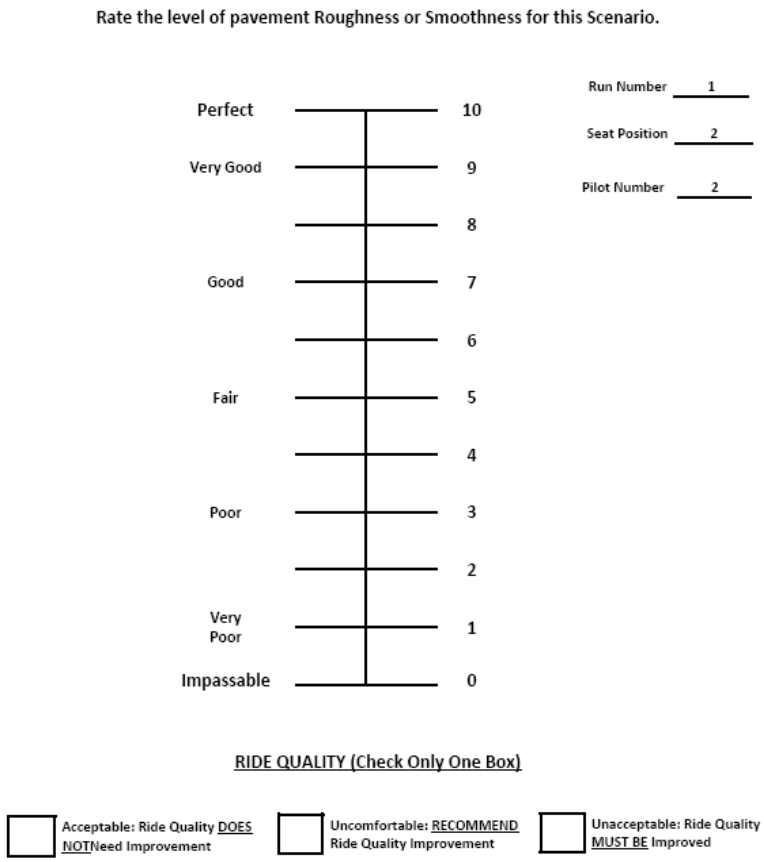


Figure 7. Preliminary Study Rideability Rating Form

Analysis of preliminary study data showed that the subjective pilot ratings of profile roughness as compared to their corresponding selection of need-for-improvement (acceptable, uncomfortable, and unacceptable) were not consistent. There was a wide range of ride quality ratings within each of the need-for-improvement categories and therefore overlapping ranges between categories.

For the final study, the pilot rating form was modified to simplify the need-for improvement ratings and to align more closely with the ASTM rating form [Ref 1]. The number of need-for-improvement checkboxes was changed from three (acceptable, uncomfortable, unacceptable) to two (acceptable, unacceptable). Additionally the need-for-improvement section label was changed from “Ride Quality” to “Need For Improvement” for clarity. The updated final study pilot rating form is shown in Figure 8 and Appendix K.

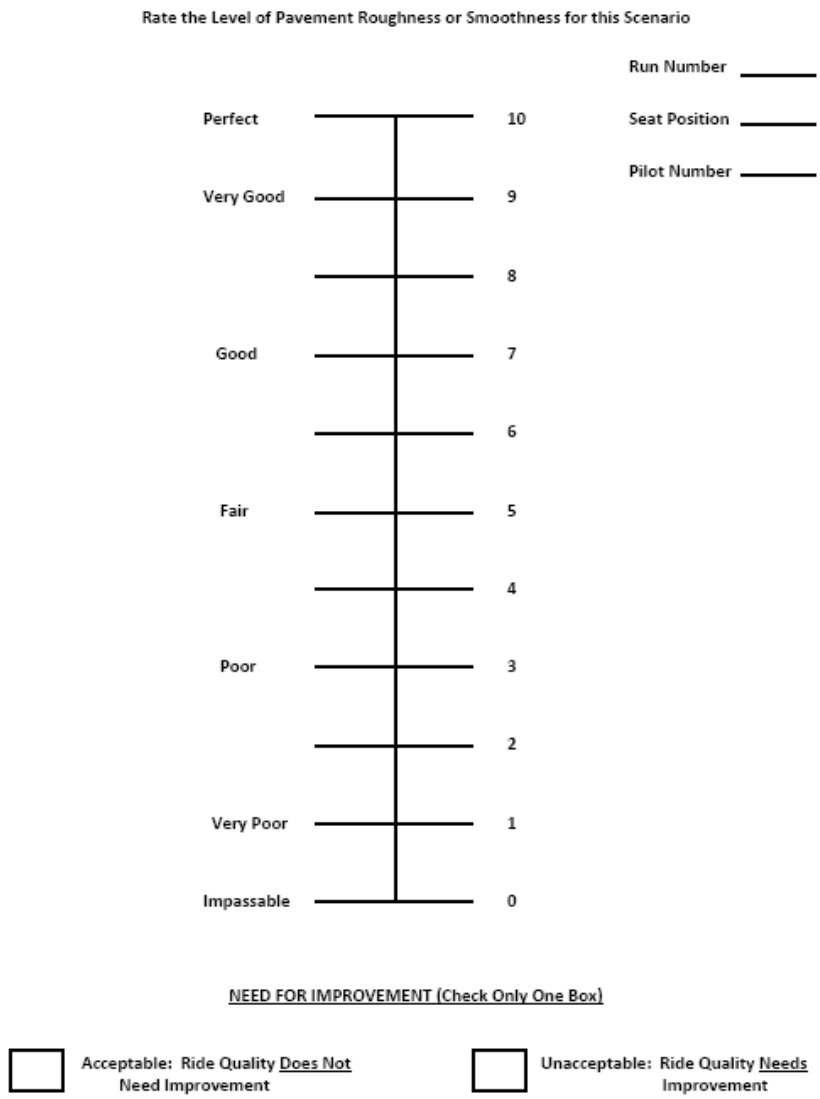


Figure 8. Final Study Rideability Rating Form

2.6 PILOT BRIEFINGS

The final study included pre-brief and post-brief sessions. The Microsoft PowerPoint presentations for the briefings are shown in Appendices H and I. The pre-brief session provided background information about the study, explained the simulator testing process, and defined the rating scale limits. The pre-brief information provided study background information and an explanation of the study test sessions. The roughness rating form instructions were read from a script to each group of pilots to ensure that the same instructions were provided to each group. A post-brief session was conducted immediately following each test session and to capture pilot feedback on the fidelity of the roughness simulation, recommendations on methods for improving the study, and perceptions of airport surface roughness in the real world.

2.7 PILOT RATING PANEL SELECTION

Thirty six pilots were recruited to participate in the final study. The original plan was to utilize commercial airline pilots with current experience in Boeing 737-800 aircraft and no previous experience in the FAA's pavement roughness studies. However, due to the limited availability of commercial Boeing 737 pilots it was necessary to include some pilots with prior study experience and also some military KC-135 pilots. The KC-135 military pilots were selected because the KC-135 and Boeing 737-800 are both Boeing aircraft with similar sized fuselages.

As part of the pilot data collection CCRC also collected the following information about each participant; current airline, current type ratings, hours of operation of each type of aircraft, rank and if they have previous military experience.

3.0 FINAL STUDY TESTING

3.1 ROUGHNESS RATING SESSIONS

For the final study, twelve roughness rating sessions were conducted. The first rating session was used for final tuning of the roughness profile intensities. The profile intensities were decreased slightly after session 1 to provide a more even distribution of roughness. A gain factor was adjusted to decrease the input profile height for each scenario. Appendix B provides a list of the surface profile gain factors. After adjustment, the profile intensities remained consistent for sessions 2 – 12. For each test session, three pilots provided roughness rideability and need-for-improvement ratings for the eighty taxiway and runway profiles. Prior to rating the profiles, two practice scenarios were presented to acquaint the pilots with the scenario format and rating process. Each test scenario presented aircraft movement along a taxiway or runway profile for 30 seconds at a fixed speed of 20 knots for taxiway profiles and 100 knots for runway profiles. Pilots were requested to limit conversations during scenarios and to evaluate each scenario with no discussion.

3.2 COLLECTED SIMULATOR PARAMETERS

Flight simulator parameters were collected to obtain objective measures of the aircraft response to the roughness profiles. Data was collected during each test scenario at a 60 Hz rate and stored in a comma delimited text file. The following parameters provided the primary data used for analysis:

- Elapsed time
- Aircraft distance along roughness profile
- Profile height at aircraft
- Cockpit vertical acceleration

An accelerometer mounted under the cockpit floor provided vertical acceleration levels during the test scenarios. The accelerometer output was captured and processed to provide a numeric average of the cockpit acceleration for correlation with subjective pilot ratings of roughness. Appendix E provides a detailed list of the collected data parameters.

3.3 INITIAL DATA EVALUATION

3.3.1 EVALUATION OF ACCELEROMETER DATA FOR CONSISTENCY

The collected accelerometer data was reviewed for validity and for the consistency of accelerations provided for the eleven test sessions. Figures 9 and 10 show the recorded accelerations for the 37 taxiway and runway scenarios with data for each test session plotted in a different color (Sessions 2 – 12). As shown by the close alignment of the eleven plots, consistent accelerations were provided for the eleven test sessions.

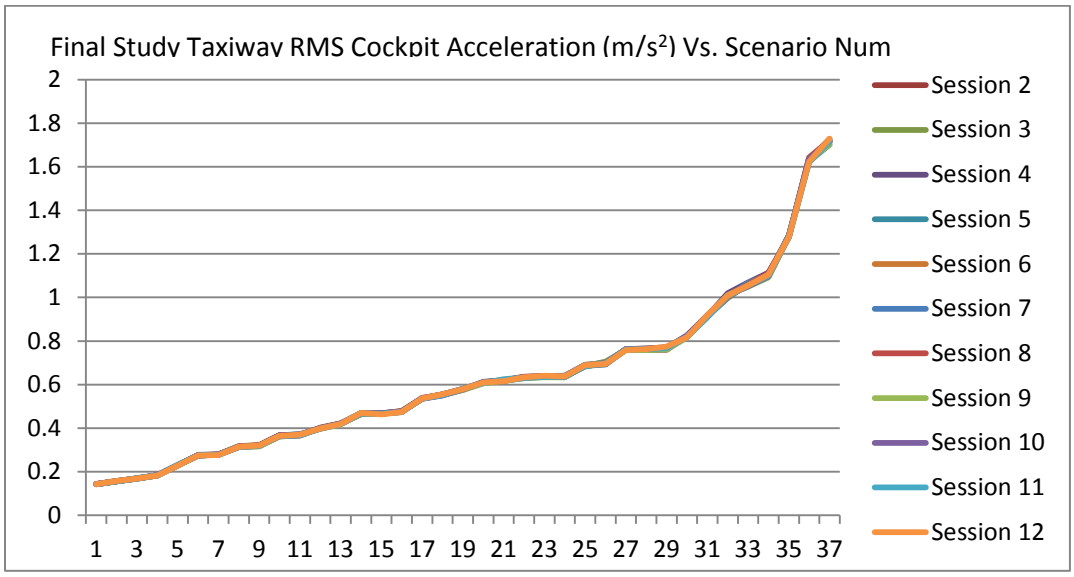


Figure 9. Overlay of Taxiway RMS Cockpit Acceleration for Sessions 2 – 12

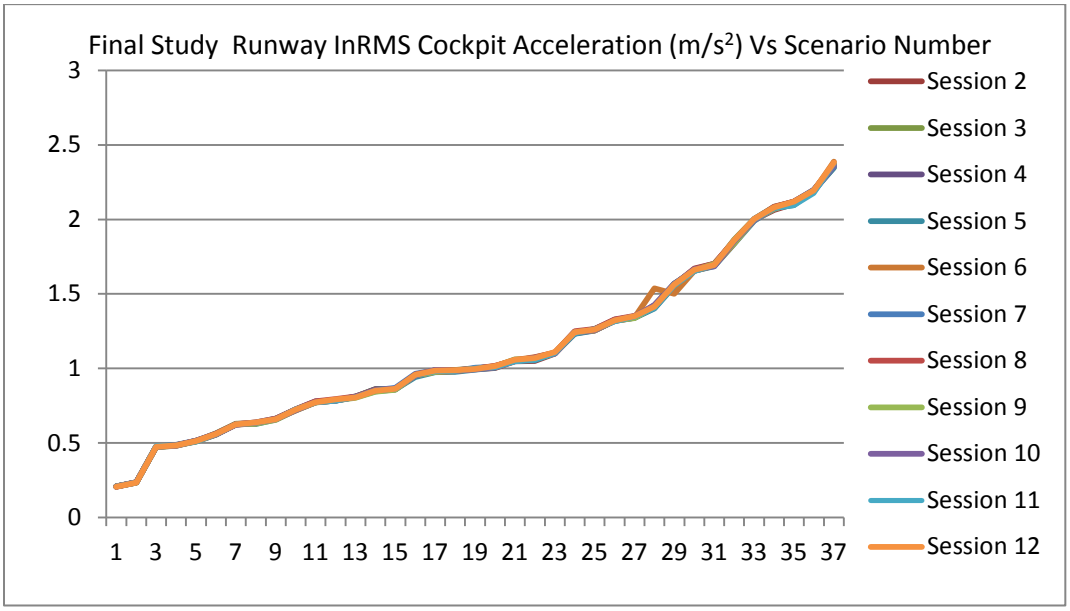


Figure 10. Overlay of Runway RMS Cockpit Acceleration for Sessions 2 – 12

3.3.2 ISO Processing of Accelerometer Data

The requirements of ISO International Standard ISO 2631, "Mechanical vibration and shock – Evaluation of human exposure to whole-body vibration," Part 1: "General Requirements" [Ref. 3] were followed to the extent possible in the collection, processing, and evaluation of the measured cockpit accelerations. The basic evaluation method contained within the standard is in terms of a "weighted root-mean-square (RMS) acceleration" as defined in subclause 6.1 by the equation:

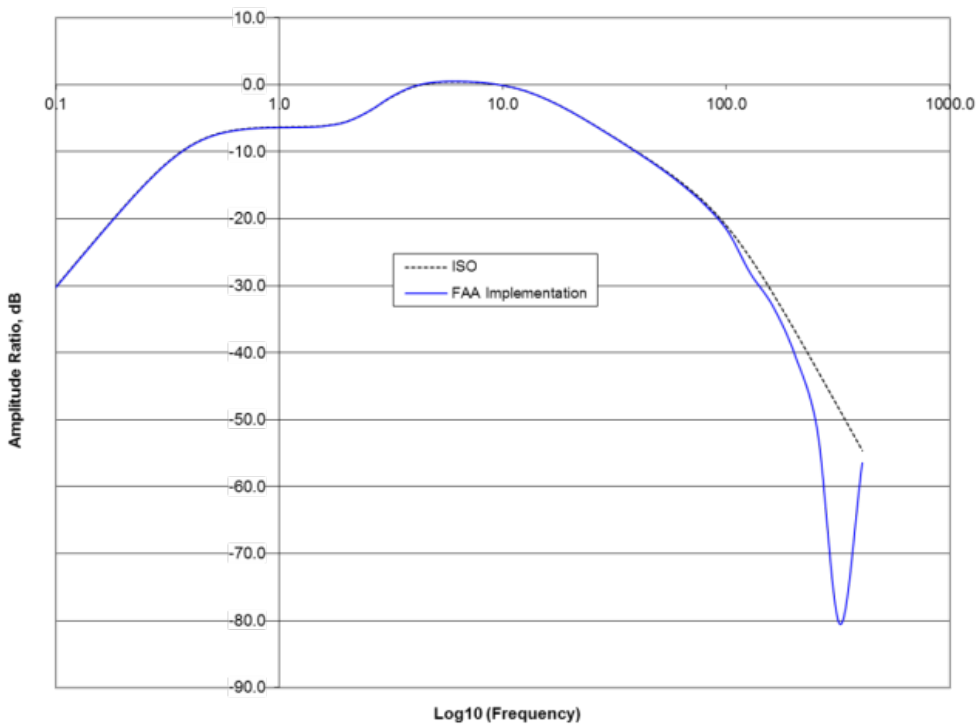
$$a_{\mathrm{W}} = \left[\frac{1}{T}\int_0^T a_{\mathrm{W}}^2(t)\,\mathrm{d}t\right]^{\frac{1}{2}}$$

Where

$a_{\mathrm{W}}(t)$ is the weighted acceleration (translational or rotational) as a function of time (time history), in meters per second squared (m/s^2) or radians per second squared (rad/s^2), respectively;

T is the duration of the measurement, in seconds.

The weighted acceleration ($a_{\mathrm{W}}(t)$) is computed from the measured acceleration by application of a set of transfer functions completely defined in the standard. The transformations have two purposes: 1) "Lower and upper frequency band limitation;" and 2) Normalization of the acceleration amplitude to mimic relative human response over the frequency range of interest. The transfer functions for ISO Weighting (W_{k}) were applied to the measured accelerations in the time domain by the method described in Appendix F. Amplitude ratio versus frequency plots are shown in Figure 11 for the ISO specified transfer functions (computed directly) and the filtering routines used in this project (computed from the ratio of output to input sine wave amplitudes). In this Figure the frequency response (absolute value of the transfer function) of the weighting procedure is compared with the ISO specified frequency response on a linear scale then and in decibels (ten times the logarithm of the response). Calculation of the transfer function is detailed in Appendix F.



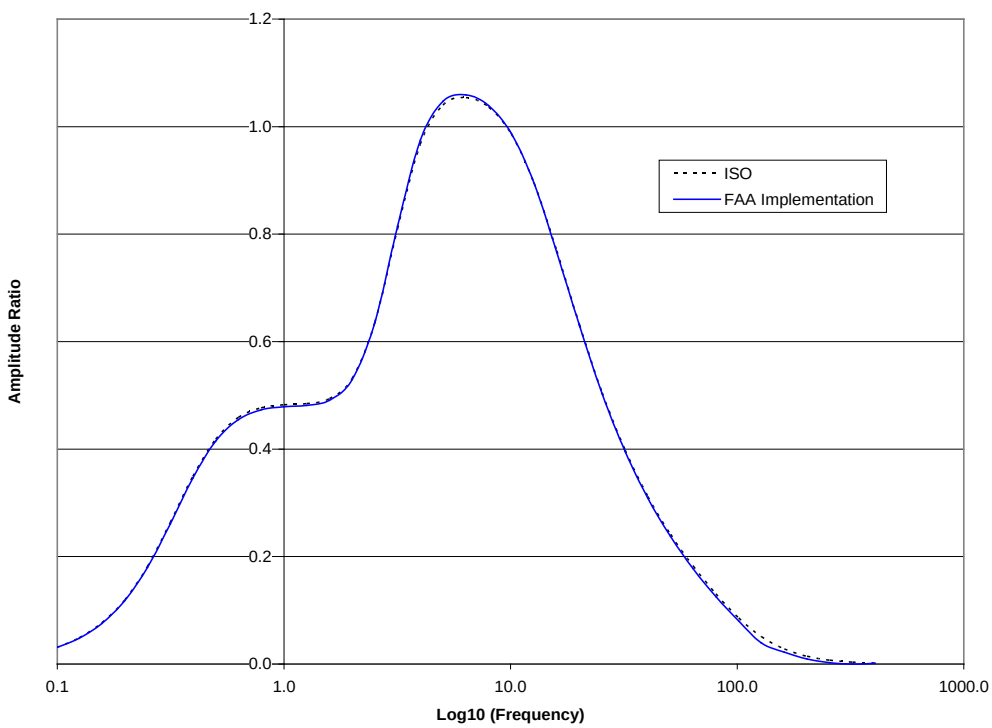


Figure 11. Frequency response of the weighting procedure compared with the ISO specified frequency transfer function for Weighting W_k. Amplitude ratio plotted on a linear scale, top, and a logarithmic scale, bottom.

Note that the sample rate of the measured cockpit acceleration was 60 Hz and the bandwidth of the measured accelerations necessarily limited to 30 Hz. For this, and other reasons, the sample rate of the measured accelerations was increased to 1,280 Hz by cubic spline interpolation before applying the weighting functions. (The cubic spline fit of data is a piecewise fit of third degree or lower polynomials that interpolates the data points and is smooth to the extent of having continuously changing first and second derivatives; that is, continuously changing slope and concavity. A fit that was less smooth would artificially introduce higher frequency accelerations that were not in the original data. From this fit the cockpit accelerations were estimated at points in-between the recorded data points in order to estimate amplitudes of vibration at higher frequencies than 30 Hz. The cubic spline or any interpolation is a necessary drawback of estimating vibrations at higher frequencies than can be sampled.) After applying the weighting functions, the sample rate of the weighted accelerations was decreased by low-pass filtering appropriately, cubic spline fitting, and sub-sampling to 160 Hz. For the results to be compatible, these sample rate transformations should be applied if the weighting and RMS calculations are applied to accelerations from other sources, such as an independent aircraft simulation or an operating aircraft with data rates running at higher than 60 Hz.

3.3.3 Additional Indices

Compared to almost all other public transportation travel-ways, airport pavements, and in particular runways, provide very short exposure times to the vibration environment. Also, within these short exposure times there tend to be a larger than normal number of short-length changes of pavement characteristics in the form of crossing pavements and patching. Adding to this the wide variation of speed that occurs during takeoff and landing means that the assumption of a stationary vibration environment existing within an aircraft is rarely to be expected, particularly on runways. The ISO standard provides for these kinds of conditions by providing two alternative indices which can be computed and reported when the weighted RMS method is suspected to be underestimating the true environment due to “shocks” in the acceleration record.
The ISO standard suggests using crest factor, a measure that shows the ratio of peak values to the average value in waveforms, to determine if the alternative indices should be used. General guidance in the standard is that, in addition to the weighted RMS, one of the alternative indices should be reported when the crest factor of the acceleration record is approximately 9.0 or greater. A third alternative is specified in ISO 2631 Part 5: “Method for evaluation of vibration containing multiple shocks.” The three alternative indices are:

1. Maximum transient vibration value from a running RMS computation, MTVV.
2. Fourth power vibration dose value, VDV.
3. Spinal response acceleration dose, DK.

Together with crest factor, all three of the alternative indices are reported for each of the cockpit acceleration records analyzed. There is not a recommended ISO index because the choice of ISO index or indices may depend upon the particular application. However:

- weighted RMS is the most popular index to indicate the average roughness of a ride in literature and has the most mention in the Standard,

- weighted VDV is sometimes preferred to weighted RMS when a ride is not uniform because occasional shocks contribute more to calculating its value (so it better estimates discomfort due to occasional shocks),
- weighted MTVV is the maximum weighted RMS value calculated over a short time period (one second was used, as recommended by ISO 2631-1) for estimation of maximum short-term discomfort, and
- DKup is a spinal dose indicator for estimation of spinal discomfort.

Complete definitions and the methods used for the computation of the alternative indices are given in Appendix F.

4.0 DATA ANALYSIS

Analysis was performed using the subjective pilot ratings (Section 2.5) and objective ISO index values (Sections 3.3.2 and 3.3.3) for vertical vibrations of the simulated Boeing 737-800 real world taxiway and runway rides compiled from the two test sets, preliminary and final. The following were used:

- 37 real world taxiway and 37 real world runway simulations with ISO index ratings and ratings by 33 pilots from final testing and
- 20 real world taxiway and 20 real world runway simulations with ISO index ratings and ratings by 12 pilots from preliminary testing

Also, three pilots who participated in the fine tuning process for the final tests received 37 slightly different taxiway and runway experiences than the other 33 pilots so their responses are not included in pilot average taxiway and runway ratings but are included in pilot individual response analysis. Since analysis of the preliminary data was published in [Ref. 7] *this report focuses on evaluation of the real world final test data and, where possible, upon analysis of the combined real world final and preliminary data*

The ISO index ratings were generated from accelerometer data for each simulated taxiway and runway. An accelerometer mounted under the cockpit floor provided vertical acceleration levels during the test scenarios. The accelerometer output was captured at a 60 Hz rate and the data was processed to provide a numeric average of the cockpit acceleration for correlation with subjective pilot ratings of roughness. For each taxiway and runway ride the acceleration indices[1] of Section 3.3:

- weighted RMS Value (m/s^2)
- weighted MTVV (Maximum Transient Vibration Value) (m/s^2)

[1] In preliminary tests other ISO indices, called weighted VDM and "DKdown", were also considered, but were dropped as superfluous for final testing analysis.

- weighted VDV (Vibration Dose Value) ($m/s^{1.75}$)
- DKup (Spinal Response Acceleration Dose, or Acceleration Dose Value) (m/s^2)

were calculated. The ISO weighted crest factor index was also computed for determination of instances in which one of the last three indices might be appropriate instead of weighted RMS. The implementation of real world profiles on the flight simulator may result in the loss of short duration jolts due to the simulator's 60 Hz iteration rate and the larger profile sample spacing. For this reason the crest factor index may be less effective for identifying profiles with a high percentage of peak waveforms in the flight simulator roughness implementation as compared to the real world. Except in the case of the Acceleration Dose Value Index, the cockpit acceleration signals were processed with a set of weighting (filter) functions before calculating the index values. The intent of the weighting is to normalize the frequency content of the accelerometer signal to account for the variation of the sensitivity of human responses over the frequency range of interest. The Acceleration Dose Value Index was computed using the raw accelerometer signal. See [Ref. 3] and Appendix F for more details.

The pilot data to be analyzed consisted of a numerical rating from 0 to 10 (Section 2.5), with fractional responses permitted, and an "acceptable" or "unacceptable" rating from each pilot for each taxiway and each runway in the final study. Taxiway and runway data used from the preliminary study consisted of a numerical rating 0-10 from each pilot[2] for each taxiway and each runway.

4.1 ANALYSIS OVERVIEW

Evaluation of the data was based upon

- Consideration of the average pilot numerical ratings of 37 taxiways and 37 runways from final testing as a function of the first four ISO indices,
- Evaluation of the percentage of pilots rating the 37 taxiways and 37 runways in final testing "unacceptable" as a function of the first four ISO indices,
- Discussion of crest factor as an indicator of large shocks,
- Computations using the combined real world numerical ratings of all pilots from both studies on taxiways or runways (33 pilots x 37 taxiways/runways in regular final testing, 12 pilots x 20 taxiways/runways in preliminary testing, and 3 pilots x 37 different taxiways/runways in final testing for a total of 1572 taxiway and 1572 runway numerical ratings) in order to evaluate probabilities of individual numerical pilot responses to taxiways or runways and to verify consistency between the preliminary and final testing.

[2] Pilots in the preliminary study also gave a rating of acceptable, uncomfortable, or unacceptable to each ride but their responses could not be combined with the acceptable, unacceptable data of the final study.

4.2 PILOT NUMERICAL RATINGS AS A FUNCTION OF FOUR ISO ROUGHNESS FACTORS

The average pilot numerical ratings (0-10) of the 37 real world taxiways and 37 real world runways of the final data set were computed and compared with the ISO roughness values: weighted RMS, weighted MTVV, weighted VDV and DKup. (Pilot average ratings from the preliminary data were not included because they were based on averages from a different number of pilots.) Correlation coefficients were evaluated relating pilot numerical ratings and the calculated ISO index values for taxiways and runways and appear in Table 3.

Table 3**.** Correlation Coefficients Relating Subjective Pilot Avg. Ratings to Objective ISO Roughness Indices

	Weighted RMS (m/s^2)	Weighted MTVV (m/s^2)	Weighted VDV ($m/s^{1.75}$)	DKup (m/s^2)
Taxiway Avg. Rating (0-10)	-.960	-.942	-.972	-.951
Runway Avg. Rating (0-10)	-.973	-.962	-.983	-.972

Correlation coefficients are negative because pilot ratings go down as each ISO index rises. These compare favorably with the correlation coefficient used in [1] for highway pavement study in which the subjective rating MPR (Mean Panel Rating[3]) for highway travel was correlated with the objective measure PI (Profile Index, used to calculate ride number) with a correlation coefficient of 0.93.

The high correlation between pilot ratings and ISO indices for taxiways and runways (with correlation coefficient -1.0 indicating perfect correlation between rising ISO index and falling pilot rating) indicates a strong linear trend between pilot ratings and ISO ratings of taxiways and runways, so the subjective pilot average numerical rating can reasonably be estimated as a function of each objective ISO index.

Curve fits of average pilot rating were therefore made as a function of each ISO index and showed that *taxiways and runways had distinct trends* as a function of any index (Figure 12 shows the weighted RMS trends computed by quadratic least squares fits).

[3] Mean panel rating is the average rating over a number of panels in which individual panel ratings have been made using a 0-5 numerical scale).

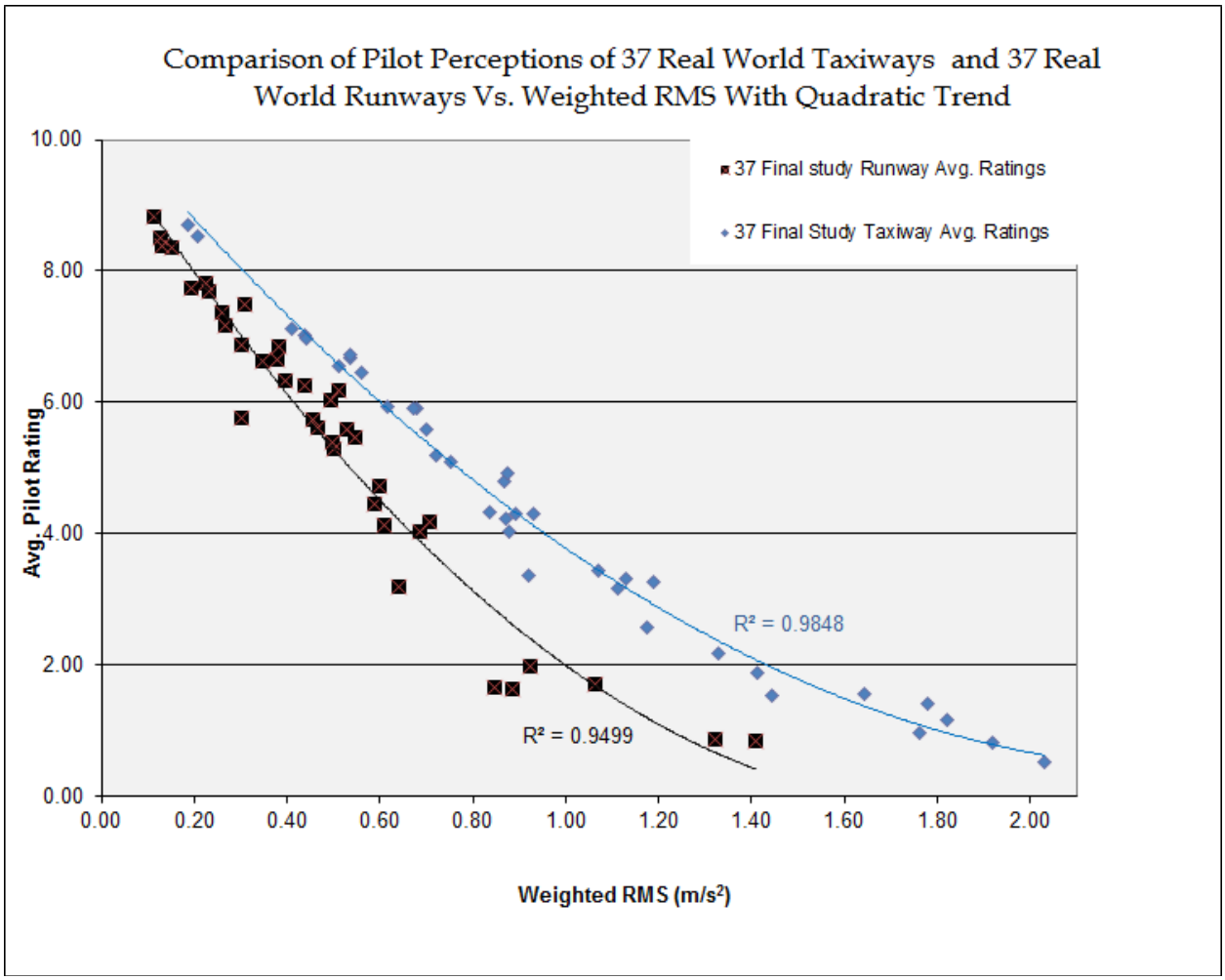


Figure 12. The Distinct Taxiway and Runway Trends for Final Study Real World Pilot Average Numerical Ratings vs. Weighted RMS.

In many applications a simple least squares fit of a line (also called a linear regression) can be applied to data; however, the data trends in Figure 12 (and in other plots to come) show substantial curvature, demanding that at least a quadratic fit is required. A quadratic is the simplest function that can be fit to data that shows curvature and fit the data within the data range; however, the quadratic graph is an arch shape (parabola) so when data is noisy the fit can trend upward on the right before reaching the x-axis (Figure 13). Hence, the quadratic fits should not be extrapolated outside of the input data range.

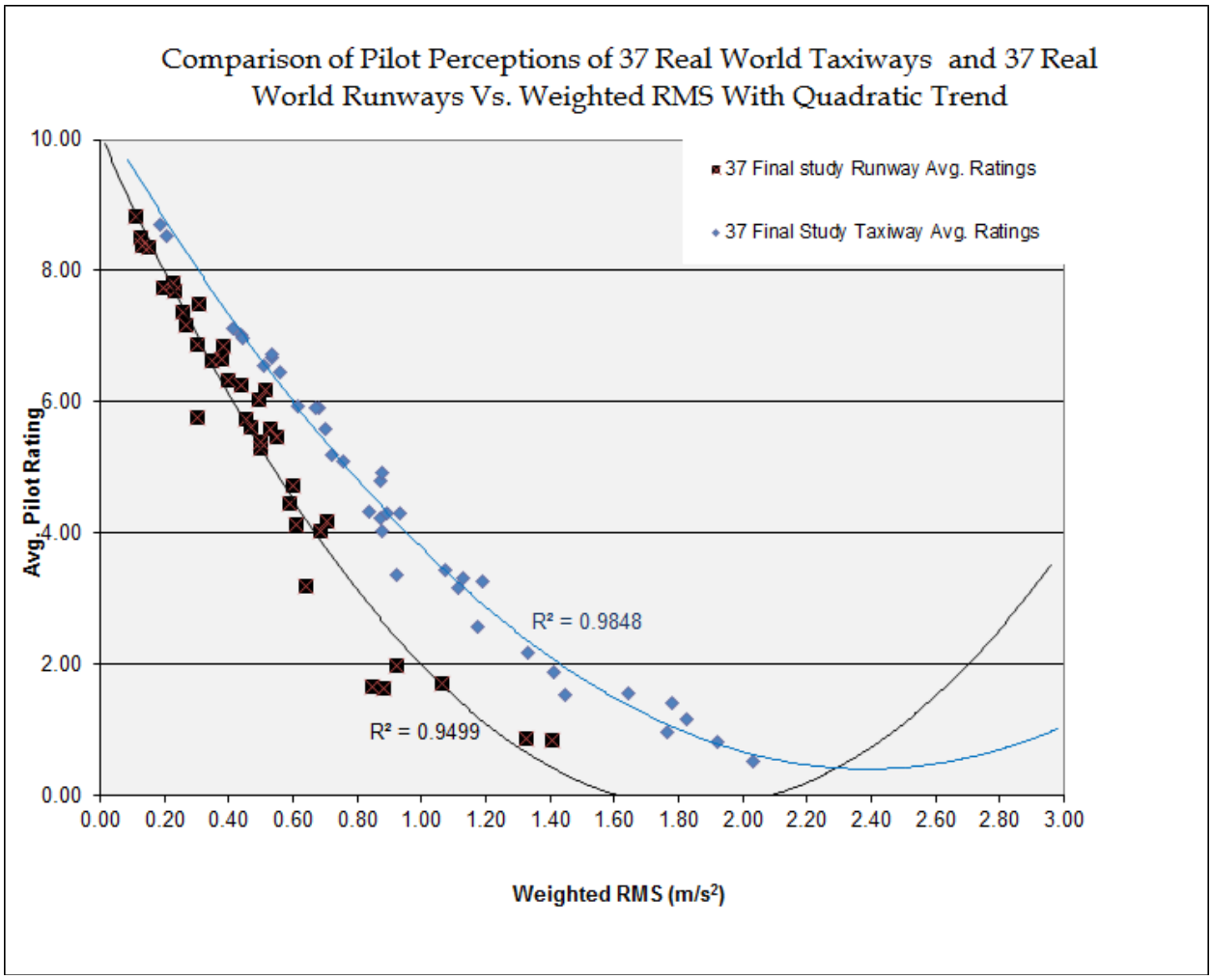


Figure 13. The Problem Extrapolating a Quadratic Fit Outside of the Data Range.

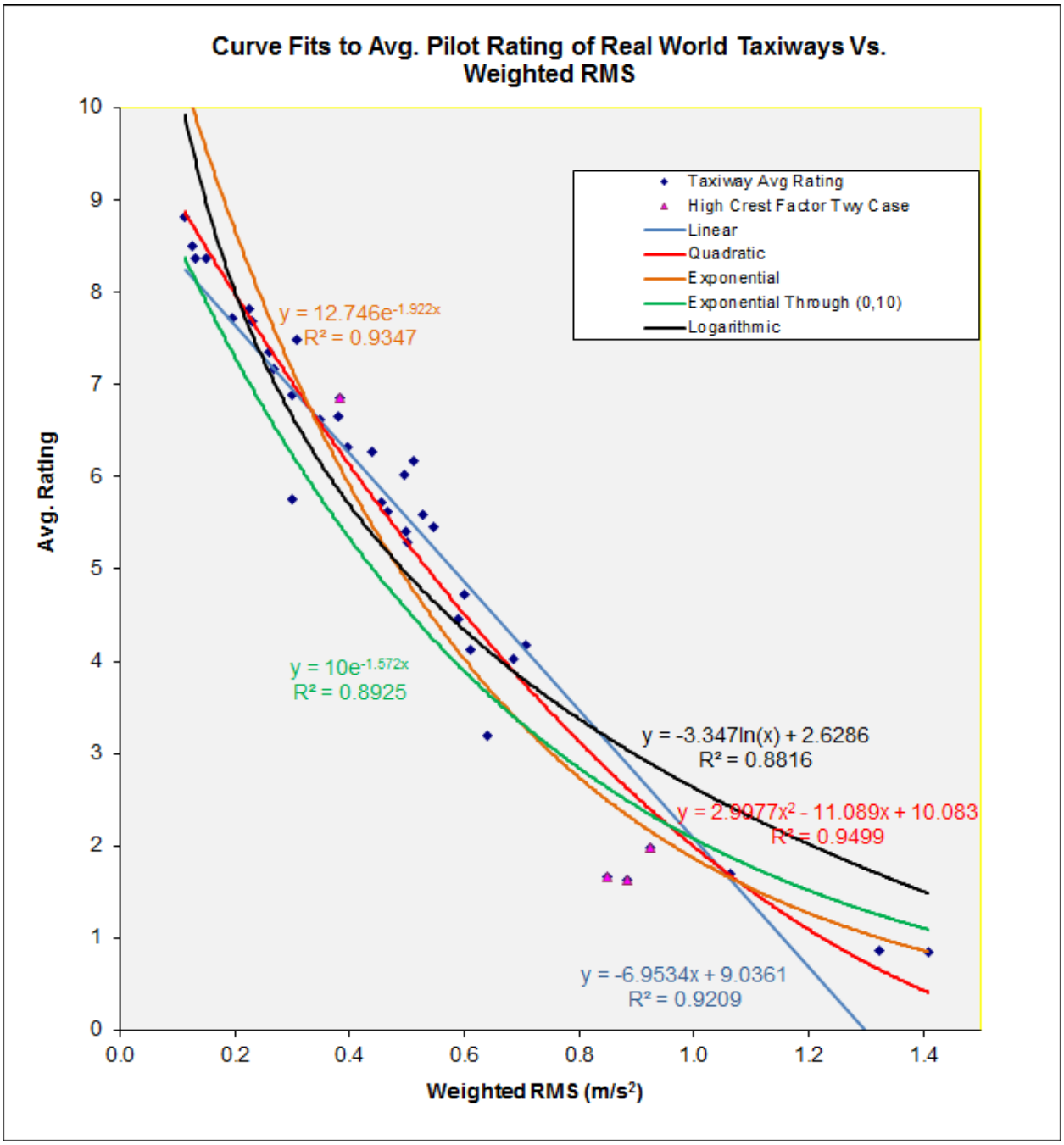
Various types of curve fits were attempted by the least squares method to the average runway ratings y versus each ISO index x, of the final data set:

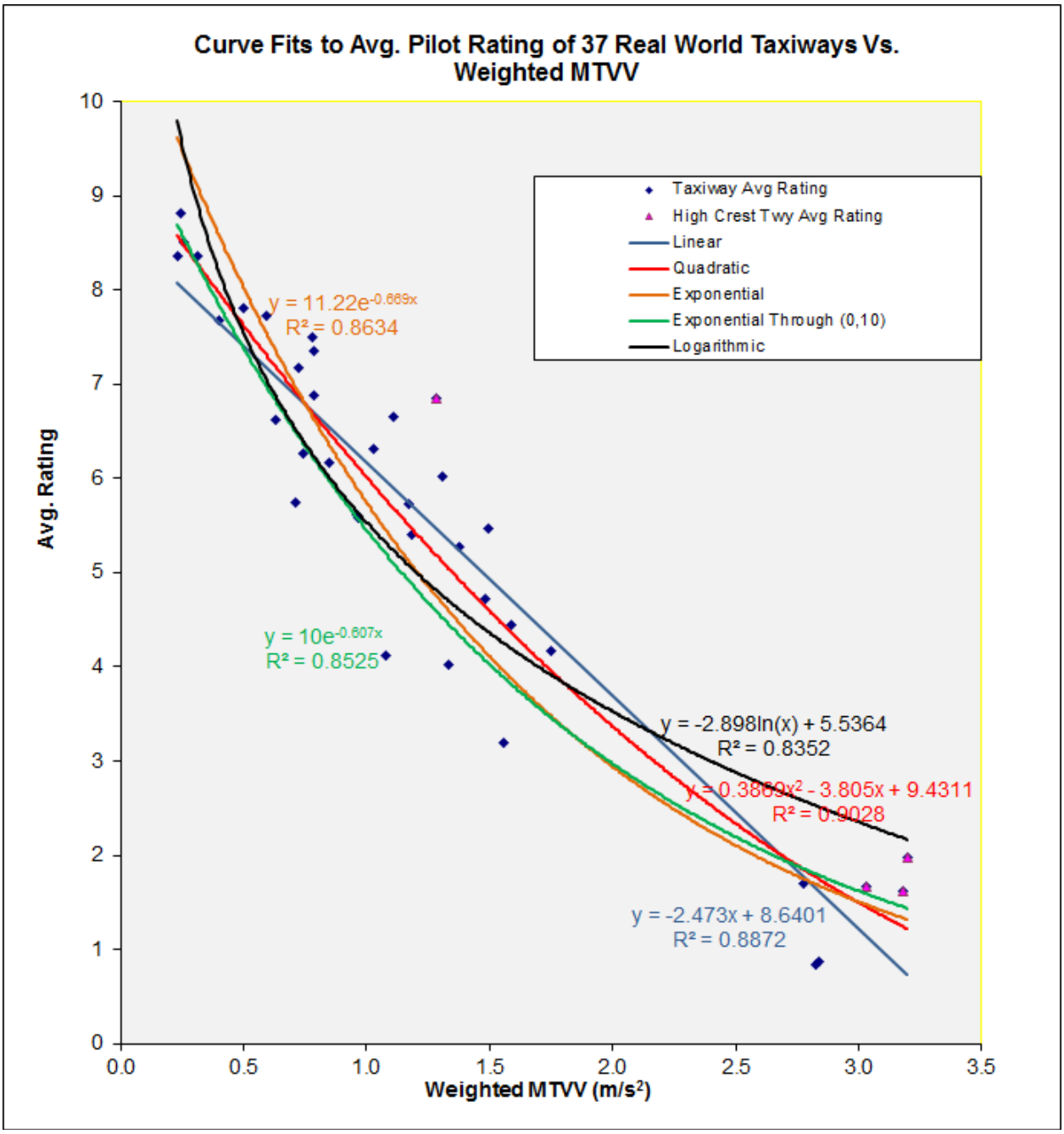
- linear ($y = ax + b$)
- quadratic ($y = ax^2 + bx + c$)
- exponential ($y = ae^{bx}$)
- exponential through (0,10) ($y = 10e^{bx}$)
- logarithmic ($y = a \ln(x) + b$)
- shifted logarithmic ($y = a \ln(x+c) + b$)

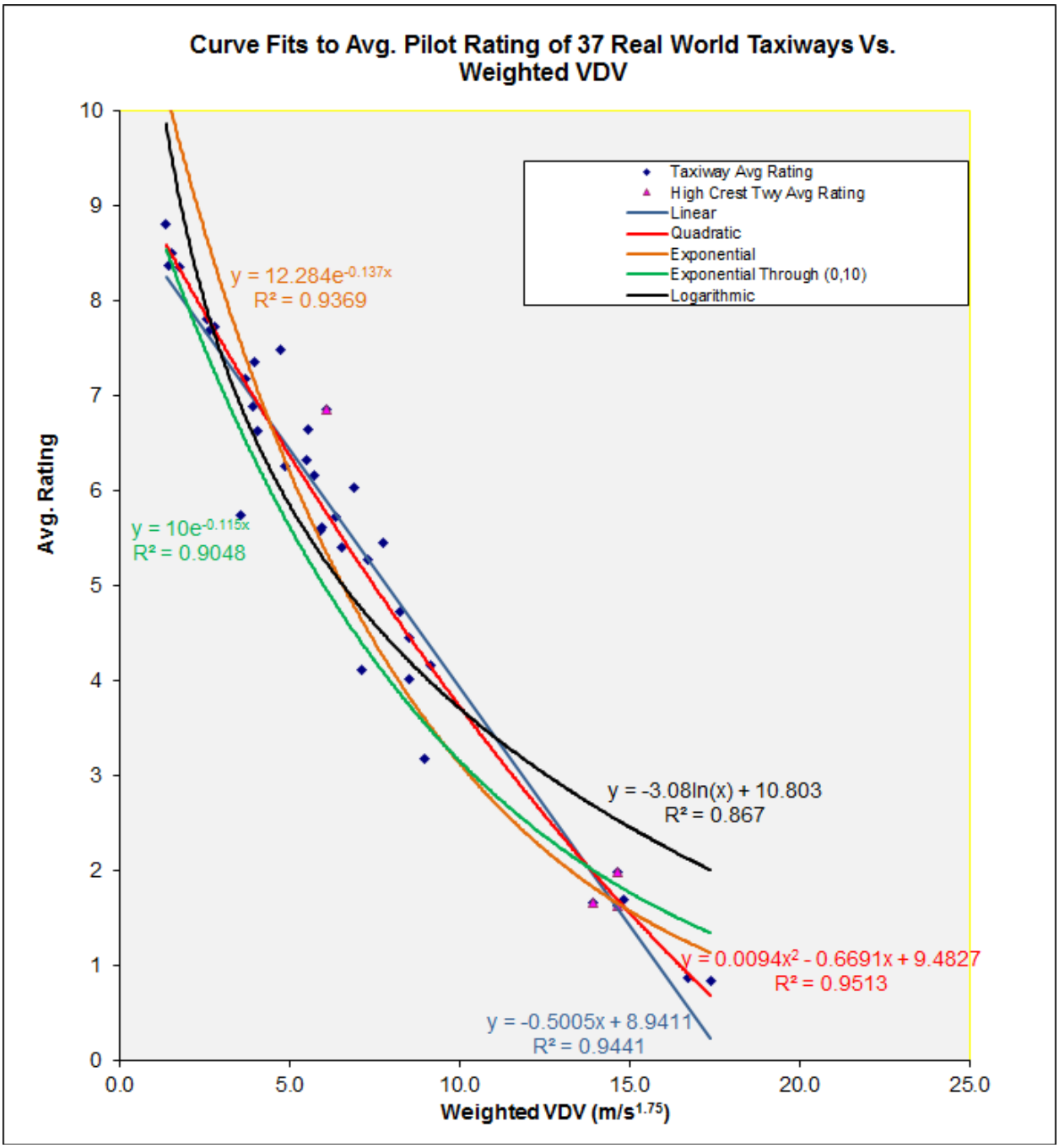
The linear fit is motivated for classical comparison purposes and because its analysis is simple and standard. The quadratic fit is provided as the simplest fit that has curvature. The exponential and logarithmic fits are motivated by previous roughness human vibration models such as [Ref. 1]. The exponential fit through (0,10) provides an exponential fit in which pilots are modeled to give exactly a ten rating when (and only when) there is no vibration. The logarithmic fit is the Fechner (or Weber-Fechner) law for general human response to external

stimulus that has been experimentally verified as a reasonable approximation for certain ranges of human perception of brightness of light and loudness of sound[4]. The fitted functions are mathematically simplistic in order to avoid accidental insertion of complicated behavior without motivation and in order that additional statistical analysis such as confidence intervals and predictive intervals can be performed. Figure 14 shows the linear, quadratic, exponential, exponential through (0,10) and logarithmic fits for the pilot average numerical taxiway ratings of the 37 runways in the final data set as a function of each ISO index and Figure 15 shows the corresponding average runway ratings.

[4] Fits by Stevens' power law for human perception ($y = ax^b$) were also attempted but found to be inaccurate and are not presented.


Curve Fits to Avg. Pilot Rating of Real World Taxiways Vs. Weighted RMS
Avg. Rating
Weighted RMS (m/s²)
Taxiway Avg Rating
High Crest Factor Twy Case
Linear
Quadratic
Exponential
Exponential Through (0,10)
Logarithmic
y = 12.746e-1.922x
R² = 0.9347
y = 10e-1.572x
R² = 0.8925
y = -3.347ln(x) + 2.6286
R² = 0.8816
y = 2.9077x² - 11.089x + 10.083
R² = 0.9499
y = -6.9534x + 9.0361
R² = 0.9209


Curve Fits to Avg. Pilot Rating of 37 Real World Taxiways Vs. Weighted MTVV
Avg. Rating
Weighted MTVV (m/s²)
Taxiway Avg Rating
High Crest Twy Avg Rating
Linear
Quadratic
Exponential
Exponential Through (0,10)
Logarithmic
y = 11.22e-0.669x
R² = 0.8634
y = 10e-0.607x
R² = 0.8525
y = -2.898ln(x) + 5.5364
R² = 0.8352
y = 0.3869x² - 3.805x + 9.4311
R² = 0.9028
y = -2.473x + 8.6401
R² = 0.8872


Curve Fits to Avg. Pilot Rating of 37 Real World Taxiways Vs. Weighted VDV
Taxiway Avg Rating
High Crest Twy Avg Rating
Linear
Quadratic
Exponential
Exponential Through (0,10)
Logarithmic
y = 12.284e-0.137x
R² = 0.9369
y = 10e-0.115x
R² = 0.9048
y = -3.08ln(x) + 10.803
R² = 0.867
y = 0.0094x² - 0.6691x + 9.4827
R² = 0.9513
y = -0.5005x + 8.9411
R² = 0.9441
Avg. Rating
Weighted VDV (m/s1.75)
0
1
2
3
4
5
6
7
8
9
10
0.0
5.0
10.0
15.0
20.0
25.0

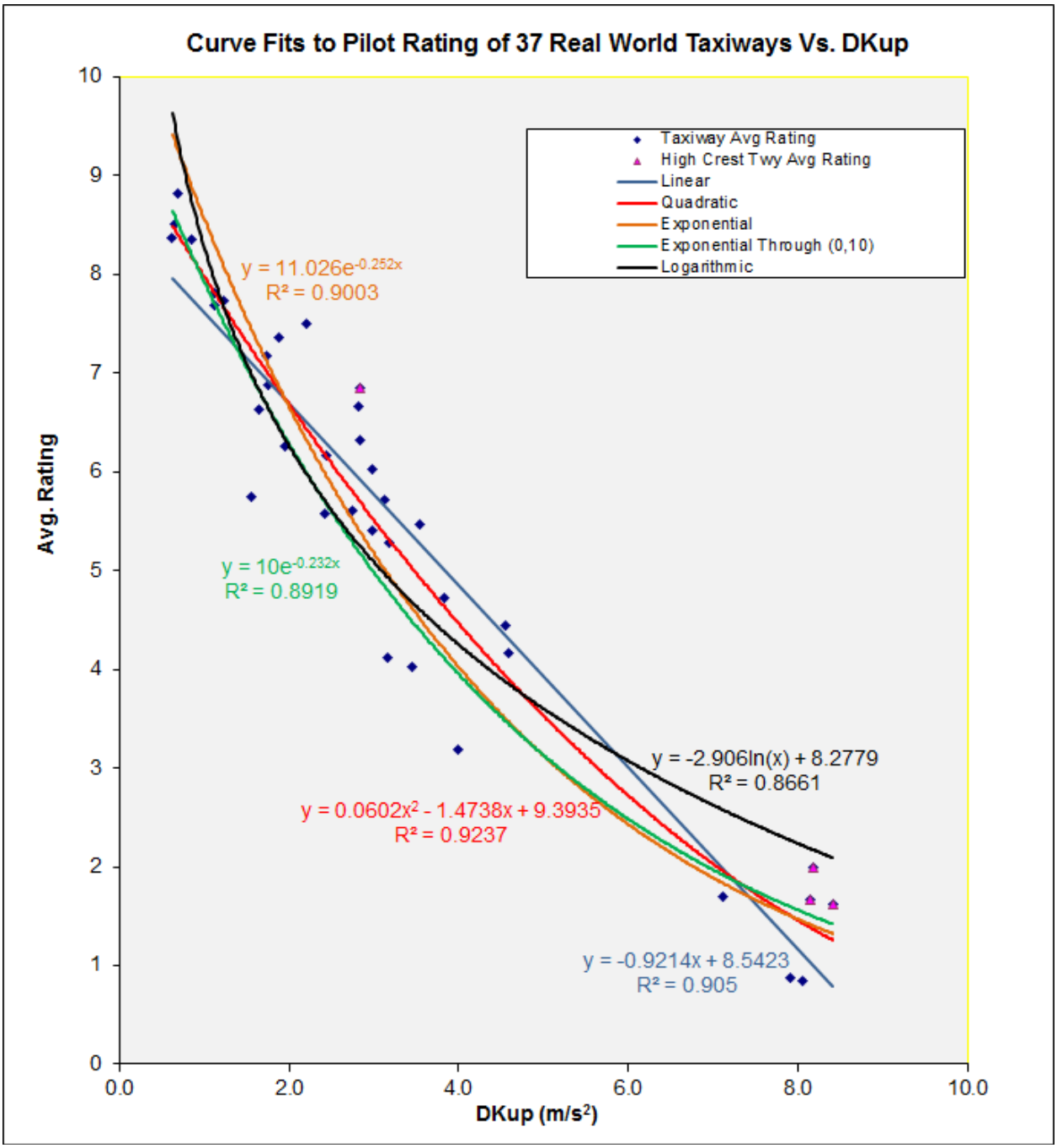
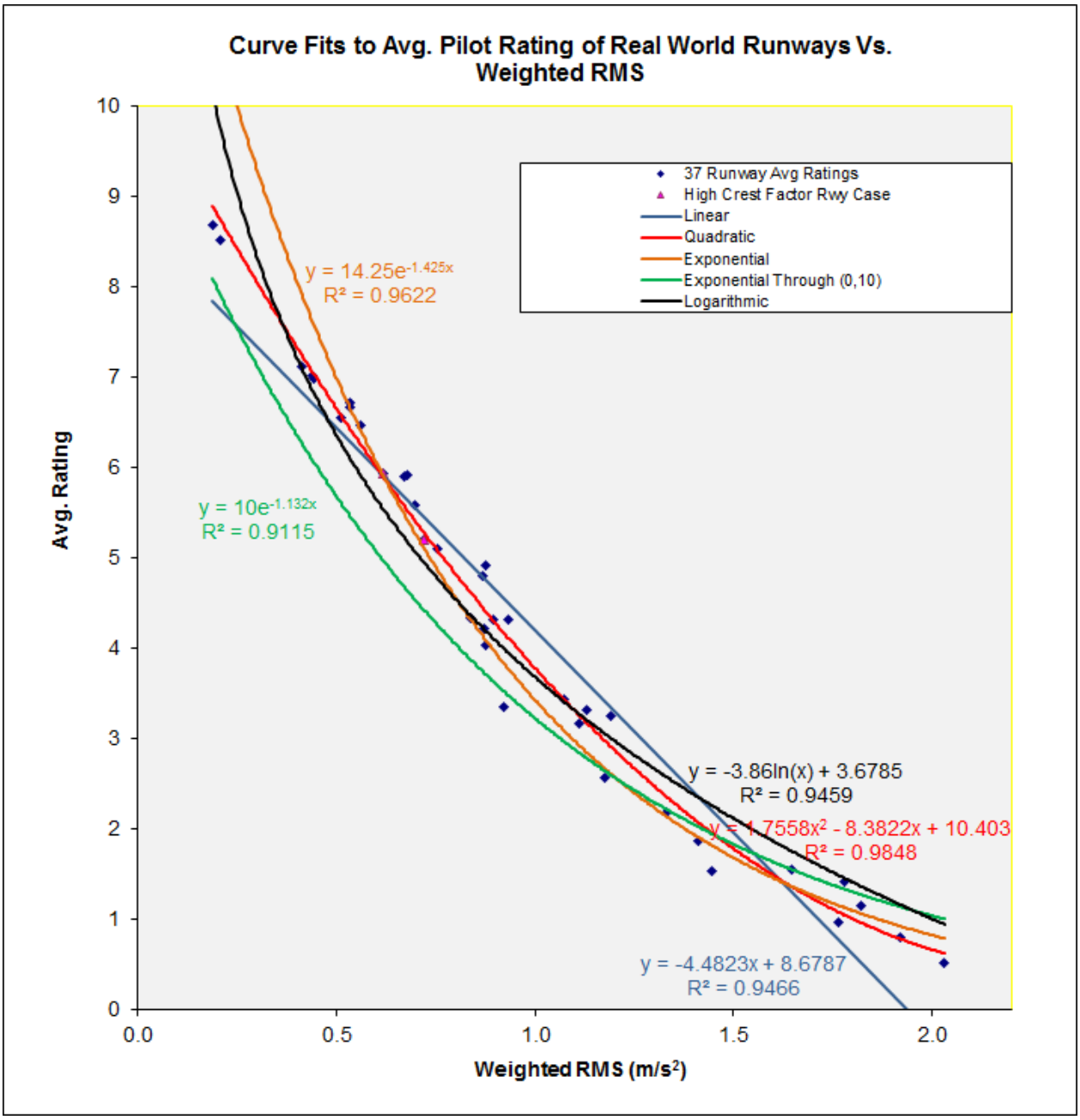
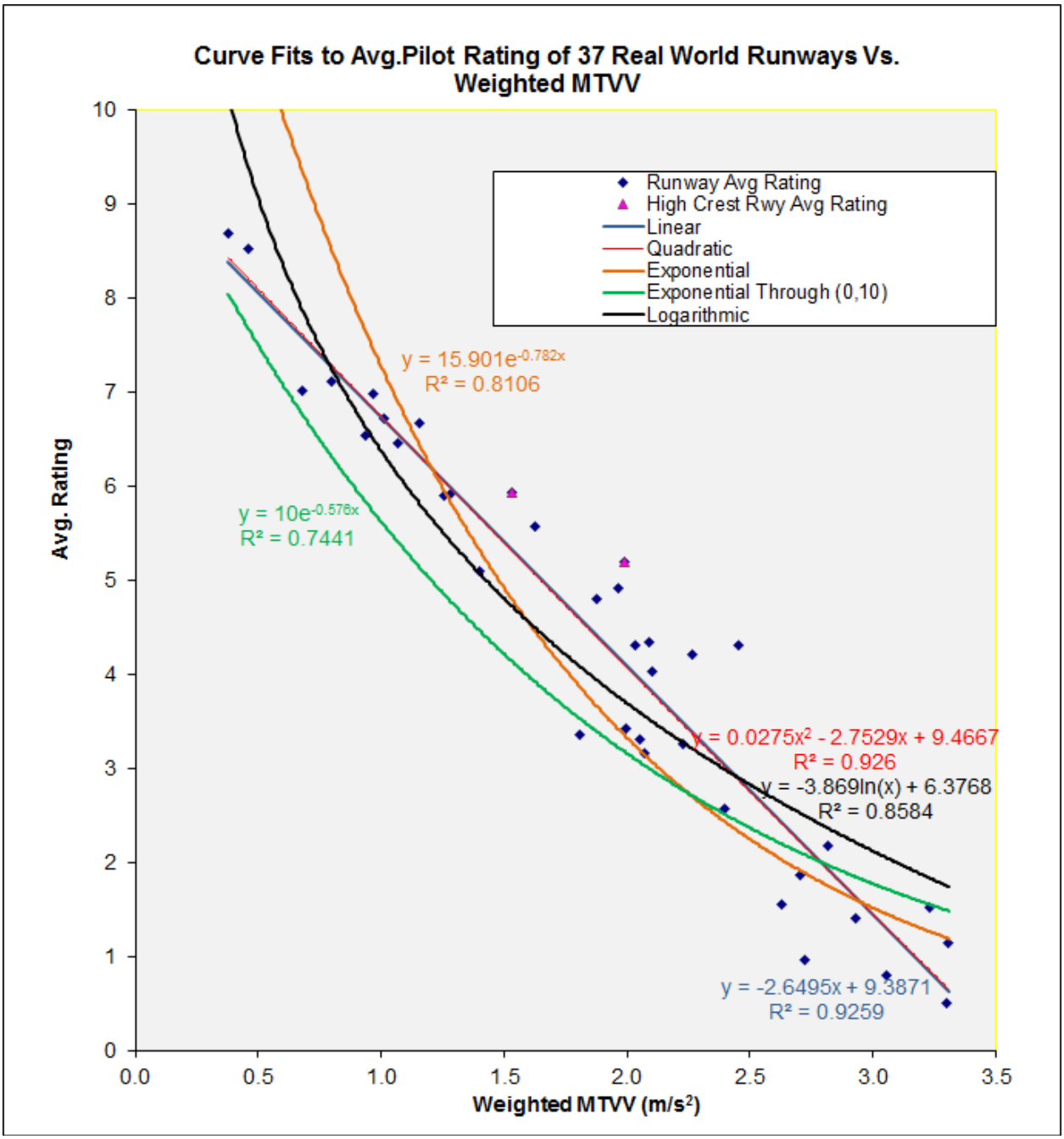
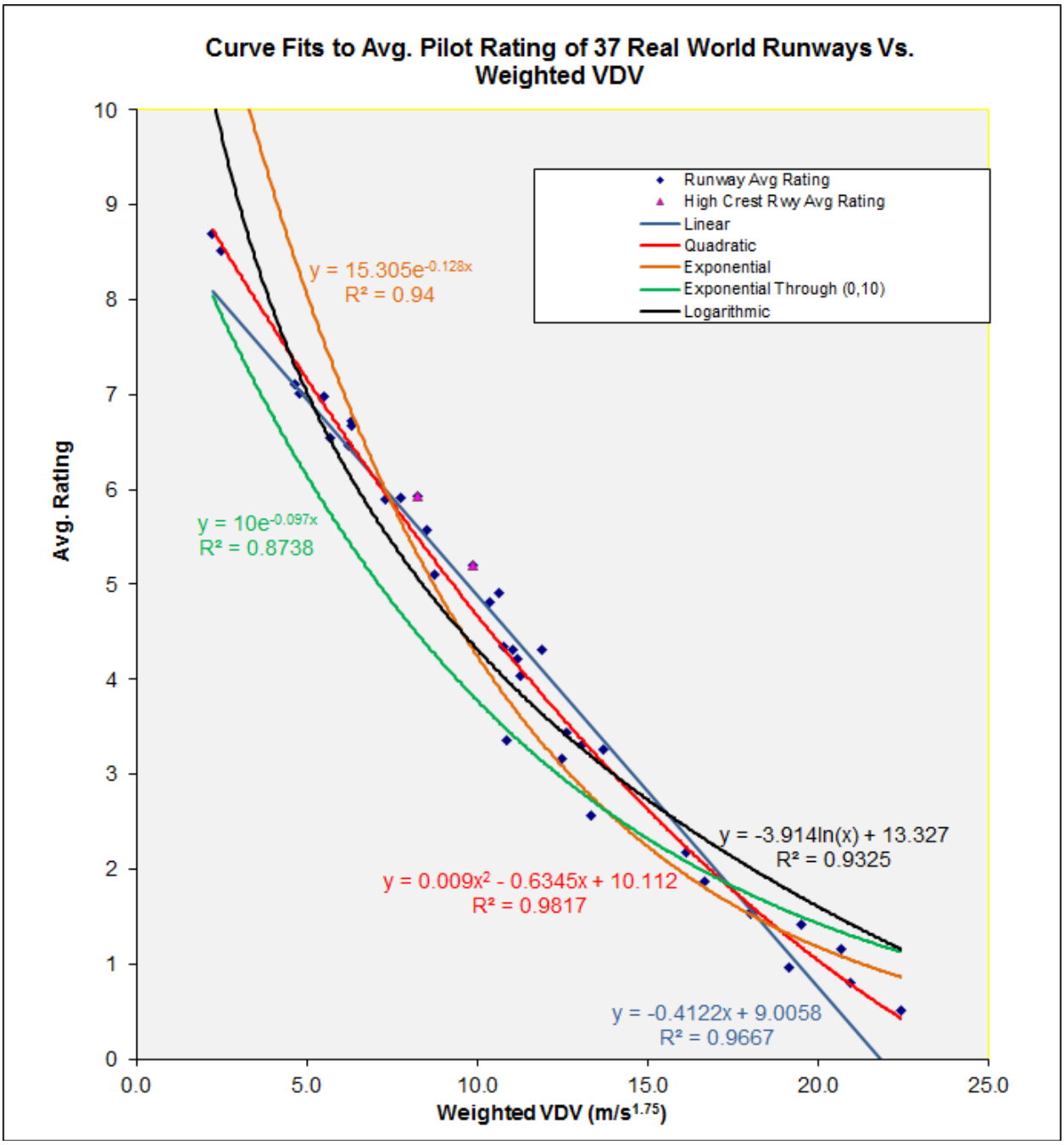


Figure 14. Various Curve Fits by Least Squares to Pilot Average Taxiway Ratings: Linear (Blue), Quadratic (Red), Exponential (Yellow), Exponential Through (0,10) (Green), Logarithmic (Black)


Curve Fits to Avg. Pilot Rating of Real World Runways Vs. Weighted RMS
Avg. Rating
Weighted RMS (m/s²)
37 Runway Avg Ratings
High Crest Factor Rwy Case
Linear
Quadratic
Exponential
Exponential Through (0,10)
Logarithmic
y = 14.25e-1.425x
R² = 0.9622
y = 10e-1.132x
R² = 0.9115
y = -3.86ln(x) + 3.6785
R² = 0.9459
y = 1.7558x² - 8.3822x + 10.403
R² = 0.9848
y = -4.4823x + 8.6787
R² = 0.9466
0
1
2
3
4
5
6
7
8
9
10
0.0
0.5
1.0
1.5
2.0


Curve Fits to Avg.Pilot Rating of 37 Real World Runways Vs. Weighted MTVV
Runway Avg Rating
High Crest Rwy Avg Rating
Linear
Quadratic
Exponential
Exponential Through (0,10)
Logarithmic
y = 15.901e-0.782x
R² = 0.8106
y = 10e-0.576x
R² = 0.7441
y = 0.0275x² - 2.7529x + 9.4667
R² = 0.926
y = -3.869ln(x) + 6.3768
R² = 0.8584
y = -2.6495x + 9.3871
R² = 0.9259
Avg. Rating
Weighted MTVV (m/s²)
0
1
2
3
4
5
6
7
8
9
10
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5


Curve Fits to Avg. Pilot Rating of 37 Real World Runways Vs. Weighted VDV
Runway Avg Rating
High Crest Rwy Avg Rating
Linear
Quadratic
Exponential
Exponential Through (0,10)
Logarithmic
y = 15.305e-0.128x
R² = 0.94
y = 10e-0.097x
R² = 0.8738
y = -3.914ln(x) + 13.327
R² = 0.9325
y = 0.009x² - 0.6345x + 10.112
R² = 0.9817
y = -0.4122x + 9.0058
R² = 0.9667
Avg. Rating
Weighted VDV (m/s1.75)
0
1
2
3
4
5
6
7
8
9
10
0.0
5.0
10.0
15.0
20.0
25.0

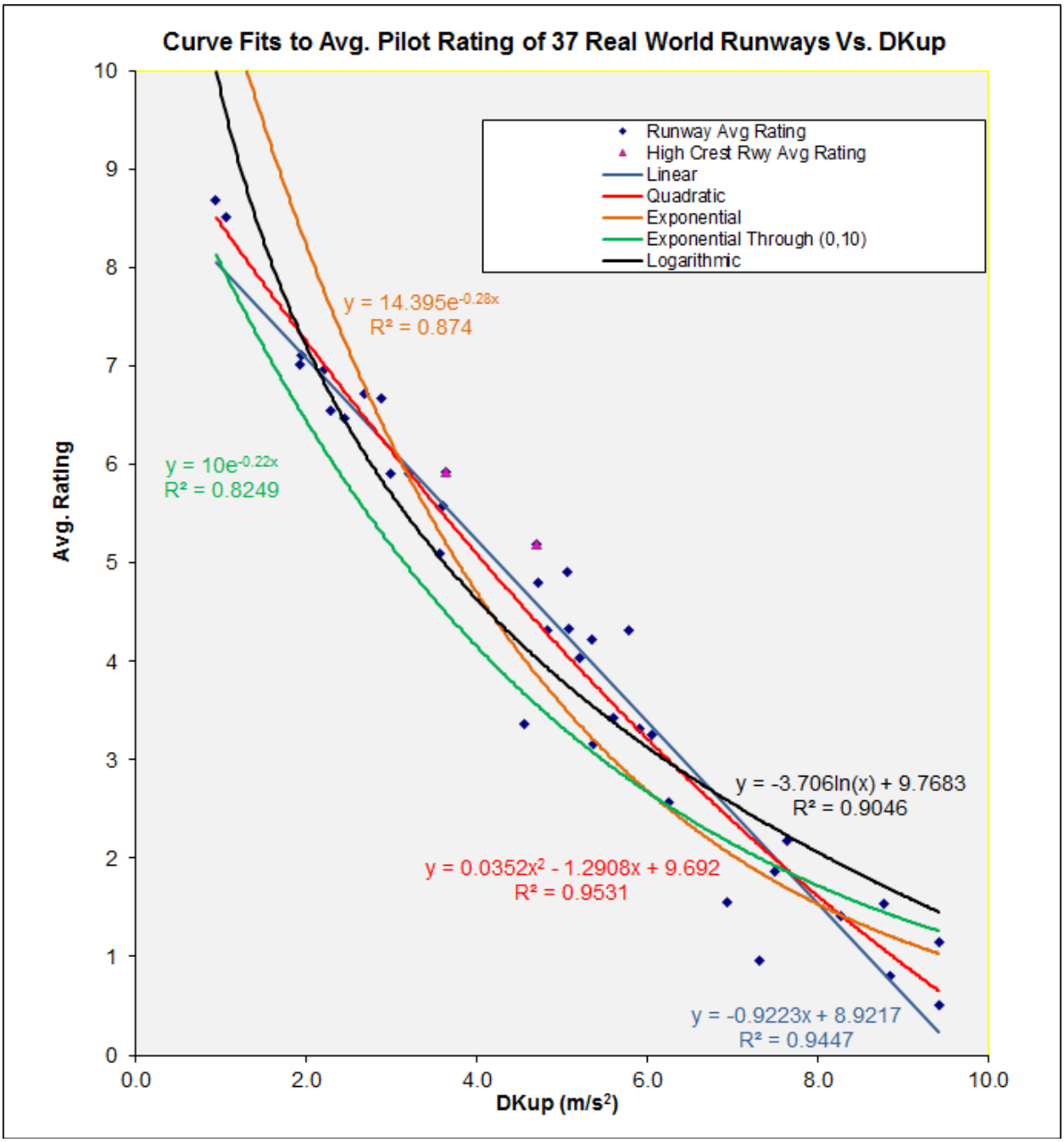


Figure 15. Various Fits of Average Runway Rating Vs. ISO Index by Least Squares Linear (Blue), Quadratic (Red), Exponential (Yellow), Exponential Through (0,10) (Green), Logarithmic (Black)

The R^2 (R squared, or coefficient of determination) value of each fit shows the goodness of fit in that the closer R^2 is to 1.0 the better is the fit, and indicates that the quadratic fit is the best of the five curve fits shown.

In order to capture the closer fit of the quadratic function while preserving the more logarithmic behavior for extrapolation at larger index values, a fit that captures the curvature capability of the

quadratic while preserving the decaying nature of the logarithmic fit on the right was introduced in the form of a shifted logarithmic fit: $y = a \ln(x+c) + b$, in which y is the pilot average numerical rating and x is an ISO index (to which "shift" c is added before taking the logarithm). The actual curve fit is unknown since subjective human evaluation is involved, but the shifted logarithmic fit provides the best combination of R^2 near one and decaying behavior for larger ISO index values from among the fits tested. An example appears in Figure 16 for weighted VDV.

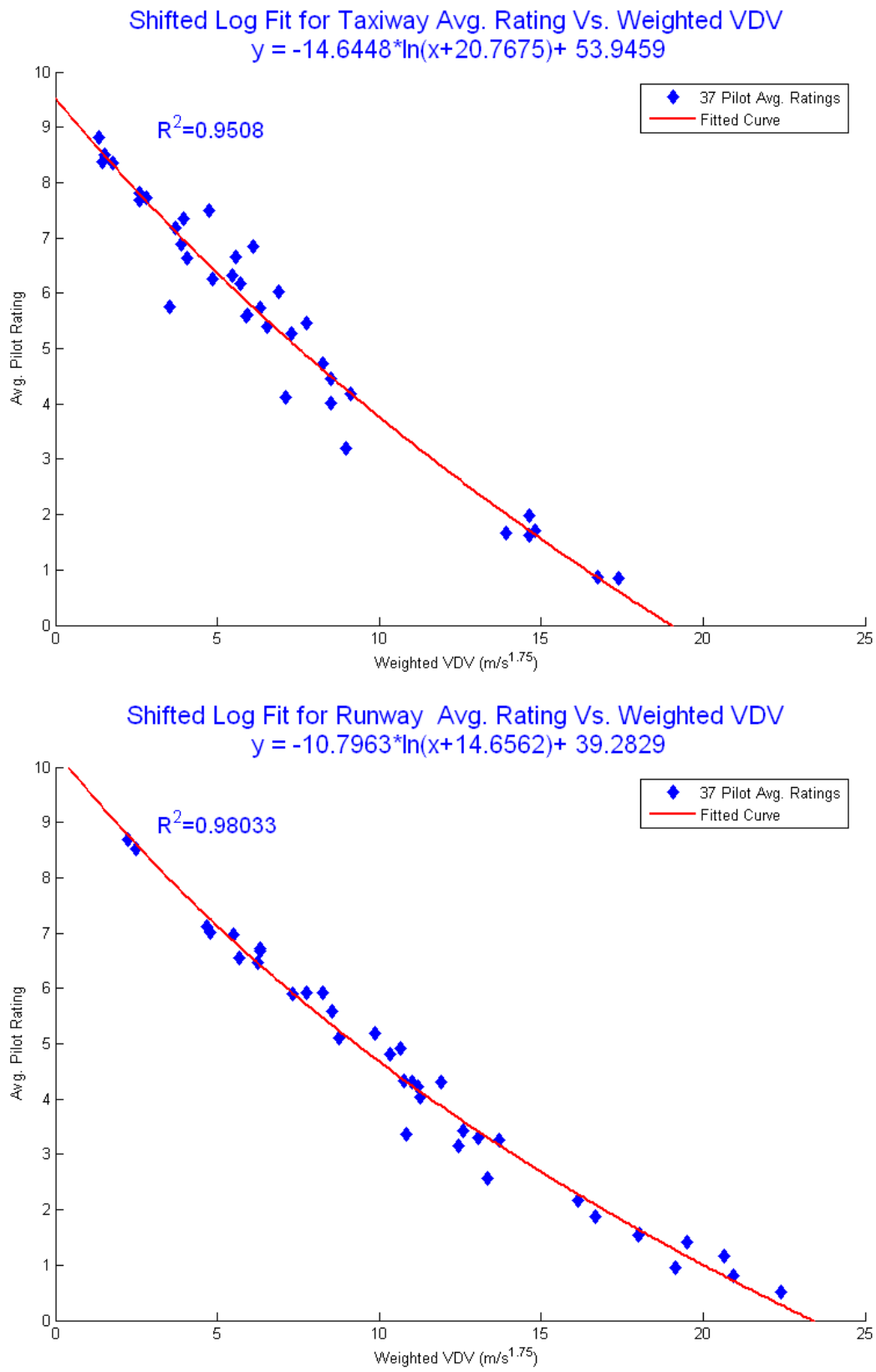


Figure 16. Shifted Logarithmic Fits of Taxiway and Runway Rating Vs. Weighted VDV

Unlike the quadratic fit, the shifted logarithmic fit can be extrapolated to y = 0 (pilot numerical rating equal to zero) on the right because it is strictly decreasing. Furthermore, a shifted logarithmic fit is still simple enough to be amenable to analysis by a confidence interval for the fitted curve and a predictive interval for the samples. The confidence interval is an interval within which a fitted curve is likely to lie with a given probability under the assumption that the data points would fit the curve type except that they include normally distributed noise. The predictive interval shows where similarly generated data points would lie with a given likelihood under the same assumptions. A Matrix Laboratory (MATLAB) program was generated to compute optimal shifts such that the least squares fit made R^2 as close as possible to 1.0 in the shifted logarithm fits.

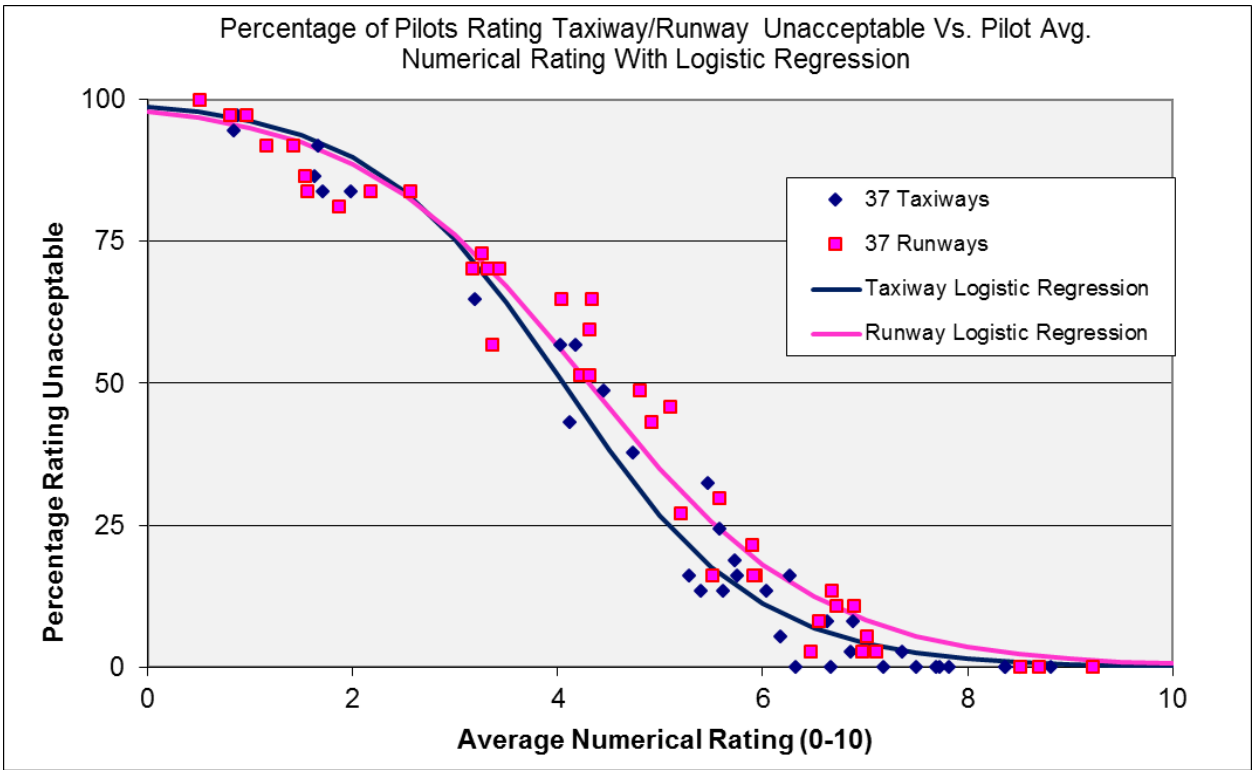
Each ISO index may be considered to be best depending upon the particular application desired, so no ISO index is preferred here – despite that the fits to weighted RMS and weighted VDV are somewhat closer (by R^2 values and [Ref. 8]). Instead it is noted that weighted RMS is the most commonly considered indicator and provides an estimate of the average roughness of a ride, while weighted VDV is often considered in the analysis of rides that contain non-uniform shocks and may provide "more cautious assessments of the limiting daily exposure duration" than weighted RMS methods [Ref. 9], weighted MTVV provides a maximum running value over a given period of weighted RMS, and DKup may be used as a spinal dose indicator.

4.3 Calculation of the Unacceptable Range for Each ISO Index

Each pilot in the final study rated each simulated taxiway and runway either as "acceptable" or "unacceptable" and this data was used to calculate acceptable values for the ISO roughness indices. (Data from the preliminary study could be combined because pilots in the preliminary study were allowed a third acceptability option of "uncomfortable.")

4.3.1 The Numerical Rating at Which a Ride Becomes Unacceptable

The percentage of pilots stating that a test ride was "unacceptable" in final testing is plotted as a function of pilot (average) numerical rating in Figure 17 with a full view and a close up view near where the percentage rating the ride as unacceptable declines to zero.



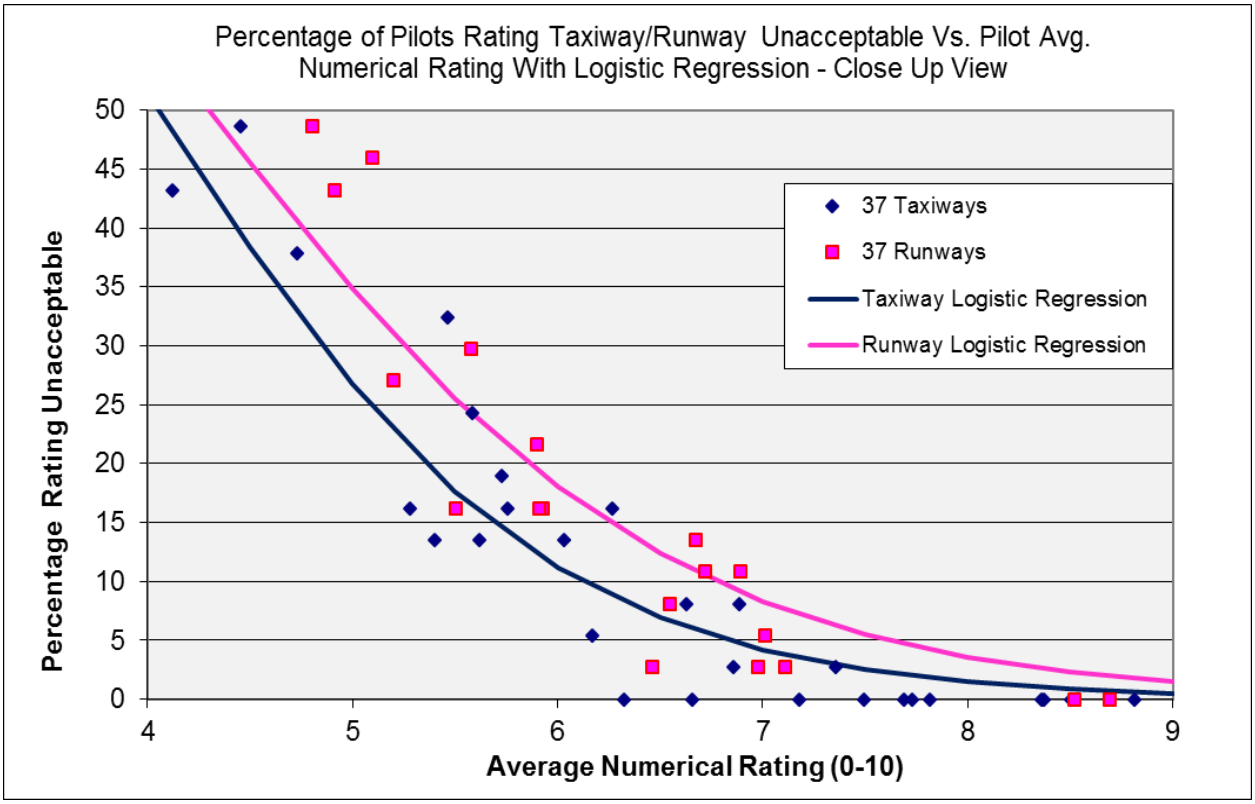


Figure 17. Percentage of Taxiways and Runways Rated Unacceptable Vs. Numerical Rating

The correlation coefficient between average taxiway rating and percentage of pilots rating the taxiway as unacceptable was -.960, and the coefficient for runways was -.973. These strong correlations show that a fit of percentage of pilots rating a ride as unacceptable as a function of pilot average numerical rating is reasonable. Consequently, logistic regression and simple polynomial fits that could accommodate the general behavior of each data trend (taxiway or runway) with y representing the percent of pilots rating a ride as unacceptable versus x the average pilot numerical rating (0-10) were made. These were subject to the requirement that y=100 (percent unacceptable) when x=0 (the lowest numerical rating). For taxiways the logistic regression was

$$y = \frac{100}{1 + e^{-4.308+1.063x}} \tag{1}$$

and for runways

$$y = \frac{100}{1 + e^{-3.831+.891x}} \tag{2}$$

Figure 17 shows that logistic regression is slightly conservative in the tails so polynomial fits are also given (Appendix A) for a less conservative (slightly slimmer tails on the right) comparison. By solving equations (1) and (2) the numerical pilot rating at which any percentage of pilots rates taxiways or runways as unacceptable can be deduced (Table 4).

Table 4. Pilot Average Numerical Rating of a Taxiway or Runway When 5%, 10%, or 50% of Pilots Rate the Ride as Unacceptable

Amount Unacceptable	Taxiway Rating (0-10)	Runway Rating (0-10)
5%	6.88	7.60
10%	6.18	6.77
50%	4.11	4.30

If taxiways and runways are considered unacceptable when 5% of pilots rate them as unacceptable then a taxiway becomes unacceptable at a pilot average numerical rating of about 6.88 and a runway becomes unacceptable at a pilot average numerical rating of about 7.60.

4.3.2 The "Unacceptable" Range for Each ISO Index

It is desired to predict the discomfort level that will be felt and the point at which a ride becomes unacceptable on an aircraft pavement as a function of each ISO index. In order to predict this, apply the curve fits of Figure 17 (equations (1) and (2), or Table 4) to find the average pilot numerical ratings 0-10 that correspond to 5%, 10%, or 50% of the pilots declaring a taxiway or runway ride unacceptable (Figure 17). Then, apply the previous curve fits of pilot average numerical rating to the ISO indices (such as in Figure 16 for the shifted logarithmic case of weighted VDV) to find the value of each index that caused 5%, 10% or 50% of pilots in the final data set to rate a ride as unacceptable. Figures 16, 18, 19, and 20 show the fits of pilot average numerical rating to each index by the shifted logarithm method that together with equations (1) and (2) determine the index values at which pilot average rating indicates taxiways and runways are unacceptable[5] (Table 4.)

5 Table 2 was generated using the shifted logarithmic fits of average pilot rating versus ISO parameter because the shifted logarithm provided a good overall fit, but a similar table can be generated using any of the curve fits. Most exponential fits have substantial unconservative deviations from the data at small pilot percentages, however, and should not be used.

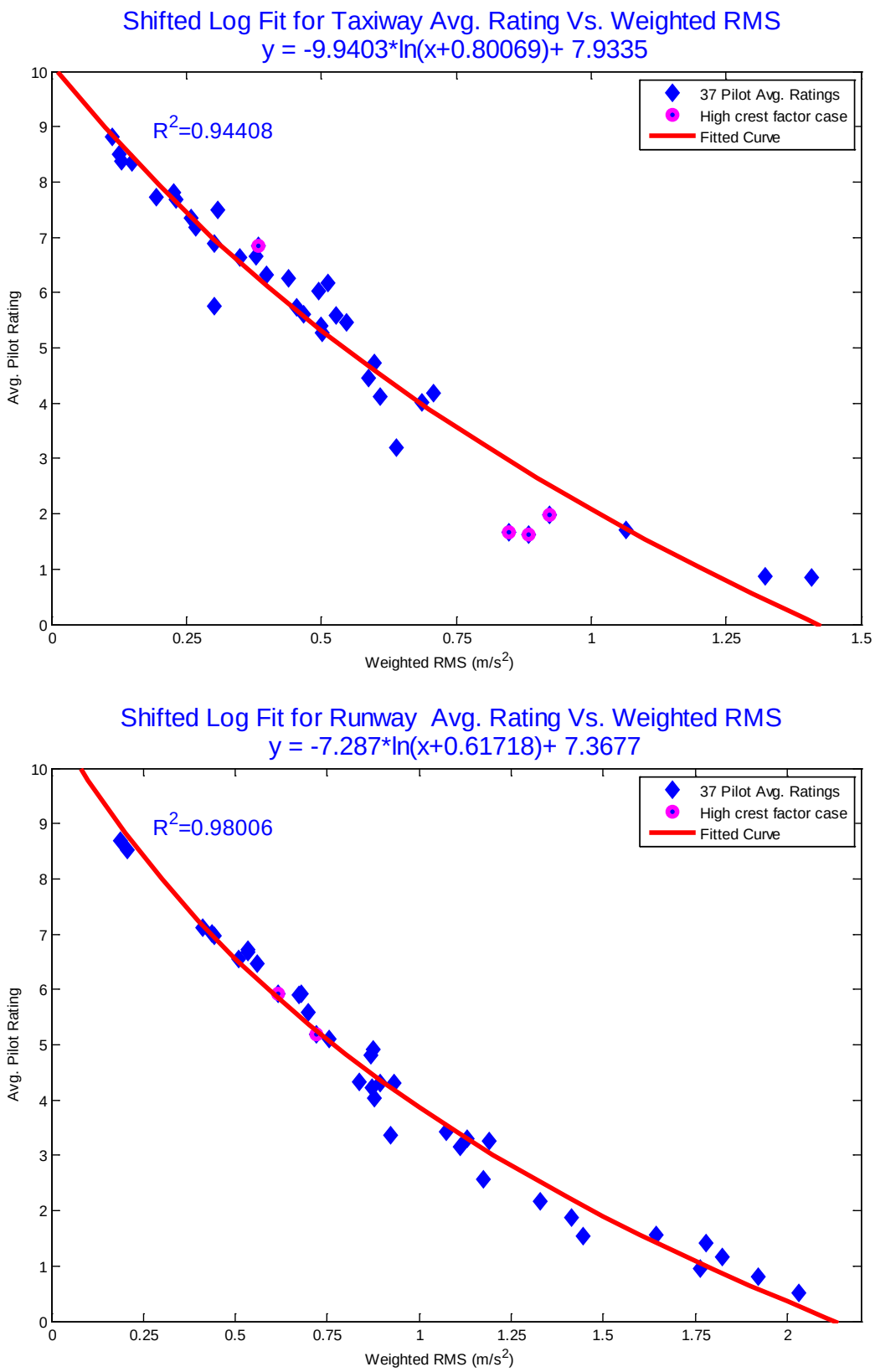
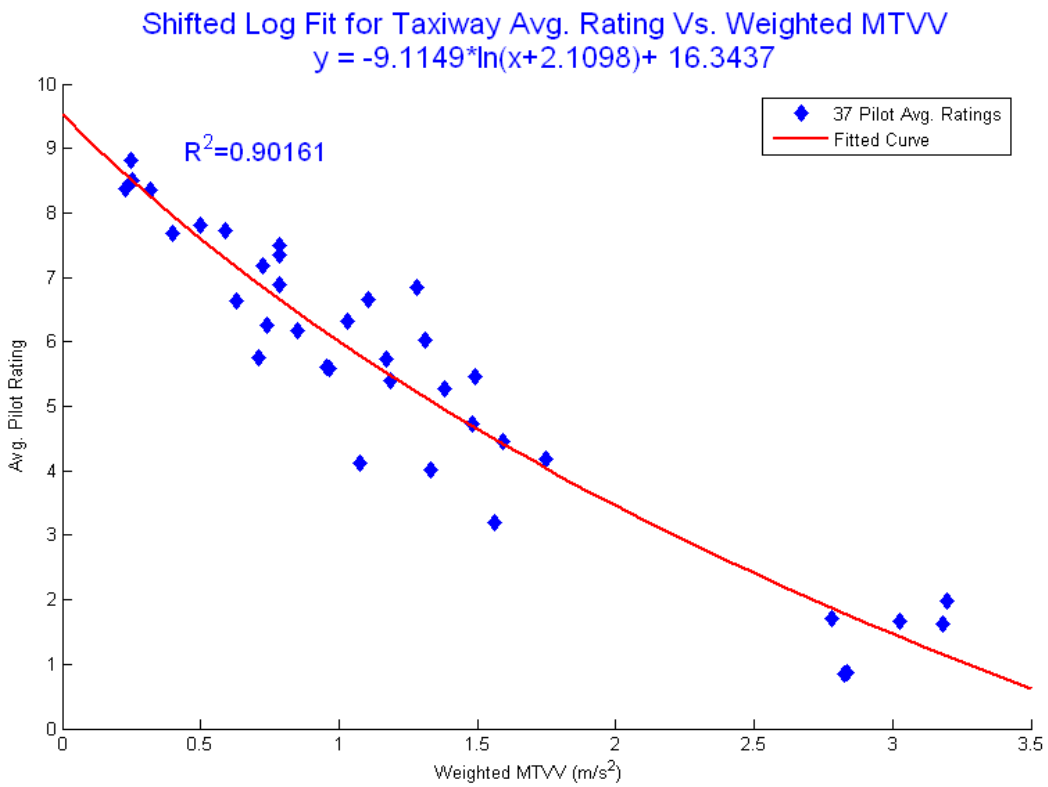


Figure 18. Average Pilot Taxiway and Runway 0-10 Rating vs. Weighted RMS



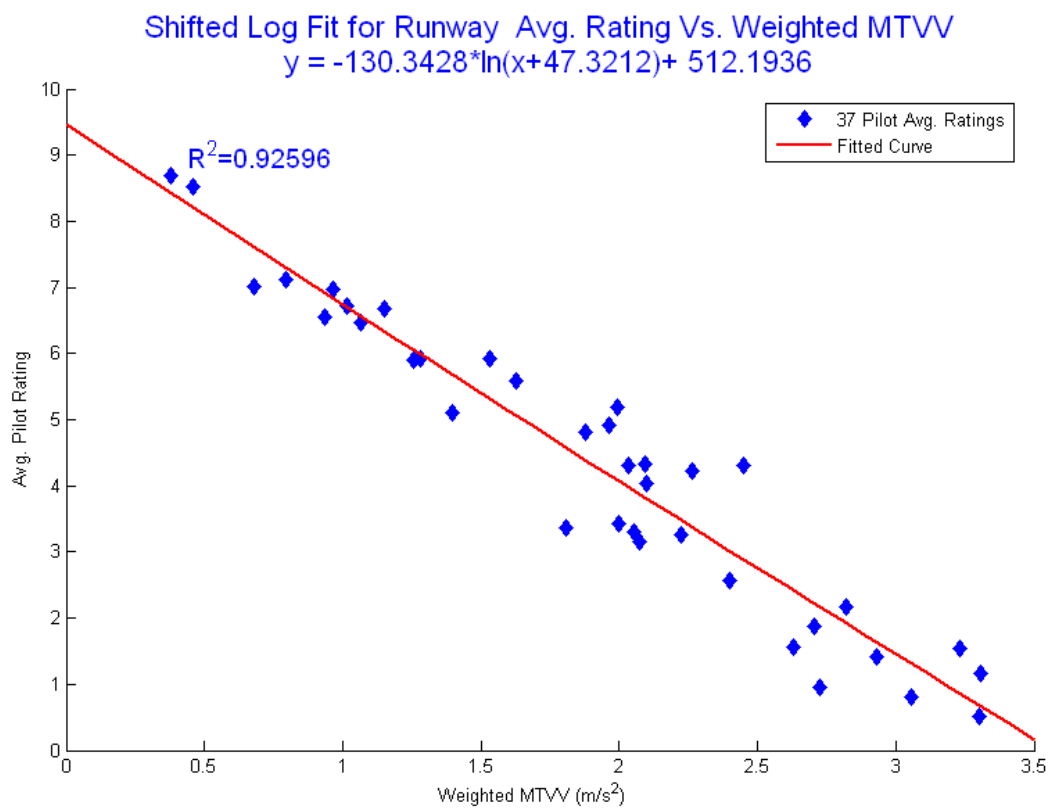


Figure 19. Average Pilot Taxiway and Runway 0-10 Rating vs. Weighted MTVV

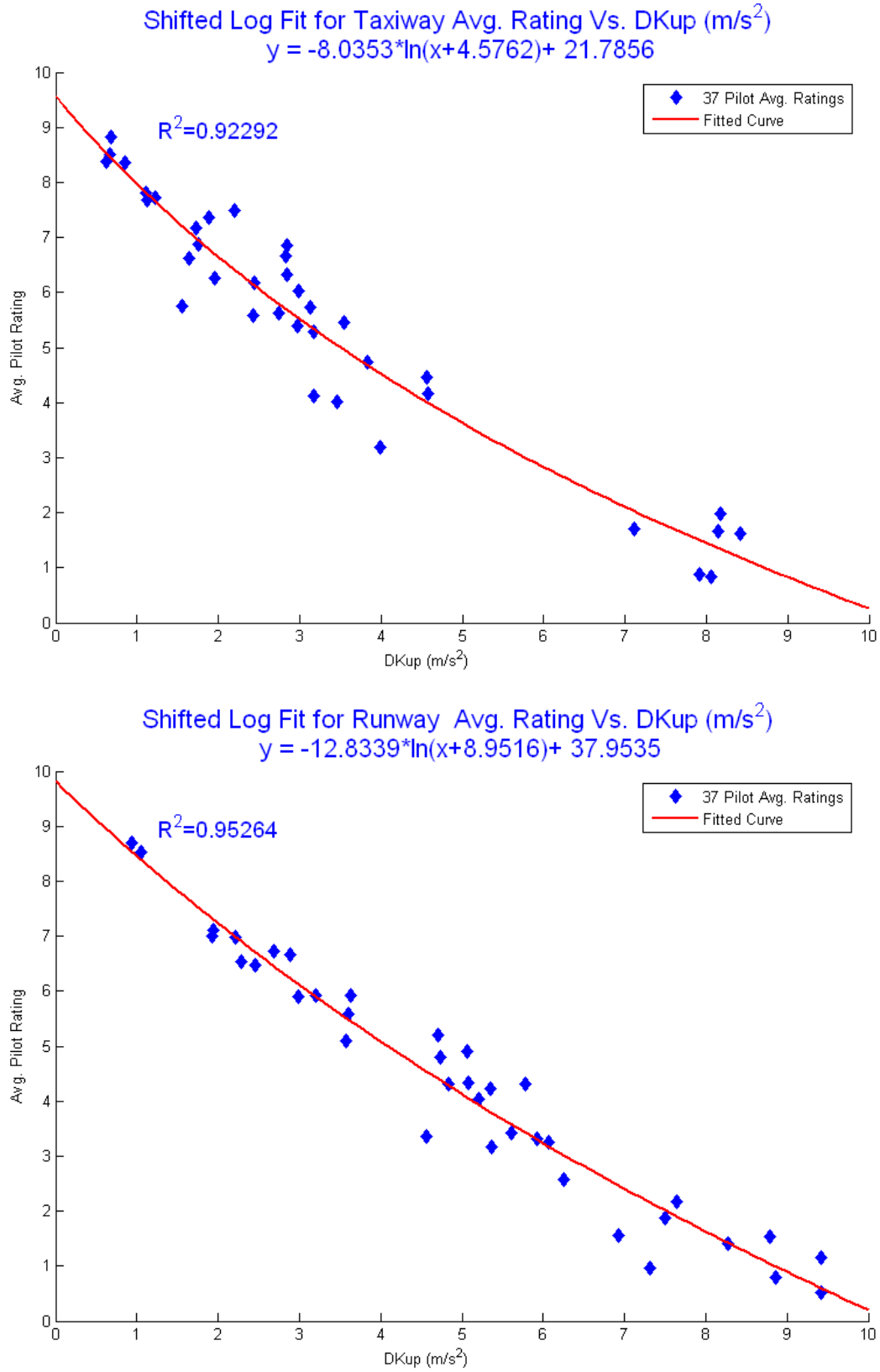


Figure 20. Average Pilot Taxiway and Runway 0-10 Rating vs. DKup

Table 5. ISO Index Values at Which 5%, 10% and 50% of Pilots Are Estimated to Rate a Taxiway or Runway as Unacceptable

ISO Roughness Index	Index Value When 5% of Pilots Rate the Taxiway as Unacceptable	Index Value When 10% of Pilots Rate the Taxiway as Unacceptable	Index Value When 50% of Pilots Rate the Taxiway as Unacceptable	Index Value When 5% of Pilots Rate the Runway as Unacceptable	Index Value When 10% of Pilots Rate the Runway as Unacceptable	Index Value When 50% of Pilots Rate the Runway as Unacceptable
Weighted RMS (m/s^2)	***0.31***	*0.39*	*0.67*	***0.35***	*0.47*	*0.91*
Weighted MTVV (m/s^2)	***0.71***	*0.94*	*1.72*	***0.68***	*0.99*	*1.91*
Weighted VDV ($m/s^{1.75}$)	***4.11***	*5.32*	*9.29*	***4.16***	*5.66*	*10.88*
DKup (m/s^2)	***1.82***	*2.40*	*4.45*	***1.69***	*2.40*	*4.81*

If the index value leading to 5% of pilots rating a ride as unacceptable is used as the threshold[6] *for calling a taxiway or runway unacceptable then Table 5 shows the unacceptable index ranges. A similar table can be generated for any percentage of pilots rating a ride as unacceptable.*

Table 5 indicates that for each index, the value at which a runway or taxiway becomes unacceptable to 5% of pilots is similar; however, the value at which 50% of pilots find a taxiway unacceptable is uniformly lower for the taxiways. Since during post-test evaluations some pilots expressed concern about flight attendants during taxiway maneuvers because they were not seat-belted, it is likely that the lower taxiway thresholds resulted at least partly from concern for flight attendant safety.

Table 6. Index Values at Which Taxiways and Runways Become Unacceptable Based Upon the Index Values at Which 5% of Pilots Rated Rides Unacceptable

ISO Roughness Index	Unacceptable Taxiway Range	Unacceptable Runway Range
Weighted RMS (m/s^2)	***≥0.31***	***≥0.35***
Weighted MTVV (m/s^2)	***≥0.71***	***≥0.68***
Weighted VDV ($m/s^{1.75}$)	***≥4.11***	***≥4.16***
DKup (m/s^2)	***≥1.82***	***≥1.69***

[6] 5% is not necessarily recommended as a criterion for accepting or rejecting a runway, but there is a lack of data to make approximations of unacceptability at much lower percentages (e.g. 1%), so 5% is chosen as the smallest possible commonly used value for statistical acceptability in hypothesis tests.

4.4 COMPARISON OF PILOT RATINGS FOR RIDES WITH HIGH CREST FACTOR

The crest factor for each taxiway and runway in the final study was evaluated. A high crest factor is an indication that a ride has some sharp jolts. There were no tests that had a crest factor greater than the ISO standard value of 9.0, but Figure 21 plots the 37 final study taxiway average ratings and highlights in magenta for the four taxiway ratings that are from taxiways with crest factors (7.63, 8.10, 8.62, 8.77) greater than seven. Similar plots for runways were constructed but do not appear here since only two runways had crest factors (of 5.67 and 6.36) greater than 5.5.

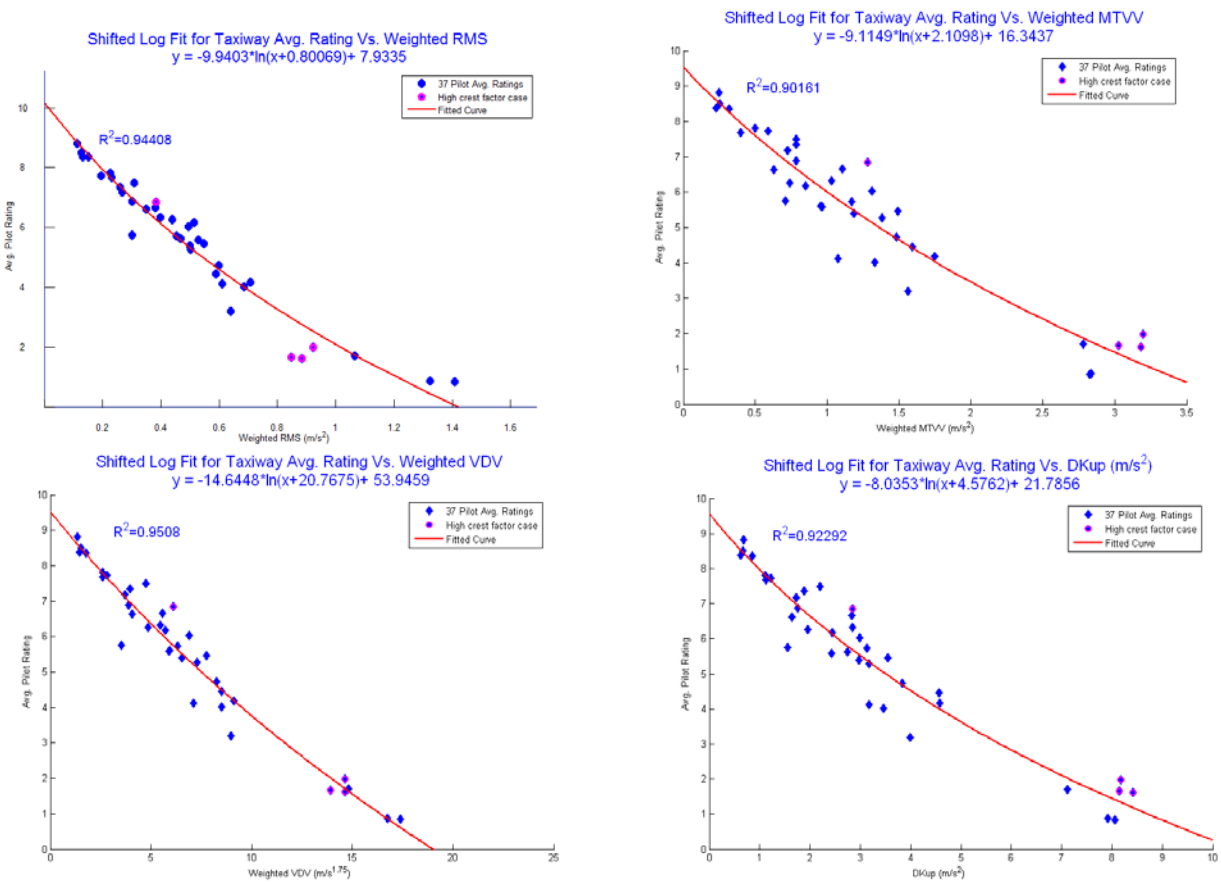


Figure 21. Shifted Logarithmic Fits Highlighting Taxiways with High Crest Factor

For the four taxiways with high crest factor ratings (magenta) it is observed that pilot rating given as a function of weighted RMS or weighted VDV provides the best fit.

4.5 MILITARY AND REPEAT PILOT BIAS

A number of pilots were classified as flying for a branch of the military and it was considered that such pilots might have different ride expectations, so responses of military pilots were separated from those of the non-military pilots and compared (see Figures 22 and 23 and Appendix A).

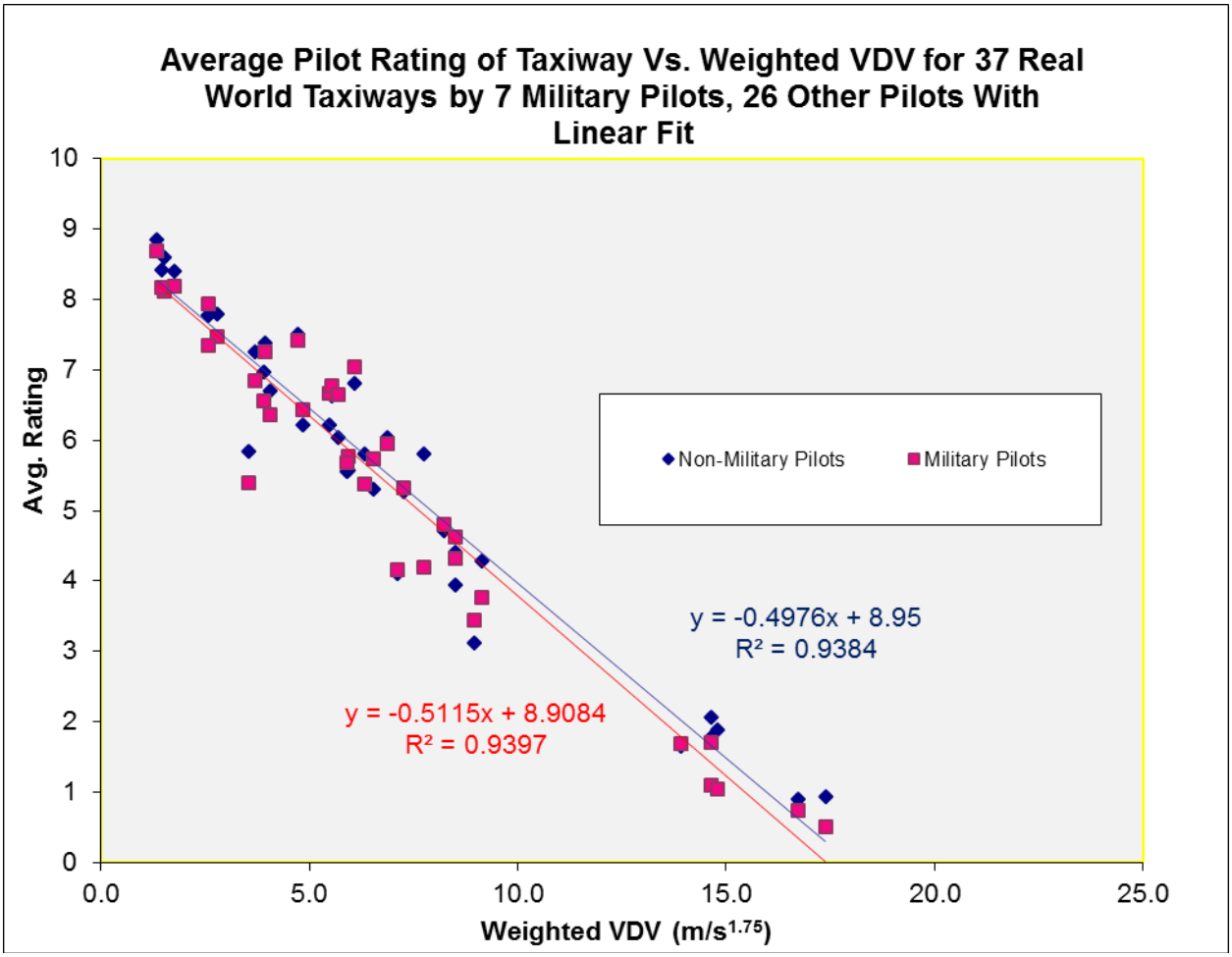


Figure 22. Linear Fits of Average Pilot Rating on 37 Taxiways to Weighted VDV for Military and Non-military Pilots.

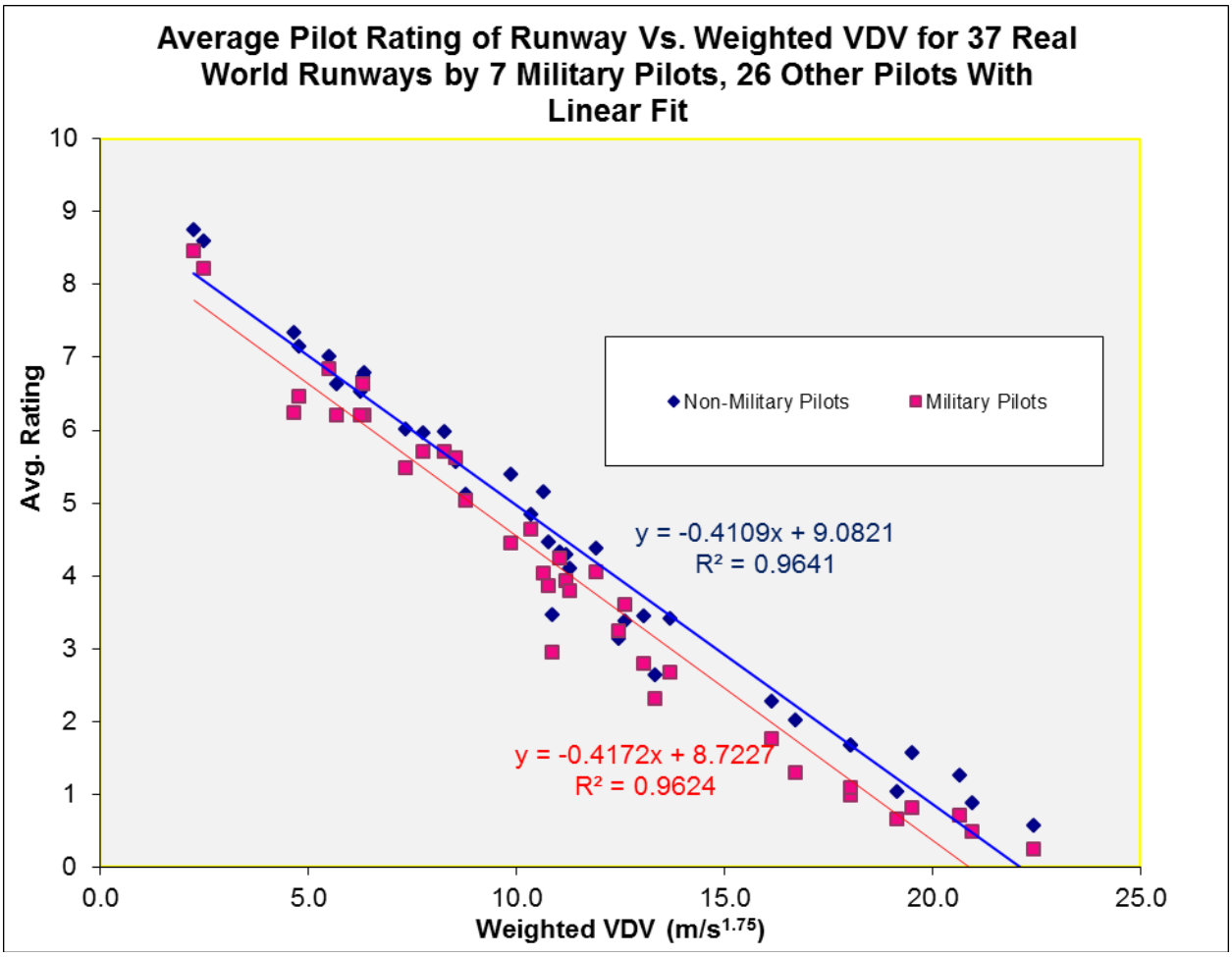


Figure 23. Linear Fits of Average Pilot Rating on 37 Runways to Weighted VDV for Military and Non-military Pilots.

Weighted VDV fits are used for this analysis because the fits are closer to linear than the weighted RMS fits and the linear fits are more easily evaluated for the possibility of bias.

It was decided that having so many military pilots among the test pilots might bias the results, so hypothesis tests at the 5% level were made. For such tests it is easiest to consider only simple linear fits. For the linear fits, in which y = mx + b is the form of the fitted line of slope m and y-intercept b, a hypothesis test that the y-intercept was the same for taxiways or runways could not be rejected at the 5% level - indicating that it is more than 5% likely that the separation between the lines in Figures 22 and 23 could have happened by chance. The hypothesis that the y-intercept was the same for runways was, however, not far from being rejected.

It is speculated, but not with 95% confidence, that there is a slight military bias on runways that occurs because many of the military test pilots in this study fly KC-135 refueling aircraft.

Similar hypothesis tests were performed at the 5% level for "repeat pilots." Repeat pilots are test pilots who had previously been involved in a flight simulator evaluation of taxiways and/or

runways. Trend lines for repeat pilots were found to cross in the middle of the rating range with the trend lines for non-repeat/non-military pilots (Figures 24 and 25 and Appendix A).

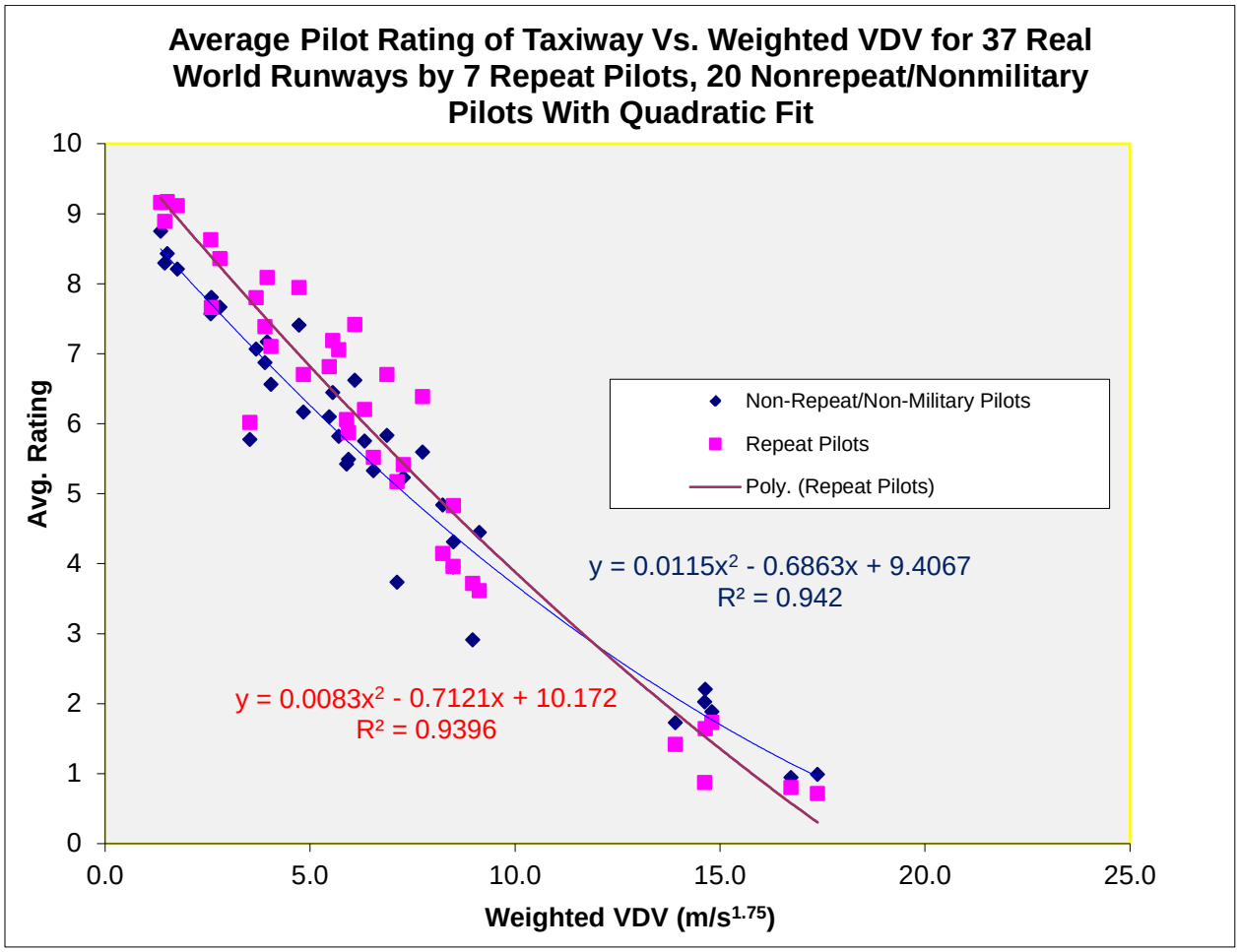


Figure 24. Linear Fits of Average Pilot Rating on 37 Taxiways to Weighted VDV for Repeat Pilots and Non-repeat/Non-military Pilots.

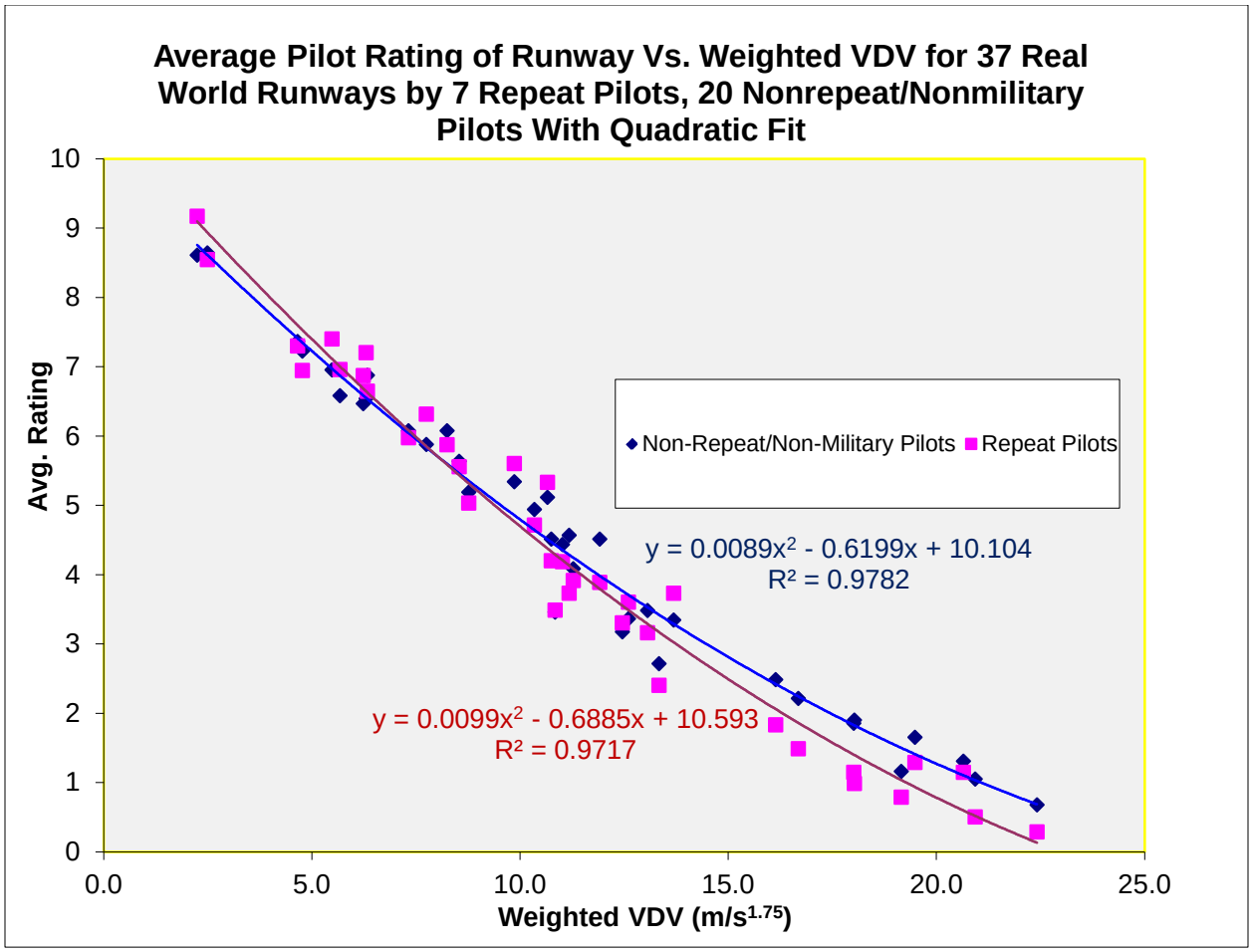


Figure 25. Linear Fits of Average Pilot Rating on 37 Runways to Weighted VDV for Repeat Pilots and Non-repeat/Non-military Pilots.

The hypothesis that the fitted lines for the average ratings of repeat pilots and for non-repeat/non-military pilots had the same slope was rejected at the 5% level for both taxiways and runways.

It is more than 95% likely that repeat pilots had some anticipation of rating the taxiways and runways that caused them to rate the smoother surfaces in the study slightly higher and the rough surfaces slightly lower than the other pilots.

The biases were not considered significant enough to remove the repeat pilots from the data set but are presented because they constitute a human factor[i] that slightly influenced the data.

4.6 Human Vibration Limits

The first row of Table 5 on page 48 (the weighted RMS line) may be compared with the numbers in Subclause 2.3 of Annex C (informative) of the ISO Standard [3], which for weighted RMS provides "approximate indications of likely reactions to various magnitudes of overall vibration total values in public transport:"

Table 7. ISO Standards for Discomfort

Weighted RMS m/s^2	**Discomfort Level**
0-0.315	not uncomfortable
0.315-0.63	a little uncomfortable
0.5-1.0	fairly uncomfortable
0.8-1.6	uncomfortable
1.25-2.5	very uncomfortable
> 2.0	extremely uncomfortable

The range of the weighted acceleration RMS values reported in this study is from 0.11 m/s^2 to 2.03 m/s^2. Bearing in mind that the cockpit accelerations were measured on the floor of the simulator cockpit below the pilot's seat, not at the seat as required by the standard, and that the measured signal bandwidth was somewhat limited compared to the standard, the range of the reported RMS values is compatible with what would be expected from the values in Annex C of the standard.

- At a weighted RMS value of 0.31 about 5% of pilots rated a taxiway ride as unacceptable. This is where the ISO standard for general vibration indicates the vibration is just becoming "a little uncomfortable." At a weighted RMS of 0.67 about 50% of pilots rated a taxiway ride as unacceptable. The Standard [3] says vibration is "fairly uncomfortable" at this level.

- Runways were rated unacceptable by 5% of pilots when their weighted RMS value was 0.35. The Standard says general vibration is "a little uncomfortable" at this level. About 50% of pilots rated a runway as unacceptable when its weighted RMS was 0.91 (m/s^2). This is where the standard classifies vibration as transitioning to from "fairly uncomfortable" to "uncomfortable."

These numbers suggest that the "a little uncomfortable" level of the Standard is closely related to the point at which 5% of pilots rate taxiways and runways as unacceptable and the "fairly uncomfortable" level is closely related to the point at which 50% of pilots rate taxiways and runways as unacceptable.

The Standard also provides a health guideline for (shock-related) vibration, stating that "caution with respect to health risks is indicated" for VDV exposures from 8.5 and 17 m/s$^{1.75}$. Tables 4

and 6 together show that *approximately half of the pilots rated taxiway and runway rides as unacceptable when vibrations reached into the lower end of the VDV caution range.*

4.7 Computation of best fit shifted logarithmic curves to data

The MATLAB code in Appendix H was used to compute a shift to be added to each ISO index in order to provide the best possible least squares logarithmic fit to numerical pilot ratings in the final study as a function of roughness index. The program computes the least squares fit, a 95% confidence interval for the fit, and 50%, 90% and 95% prediction intervals. The confidence intervals represent intervals within which it is 95% likely the true trend points should lie, given that the fitted curve is of the type assumed: linear, quadratic, etc. when noise (errors in measurement) has been removed. The prediction intervals provide ranges within which about 95% of data points should fall when 33 pilots rate taxiways or runways and their average numerical rating is calculated.

The shifted logarithmic intervals assume that

- the data follows a shifted logarithmic trend with the given shift, and
- the difference between the taxiway or runway average rating and the actual trend has a (bell-shaped) normal distribution with its mean on the trend and a standard deviation that is constant overall index values.

Another notable feature of the shifted logarithmic fit is that it is very close to that of the quadratic fit and that these two curves provide the smallest R^2 values over all fits attempted (which is not surprising since these fits have three free constants but the others have only two or one).

The closeness of the quadratic and shifted logarithmic fits as best fits suggests that the true trend of the data is very close to these curves (Figure 26).

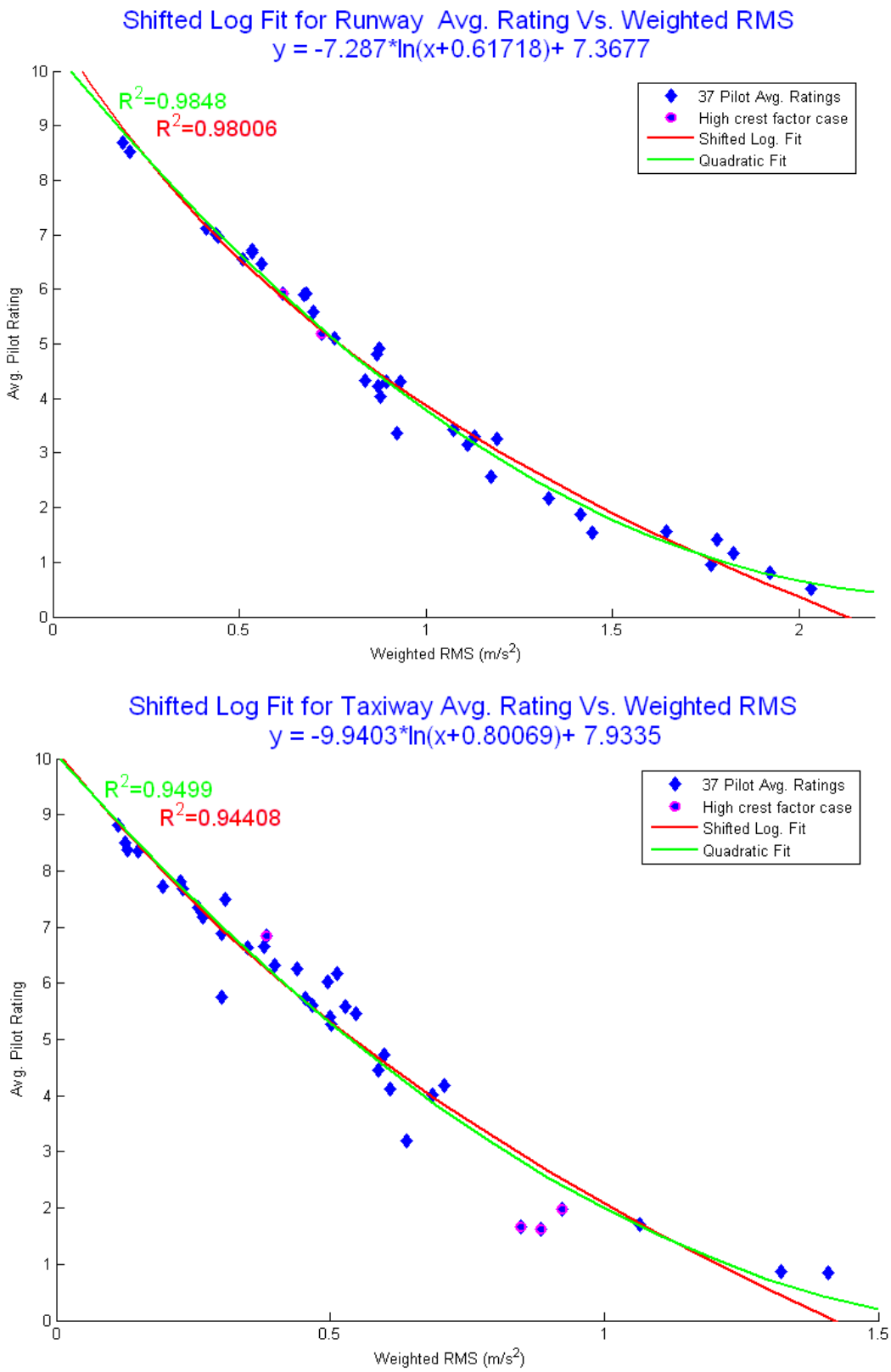


Figure 26. Comparison of the Quadratic and Shifted Logarithmic Fits to Pilot Rating Vs. Weighted RMS

The assumptions for the shifted logarithmic fits and intervals cannot be precisely true. For example, when the trend curve for taxiways or runways is near a rating of 0 or 10 the rating uncertainty cannot vary symmetrically about the trend (fit) as the model requires. However, the fact that the 95% prediction intervals contain about 95% of the averages in the data above suggests that the normality assumption of the model is not too far from correct (Figures 27-34).

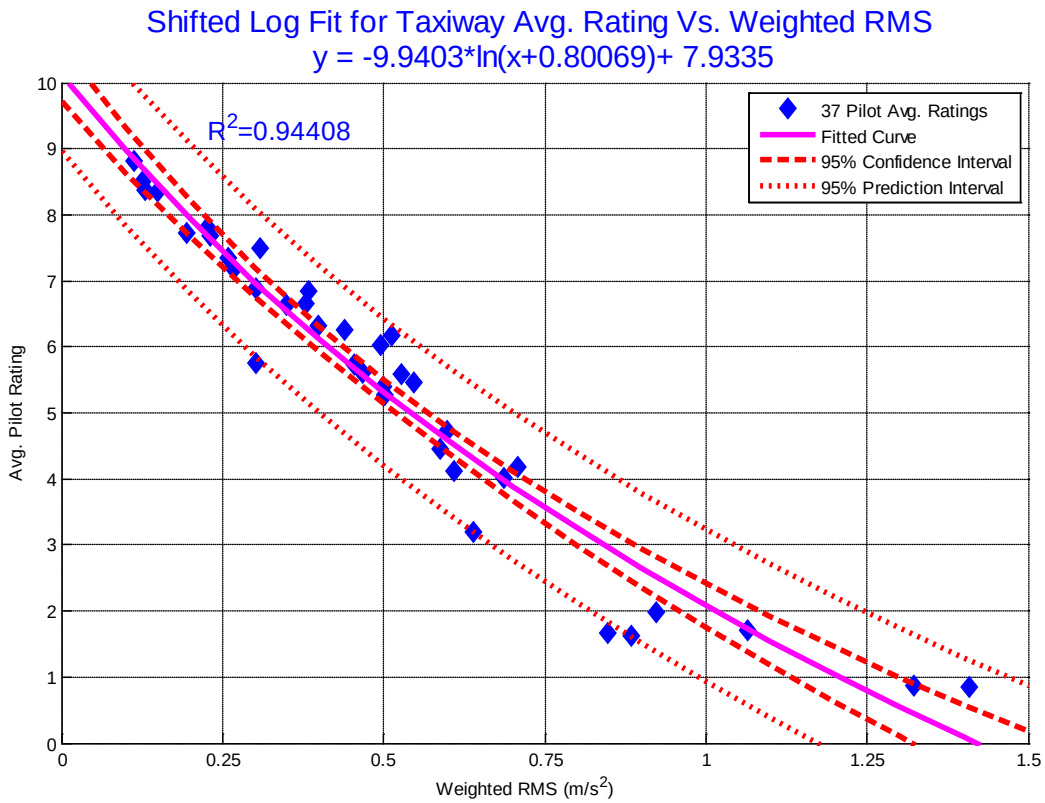


Figure 27. Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. Weighted RMS

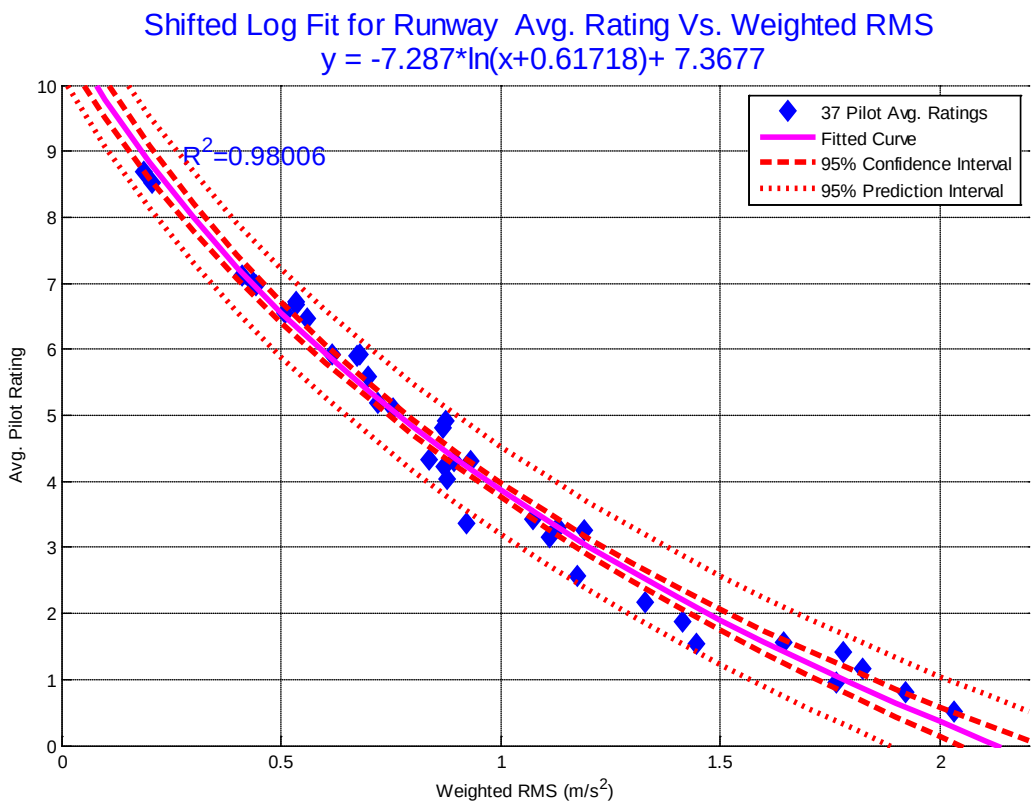


Figure 28. Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. Weighted RMS

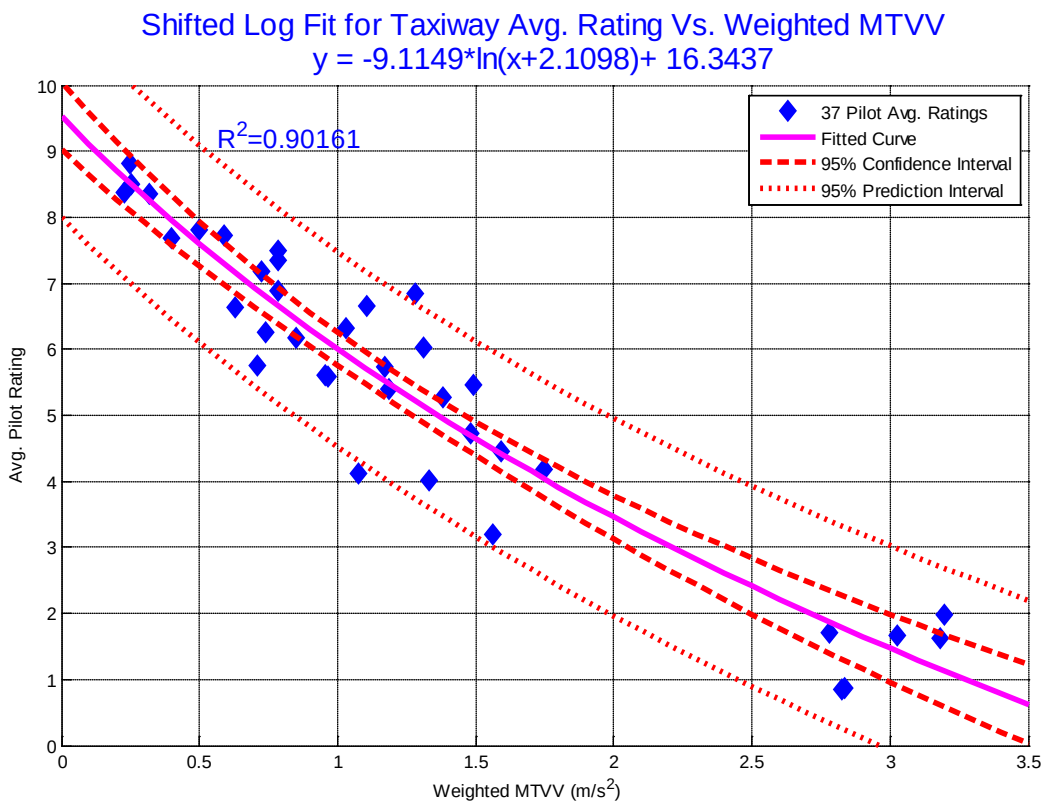


Figure 29. Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. Weighted MTVV

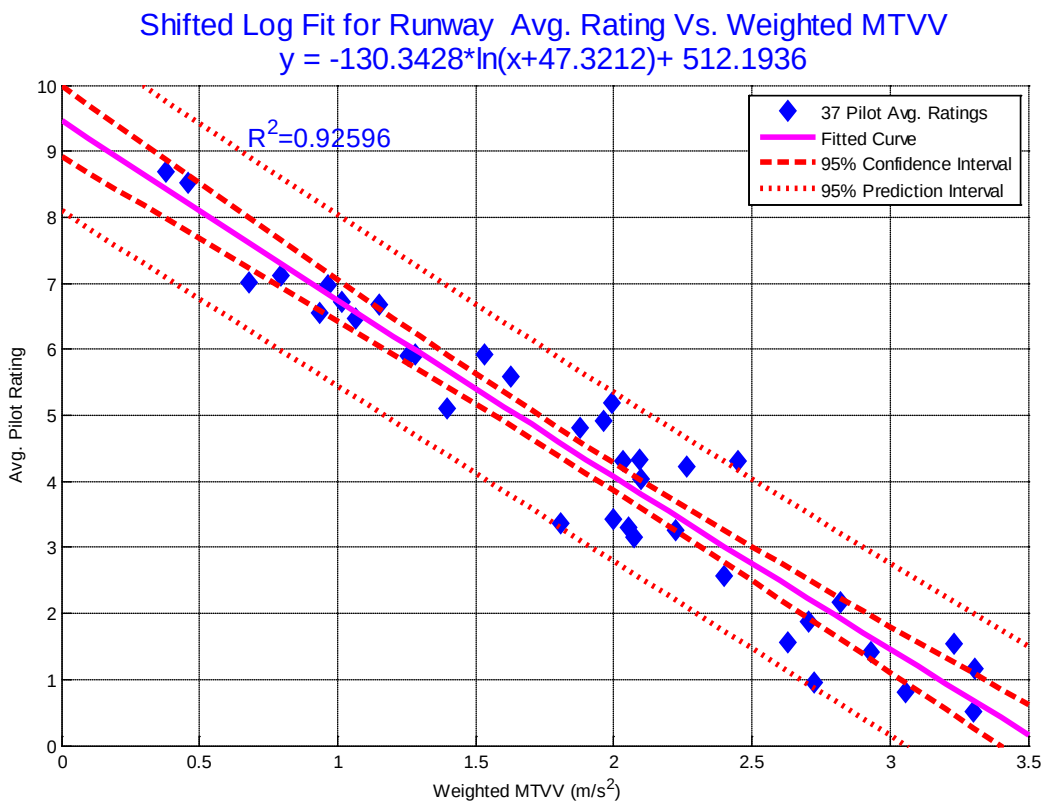


Figure 30. Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. Weighted MTVV

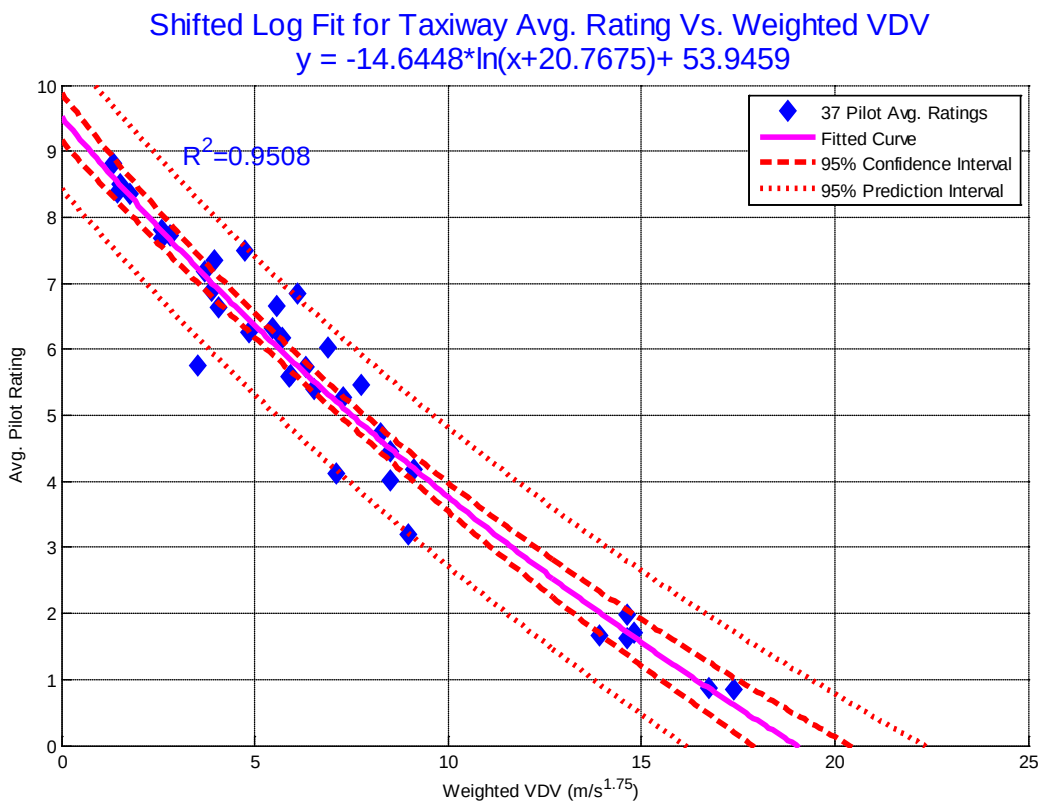


Figure 31. Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. Weighted VDV

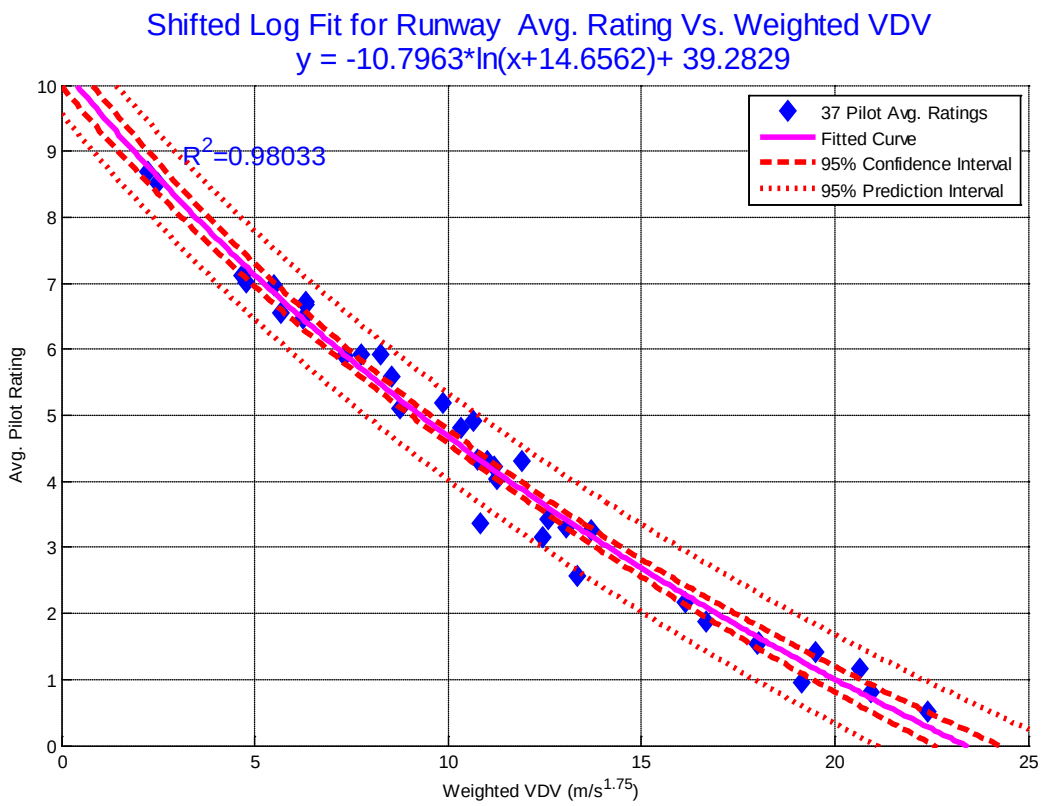


Figure 32. Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. Weighted VDV

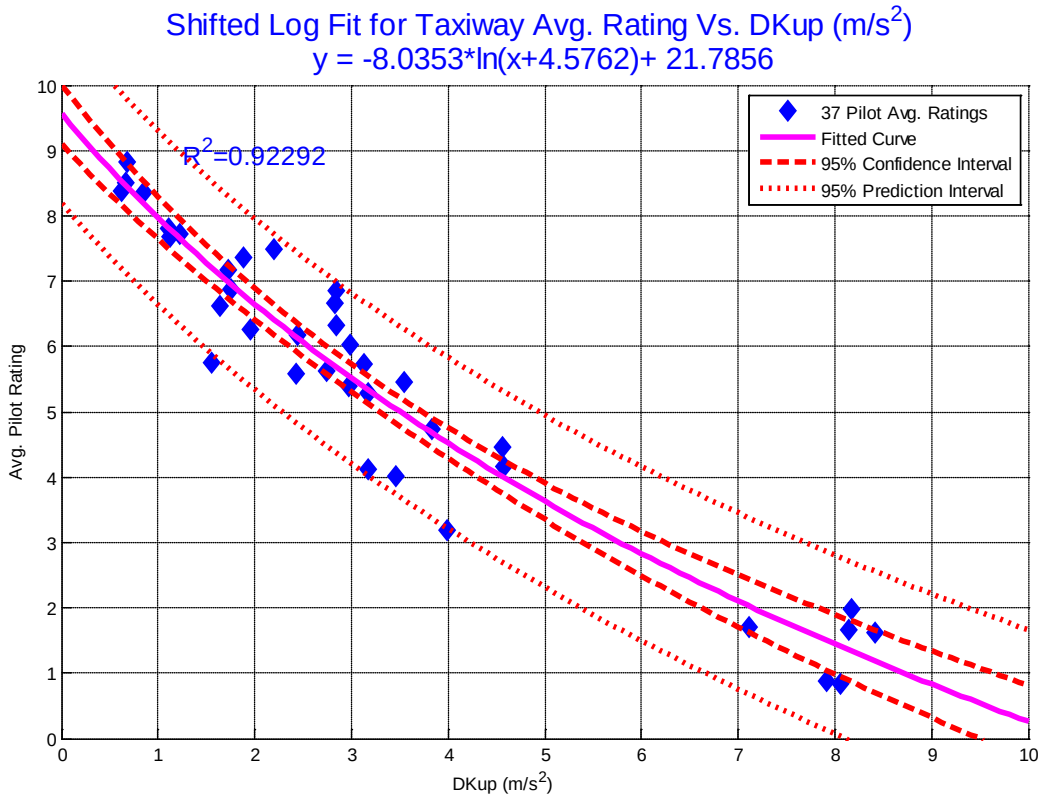


Figure 33. Shifted Logarithmic Fit for Pilot Average Taxiway Rating Vs. DKup

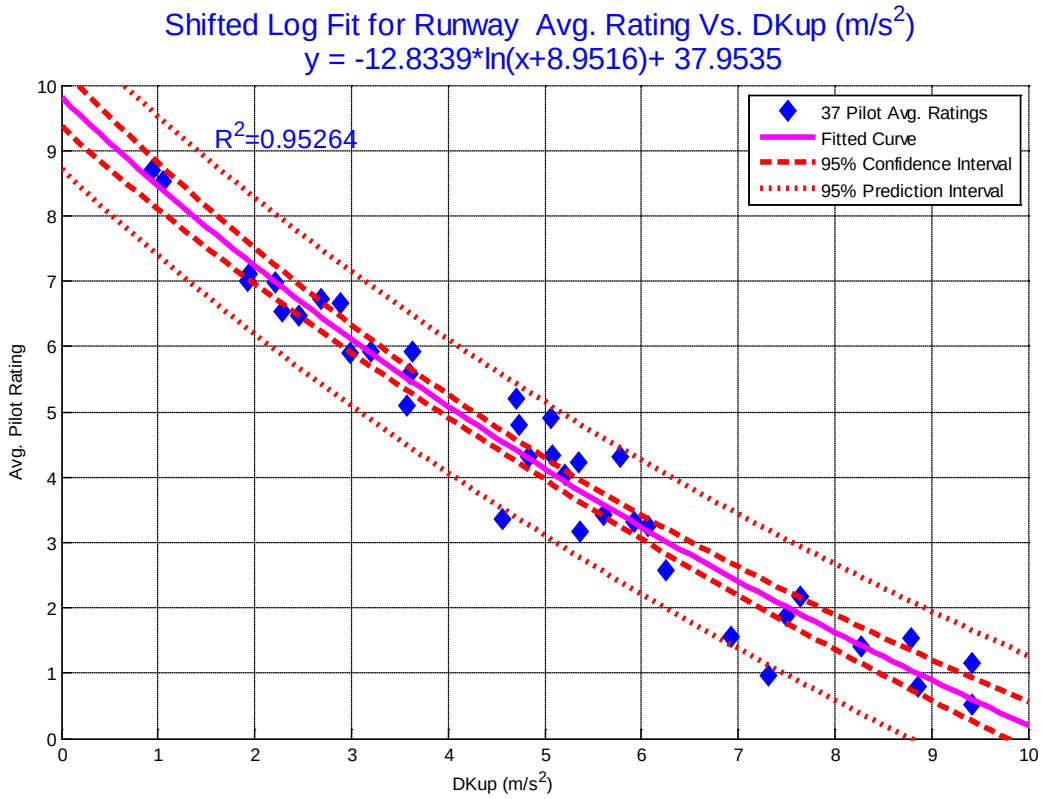


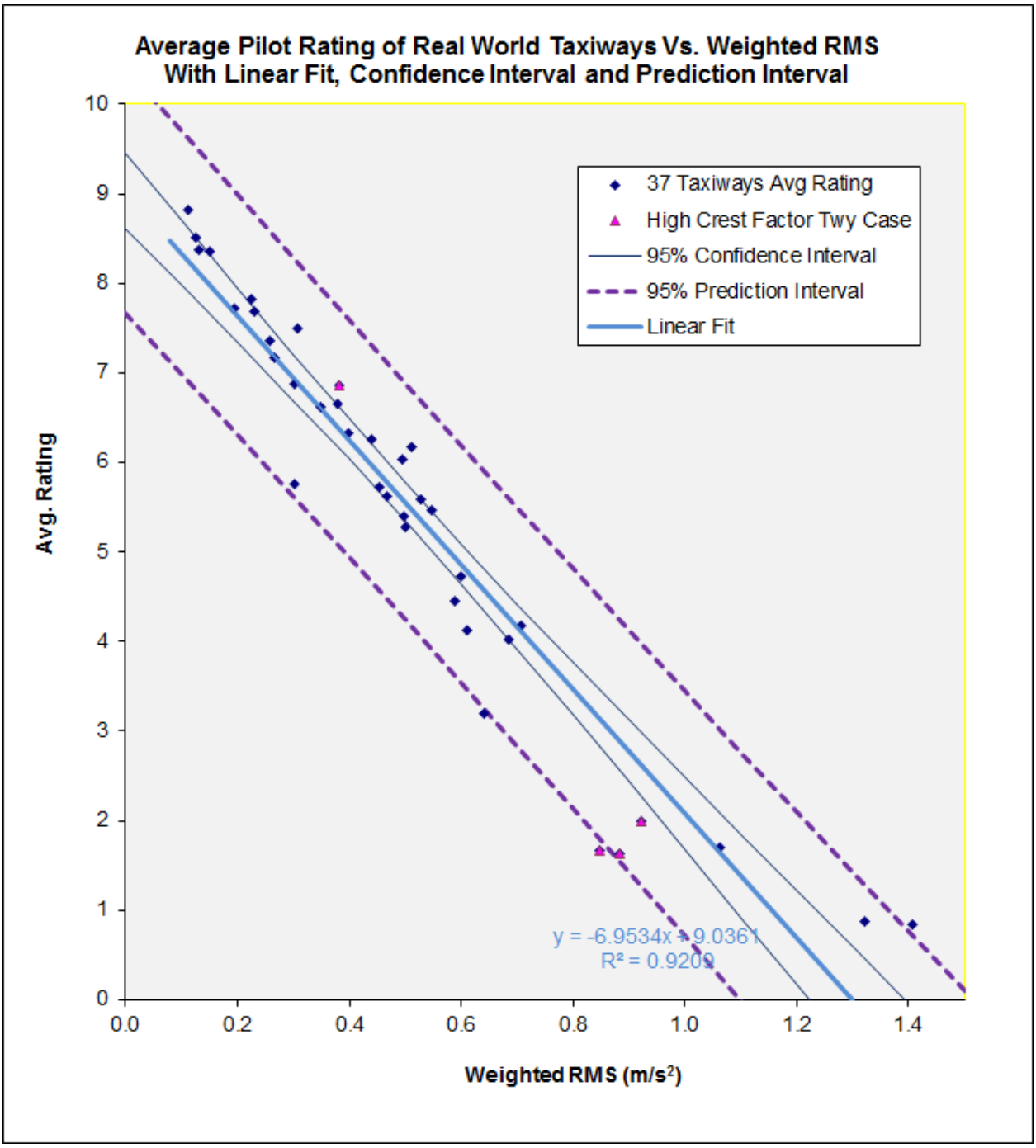
Figure 34. Shifted Logarithmic Fit for Pilot Average Runway Rating Vs. DKup

In these plots,

- *the narrower confidence and prediction intervals in the runway cases indicate that pilots were more uniform in their ratings of runways,*
- *the narrower intervals in the fits vs. weighted RMS and weighted VDV indicate that pilot responses follow a more well-defined trend when plotted against these indices (as anticipated from the correlation coefficients of Section 4.2), and*
- *the large shift for the runway weighted MTVV fit occurs because the best fit is nearly linear (and the shifted logarithmic fit is not appropriate in such a case).*

4.8 Comparison of the best shifted logarithmic fit to other fits

Confidence intervals and prediction intervals were also calculated for the linear, exponential, and logarithmic fits with the final data set. These intervals for fits of average pilot rating as a function of weighted RMS appear in Figures 35, 36, and 37.


Average Pilot Rating of Real World Taxiways Vs. Weighted RMS
With Linear Fit, Confidence Interval and Prediction Interval
37 Taxiways Avg Rating
High Crest Factor Twy Case
95% Confidence Interval
95% Prediction Interval
Linear Fit
y = -6.9534x + 9.0361
R² = 0.9209
Avg. Rating
Weighted RMS (m/s²)
0
1
2
3
4
5
6
7
8
9
10
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

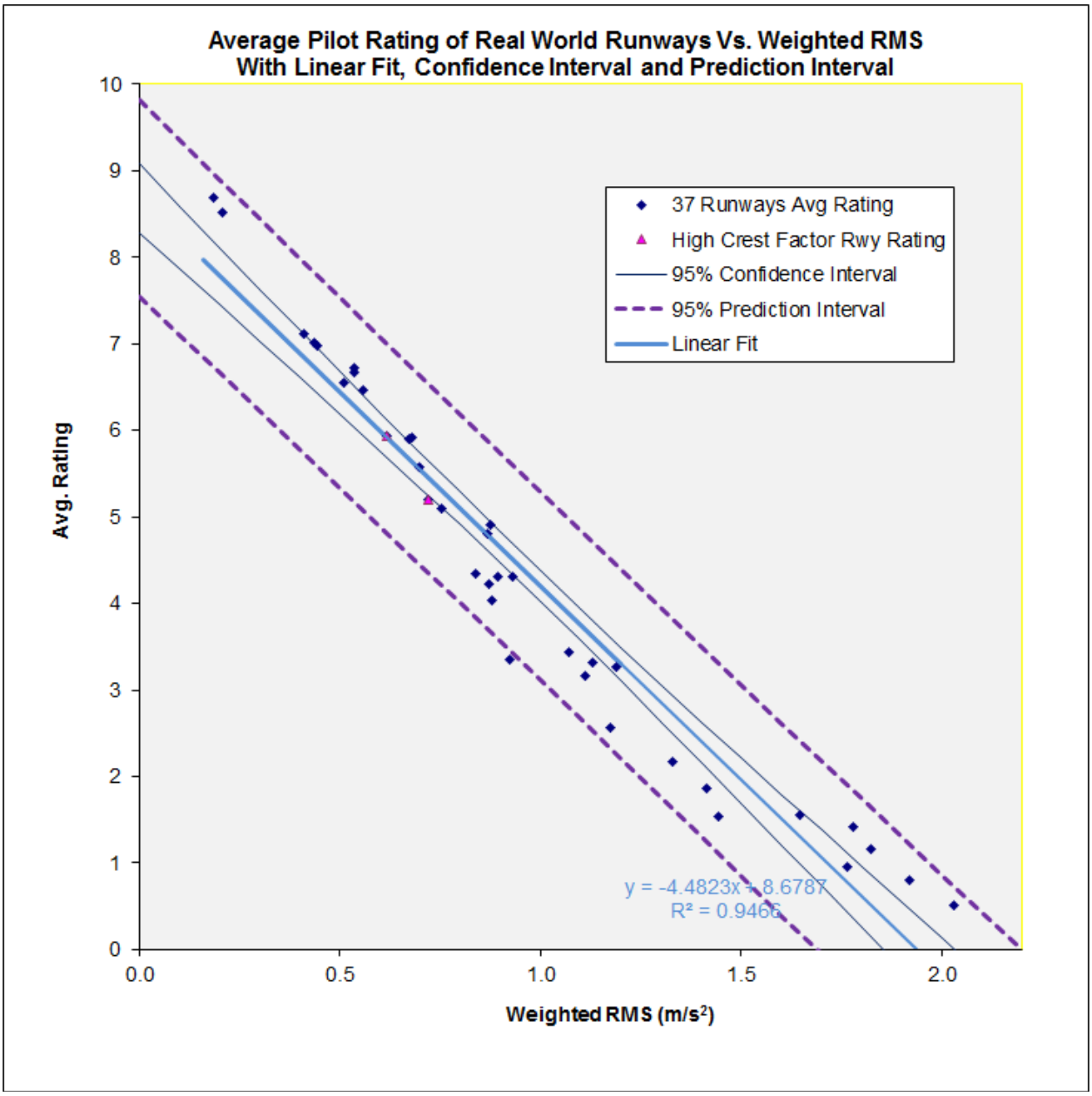
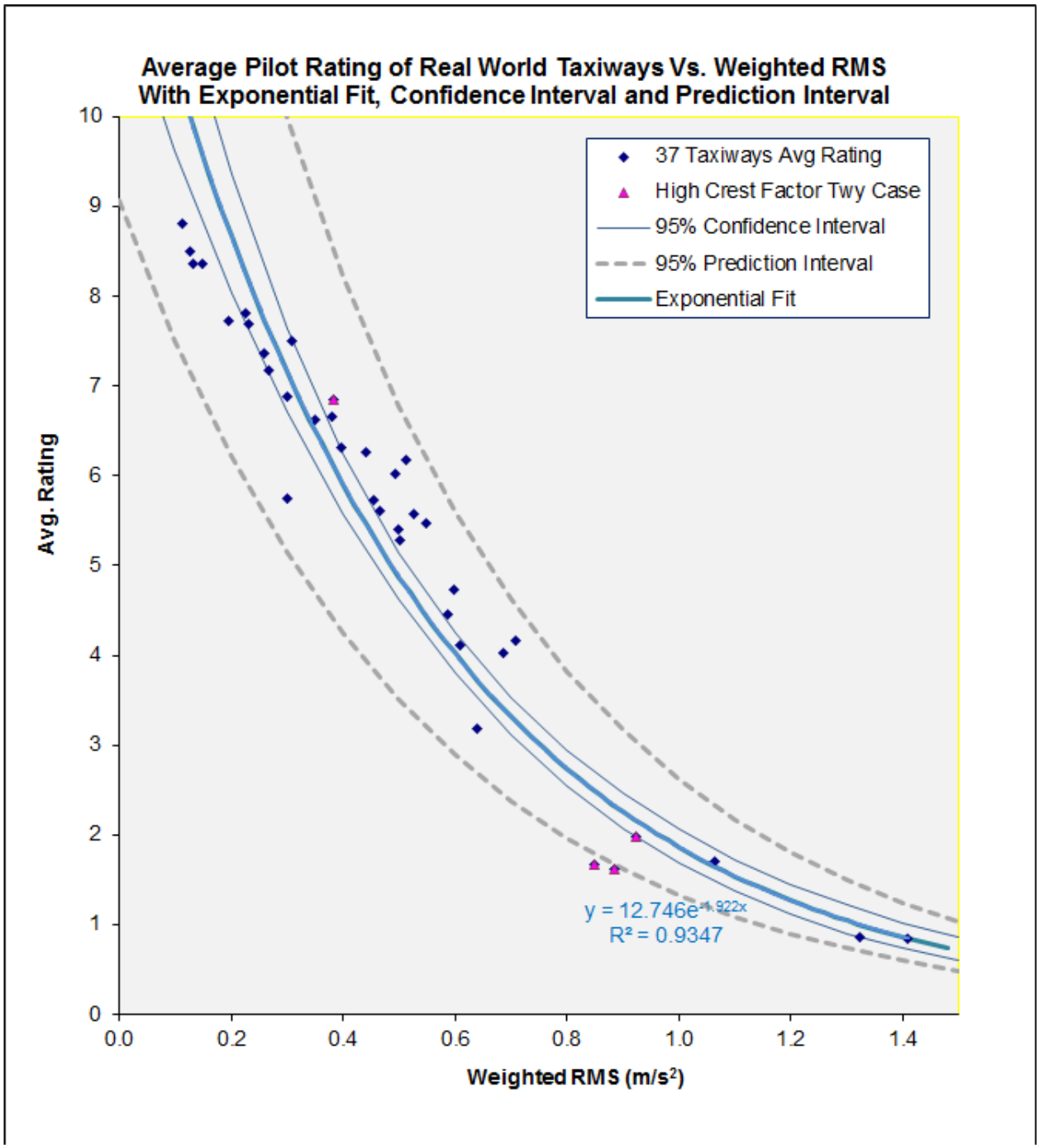


Figure 35. Linear Fit of Rating vs. Weighted RMS with Confidence & Prediction Intervals


Average Pilot Rating of Real World Taxiways Vs. Weighted RMS
With Exponential Fit, Confidence Interval and Prediction Interval
37 Taxiways Avg Rating
High Crest Factor Twy Case
95% Confidence Interval
95% Prediction Interval
Exponential Fit
y = 12.746e^-1.922x
R² = 0.9347
Avg. Rating
Weighted RMS (m/s²)
0
1
2
3
4
5
6
7
8
9
10
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

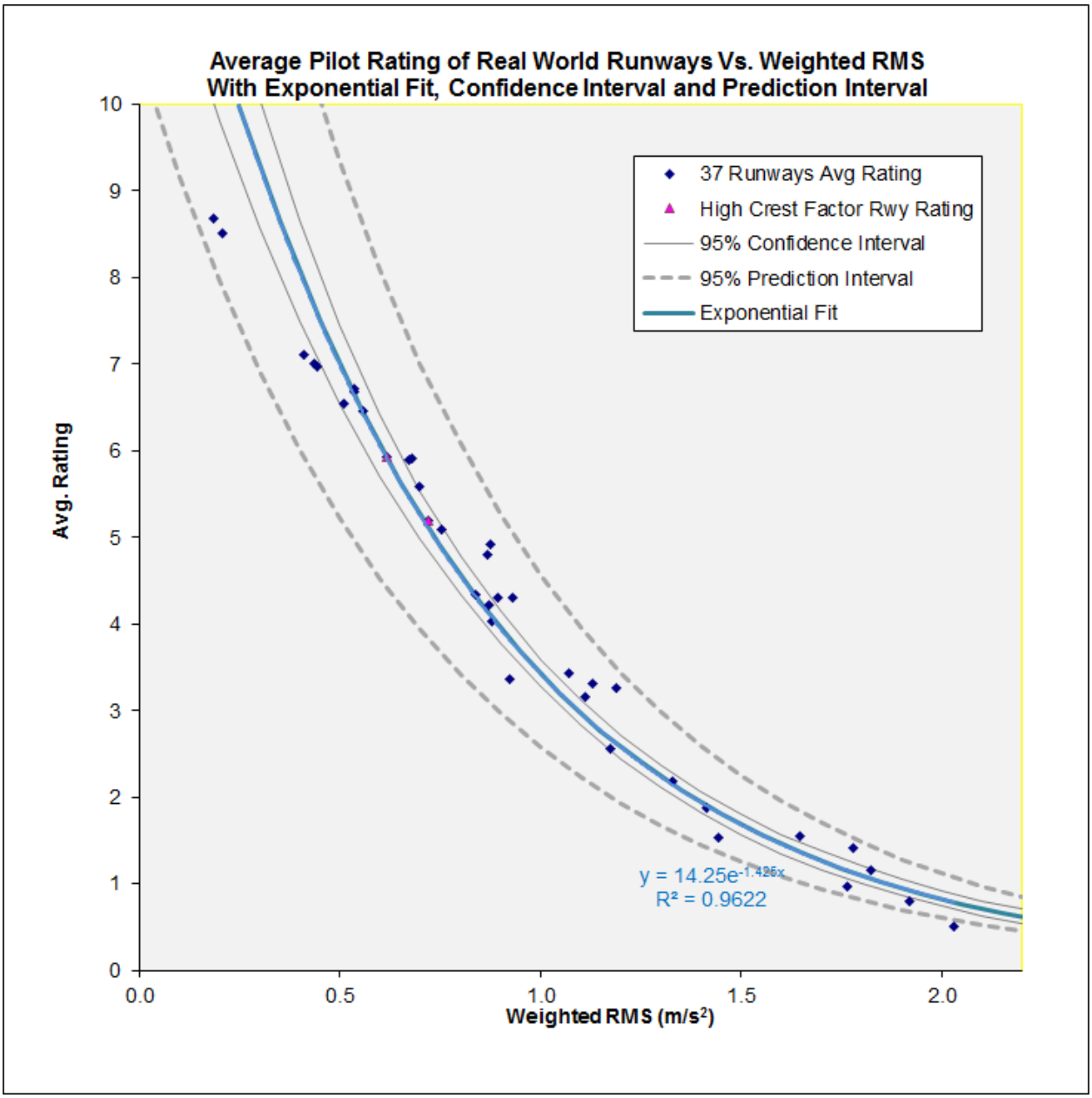
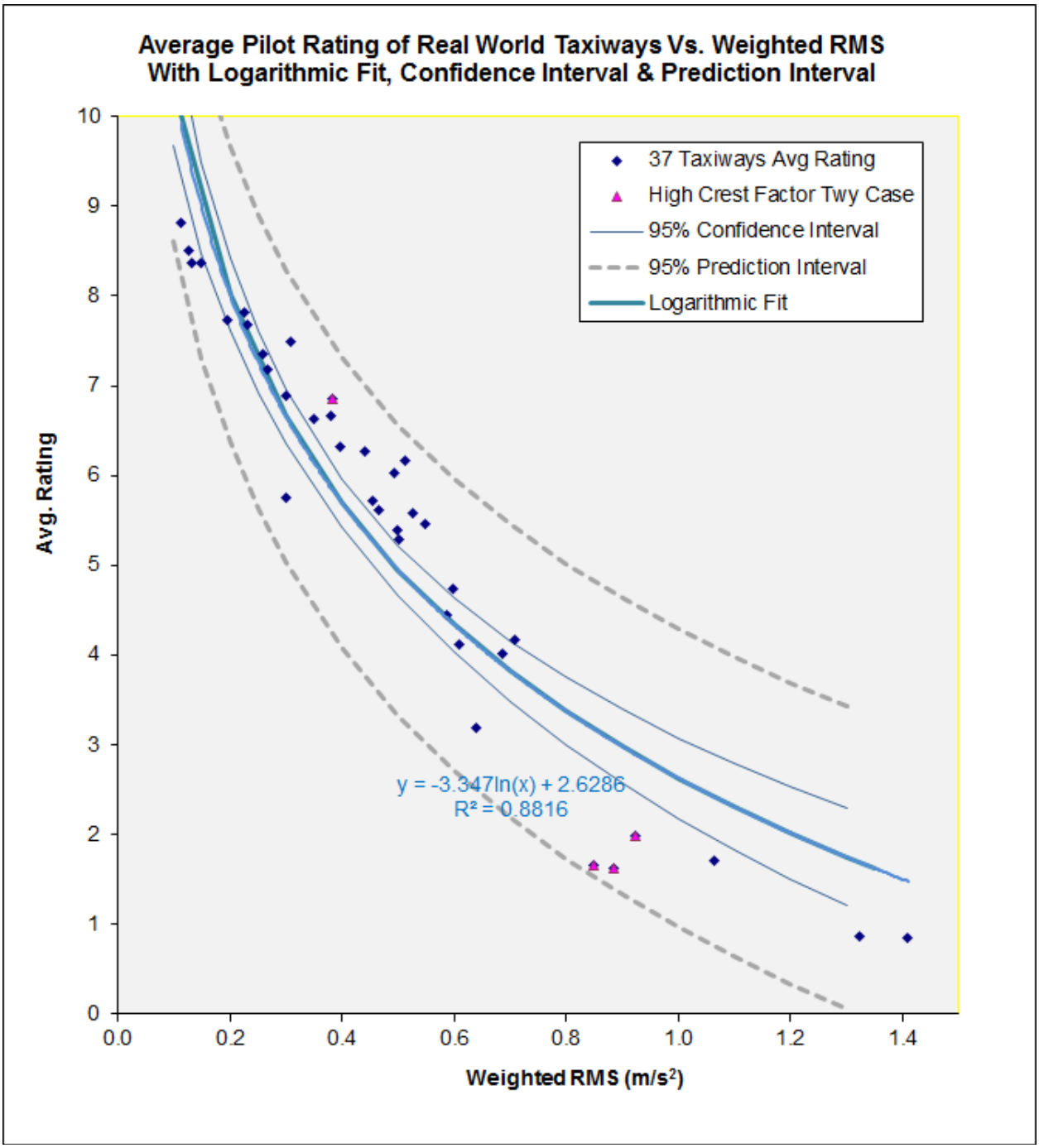


Figure 36. Exponential Fit of Rating vs. Weighted RMS With Confidence & Prediction Intervals


Average Pilot Rating of Real World Taxiways Vs. Weighted RMS
With Logarithmic Fit, Confidence Interval & Prediction Interval
37 Taxiways Avg Rating
High Crest Factor Twy Case
95% Confidence Interval
95% Prediction Interval
Logarithmic Fit
y = -3.347ln(x) + 2.6286
R² = 0.8816
Avg. Rating
Weighted RMS (m/s²)
0
1
2
3
4
5
6
7
8
9
10
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

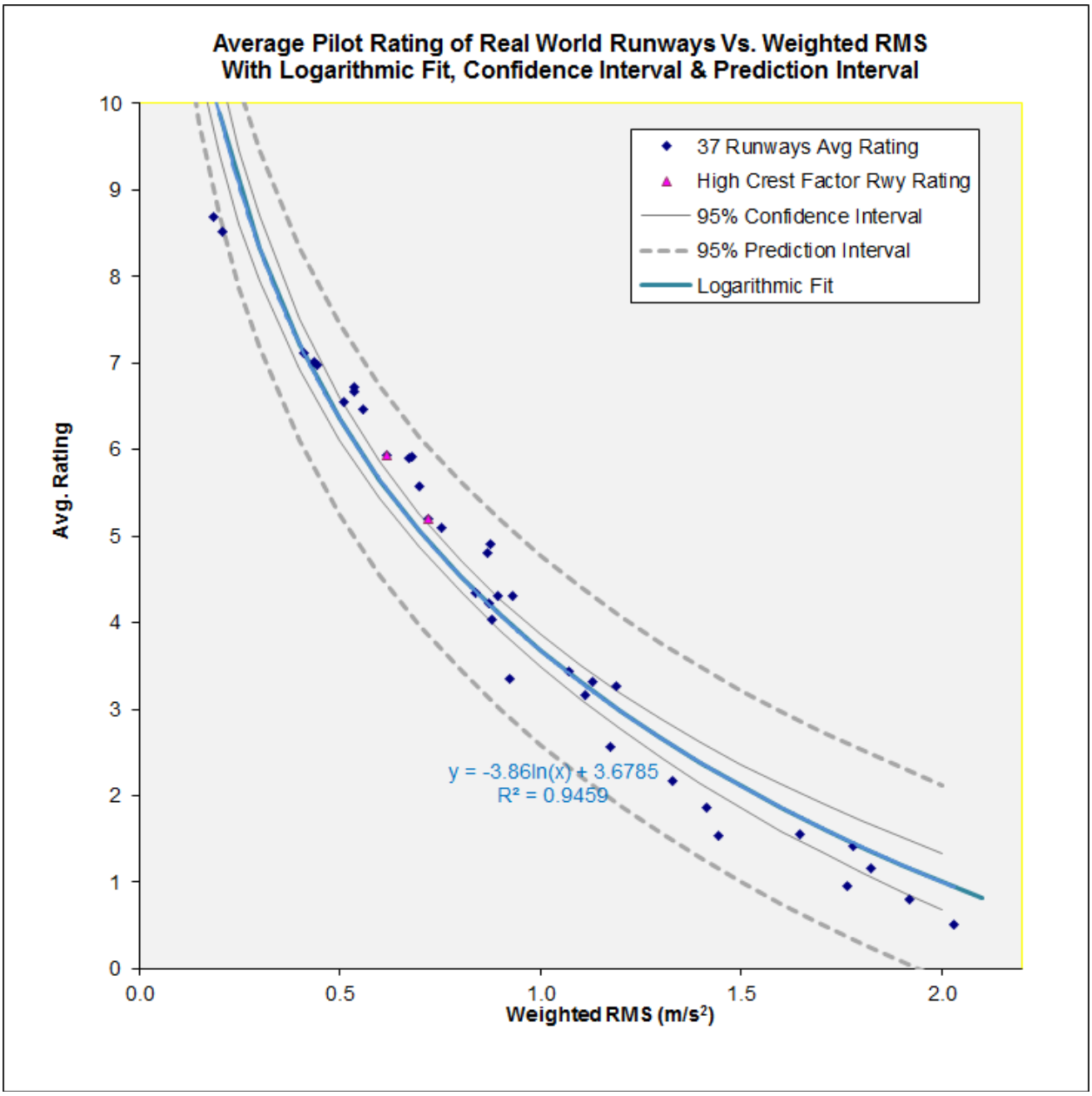


Figure 37. Log. Fit of Rating vs. Weighted RMS With Confidence & Prediction Intervals

The prediction intervals help determine where a fit is insufficient by indicating when data points are substantially off to one side of the trend. For example, the linear fits have many data points well above the trend line at small and large values of weighted RMS and well below the trend line in the middle. The exponential and unshifted logarithmic fits similarly display too much curvature for small values of weighted RMS.

4.9 Comparison of best fits when data from the preliminary study is included

Figure 38 shows the combined quadratic trends[7] for runways as a function of weighted RMS when the 20 average numerical ratings from runways in the preliminary tests and from 37 runways in final tests are all fit.

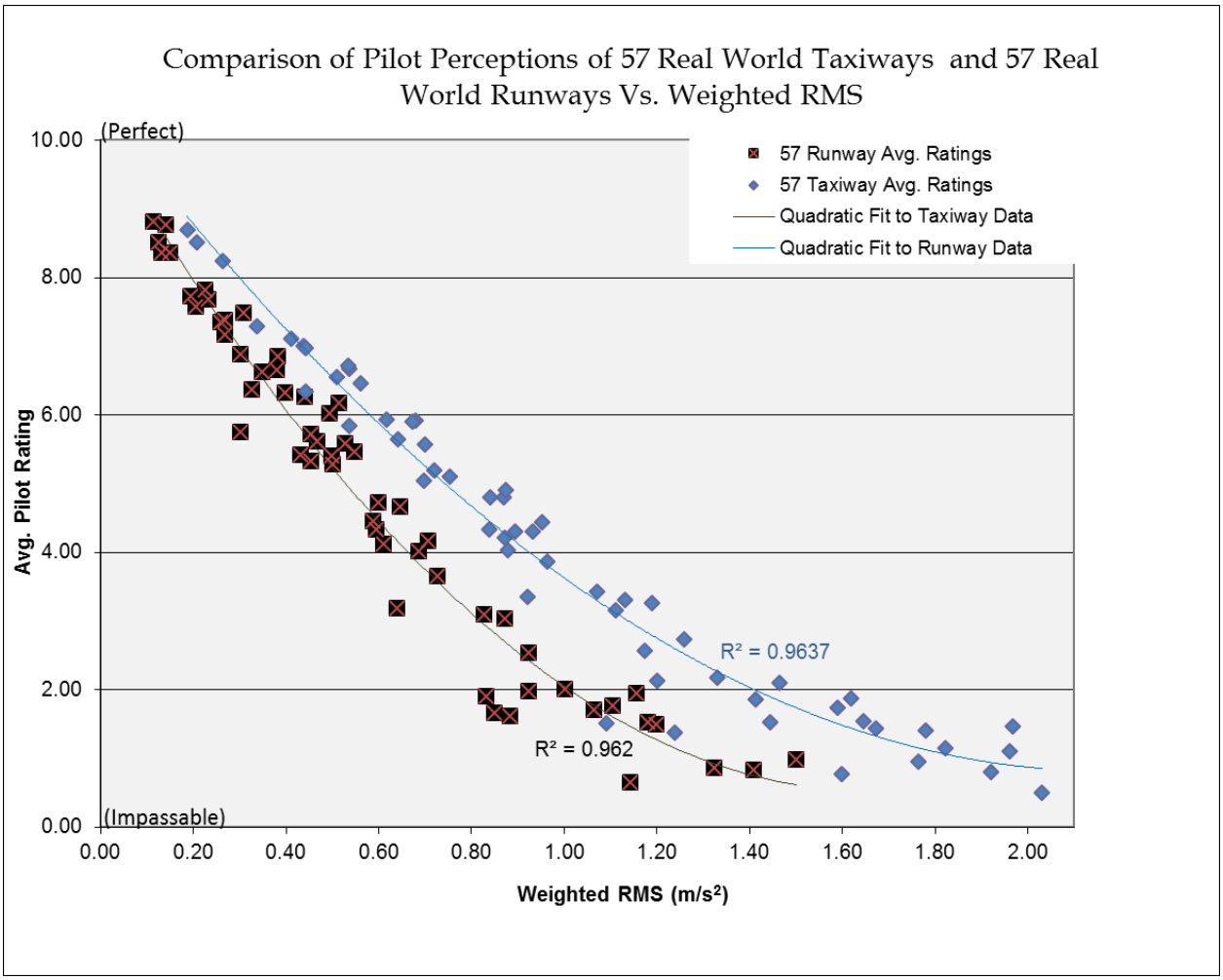


Figure 38. Combined Preliminary and Final Runway Average Ratings vs. Weighted RMS

This Figure illustrates that combining runway average ratings from final and preliminary testing yields a single trend line for taxiways and for runways, indicating consistency in the two test sessions.

When all 1572 individual pilot numerical ratings of taxiways or runways (from 36 pilots x 37 taxiways/runways plus 12 pilots x 20 taxiways/runways) are combined the confidence intervals and prediction intervals for individual pilot responses are obtained, as shown in Figures 39 through 42 for the shifted logarithm fits.

[7] The trend equations are not given here because the fits shown used equal weighting for averages computed with 33 pilots from the final study and averages computed using 12 pilots from the preliminary study and are not valid.

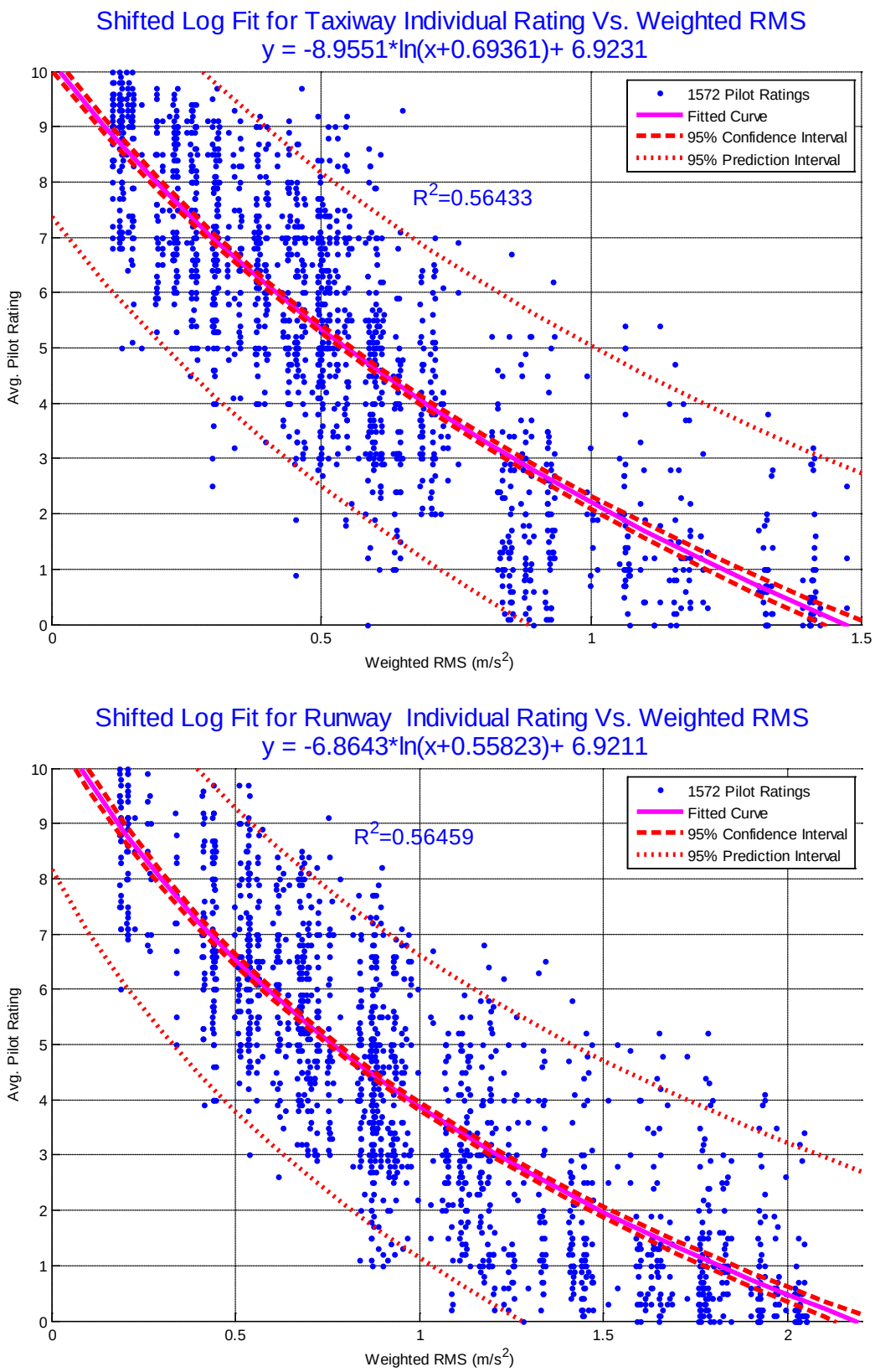


Figure 39. Confidence & Prediction Intervals for Logarithmic Fits with Weighted RMS

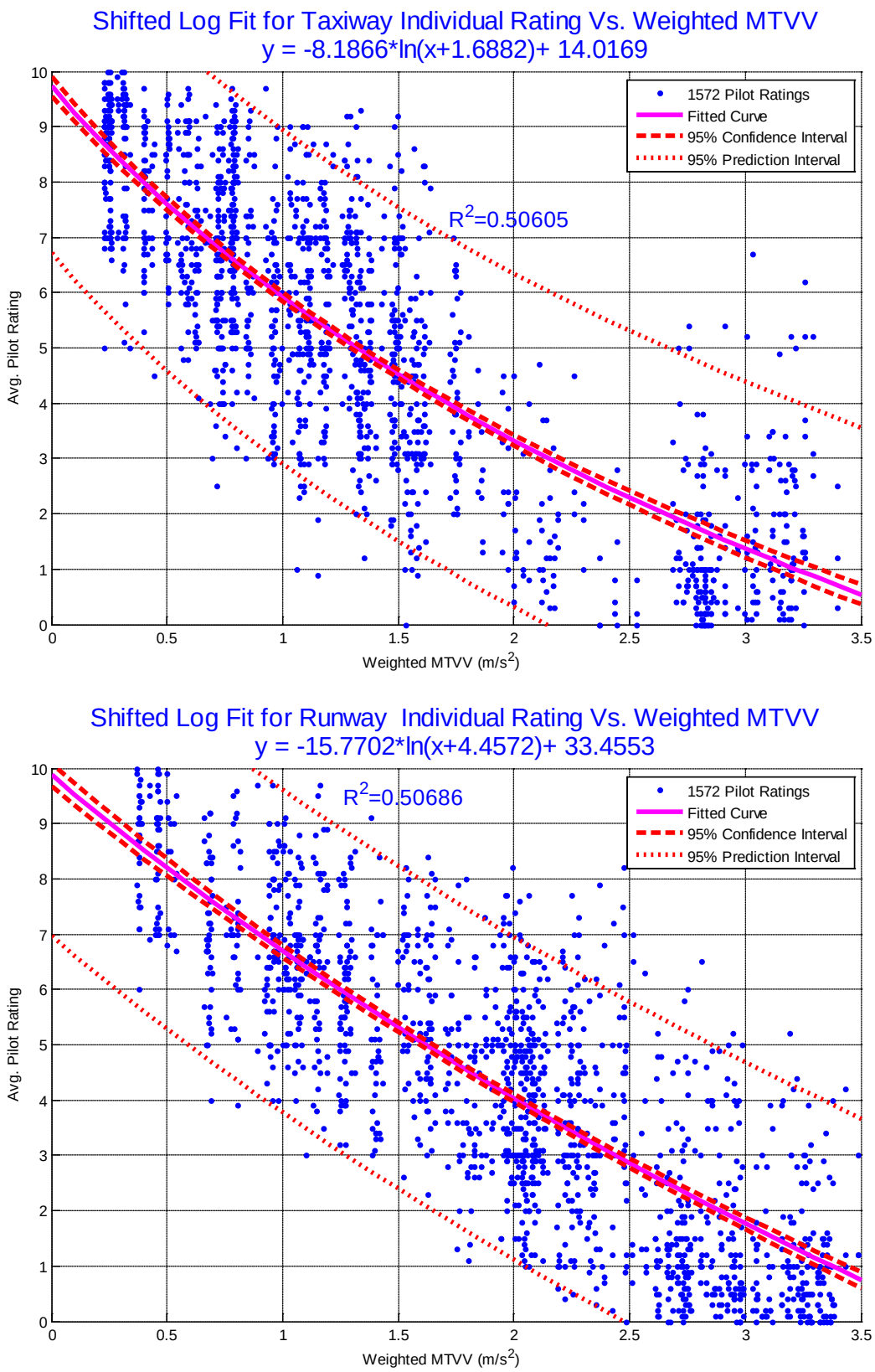


Figure 40. Confidence & Prediction Intervals for Logarithmic Fits with Weighted MTVV

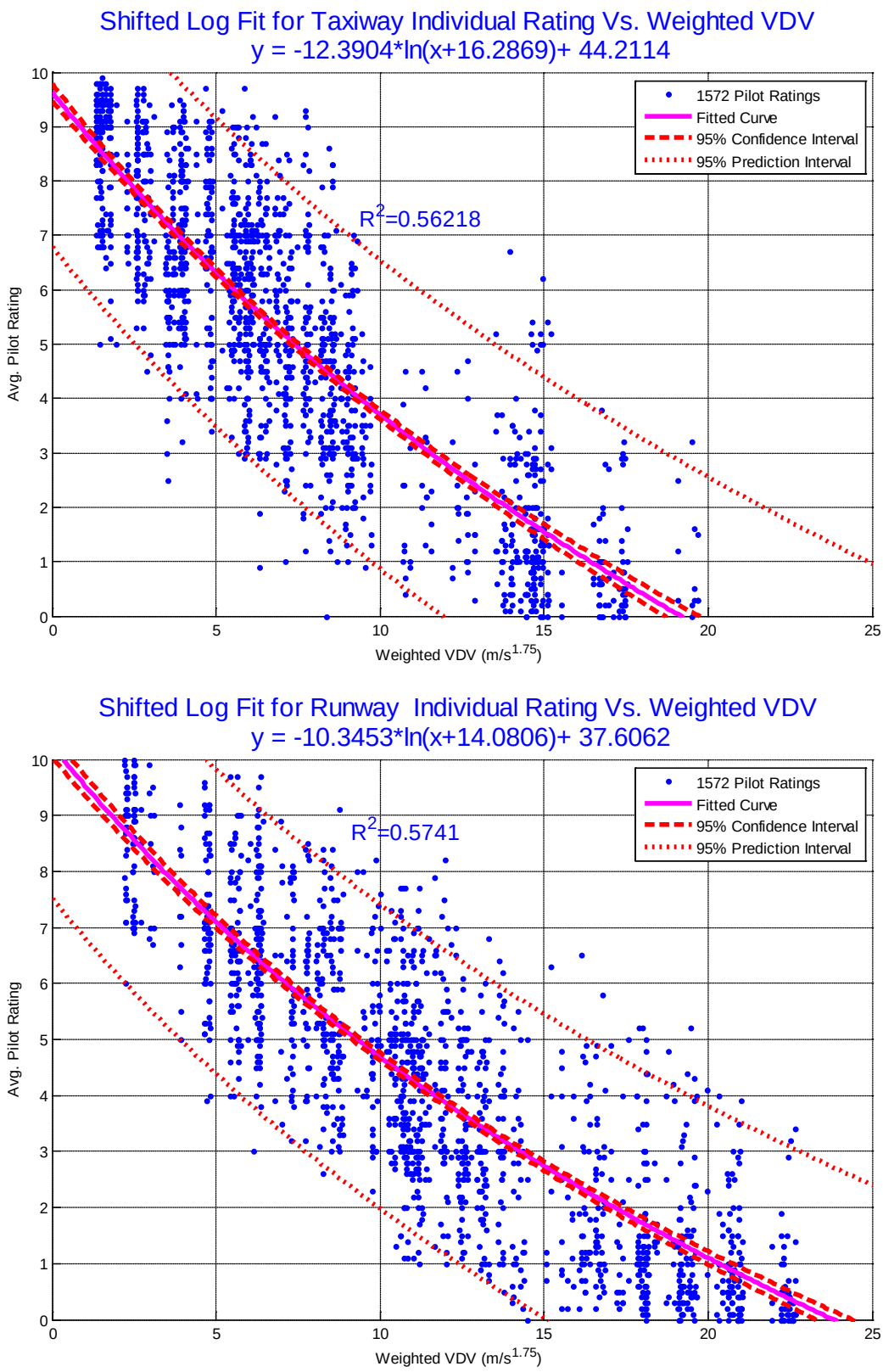
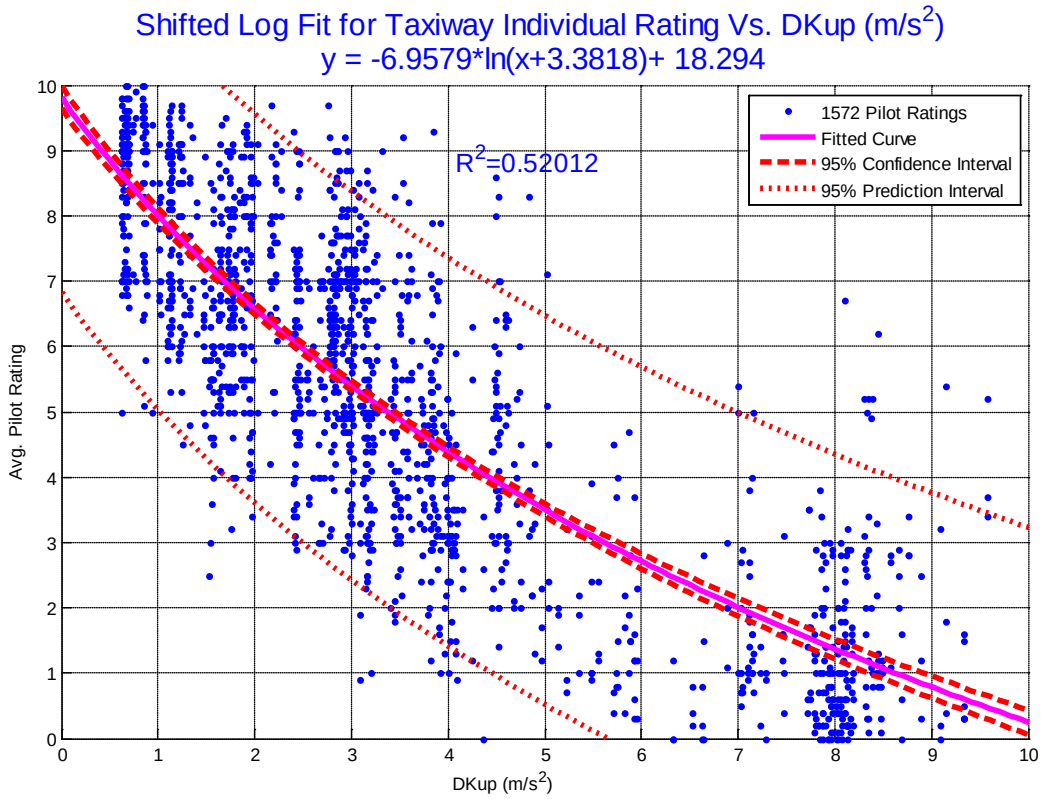


Figure 41. Confidence & Prediction Intervals for Logarithmic Fits with Weighted VDV



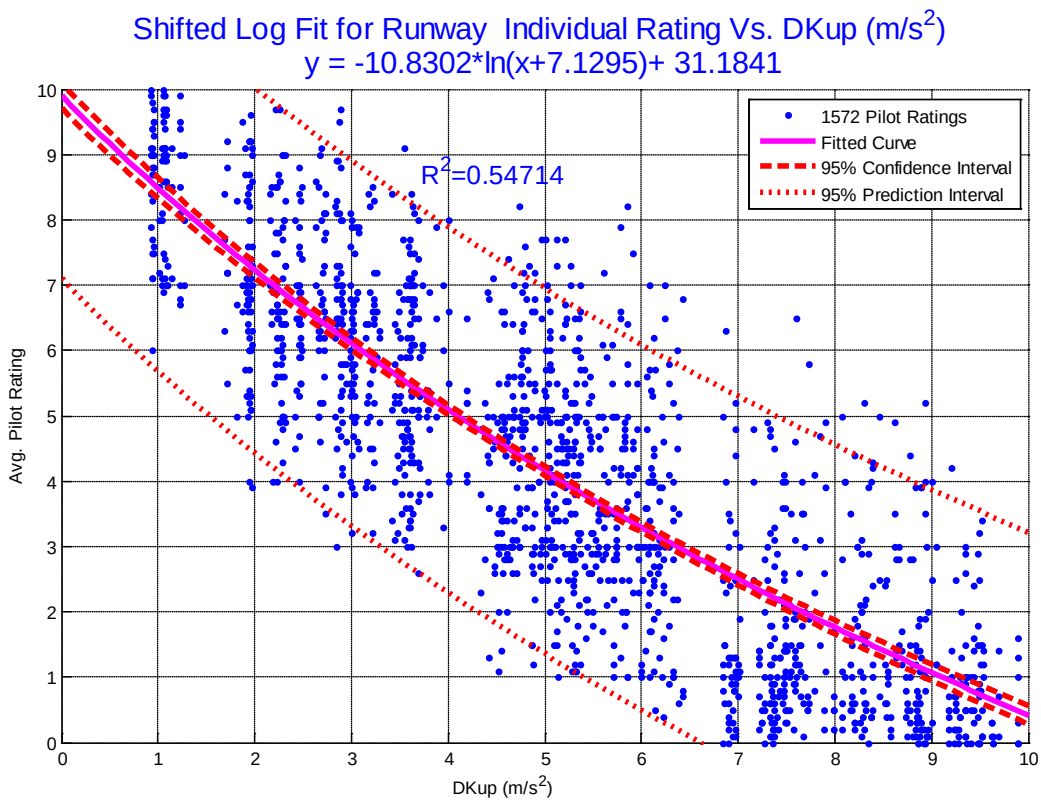
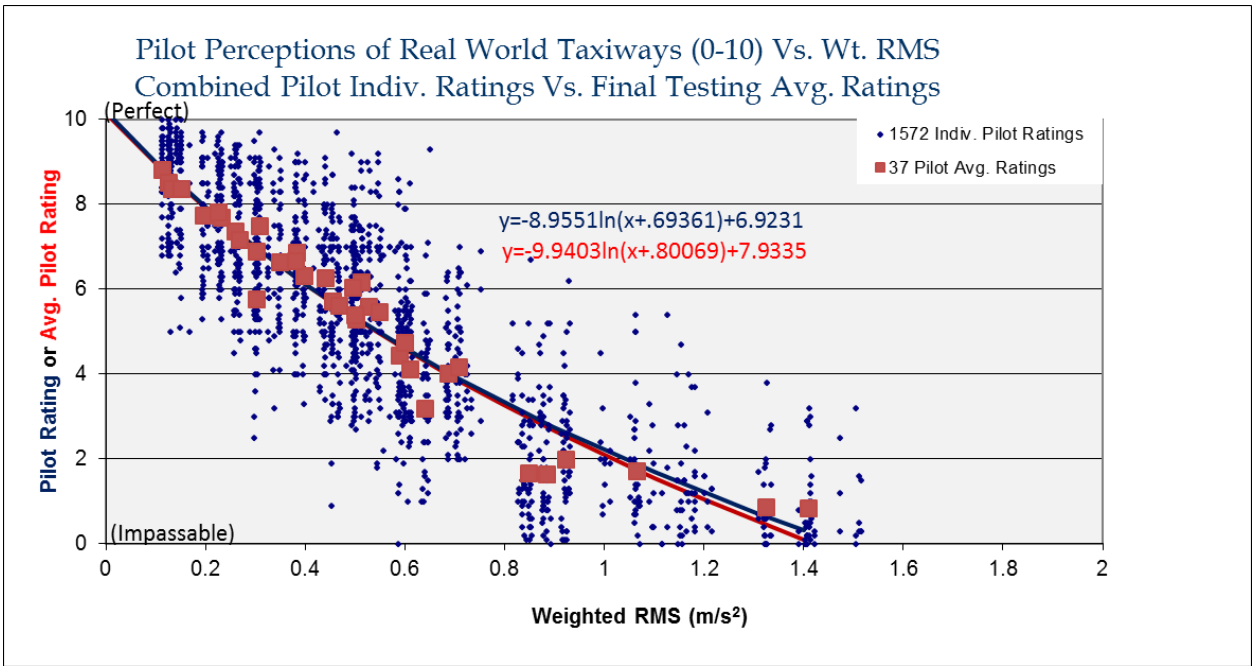


Figure 42. Confidence & Prediction Intervals for Logarithmic Fits with DKup

Each dot in these figures represents an individual pilot response. The confidence intervals are narrow because of the large number of samples. Many dots coincide - especially in the cases of pilot responses of 0 or 10, so it is not possible to count 1572 points or to determine precisely how many are on each side of the trend from the picture alone.

The least squares shifted logarithmic fits to the 1572 individual pilot numerical ratings and to the 37 taxiways/runways of final testing with pilot average ratings nearly coincide. Figure 43 compares such trends for the case of fits using weighted RMS and shows that the blue and red trends (for individual and average pilot responses) nearly coincide.



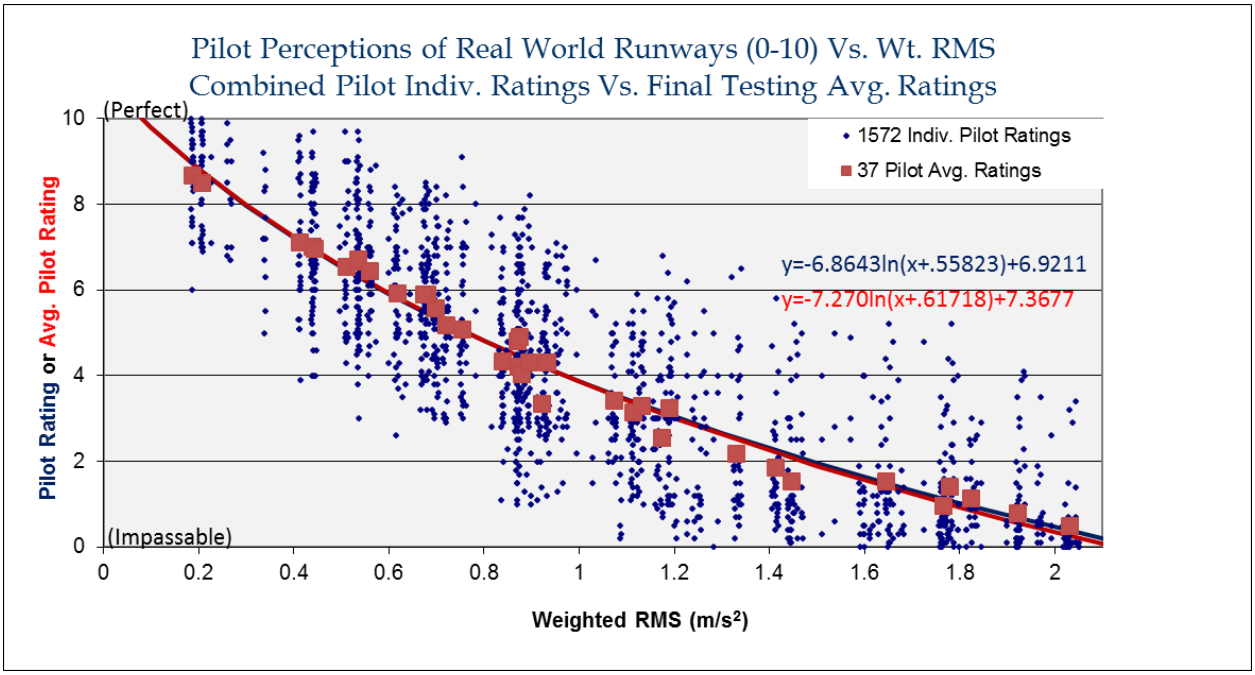
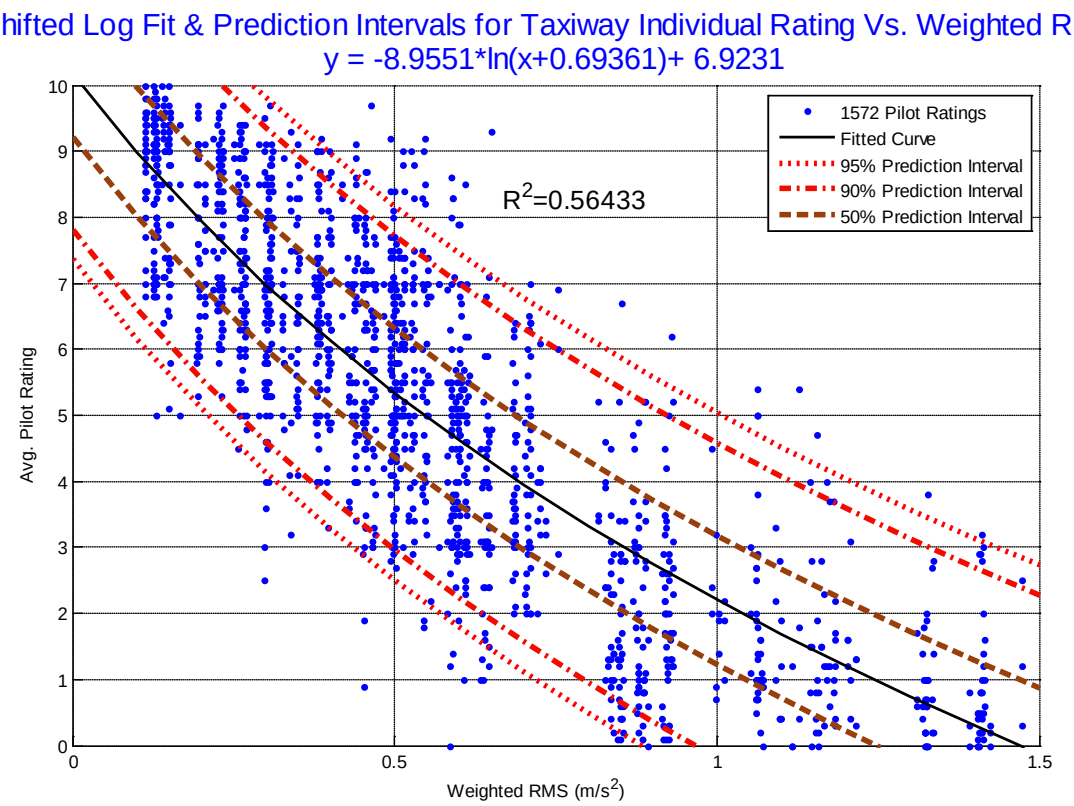


Figure 43. Comparison of Shifted Logarithmic Trends Using All Pilot Numerical Ratings vs. Pilot Average Ratings from Final Simulator Runs

The scatter in individual pilot responses, the same that appears in Figure 39, has a trend with very little uncertainty - as was indicated in that figure by the narrow confidence interval, despite the wide range of individual responses. Figure 43 shows agreement of the trend of individual responses with the trend of the average pilot responses. The wide range of responses is in agreement with the wide prediction interval shown in Figure 39.

4.10 Variation in Pilot Ratings

Similar prediction intervals to the 95% intervals in Figure 42 can be constructed for any percentage. For example, Figure 44 shows the 95%, 90% and 50% intervals as a function of weighted RMS.



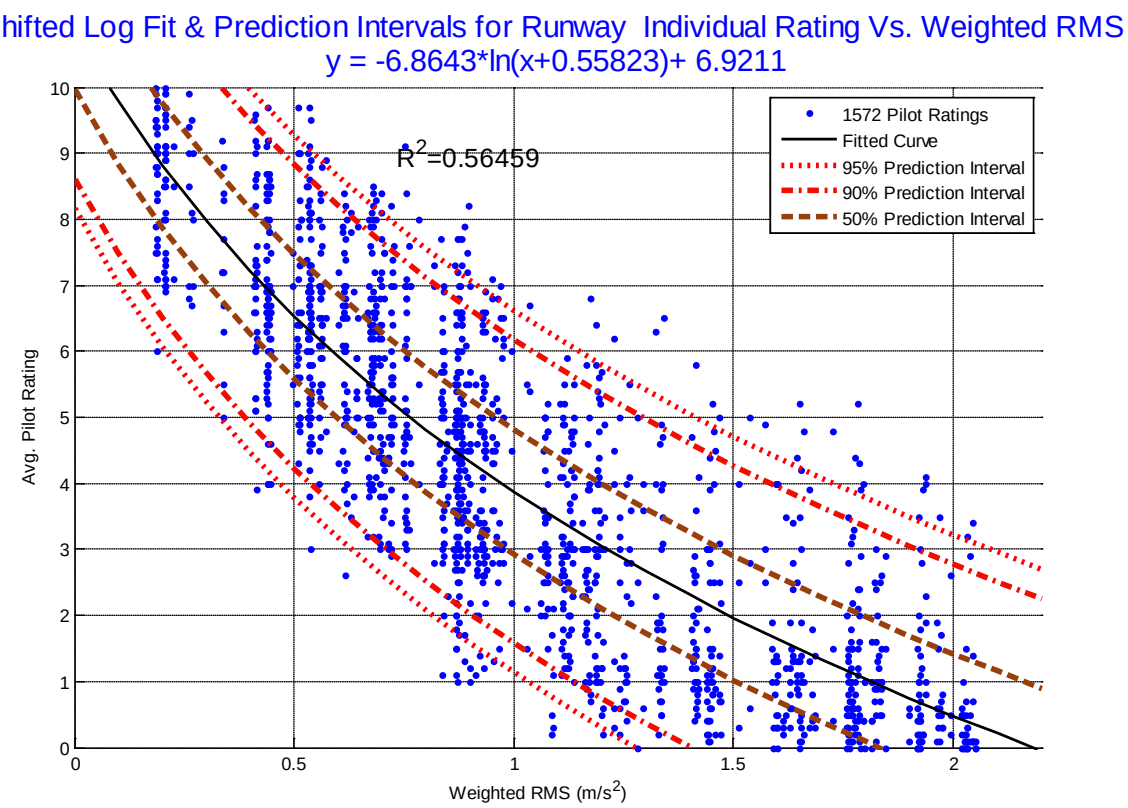


Figure 44. 95%, 90% and 50% Prediction Intervals for Logarithmic Fits vs. Weighted RMS

These prediction intervals provide, for each value of an ISO index, a response range within which approximately 95%, 90% and 50% of individual pilots are expected to rate a ride given its weighted RMS value.

For example, a taxiway with a weighted RMS rating of 0.5 will have about 95% chance of a random pilot rating it between 2.5 and 8.2 on the 0-10 scale, a 90% chance of a rating between 3.0 and 7.7 and a 50% chance of a rating between 4.4 and 6.4; and similarly, a runway with a weighted RMS rating of 0.5 will have a 95% chance of a random pilot rating it between 3.8 and 9.2 on the 0-10 scale, a 90% chance of a rating between 4.2 and 8.8 and a 50% chance of a rating between 5.6 and 7.5.

4.11 Comparison of simulator and ProFAA roughness model results

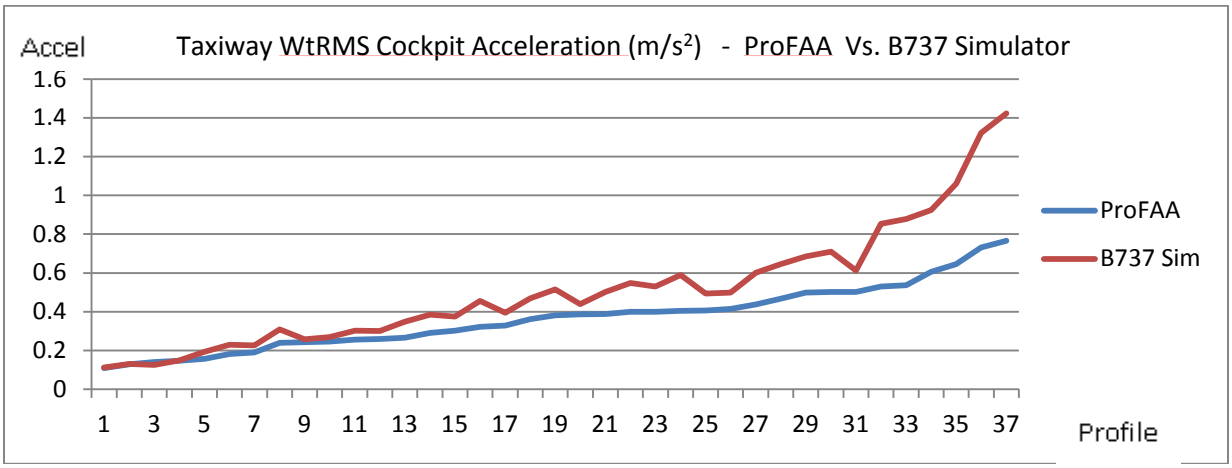
Boeing 737-800 simulator cockpit vertical accelerations were compared with predicted ProFAA cockpit accelerations using roughness final study taxiway and runway surface profiles. The ProFAA 737 surface roughness model was created by modifying the existing ProFAA B727 roughness model to adjust the wheelbase and aircraft weight to match those of a Boeing 737-800. Otherwise the ProFAA 737 model is the same as the ProFAA 727 model. Major differences between the ProFAA and B737 simulator roughness models are:

1. The ProFAA profile sample spacing is 250 mm (0.82 feet) compared with the simulator's 2.0 feet runway and 0.4 ft taxiway sample spacing

2. The ProFAA strut force calculation update rate is 400 Hz compared with a 60 Hz update rate used in the flight simulator.

3. The strut and tire-ground contact models are different.

Weighted RMS acceleration indices were used for the comparisons. As shown in Figure 45, the Boeing 737-800 simulator cockpit accelerations are higher than those modeled by ProFAA with the difference in acceleration increasing with the increase in profile roughness.



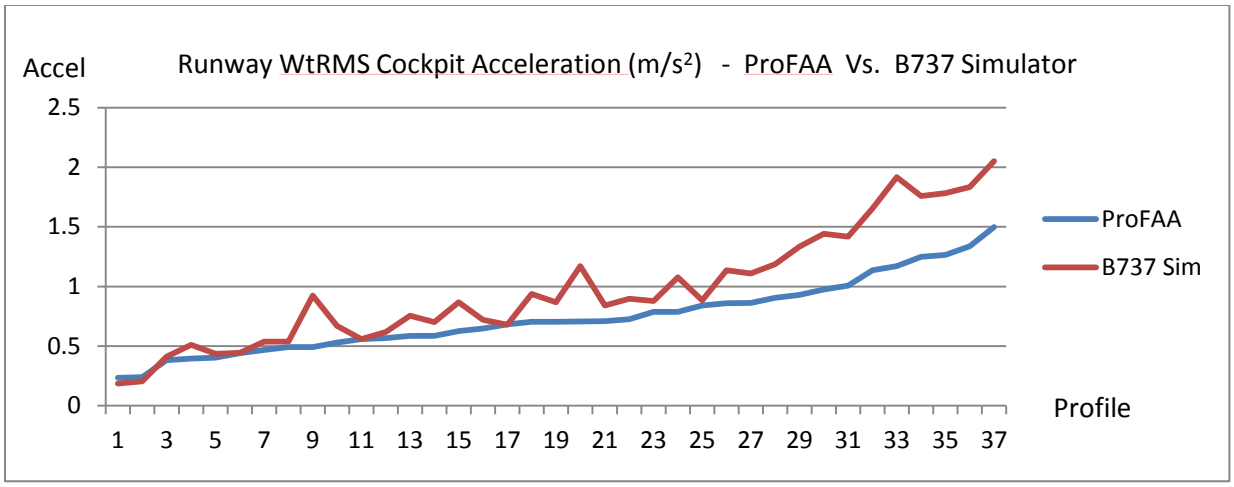
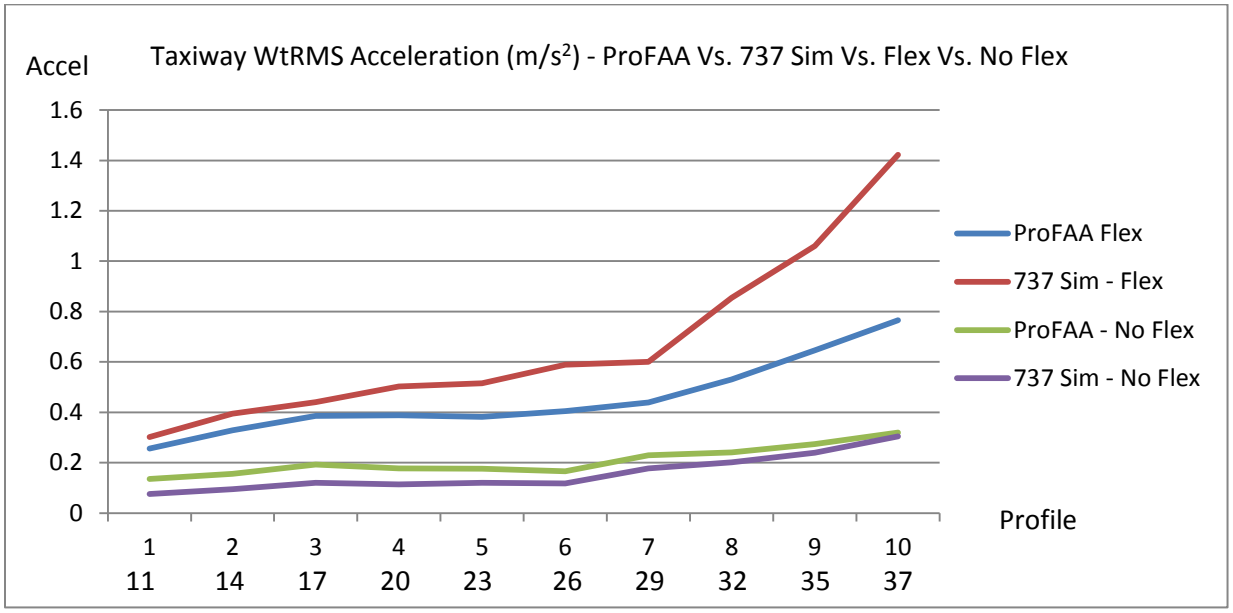


Figure 45. Cockpit Weighted RMS Acceleration – ProFAA Vs. Boeing 737-800 Flight Simulator

ProFAA and the Boeing 737-800 simulator incorporate rigid body and flexible mode simulation of the aircraft reaction to surface roughness. Additional ProFAA vs. Boeing 737-800 simulator comparisons were performed using cockpit acceleration data collected with the Boeing 737-800 simulator flexible mode model disabled to determine the relative contribution of Boeing 737-800 simulator rigid and flexible mode models toward the increase in cockpit acceleration over that predicted by ProFAA. The 'no flex' simulator data was collected for 10 real world taxiway and 8 real world runway profiles. Figure 46 shows the relative cockpit accelerations for ProFAA and Boeing 737-800 simulator with flexible mode models enabled and disabled.



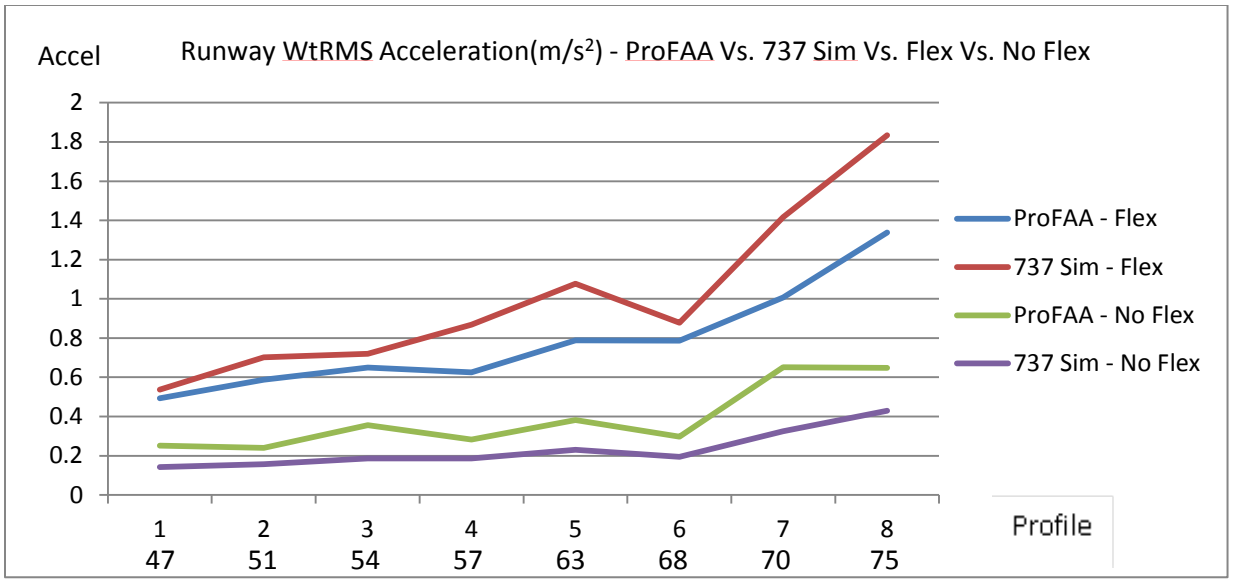


Figure 46. Cockpit Weighted RMS Acceleration - ProFAA Vs. Boeing 737 Sim Vs. Flex Vs. No Flex for a sample of 10 Final Study Taxiways and 8 Final Study Runways

The comparisons between the ProFAA and Boeing 737 flight simulator cockpit accelerations show the following:

- For rigid body simulations (no flex modes), the Boeing 737 simulator cockpit accelerations are at the same level or lower than the ProFAA cockpit accelerations.
- For combined rigid body and flexible mode simulations, the Boeing 737 simulator cockpit accelerations are higher than the ProFAA cockpit accelerations.
- The variance between the Boeing 737 simulator and ProFAA accelerations increases for rougher profiles.
- For rigid body simulations, the Boeing 737 simulator and ProFAA accelerations are more closely aligned for taxiway profiles than for runway profiles.

The differences in cockpit acceleration between the Boeing 737 simulator and ProFAA may be due to the differences in strut models and roughness profile sample spacing.

5.0 CONCLUSIONS

Roughness testing was completed in a final data collection effort conducted at the Mike Monroney Aeronautical Center (MMAC) in Oklahoma City. Measured real world airport taxiway and runway surface profiles were utilized to create 37 taxiway and 37 runway roughness scenarios on the FAA's Boeing 737-800 flight simulator. Thirty-three commercial pilots provided subjective ratings of surface rideability (on a scale of 0 - 10) rating and an acceptable/not acceptable rating for each ride.

A numerical score for the cockpit vertical accelerations resulting from each ride was calculated in terms of four ISO standard estimators of ride roughness: weighted RMS, weighted MTVV, weighted VDV and DKup. The following comparisons of pilot ratings vs. cockpit acceleration ISO indices were performed:

1. Average pilot ratings vs. weighted RMS acceleration allowed comparison of ratings with overall vibration discomfort levels
2. Average pilot ratings vs. weighted VDV acceleration provided a comparison with a shock-related health standard
3. Average pilots ratings vs. weighted MTVV acceleration compared ratings with the highest running value of weighted RMS acceleration.
4. Average pilot ratings vs. DKup acceleration constituted a comparison of ratings with a spinal response acceleration dose.

Analysis of the data provided the following observations:

- High correlations were found between the pilot ratings and the four ISO measures of total acceleration experienced, with different trends for taxiways and runways,
- The numerical 0 to 10 ratings were highly correlated with ride acceptability ratings using a ratings sheet with strong similarity to a previous effort [Ref. 1] for evaluation of highway pavement,
- Objective indicators of subjective human ratings of rideability were deduced in terms of functions of the ISO indices with confidence intervals for the fits,
- Limits for cockpit vibration were suggested by identifying index values at which it was estimated that a desired percentage of pilots would rate a taxiway or runway as unacceptable.

This approach to data collection differs from many past efforts to evaluate the discomfort to humans imparted by whole body vibration, such as NCHRP tests [Ref. 1], in which IRI (the International Roughness Index) or RN (the Ride Number) value was used as a roughness indicator. This effort is also the first to apply ISO standard indices to determine airport pavement roughness limits for in-service pavement.

It was found that the pilot numerical ratings and computed ISO roughness index values were highly correlated (correlation coefficients ranging from -0.942 to -0.983, with negative coefficients because pilot rating decreases corresponding to roughness index increases). Because of these strong correlations, reasonable fits could be made to express pilot average numerical rating as a function of each of the four ISO roughness indices.

The least scatter occurred in fitting the pilot ratings as a function of weighted RMS or weighted VDV by a quadratic fit or a shifted logarithmic fit as shown in Figures 47 and 48. Of these two fits, the shifted logarithmic fit had the better extrapolation capability on the right because it was strictly decreasing. For both of these fits the coefficient of determination R^2 ranged from .941 (for the shifted logarithmic fit of taxiway rating vs. weighted RMS) to .985 (for the quadratic fit of runway rating vs. weighted RMS) indicating very reasonable approximations since $R^2 = 1$ indicates a perfect fit.

Possible biases for military and repeat pilots were tested by hypothesis tests at the 95% level. It was found that pilots who had previously participated in rating taxiway/runway roughness were more than 95% likely to have slight biases towards rating less rough surfaces better and rough

surfaces lower as compared with the ratings of the other pilots; however, the biases were not large, so the repeat pilots were included in the analysis. It was found reasonable to think, but with less than 95% confidence, that military pilots had a slight bias to rate the runway rides lower than other pilots.

The ISO crest factor was computed for each taxiway and runway profile because a large crest factor indicates large individual jolts in the ride. For the four taxiway rides with high crest factor it was found that the best fit of pilot average numerical rating was achieved using weighted VDV as the roughness indicator. It is therefore reasonable to recommend the use of the weighted VDV for surfaces with single events. Runway profiles did not contain examples of crest factors significant enough to warrant analysis.

From the acceptable/unacceptable ratings given by pilots, fits were made to estimate the percentage of pilots rating each taxiway or runway acceptable as a function of the ISO roughness indices. It was found that if at least 5% of pilots rated a taxiway or runway as unacceptable then the corresponding ISO roughness index unacceptable values are those in Table 8.

Table 8**.** ISO Index Values for Which Taxiways and Runways Are Unacceptable to 5% of Pilots

ISO Roughness Index[8]	Unacceptable Taxiway Limits	Unacceptable Runway Limits
Weighted RMS (m/s^2)	*≥0.31*	*≥0.35*
Weighted MTVV (m/s^2)	*≥0.71*	*≥0.68*
Weighted VDV ($m/s^{1.75}$)	*≥4.11*	*≥4.16*
DKup (m/s^2)	*≥1.82*	*≥1.69*

It was also noted that the RMS values shown in this table correspond roughly to the standardized RMS values found in [3] at which vibration is considered "a little uncomfortable" and that the VDV values are well below the 8.5-17 range at which caution with respect to health risks is indicated (for persons subjected to strong vibration shocks).

Additional pilot numerical ratings of real world taxiways and runways from a preliminary collection effort for this study together with some final study pilot ratings that were not usable in the final study average ratings were combined with the final study data to generate a total of 1572 taxiway ratings and 1572 runway ratings from the test pilots. From these, statistics for individual pilot ratings (as opposed to pilot average ratings) were calculated. The fits to the combined individual pilot ratings were found to closely match the fits to average pilot ratings of rides in the final test data set (Figures 47 and 48).

[8] Recall that general guidance in the standard is that, in addition to the weighted RMS, one of the alternative indices should be reported when the crest factor of the acceleration record is approximately 9.0 or greater.

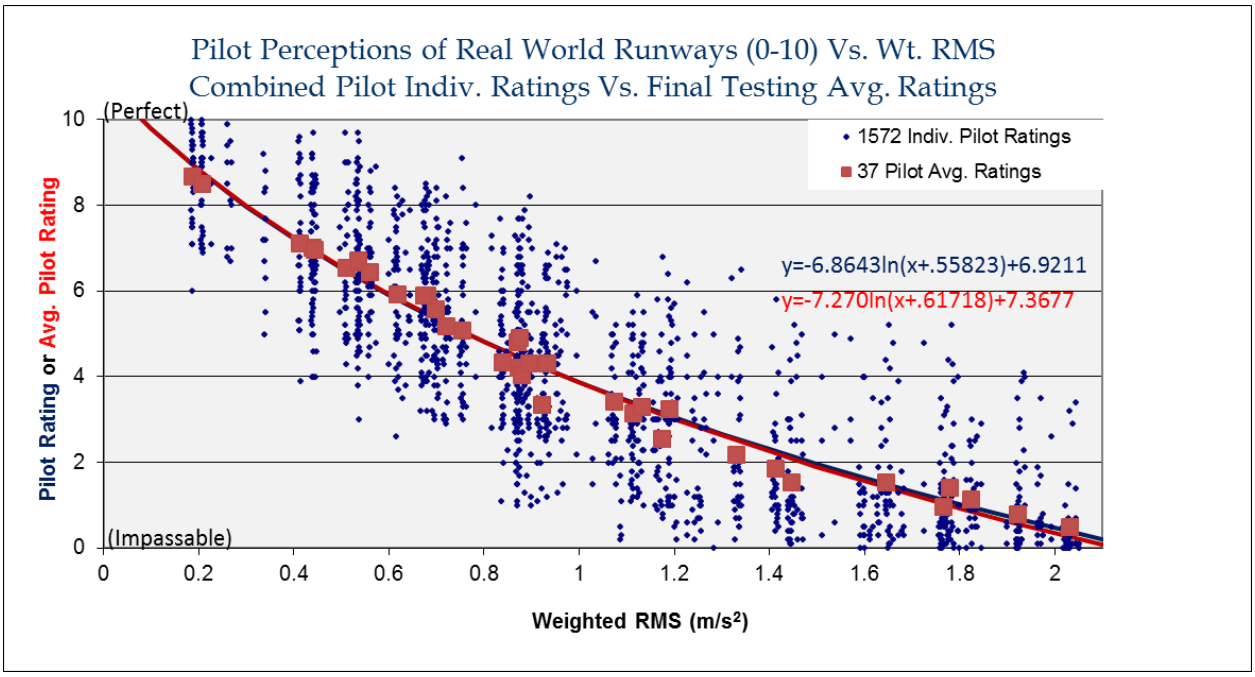


Figure 47. Shifted Logarithmic Trend Comparison: All Pilots' Runway Ratings Vs. Final Testing Average Runway Ratings

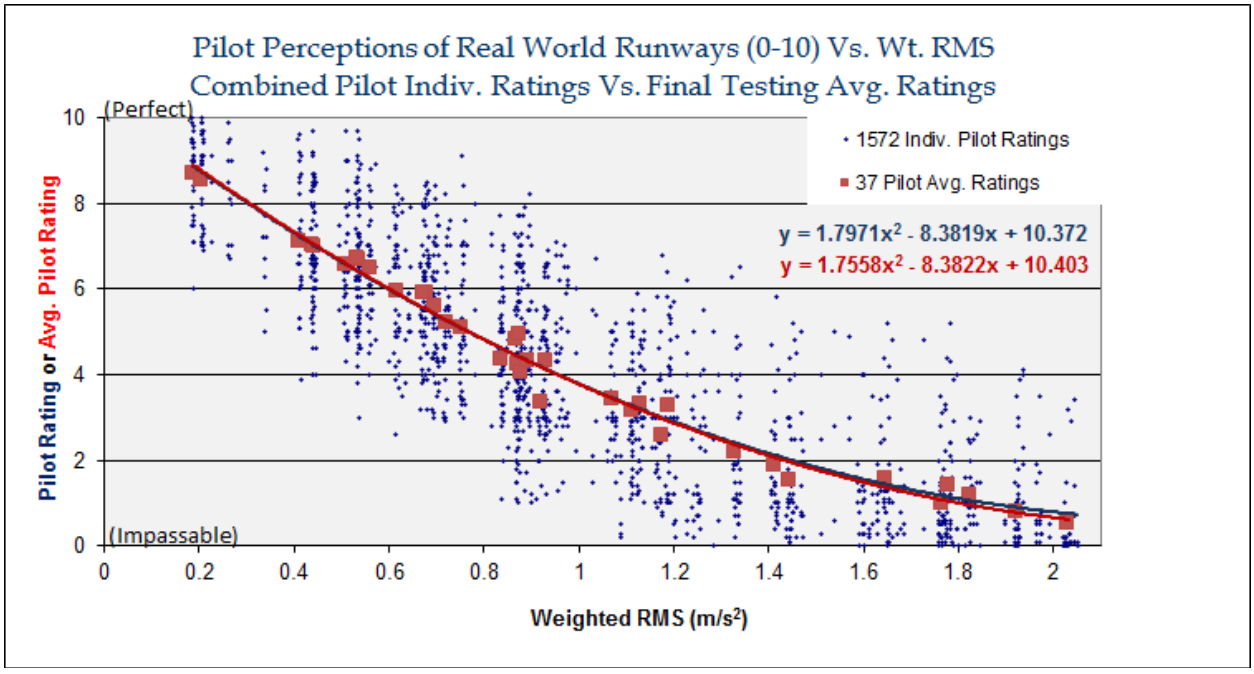


Figure 48. Quadratic Trend Comparison:
All Pilots' Runway Ratings Vs. Final Testing Average Runway Ratings

Confidence intervals for the fits and predictive intervals for individual pilot numerical ratings were computed as a function of each roughness parameter and show that despite the substantial variation in individual pilot ratings the confidence interval for average pilot rating is narrow and

various percentage levels (such as 5%) of pilots rating surfaces as unacceptable can be reasonably approximated.

No recommendation of a best ISO roughness index was made because different indices may be better indicators of roughness in different circumstances.

However, one reasonable strategy for maintaining in-service pavement would be to service the pavement when any one of the four indices (Weighted RMS, Weighted MTVV, Weighted VDV or DKup) exceeded some threshold value, such as the 5% level in Table 8.

Another reasonable strategy would be to service pavement only when a threshold value for weighted RMS or weighted VDV was exceeded, because these two indices have the best fit statistics, appear more frequently in literature, and estimate discomfort from overall vibration and from occasional shocks, respectively.

6.0 FUTURE WORK

The following areas have been identified for future research:

- Simulator roughness testing for other aircraft types
- Incorporation of pilot seat accelerometer for future testing
- Correlations of pilot ratings with standard roughness indices (Boeing Bump, IRI, etc.)
- Evaluation of new pavement construction limits for roughness
- Analysis of rideability ratings for asphalt vs. concrete surfaces
- Analysis of Recommended ISO Indices

7.0 ACKNOWLEDGEMENTS

The work described in this report was supported by the FAA Airport Technology Research and Development Branch, ANG-E260, United States of America (USA), Dr. Michel Hovan, Manager. Gordon Hayhoe provided assistance with the simulator roughness model development and accelerometer signal processing. CAE, the Boeing 737-800 flight simulator manufacturer, provided engineering support for the flight simulator ground model and motion system updates for this project. Boeing and Airbus provided some of the airport surface test profiles used in the studies.

The contents of the paper reflect the views of the authors, who are responsible for the facts and accuracy of the data presented within. The contents do not necessarily reflect the official views and policies of the FAA, USA. The report does not constitute a standard, specification, or regulation.

8.0 REFERENCES

1. National Cooperative Highway Research Program (NCHRP) 308 "Pavement Roughness and Rideability Field Evaluation."
2. American Society for Testing and Materials (ASTM) 1927-98 "Standard Guide for Conducting Pavement Ride Quality Ratings."
3. International Organization for Standardization (ISO) 2631-1 1995 "Mechanical Vibration and Shock – Evaluation of Human Exposure to Whole-Body Vibration, Part 1: General Requirements."
4. ProFAA The Federal Aviation Administration's Computer Program for Roughness Index Analysis User's Manual (Draft), FAA AFS-440, 6500 S. MacArthur Blvd., Oklahoma City, OK.
5. Runway Roughness Measurement, Quantification, and Application – The Boeing Method, Boeing Commercial Airplane Group Airport Technology Organization (B-B210), Doc. No. D6-81746.
6. Measurement and Evaluation of Human Exposure to Whole-Body Mechanical Vibration and Repeated Shock, British Standards Institution (BSI), BS 6841, BSI, London (2008).
7. Cherokee CRC, LLC, "FAA Surface Roughness Preliminary Study Data Collection Report [Draft], March 1, 2012.
8. Comparison of Subjective Responses to Vibration and Shock With Standard Analysis Methods and Absorbed Power, Mansfield, N., Holmlund, P., Lundstrom, R., Journal of Sound and Vibration, Vol. 230, Issue 3, February 24, 2000, pp 477-491.
9. A Comparison of Evaluations and Assessments Obtained Using Alternative Standards for Predicting the Hazards of Whole-Body Vibration and Repeated Shocks, Lewis, C., Griffin, M., Journal of Sound and Vibration, Vol. 215, Issue 4, August 27, 1998, pp 915-926.

---

[i] Seat number in the simulator was also considered as a human factor, but not found to have influenced pilot responses. (Graphs appear in Appendix A.)