## STATISTICAL ANALYSIS PROGRAM FOR GENERATING MATERIAL ALLOWABLES

Suresh Keshavanarayana Department of Aerospace Engineering Wichita State University 1845 Fairmount Wichita KS 67260-0044

### ABSTRACT

A computer program to conduct statistical analyses and generate tolerance values per the CMH-17 guidelines has been assembled. The computer program is written using Visual Basic language and has an Excel spreadsheet based user interface. The program combines the features of the pooling and single point methods to generate the material allowables/tolerance factors per the CMH-17 guidelines. The program also allows the user to discriminate the data for presence of outliers, between-batch variability, appropriateness of statistical distributions, equality of variances, etc., using standard statistical tests and graphical tools. The details of the computer program, the status of the ongoing work and results from verification runs on data presented in CMH-17 handbook are outlined in this paper.

### **1.INTRODUCTION**

The variability associated with the composite material properties is well known. The sources of variability include run-to-run variability in fabrication, batch-to-batch variability of raw materials, testing variability, and variability intrinsic to the material. The generation of statistically based material allowables (basis values/tolerance factors), which account for some of the variability, is a key milestone in the insertion of materials into the design process for airframe structures and their certification. After the test data for material properties are obtained using standard test procedures, the test data has to be analyzed using statistical methods to generate the allowable values.

Currently, three statistical methods are available for generation of allowable values and these methods are discussed in detail in the CMH-17 handbook[1]. The three methods are the Single Point (SP) method[1], the Regression Analysis (RECIPE)[1] and the Pooling (AGATE) method [2,3]. The SP and Pooling methods have been utilized through Microsoft Excel based Visual basic computer programs- STAT17 and ASAP respectively. The regression analysis method is implemented using a Fortran code.

The current version of the ASAP program used for the Pooling method can accommodate test data at five different environmental conditions, with a data limit of 200 per test environment. The program assumes that the data at each environment follow a normal distribution and uses statistical and graphical methods to judge the normality of the test data, prior to generating the allowable/basis values. In addition to the statistical analysis, the program highlights problems with the data (outliers), generates plots and conducts statistical tests to aide engineering judgment if/when necessary and recommends alternate analysis methods if required. The excel based STAT-17 program used for generating allowables based on SP method handles test data at

a single environmental condition individually. The program features discriminatory tests for statistical distributions such as Normal, Weibull and LogNormal, and uses non-parametric methods for estimating allowables in the event the distributions do not fit the data satisfactorily. In the present work, an Excel-Visual basic computer program which combines the features of the Pooling and the SP methods per the flowchart approved by the Statistics Working Group (SWG) of the CMH-17, has been assembled. The details of the computer program, its capabilities and features, and results from verification studies are summarized in this paper.

### 2. GENERATION OF ALLOWABLES USING CMH-17 METHOD

The CMH-17 method utilizes both the SP and Pooling methods to generate the allowables from test data obtained at different environmental test conditions. The test data generated with multiple batches of data tested across different environmental test conditions (e.g., Room temperature dry, elevated temperature wet, etc.) are screened for contiguity of test conditions (ref CMH17) and other engineering considerations prior to the statistical analysis. The statistical analysis begins with the test data being screened for acceptable and consistent failure modes across material batched and environmental test conditions. The data sets are then analyzed for the minimum requirement for number of batches and batch sizes to determine if the SP or Pooling method has to be used per the flowchart in Figure 1. The test data sets are then screened for outliers at the batch level and at the test condition level using the MNR statistic test [1, 3]. The outliers are then dispositioned and new data points added (if needed) to satisfy the minimum size requirements for the batch/sample.



Figure 1. CMH-17 Statistical analysis flowchart for generating allowables



Figure 1 contd. CMH-17 Statistical analysis flowchart for generating allowables



Figure 1 contd. CMH-17 Statistical analysis flowchart for generating allowables

# **3. EXCEL VISUAL BASIC PROGRAM**

#### 3.1 Organization of the VBA project

The Excel based Visual basic computer program consists of Excel worksheets as the front end to interact with the end user. The workbook consists of 12 worksheets which contain the control/execution sheet, data input/output areas, plots from diagnostic tests, and summaries from the analysis and diagnostic statistical tests. The Visual basic code is made up of different modules which use information from these worksheets and writes the results from the analysis into the same. The organization of the VBA project is illustrated in Figure2.



Figure 2. Organization of the VBA project

### 3.1.1 Description of Worksheets

The functionality of individual worksheets is described in the following paragraphs.

**INSTRUCTIONS:** This worksheet contains instructions for operating the program and a brief description of all the worksheets in the excel workbook.

**CMH-17 MAIN:** This worksheet contains the 'macro' buttons which will execute the program /specific modules to analyze the data. A screenshot of this worksheet is shown in Figure3 below. The worksheet lets the user select the options for the different statistical tests which are used for diagnosing the data sets to be used in single-point and pooling methods. In addition, the user has the option of running independent diagnostic tests on select data sets prior to running the entire analysis. The reading of input data from other spreadsheet files, printing of reports can be accomplished using the macro buttons on this worksheet.



Figure 3. Layout of program control /execution sheet

**INPUT DATA:** This worksheet contains the columns for input of test data. The test data may be input using the keyboard or copied from another worksheet. Alternatively, one may read the data from an input file which contains the data in a specific format using the 'Import data from .xls file' macro button on the CMH17-MAIN worksheet to read the data from a file. The input sheet contains pre-allocated cells for data, batch identifiers and specimen numbers as shown in Figure 4. The current input sheet can accommodate up to 1000 test data points at 10 environmental test conditions. The test condition identifiers (i.e., RTD, CTD, ETW, ETD) are available for selection from a drop down list corresponding to each column of data. The user is required to enter the temperatures (alphanumeric identifier) corresponding to the selected condition. In addition to the user input areas, this worksheet contains 3 additional columns associated with each data set to identify and disposition outliers. The VBA program identifies outliers at the 'batch' and 'pooled data set' levels by placing a check mark against the data point identified as outlier. The user is then allowed to either retain or delete it for subsequent analysis. The outlier can be used for subsequent analysis by placing the letter 'U' on the corresponding row. The data point may be physically deleted or deactivated for subsequent analysis by placing a 'D' on the corresponding row.



The INPUT DATA sheet contains the datasets for generating B-basis & A-basis values. The data sets may be input manually using keyboard entry or copied and pasted from another spreadsheet. In addition, the data sets may be imported from excel workbooks using the option button on the CMH17-MAIN worksheet. The user may select appropriate test condition description and add an alphanumeric to it under the temperature column.

	CONDITION		TEMPERATURE		CONDITION		TEMPERATURE		CONDITION		TEMPERATURE		CONDITION		TEMPERATURE		CONDITION		TEMPERATURE		COND	TION	TEMPE	RATL	JRE					
	CTD	-	-65F			RTD	-		75F		ETW	-	200F (HW)		)	USER 1	-	ETW2		ETW 150F		ETW	-	1	80F					
	Do not use for pooling					Do no	t use dat	a for pooling			L Do no	t use data	a for pooling			Do no	t use dat	a for pooling			L. Do no	t use dat	a for pooling			L Do no	t use data	for pooling		
									0	411 a.v.					ALL AND					dill an									-	
	Batch	Spec.	Data	0	utiler	Batch	Spec.	Data	ou	tiler	Batch	Spec.	Data	ou	tiler	Batch	Spec.	Data	0	utiler	Batch	Spec.	Data	0	utiler	Batch	Spec.	Data	- Ou	ther
No.	I.D	No.		ΒF	⊃ U/D	I.D	No.		ΒP	U/D	I.D	No.		ΒP	U/D	I.D	No.		ΒF	> U/C	I.D	No.		ΒF	P U/D	I.D	No.		ΒР	U/D
1	1	1	136.64																						_				⊢⊢	
2	1	2	125.91																	_					_				⊢⊢	_
з	1	3	144.45	х	х					-										_					_				⊢⊢	
4	2	1	107.79							_										_					_				⊢⊢	-
5	2	2	114.58	-										$\vdash$					$\vdash$	_				$\left  \right $	_				⊢⊢	
5	2	3	110.7											$\vdash$																
		2	119 79																	_					-				+	-
9		2	131 24																											+
10	4	1	125.91																											-
11	4	2	127.86	х																									rt-	1
12	4	3	125.91																											
13	5	1	134.41																											
14	5	2	124.6																											
15	5	3	127.54																											
16	6	1	139.35																										⊢⊢	
17	6	2	119.03																										⊢⊢	
18	6	3	125.81																	_					_				⊢⊢	_
19	7	1	120							_										_					_				⊢⊢	
20	7	2	121.94										-					-		_					_				⊢⊢	-
21	7	3	132.58																										⊢⊢	
22	8	1	119.28	_																_					-				⊢⊢	-
23	8	2	118.3							-			-	++	-				++	_					-				++	
24	8	1	109.5	$\vdash$	_				$\vdash$	1				$\vdash$	1				$\vdash$		1			++					r+	+
20	9	2	121 23						$\vdash$	1				⊢⊢	+				++					++	+				$\vdash$	+
27	9	3	130.03																$\square$											+

Figure 4. Data input sheet layout

**RESULTS SUMMARY:** The first half of this worksheet contains the summary of the statistics and basis values obtained using pooling and single-point approaches. A comments section on the sheet indicates any problems with the data set(s), failure of diagnostic test(s), validity of basis values, etc. The second half of the worksheet contains a graphical display of the average values and basis values for all the data sets analyzed. The screenshots of the results summary sheet is shown in figures 5a and 5b.

TRIAL VERSION;	DO NOT USI	FOR ALLC	WABLE GI	ENERATIO	N. DECEME	ER 2009							
COMPANY: ACME Composites MATERIAL: T900-PX Graphite-Enoxy													
MATERIAL: T900-PX Graphite-Epoxy PROPERTY: 0 deg Tensile Strength													
PROPERTY: 0 deg Tensile Strength TEST METHOD: ASTM D3039													
TEST METHOD: ASTM D3039 COMPOSITE MATERIALS HANDBOO													
DATE	E: May 5, 2010	)											
	1			DATA									
STATISTIC			FTW 200F			TEST CON		1	1	1			
	CTW -65F	RTD 75F	(HW)										
Sample Size	20	20	34										
No. of Batches	3	3	3										
Mean	55.31	41.44	19.30										
Std.Dev	2.87	1.86	1.27										
% Co.V	5.19	4.50	6.56										
Minimum	50.41	38.79	16.78										
Maximum	61.39	46.40	23.32										
										1			
SUMMARY OF BASIS VALUES FOR DATA POOLED ACROSS ENVIRONMENTAL TEST CONDITIONS													
ruuleu % cu. v 5.63													
к <sub>ь</sub>	2,9720	1.7238	1.04/5										
г <sub>а</sub>	2.8730	2.8730	Z.811Z	Ct Day Basi		inin al Data)							
Pooled St.Dev. Basis Values (Original Data)													
B-Basis Value	49.87	37.36	17.48										
A-Basis Value	46.24	34.64	16.20	Day, Basis )	(aluea (Medii	ind C V date	- )						
Madified %C \/	0.00	0.05			alues (Mouli		a)		1				
Modified Pooled %	0.00 Co.V	0.25	7.20		6.	74							
R-Rasis Value	48 80	26 FF	17 13	r –	I	, ,	1	r –	T				
	40.00	33 30	17.13										
	2010	55.50	15.57										
		SUMMARY	OF BASIS VA	LUES USING	SINGLE POIN	T APPROACH	(Original Da	ta)					
Distribution	Normal	Normal	Normal										
B-Basis Value	49.78	37.85	17.10										
A-Basis Value	45.84	35.29	15.50										
COMMENTS													
(1) Test data has b	een provided	at 3 test co	on dition (s)										
(2) 1 outliers dete	cted at the ba	tch level. S	ee INPUT SH	EET for outli	ers								
(3) 1 outliers dete	cted at the po	oled data l	evel. See IN	PUT SHEET f	or outliers								
(4) Disposition out	tliers and reru	ın analysis											
(5) Batch equivale	nce not staisf	ied at RTD	75F; conditi	ons									
(6) Batches are not	t poolable at	1 test condi	tions. Use e	ngineering.	udgementf	or pooling n	nethod or us	se Single Po	int method				
(7) Pooled Conditio	onsare:CTW	-65F; RTD 7	5F; ETW 200F	-(HW);									
(8) Pooled dataset	s pass Levene	's test !!											
(9) Single Point An	aiysis : Norma	il Distributi	on is indica	ted for CTW	-65F Note :	vormal Dist	bn Overridii	ng factor sel	ected = 10				
(10) Single Point Analysis : Normal Distribution is indicated for RTD 75F Note : Normal Distbn Overriding factor selected = 10													
(11) Single Point Analysis : Normal Distribution is indicated for ETW 200F (HW) Note : Normal Distbn Overriding factor selected = 10													
(12) Pooled Condit	od data sets a	-05r; RID	/5r; EIW 200	)F (HVV);									
(13) Woutted Pool	eu ualasets p	ass Levene	s test !!										

Figure 5a. Results summary sheet showing statistics, basis values and comments



Figure 5b. Results summary sheet showing graphical comparison of basis values

**STATISTICS SUMMARY:** This worksheet contains a detailed summary of the different diagnostic tests per the CMH-17 handbook. In addition, space for user comments to override failure of tests based on engineering judgement, is provided on this sheet. The statistical summaries are split into two parts – one containing the summaries for the pooling method and the other containing the same for the SP method, as shown in figures 6a and 6b respectively

SUMMARY OF DIAGNOSTIC TESTS													
Test Condition	1	2	3	4	5	6	7	8	9	10			
Condition I D	ст₩-65∓	R TD 75F	E TW 200F										
Maximum Normed Re	l sidual Test fi	n Outliers	(111)										
Batch Level	0	1	0										
Pooled Data	0	0	1										
Note : See INPUT DATA sheet for specific data points which have been identified as outliers													
_													
k-Sample Anderson Darling Test for Batch Equivalence ADK < ADC for equivalence $regimerrow$ *Significance Level $\alpha$ =													
ADK	1.698	2.444	0.917										
ADC (0. = 0.05)	1.930	1.930	1.976										
ADC (0. = 0.025)	2.233	2.233	2.294										
ADC (a = 0.01)	2.635	2.635	2.716										
Same Population ?*	YES	NO	YES										
Modified CV Data - for poolin	g method												
ADK	1.589	1.144	0.622										
Same Population ?(🛛=0.025)	YES	YES	YES										
Summary of Diagnostic Statistics for Pooling Method													
O.S.L. (original data)	0.449	0.115	0.104										
Normality acceptable ?	Yes	Yes	Yes										
O.S.L. (Modified data)	0.251	0.440	0.417										
Normality acceptable ?	YES	YES	YES										
Check for Normality Based of Pearson Coefficient r	n <b>G raphical I</b> 0.984	Vlethod 0.956	0.961										
Normality acceptable ?	Yes	Yes	Yes										
Anderson, Darling Test for No	umality of P	noled Data											
O.S.L. for Original Data		Joica Data			0.1	576							
O.S.L. for Modified Data					0.5	424							
Louisne's Test for Equality of	Variations (	ECALCULATE	р < горпки	) for an util	(ance)								
		0.050	0.025		incince)								
FCRITICAL	2,379	3 1 2 6	3 887	4 917									
FCALCULATED	1.260	5.120	2.007										
FCALCULATED - Modified CV	0.546												
USER COMMENTS													

Figure 6a. Statistics summary sheet showing results for pooling method.

SUMMARY OF DIAGNOSTIC TESTS													
		Summ	Summary of Diagnostic Statistics for Single-Point Method										
Distribution	Statistic	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10		
Normal		0.449	0.115	0.104									
Log Normal	Observed Significance	0.398	0.161	0.125									
Weibull		0.359	0.006	0.001									
	Mean	55.31	41.44	19.30									
Normal	Stdev	2.87	1.86	1.27									
	%C.V.	5.19	4.50	6.56									
log Normal	Log Mean	4.01	3.72	2.96									
	Log Stdev	0.052	0.044	0.065									
Weibull	Scale parameter	56.65	42.36	19.90									
Weibuli	Shape Parameter	20.70	20.32	13.82									
	NON-PARAMETRICS	τατικτια		H-K~Hai	nson-Kooni	mans							
	B-Basis Method	H-K	Н-К	Pank		lians	1						
	Δ-Basis Method	н-к	н-к	H-K									
	B-Bacic Pank	n/ə	n/ə	1									
	A-Basis Rank	n/a	n/a										
	B-Basis Hans-Koon k Factor	1 253	1 253	n/a									
	A-Basis Hans-Koop k Factor	2.367	2.367	1872									
		2.007	2.007	1.072			1						
	EQUALITY OF VARIA	NCES TEST	-										
	Fcalculated	0.097	1.973	2.483									
	Fcritical	4.619	4.619	4.165									
	Variances Equal ?	yes	yes	yes									
	ANALYSIS OF VAR	RIANCE (A	NOVA) ST.	ATISTICS									
Sample Bet	tween-batch Mean Sq. (MSB)	22.00	10.84	1.10									
	Error Mean Square (MSE)	6.63	2.61	1.63									
Estim	ate of Pop. Std. Deviation (S)	2.99	1.96	1.26									
B-Basi	s Tolerance Limit Factor (TB)	1.98	1.99	1.74									
A-Basi	s Tolerance Limit Factor (TA)	3.38	3.39	3.01									
	B-Basis Value	49.39	37.53	17.11									
	A-Basis Value	45.20	34.78	15.51									
	SUMMARY OF BASI	S VALUES											
	WEIBULL	47.89	35.70	15.91									
	NORMAL	49.78	37.85	17.10									
B-Basis	LOGNORMAL	49.97	38.02	17.21									
	NON-PARAMETRIC	49.16	38.27	16.78									
	ANALYSIS OF VARIANCE	49.39	37.53	17.11									
	WEIBULL	40.71	30.26	12.77									
	NORMAL	45.84	35.29	15.50									
A-Basis	LOGNORMAL	46.53	35.78	15.85									
	NON-PARAMETRIC	38.50	30.36	12.59									
	ANALYSIS OF VARIANCE	45.20	34.78	15.51					L				

Figure 6b. Statistics summary sheet showing results for the single point method

**BATCH PLOTS:** This worksheet contain the batch plots of the data sets on the input sheet. The batch plots are generated during the execution of the program. The number of batch plots will equal the number of test conditions at which input data is available. The batch plots contain individual data points in each batch, the mean of each batch and lines corresponding to the pooled average plus standard deviations. These plots are used to aide engineering judgment when the statistical test for batch equivalence is not satisfied. The program automatically adjusts the scaling for better viewing of the plots. However, the user is allowed to change the scales on individual plots if desired. The typical batch plots generated by the program is illustrated in Figure7



Figure 7. Typical batch plots generated by the program

**QUANTILE PLOTS:** This worksheet contains the Quantile Box plots for data at individual environmental test conditions and for the pooled data as illustrated in Figure8.



Figure 8. Quantile Plots generated by the program

**NORMAL PLOTS:** This worksheet contains the Normal plots and Normal Scores plots for individual environments and the pooled data as shown in Figure 9. These plots are used for visualizing normality of the test data.



Figure 9. Normal Plots generated by the program

**TABLES**: This worksheet contains the tables for Normal, Weibull and NonParametric methods. The worksheet is locked and the user is not allowed to edit any portion of these worksheets as it will be detrimental to numerical values generated from the analyses.

#### 3.2 Visual basic program

The statistical analysis per the CMH-17 guidelines is carried out by a VBA macro/program in the background. The macro(s) are activated by using the 'macro buttons' in the CMH17-MAIN worksheet described previously. The sequence of operation of the program and its interaction with the various worksheets is summarized in the flowchart presented in Figure10. In addition to reading/writing data from/to worksheets, the program also flashes error/warning messages to make the user aware of problems with data sets or failure of certain statistical diagnostic tests. In

certain case, the user will have the ability to abort the program to further investigate the source of errors in the data sets, or simply continue with the analysis.



Figure 10. VBA program for generating allowables

The VBA program is made up of multiple modules, each of which accomplishes a specific task per the flowchart presented in Figure1. Each module further consists of functions and procedures which accomplish certain repetitive task, e.g., computing the means, detecting outliers, etc. Most of these functions and procedures are public in nature, i.e., they can be called by any other procedure in the other modules. The procedures are grouped into modules to facilitate ease of debugging and editing. In addition, the structure of the program lends itself to future expansions, or debugging by users.

## **4. VERIFICATION TESTS**

The VBA program was verified using data published in CMH-17 handbook example problems. The statistics and basis values computed using the VBA program matched the results published in CMH-17 handbook examples. Additional data sets (especially large ones) are being generated to further exercise the capabilities of the program and identify any problems in the program execution and computation of basis values.

## **5. SUMMARY OF PROGRAM CAPABILITIES**

The capabilities/features of the VBA program are summarized below

- 1. The program can analyze test data at 10 environmental test conditions. At each test condition, there is no limit on the number of batches or the size of batches, except that they have to meet the CMH-17 guidelines.
- 2. The datasets for analysis may be entered into the existing worksheet or read from another excel workbook (typical for batch analysis/ multiple data set analysis mode)
- The program has the capability to conduct specific diagnostic tests on selected data sets (e.g., at test condition RTD) prior to conducting the entire analysis.
- 4. The program displays error messages and status messages to the user. The user is further provided with the option of exiting the program or continuing using interactive message boxes.

- 5. The number of batch plots generated by the program is the same as the number of data sets that have been input to the program.
- 6. All plots are created during program execution and thus do not contribute much to the initial file size.
- 7. The program provides the user with multiple options for setting the significance levels for diagnostic tests using a scroll list, in addition to the default values corresponding to the CMH-17 recommendations. This allows the user to explore '*what-if*' scenarios in case some of the statistical tests are not favorable.
- 8. The summary sheets and plots may be printed to a \*.xps file which can be viewed using internet explorer and does not require document editors (word/Adobe Acrobat, etc)

## **6. PROGRAM STATUS**

The coding of the program and the formatting of the excel worksheets has been completed. Limited trials have been conducted on data sets presented in CMH-17 handbook and data sets provided by NIAR. A beta-version of the program will be presented and released to the members of the statistics working group during the upcoming CMH-17 meeting (July 2010). In addition, the program documentation and development of a *User's Guide* is currently under progress.

### 7. SUMMARY

An Excel based Visual Basic computer program to conduct statistical analyses and generate tolerance values per the CMH-17 guidelines has been assembled. The program combines the

single point and pooling methods to generate allowables per the CMH-17 guidelines. A trial version of the program has been tested using published data and the results have been satisfactory.

# 8.REFERENCES

- 1. The Composites Materials Handbook, CMH-17
- Shyprykevich, P, "The Role of Statistical Data Reduction in the Development of Design Allowables for Composites," *Test Methods for Design Allowables for Fibrous Composites: 2nd Volume, ASTM STP 1003*, C.C. Chamis ed., American Society for Testing and Materials, Philadelphia, PA, 1989, pp. 111-135.
- John S. Tomblin, Yeow C. Ng and K. Suresh Raju, "Material Qualification and Equivalency for Polymer Matrix Composite Material Systems," DOT/FAA/AR-00/47, April 2001