DOT/FAA/TC-12/57

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

Statistical Analysis Program for Generating Material Allowables

September 2013

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/TC-12/57	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle STATISTICAL ANALYSIS PROGRAM FOR GENERATING MATERIAL ALLOWABLES		5. Report Date September 2013 6. Performing Organization Code
7. Author(s) Suresh Keshavanarayana		8. Performing Organization Report No.
9. Performing Organization Name and Address Wichita State University 1845 Fairmount Wichita, KS 67260-0044		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Northwest Mountain Region—Transport Airplane Directorate 1601 Lind Avenue, SW Renton, WA 98057		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code ANM-115

15. Supplementary Notes

The Federal Aviation Administration William J. Hughes Technical Center Aviation Research Division COR was Allan Abramowitz

16. Abstract

A Visual Basic application with an Excel® user interface was assembled to implement the Composite Materials Handbook (CMH)-17 procedure for computing statistically based material allowables. The computer program implements both the Single Point and Pooling methods involved in the CMH-17 procedure and is capable of following the decision tree outlined. The program is capable of analyzing single data sets or batch processing of multiple data sets. The computer program has user options that allow the selection of significance levels for statistical tests as well as overriding the failure of statistical tests. The program generates summary sheets comprised of numerical results from the statistical tests and plots to facilitate visual examination of the data to aid engineering judgment. The computer program has been verified and validated against benchmark data sets published in the CMH-17 Handbook.

17. Key Words Composite material properties. Material allowables, Pooling method (ASAP), Single Point method (STAT17), CMH-17, Statistical analysis	18. Distribution Statement This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 53	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGEMENTS

The author acknowledges Allan Abramowitz, Federal Aviation Administration William J. Hughes Technical Center technical monitor, for his support. The author is grateful to Dr. Beth Clarkson, National Center for Advanced Materials Performance, for her support during the development and verification of the computer program. The input from members of the Composite Materials Handbook-17 Statistics Working Group and volunteers from industry were valuable in shaping the computer program.

TABLE OF CONTENTS

			Page
EXECUTIVE SUMMARY			xi
1.	INTRODUCTION		1
2.	GENERATION OF ALLOWABLES USING CMH-17 PROCEDURE		1
3.	EXCEL VISUAL BASIC FOR APPLICATIONS		5
	3.1	Microsoft Excel Workbook	7
		3.1.1 Instruction Worksheet	8
		3.1.2 The CMH-17 Main Worksheet	8
		3.1.3 Data Input Worksheet	9
		3.1.4 Results Summary Worksheet	10
		3.1.5 Statistics Summary Worksheet	10
		3.1.6 Batch Plots Worksheet	10
		3.1.7 Quantile Plots Worksheet	11
		3.1.8 Normal Plots Worksheet	11
		3.1.9 Tables Worksheet	11
		3.1.10 Plot Data Worksheet	12
		3.1.11 Sort Worksheet	12
		3.1.12 Test Sheet Worksheet	12
		3.1.13 Messages Worksheet	12
		3.1.14 Revision Worksheet	12
	3.2	Visual Basic Program	12
4.	USER GUIDE		17
	4.1	Input Data	17
	4.2	Analysis of Single Data Set	19
	4.3	Analysis of Multiple Data Sets (Batch Processing)	23
5.	EXAMPLE PROBLEMS		24
	5.1	Example Problem I	24
	5.2	Example Problem II	32
6.	SUMMARY		32
7.	REFERENCES		33

APPENDICES

A—Samples of Worksheets
B—Example Data Sets

LIST OF FIGURES

Figure		Page
1	The CMH-17 Procedure for Generating Allowables	3
2	Overall Organization of the Excel VBA for Generating Allowables	6
3	A Screenshot of the CMH-17 MAIN Worksheet	8
4	A Partial Screenshot of the Data Input Worksheet	9
5	Typical Batch Plots Generated by the Program	11
6	Definition of a Single Data Set and Multiple Data Sets	13
7	Flow Chart Showing the Initial Stage of the Analysis	15
8	Flow Chart Showing the Second Phase of the Analysis	16
9	Flow Chart Showing the Analysis of CV Modified Data	17
10	Template for Input Data	18
11	User Input of Information to Identify Data Set	20
12	Selection of Options on CMH-17 MAIN Worksheet for Statistical Analysis	21
13	Program Window to Select Input Data File	21
14	Drop-Down Menu for Assigning Alphanumeric Identifier for Test Condition	22
15	Message Windows Generated by Program After Conducting Diagnostic Tests	22
16	Status Window Indicating the Progress of the Analysis	23
17	Options Chosen for Example Problem I	24
18	Input Data Sheet Indicating the Outliers at ETW1 Test Condition	26
19	Results Summary Sheet Showing the Statistics and Basis Values Based on SP and Pooling Analysis	27
20	Comments and Warning Messages Generated by the Program	28

21	Summary of Statistical Diagnostic Tests for Pooling Analysis	29
22	Summary of Statistical Diagnostic Tests for SP Analysis	30
23	Batch Plots for the Test Data and Corresponding Average and Basis Value Lines	31
24	Quantile Box Plots for Raw Data and Normalized Pooled Data in Example Problem I	31
25	Normal Plots and Normal Scores Plots for Data in Example Problem I	32

LIST OF TABLES

Table		Page
1	Summary of Worksheets in the Workbook	7

LIST OF ACRONYMS

AGATE	Advanced General Aviation Transport Experiments
ASAP	AGATE Statistical Analysis Program
CMH	Composite Materials Handbook
CV	Coefficient of variation
MNR	Maximum normed residuals
OSL	Observed significance levels
RECIPE	Regression Confidence Intervals for percentiles
SP	Single Point
SWG	Statistics Working Group
VBA	Visual Basic for Applications

EXECUTIVE SUMMARY

The generation of statistically based allowables for composite material properties includes the use of both Pooling and Single Point methods. Currently, individual computer programs exist to implement the Pooling and the Single Point methods (STAT17). A detailed procedure that uses a computer program for both methods was implemented by the Composite Materials Handbook (CMH)-17 Statistics Working Group (SWG). However, in addition to computing the statistics and allowables, the program has to provide the user with the necessary help in making decisions related to the validity of data sets, outliers, etc., in accordance with the guidelines set forth by the SWG. (The current work involved the development of a computer program to address these issues.) The new program incorporates the features of the existing programs and provides a user-friendly environment that aids and assists in the data decision-making process that generates composite material allowables per the CMH-17 guidelines.

A Visual Basic for Applications with the Microsoft® Excel® user interface was assembled to implement the CMH-17 statistical analyses for computing A- and B-basis allowables for material properties. The program accepts data manually or from a data file. The program is capable of analyzing up to 1000 data points each for ten individual test conditions. There are no limitations on the number of batches as long as the batch sizes are consistent with CMH-17 guidelines. The macro program analyzes the data and interacts with the user (except during batch processing) in the presence of outliers, errors in data, etc. The program conducts both Single Point and Pooling analyses on the given data sets and reports the basis values. The program may be run in a compliant mode with default options in which the CMH-17 guidelines are strictly enforced, or in a manual override mode using the options provided in the program. With each analysis, the program creates a Microsoft Excel workbook report consisting of the input data, results, summary sheets, and plot sheets. The program has been verified and validated against STAT17 and Advanced General Aviation Transport Experiments (AGATE) Statistical Analysis Program (ASAP) programs using the example data sets reported in this document and independently by the SWG.

1. INTRODUCTION.

The variability associated with 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 and 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 certifications. 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 the generation of allowable values, and these methods are discussed in detail in the Composite Materials Handbook (CMH)-17 [1]. The three methods are the Single Point (SP) [1], the Regression Confidence Intervals on Percentiles (RECIPE) [1], and the Pooling (Advanced General Aviation Transport Experiments (AGATE)) [2 and 3]. The SP and Pooling methods have been implemented through Microsoft® Excel®-based, Visual Basic® computer programs STAT17 and AGATE Statistical Analysis Program (ASAP), respectively. The RECIPE method is implemented using a Fortran code.

The current version of the ASAP program (version 2008) 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 it 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, conducts statistical tests to aid engineering judgment if or when necessary, and recommends alternate analysis methods if required. The Excel-based STAT17 program used for generating allowables based on the SP method handles test data at a single environmental condition. The program features discriminatory tests for statistical distributions, such as Normal, Weibull, and LogNormal, and uses nonparametric 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 that combines the features of the Pooling and SP methods pursuant to the flow chart approved by the CMH-17 Statistics Working Group (SWG) has been assembled. The details of the computer program, results from verification runs on test data presented in the CMH-17 Handbook, and a user guide are presented in this report.

2. GENERATION OF ALLOWABLES USING CMH-17 PROCEDURE.

The detailed description of the CMH-17 procedure for computing allowables, along with the statistical tests and corresponding equations, tables, etc., are found in CMH-17 [1]. An overview of the procedure is presented in this section for completeness and is shown in figure 1. The CMH-17 method uses both the SP and Pooling methods to generate the allowables from test data obtained at different environmental test conditions. The test data consists of multiple batches of specimens tested across different environmental test conditions (e.g., room temperature dry, elevated temperature wet). The test data sets are screened for contiguity of test conditions [1] 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 checked to see if the minimum requirements are met for the number of test conditions and batches, batch size, and total sample size. If any of these conditions are violated, the SP method is used to treat the data set at each specified test condition.

The test data sets are then screened for outliers at the batch level using the maximum normed residuals (MNR) statistical test [1 and 3]. The outliers are then dispositioned [1] and new data points are added (if needed) to satisfy the minimum size requirements for the batch or sample. Although this is not included in the flow chart, it is necessary that the batch statistics, such as the batch average and standard deviations, be computed at this stage. These values are then used in the statistical tests. The batches of data under individual test conditions are checked for between-batch variability using the k-sample Anderson-Darling test [1]. If the between-batch variability is significant, engineering judgment and experience (see Section 8.3.10.1 of CMH-17 [1]) could be exercised to ignore the batch variability and combine the batches. Otherwise, the data set at the particular test condition should be analyzed separately using the SP method. The remaining test conditions are then analyzed using the Pooling method. First, a check for outliers at the pooled level (pooled batches) is made using the MNR test on data sets at each test condition. If any outliers are detected, they must be dispositioned and new data points may be added, if necessary. The data at this test condition must then be reanalyzed for minimum size requirement and batch variability prior to being used in the pooling analysis.

The test conditions at which the data sets pass the statistical tests mentioned above are then normalized using their respective average values. Such statistics as standard deviations and coefficient of variations (CV) are computed for the pooled data sets. The equality of variances (and/or CV) within the data sets to be pooled are checked using Levene's test. If the data sets pass the Levene's test, the data sets are pooled. The appropriateness of using a normal distribution for the pooled data is then checked. If the normal distribution is found to fit the data, the pooled statistics and allowables for each of the pooled test conditions are calculated. In the event that the equality of variances is not satisfied, certain data sets may have to be ignored during the pooling process and analyzed using the SP method for generating allowables.

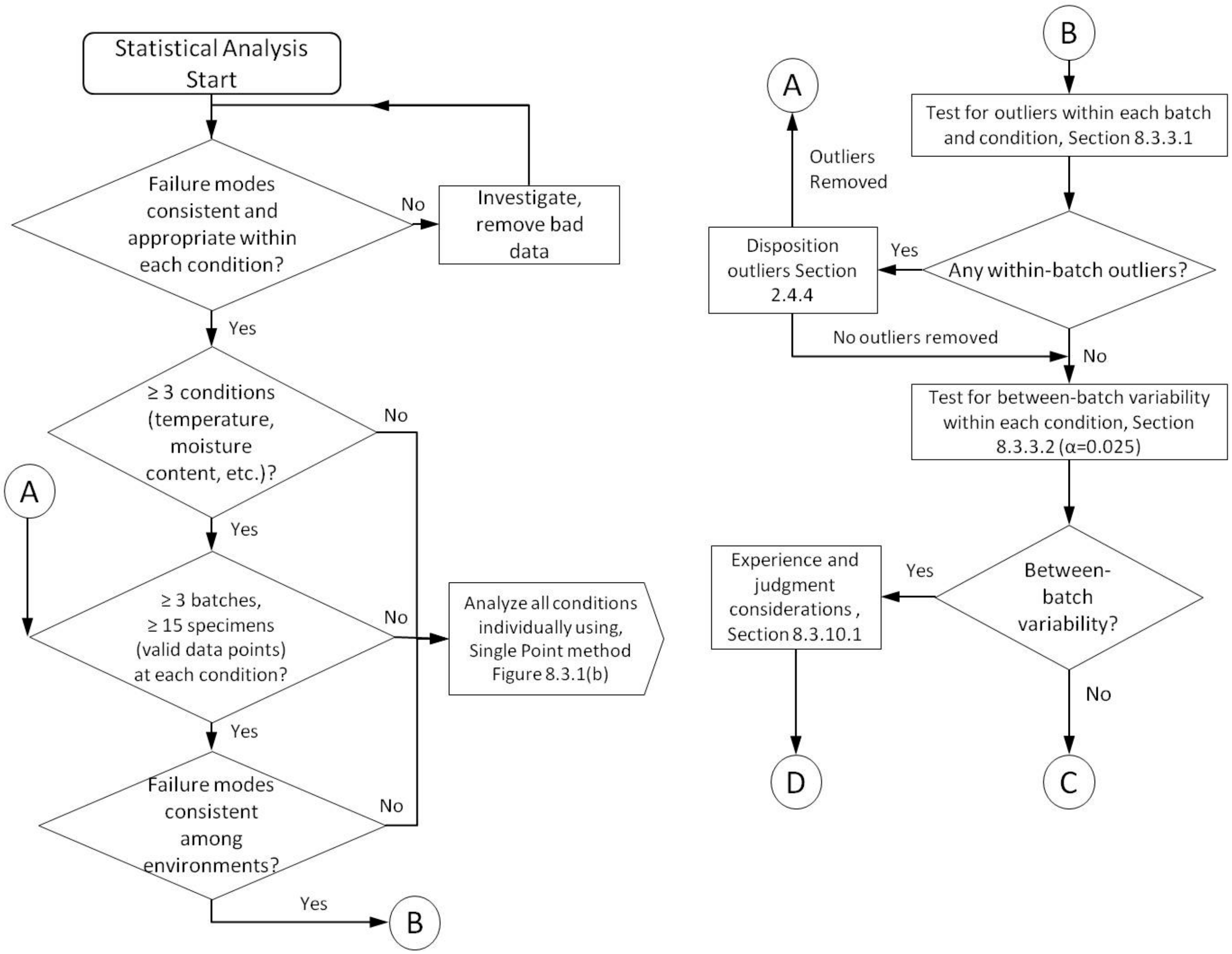


Figure 1. The CMH-17 Procedure for Generating Allowables

*All references to figures and section are from CMH-17

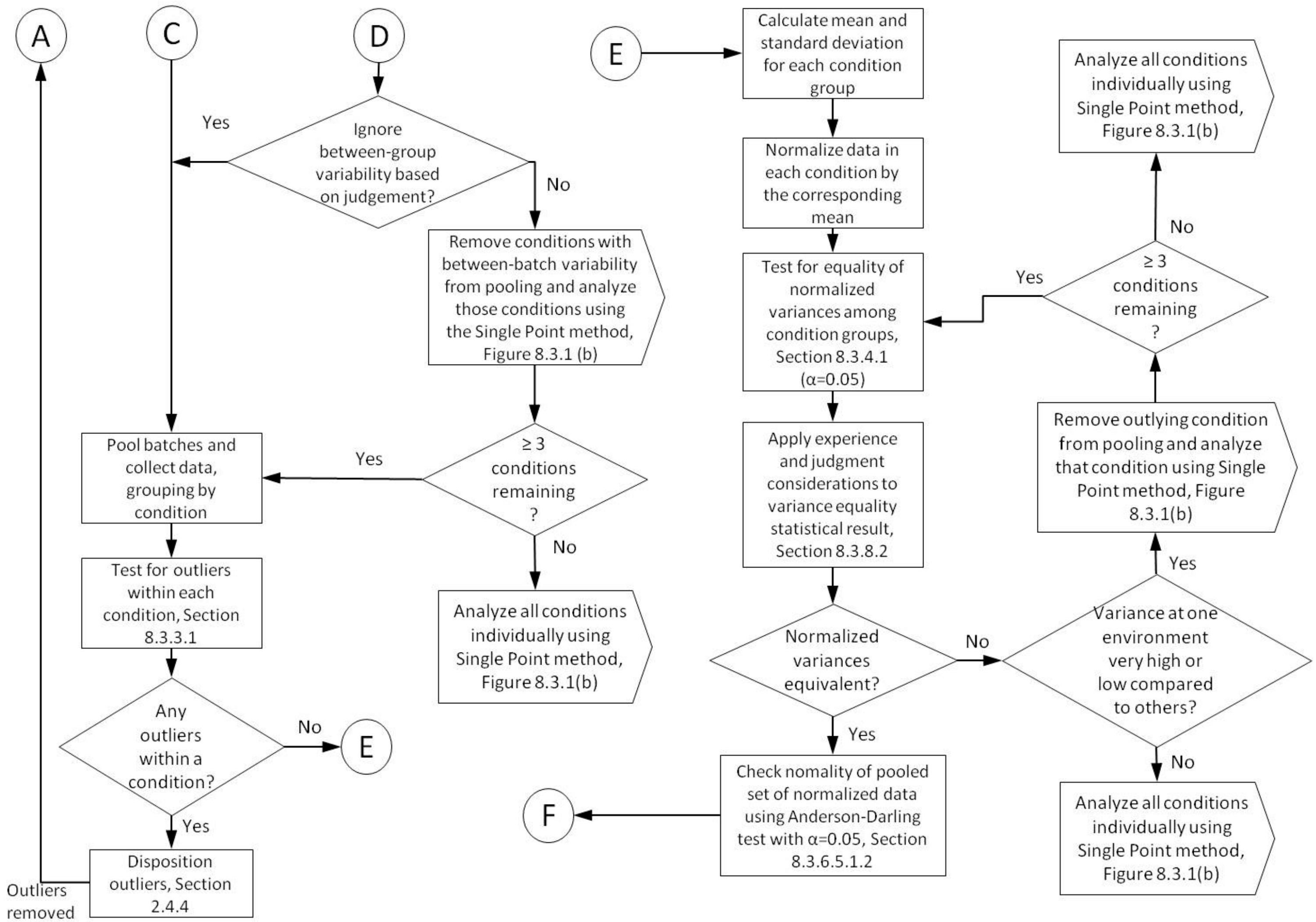
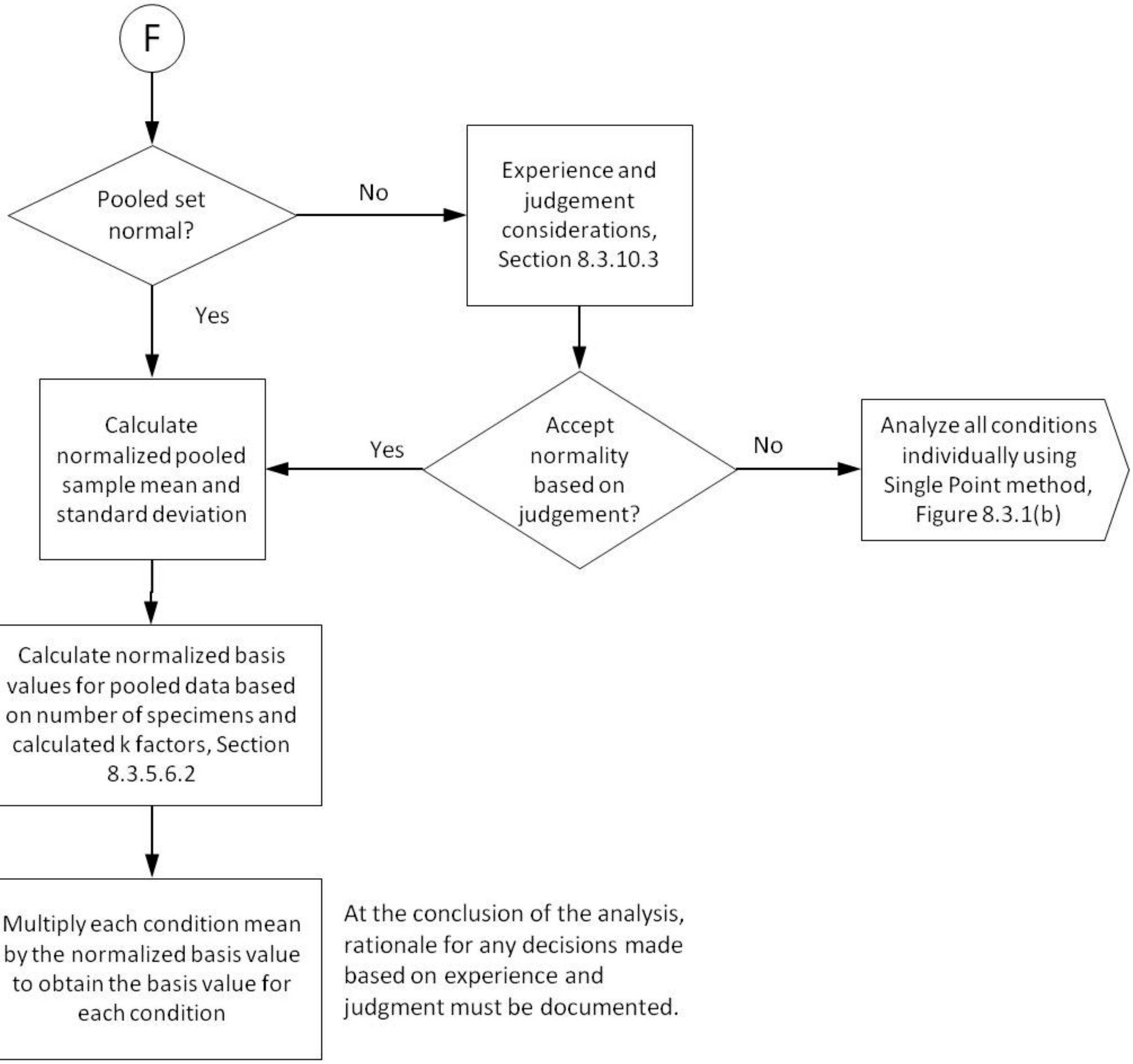


Figure 1. The CMH-17 Procedure for Generating Allowables (continued)

*All references to figures and section are from CMH-17



*All references to figures and section are from CMH-17

Figure 1. The CMH-17 Procedure for Generating Allowables (continued)

3. EXCEL VISUAL BASIC FOR APPLICATIONS.

A Visual Basic for Applications (VBA) with Microsoft Excel as the user interface was assembled to implement the CMH-17 procedure for generating statistically based material allowables. The organization of the Excel VBA program is illustrated in figure 2. The Excel VBA consists of a macro- (code) enabled workbook (file with .xlsm extension in the 2007 and .xls versions 97 through 2003). The workbook consists of worksheets that are used for input/output of data, presentation of results and charts, and execution and control of the program. The VBA macro consists of the program instructions to implement the CMH-17 procedure for generating allowables. The details of the workbook and VBA program are discussed in the following paragraphs.

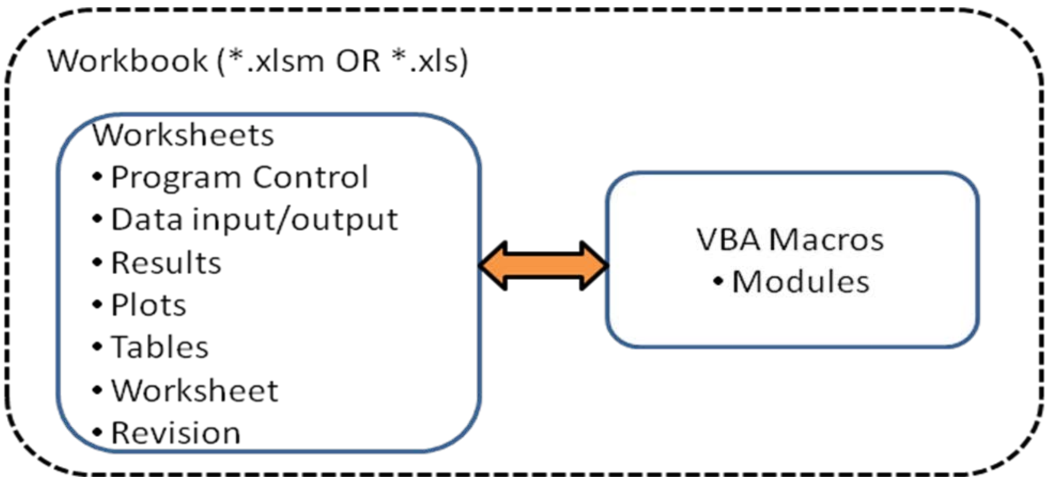


Figure 2. Overall Organization of the Excel VBA for Generating Allowables

The capabilities of the program are as follows:

- Can analyze data sets at ten environmental conditions with up to 1000 data points each.
- There is no limit on number of batches or batch size as long as they are compliant with the CMH-17 requirements.
- Reads data from files or manual input.
- Can process multiple sets of data (batch processing).
- Program may be run in "free" mode to override failure of statistical tests. Here, allowables using both Pooling and SP methods are generated. In the "compliant" mode, the CMH-17 flow chart is strictly enforced.
- Can perform diagnostic tests on data sets at individual test conditions.
- Checks the appropriateness of grouping data across different test environments.

The features of the program are as follows:

- Can perform double-precision math.
- Provides dynamic memory allocation for variables.
- Uses Microsoft Excel library for statistical functions (e.g., F-test).
- With each run of the program, a Microsoft Excel workbook (*.xls file) containing input data and results is created. Macros are not attached to this file.

- The number of batch plots generated is equal to the number of test conditions at which data are provided.

- Provides user interaction when processing a single data set (disabled for batch processing).

- User may opt to ignore the outliers for analysis.

- Provides error handling.

- Provides user options for controlling statistical tests (e.g., selection of significance levels).

3.1 MICROSOFT EXCEL WORKBOOK.

The interface between the user and the VBA program is facilitated through Microsoft Excel. The Microsoft Excel workbook (CMHSTATS.xlsm or CMHSTATS.xls) consists of worksheets and the VBA macro (running in the background). The worksheets available in the workbook are summarized in table 1. The images of individual worksheets are presented in appendix A. The table summarizes the utility/features of each worksheet along with the amount of user control over the same.

Table 1. Summary of Worksheets in the Workbook

Worksheet	Use	User Privileges
Instruction	Abbreviated instructions for user	None
CMH-17 main	Test control and execution	Options and control buttons
Data input	Input of test data	User allowed to change input regions
Results summary	Display of analysis results	None
Statistics summary	Display of statistical test results	Editing to substaniate engineering judgment and experience
Batch plots	Display of batch plots	None
Quantile plots	Display of quantile plots	None
Normal plots	Display of normal plots	None
Tables	Statistics tables for analysis	None
Plot data	Tabular data for plotting	None
Sort	For program use	None
Test sheet	Test new features and revisions	None
Messages	List real-time messages	None
Revisions	Compilation of revisions and changes	None

3.1.1 INSTRUCTION WORKSHEET.

The instruction worksheet contains an abridged version of the user guide for the program. The program contents include capabilities and instructions for the user to attain a quick knowledge of the program usage. This worksheet cannot be edited.

3.1.2 THE CMH-17 MAIN WORKSHEET.

The CMH-17 main worksheet is used for controlling program execution. This worksheet permits the program and user to:

- Input information about the organization, material, property, etc.
- Read input data from files.
- Select options for the program execution.
- Conduct diagnostic tests on specific data sets.
- Print copies of the analysis report.
- Facilitate batch processing of data sets.
- Provide a path for input data files and report files.
- Execute the program.

A screenshot of the CMH-17 main worksheet is shown in figure 3 below. The worksheet contains macro buttons that initiate specific program modules to conduct the tasks indicated by the labels on the macro buttons. The worksheet provides drop-down lists for selecting options for statistical tests as well as room for the user to input information about the material system, test method, etc. This information is used as part of the filename for storing the results from the analysis.

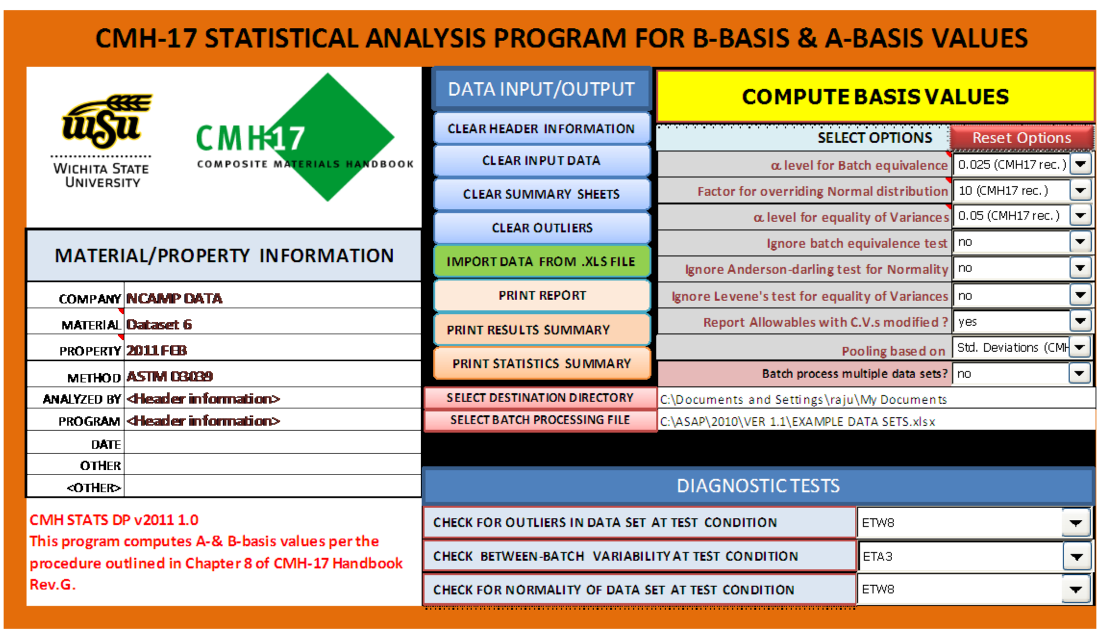


Figure 3. A Screenshot of the CMH-17 Main Worksheet

3.1.3 DATA INPUT WORKSHEET.

The worksheet provides space for the input of data for the program. The data may be entered manually via the keyboard or by copying from another Excel worksheet. For entering large data sets, it may be more convenient to read the data from a template file. The template file provided with the program may be filled with several worksheets and the data set of choice may be read into the input sheet. This can be accomplished by using the IMPORT DATA FROM .XLS FILE button on the CMH-17 main worksheet. During batch processing, the data sets are read successively into this sheet and cleared after the analysis. A screenshot of the data input worksheet is shown in figure 4. The user can define the test condition using a combination of alphanumeric identifiers (CTA, CTW, RTA, etc.) in a drop-down list and provide a description of the same in the space provided. The worksheet also provides the user with options to ignore a particular test condition during pooling. Corresponding to each data point, two adjacent cells (with columns labeled as B~outlier at batch level and P~outlier at pooled data level) will indicate if this data point is an outlier after each analysis run. The user may choose to analyze the data by simply ignoring the outlier or by retaining it in the analysis.

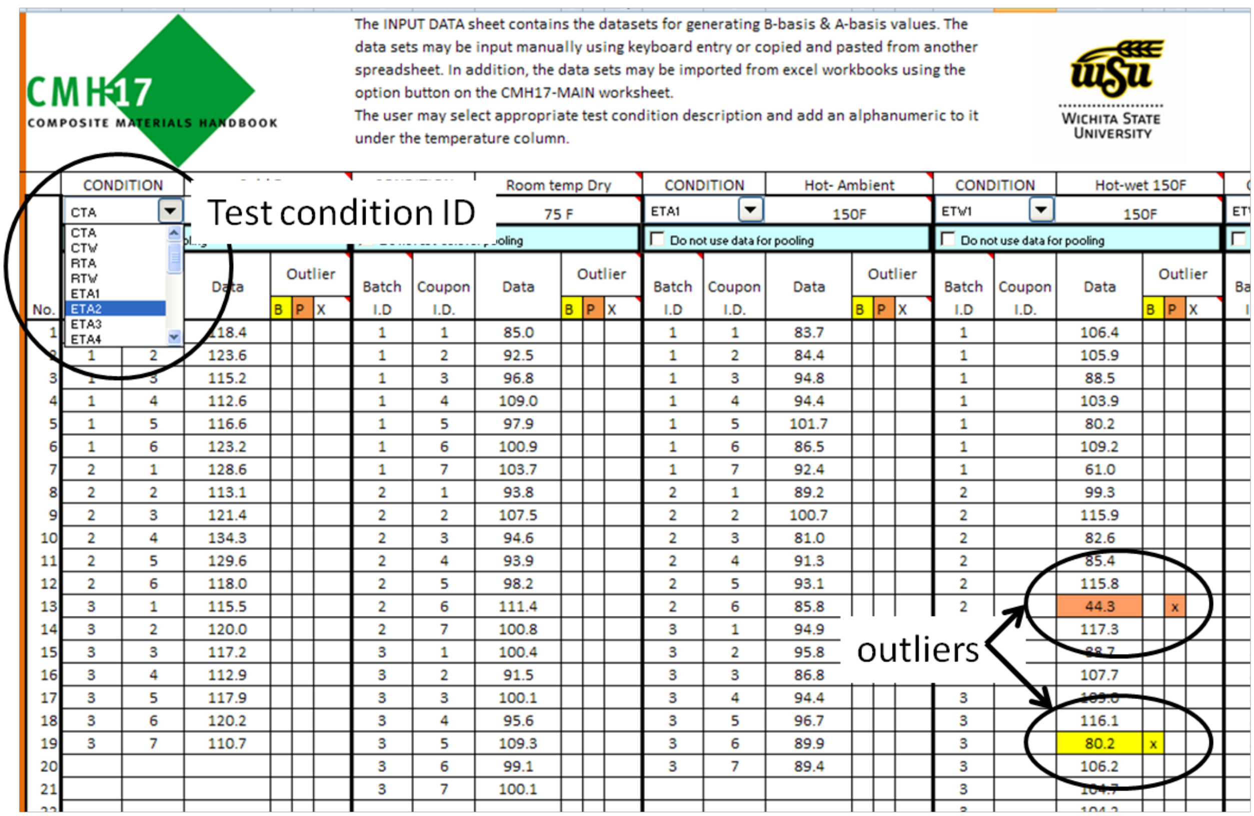


Figure 4. A Partial Screenshot of the Data Input Worksheet

3.1.4 RESULTS SUMMARY WORKSHEET.

A tabulated summary of the basic statistics (average, standard deviation, etc.) and allowable values associated with each test condition are summarized in this worksheet. A list of program-generated messages consistent with the CMH-17 flow chart is printed below the tabulated summary. The messages indicate potential problems with data sets and if the data sets passed statistical tests. A graphical display of average values and of basis values computed using SP and Pooling methods are also presented on this sheet. During batch processing of multiple data sets, any errors encountered during analysis and problems with data sets will be listed on this sheet.

3.1.5 STATISTICS SUMMARY WORKSHEET.

The detailed results from individual statistical tests are tabulated for each test condition on this sheet. In addition, space is provided for the user to key in statements of "engineering judgment/experience" based on the data presented on this sheet. The summary sheet presents the test results for the SP and Pooling methods separately.

3.1.6 BATCH PLOTS WORKSHEET.

This worksheet contains the plots of data set number versus the data set values. Reference lines corresponding to the group averages, standard deviations, and basis values from SP and Pooling methods are plotted on top of the data points for user reference as shown in figure 5. These plots may be used for presenting the input data and corresponding basis values graphically. These plots also aide in the engineering judgment to override failure of batch equivalence tests. The worksheet also generates a table that identifies the data set numbers with the batch identifiers provided by the user. In addition, the batch averages and standard deviations are listed in this table.

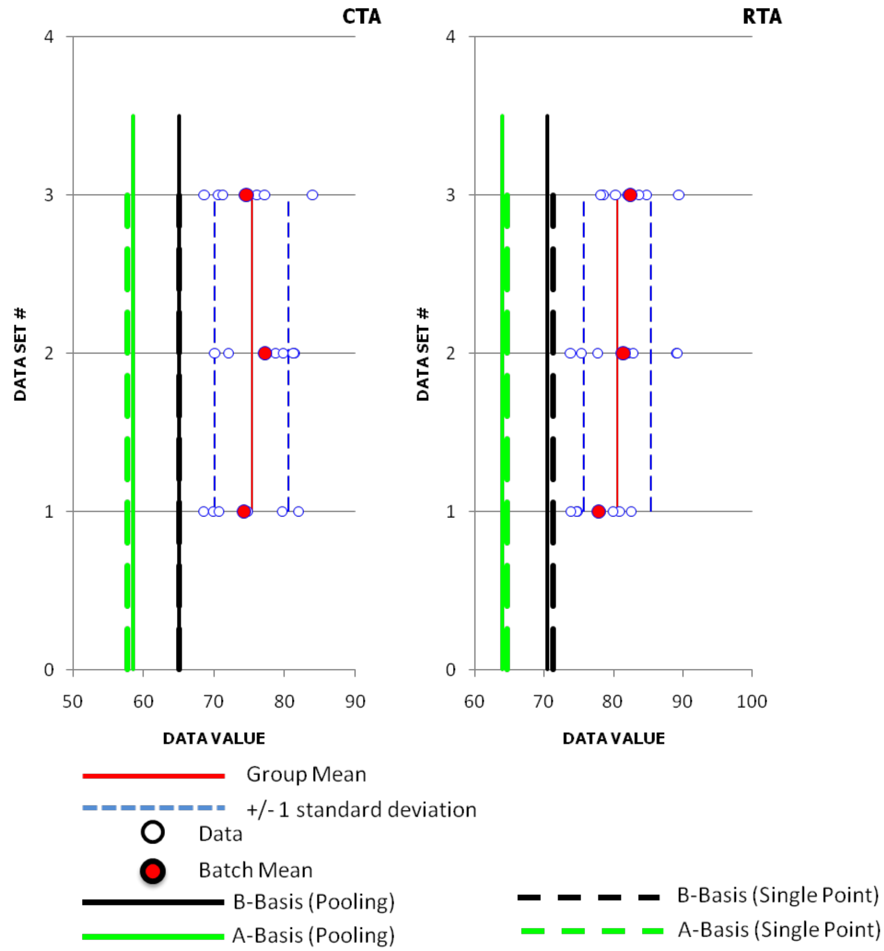


Figure 5. Typical Batch Plots Generated by the Program

3.1.7 QUANTILE PLOTS WORKSHEET.

This worksheet contains the quantile plots for test data at individual conditions and for the pooled data. These plots illustrate the symmetry, tail sizes, and median value of the sample and indicate the possible existence of outliers and inhomogeneous data [1].

3.1.8 NORMAL PLOTS WORKSHEET.

The normal plots and normal scores plots [4] for individual test conditions and pooled data are presented on this worksheet. The best-fit normal curves are superposed on top of the data points to facilitate engineering judgment in the event that one or more data sets fail the statistical tests.

3.1.9 TABLES WORKSHEET.

The tolerance factors for normal and Weibull distributions and non-parametric basis factors for small sample sizes are tabulated on this worksheet. The VBA program reads the values from this table when small sample sizes are analyzed.

3.1.10 PLOT DATA WORKSHEET.

The x-y data pairs for plotting the batch plots and normal plots are tabulated on this worksheet.

3.1.11 SORT WORKSHEET.

This sheet is used by the VBA program for sorting data sets using the Excel sorting function.

3.1.12 TEST SHEET WORKSHEET.

This sheet is used for testing the program by the developer whenever new features are added and to fix bugs reported by the users.

3.1.13 MESSAGES WORKSHEET.

The real-time messages generated by the VBA code during the program execution are listed here. The messages appear on a textbox during each program run.

3.1.14 REVISION WORKSHEET.

A table of program revisions and errors are maintained on this worksheet.

3.2 VISUAL BASIC PROGRAM.

The VBA program is a collection of program modules that contain the individual procedures and functions. The program has been organized into modules to facilitate easy debugging and changes of functionalities. The individual modules consist of procedures and functions, which can be called by other procedures across the modules during program execution.

In the following paragraphs, references to analysis of single data sets and of multiple data sets using batch processing will be made. A single data set consists of test data for a single material property (e.g., tension or compression) at multiple environmental test conditions (CTA, CTW, RTA, etc.). Under each environmental test condition, multiple batches of test data may be present. A multiple data set using batch processing implies a collection of single data sets (i.e., test data for single or multiple materials), for single or multiple material properties. The differences between a single data set, and multiple data sets using batch processing is captured in figure 6.

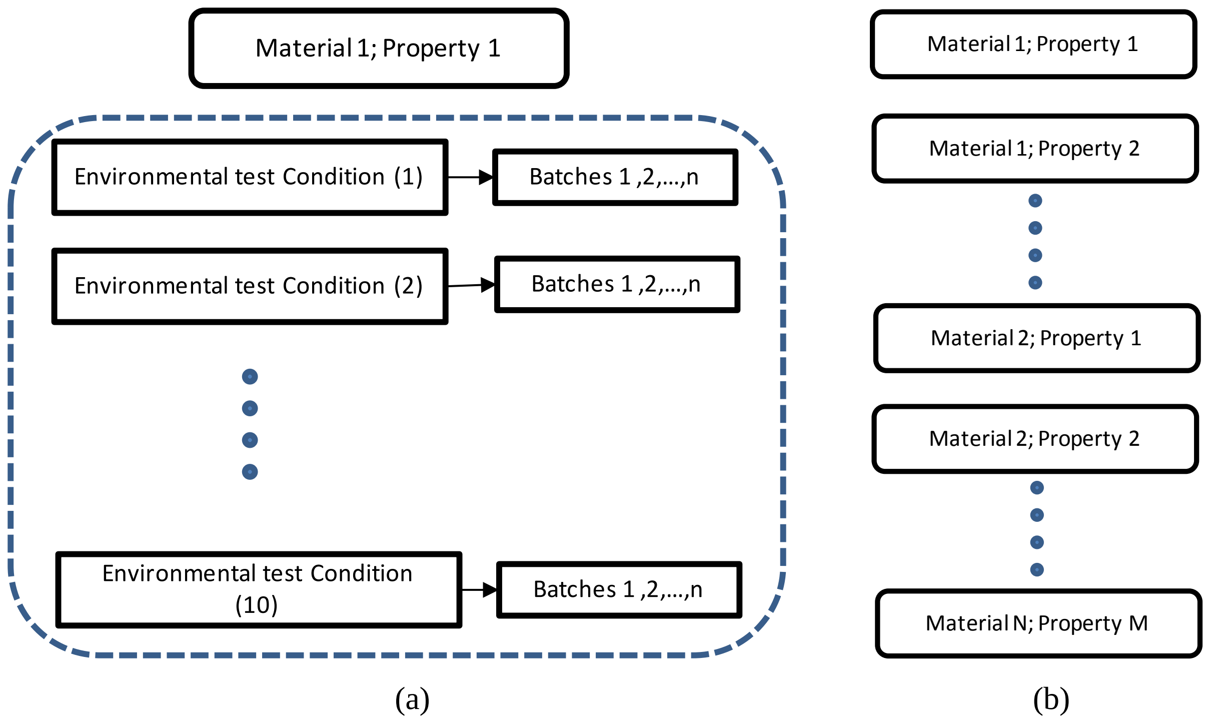


Figure 6. Definition of (a) a Single Data Set and (b) Multiple Data Sets

The analysis of a single data set is summarized in the flow charts presented in figures 7 through 9. The analysis begins with the program reading the user options for different statistical tests. To analyze single data sets, the program reads the data from the INPUT DATA sheet. For processing multiple data sets, the program first reads the data set from the batch processing file and writes it into the INPUT DATA sheet. The data set entered into the INPUT DATA sheet is consolidated to eliminate blank rows and noncontiguous batches of data. The data are then loaded into the arrays for analysis. The data set is first scrutinized for erroneous or bad data points (i.e., zeros, negative numbers, alphanumeric, all data points with same value). If bad data points or a set are identified, an error message window appears during the analysis of a single data set and the program is aborted. In the case of batch processing, the error message is written into the report file and the next data set is read into the INPUT DATA sheet for analysis.

The data set is subsequently checked for sizes (minimum of 15 samples and 3 batches per test condition) at each environmental test condition. The program generates error messages if the data does not conform to the minimum requirements and continues with the analysis. The error messages are written into the results summary worksheet of the report file for the user to address the deviation from the CMH-17 guidelines. The data at each test condition are then analyzed for the presence of outliers at the individual batch levels and after the pooling of batches. If an outlier is detected, the program displays an error message indicating the presence of outliers and prompts the user to abort the analysis or continue with outliers. If the analysis is continued, a warning message is written on the results summary worksheet for the user to address the presence of outliers per the guidelines in CMH-17. If the user chooses to abort the program and rerun it after dispositioning the outliers, the user has the option of retaining the outlier data point in the data set, but ignoring it during analysis. During batch processing, the user interaction is disabled and the program continues the analysis with outliers. After the test for outliers is completed, the program analyzes the different batches of data at each test condition for equivalence. The k-sample Anderson-Darling test [1] is used to check if the different batches belong to the same population. In the event that the batches do not pass the test, the program will recommend the use of the SP statistics and will not include the data at this particular condition for pooling analysis. The user, however, may use the option of overriding the failure of this test at the beginning of the analysis. The batches of data passing the batch equivalence test are pooled and checked using the Anderson-Darling test [1] for normality. If the data set passes the normality test, it is added to the pooled data set (pooling across test conditions). The process is repeated for all test conditions at which the data are provided.

The test data at individual test conditions are then subjected to SP and Pooling analyses as shown in the flow chart in figure 8. The SP analysis is conducted on all test conditions, whether the data at each condition qualifies for pooling or not. The statistics associated with normal, Weibull, log-normal and non-parametric methods are computed. In addition, Levene’s test [1] for equal variances across batches is conducted prior to ANOVA analysis of the data. Using the Observed Significance Levels (OSL) and depending on whether the data is structured or unstructured, the basis values are recommended per the guidelines in the CMH-17 handbook.

Using the Pooling method, the data is pooled across different test conditions to generate allowables. The data sets are first analyzed for equality of variances (or equality of CVs) to establish the validity of pooling the data across test conditions. Next, the pooled data set is checked for normality. In addition, the grouping of test conditions is checked per the CMH-17 guidelines. If the grouping is found to be unacceptable, the program writes a warning message in the results summary sheet. If the grouping is found to be acceptable and the pooled data passes the tests for equality of variances and normality, the pooled variance (or CV) is used to compute the allowables at each test condition.

The CMH-17 procedure also provides an option for modifying the data sets at individual test conditions using an altered CV. The alteration of CV is based on the observation that most composite material properties have a CV between 4% and 10%. The individual batches are modified using the rule [1] specifying the CV change. If the modified batches of data pass the batch equivalence test, then the process is continued until basis values are computed using the Pooling method. The flow chart in figure 9 shows this process.

After generating the allowables for original and modified data (if necessary), the program generates plots to enable the visualization of test data. These plots are used for exercising engineering judgment to override any failed statistical test. Finally, the program creates a report file containing the summary of the analysis along with the plots. In the case of batch processing, the program reads in the next data set and the entire process is repeated.

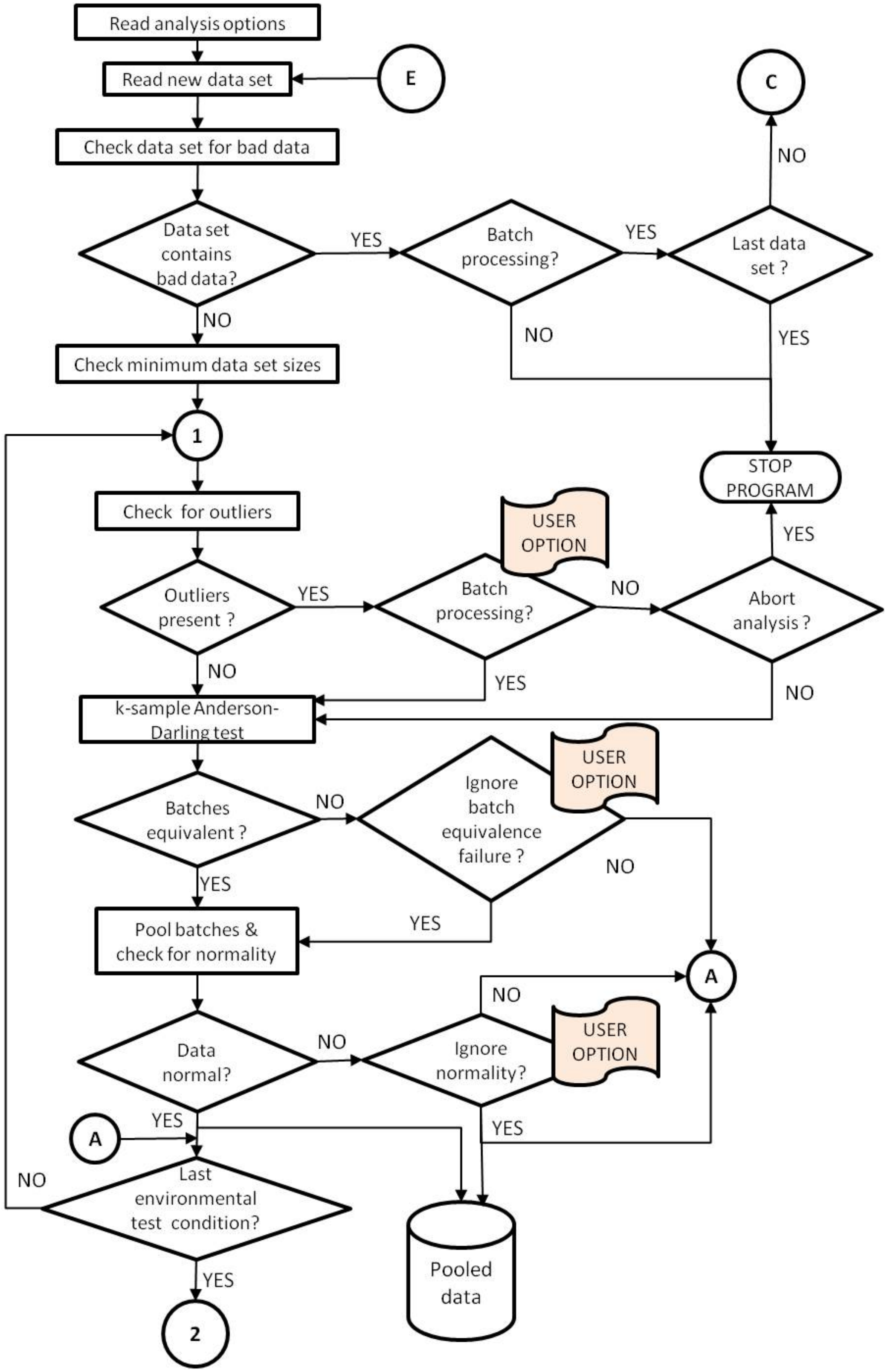


Figure 7. Flow Chart Showing the Initial Stage of the Analysis

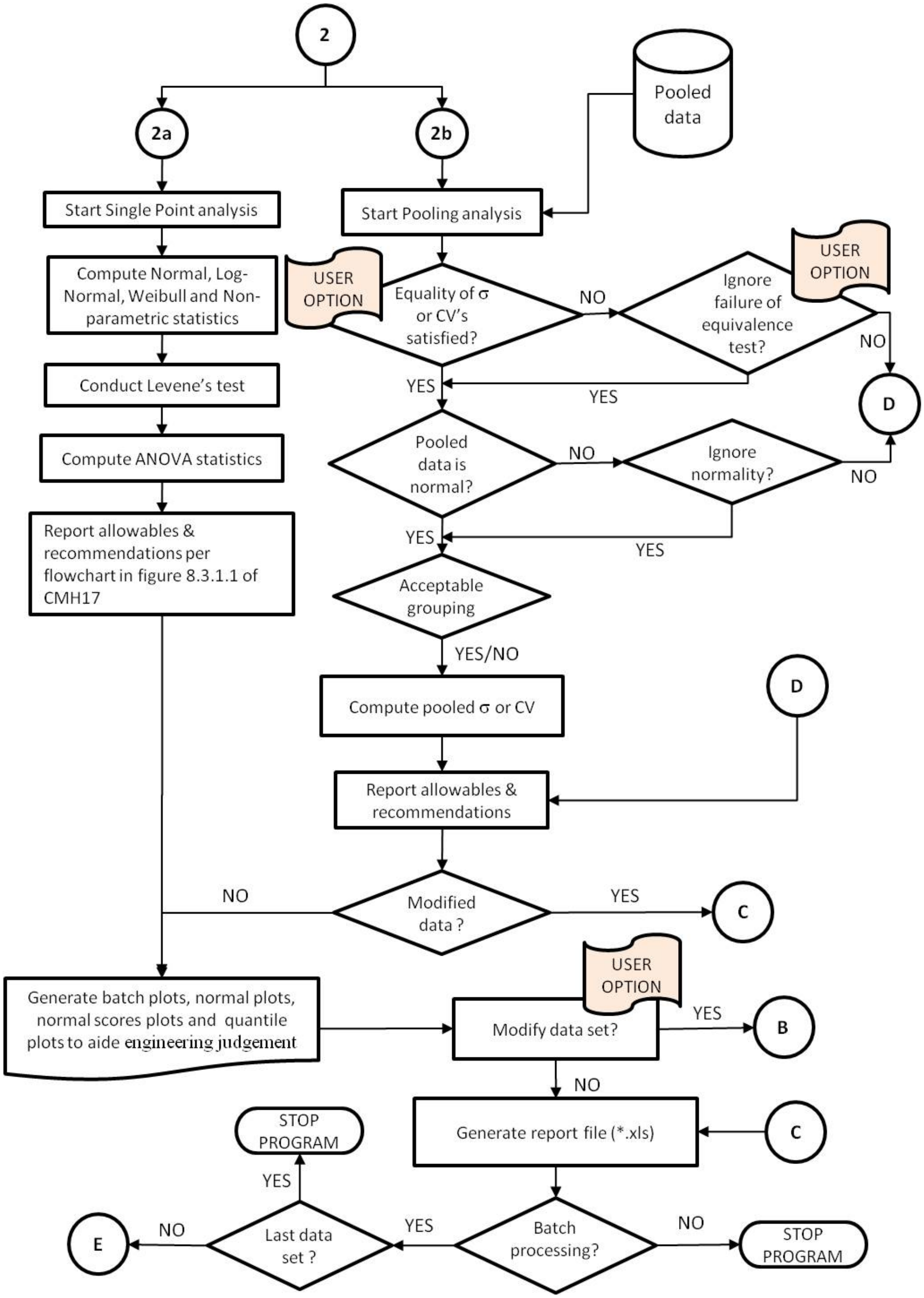


Figure 8. Flow Chart Showing the Second Phase of the Analysis

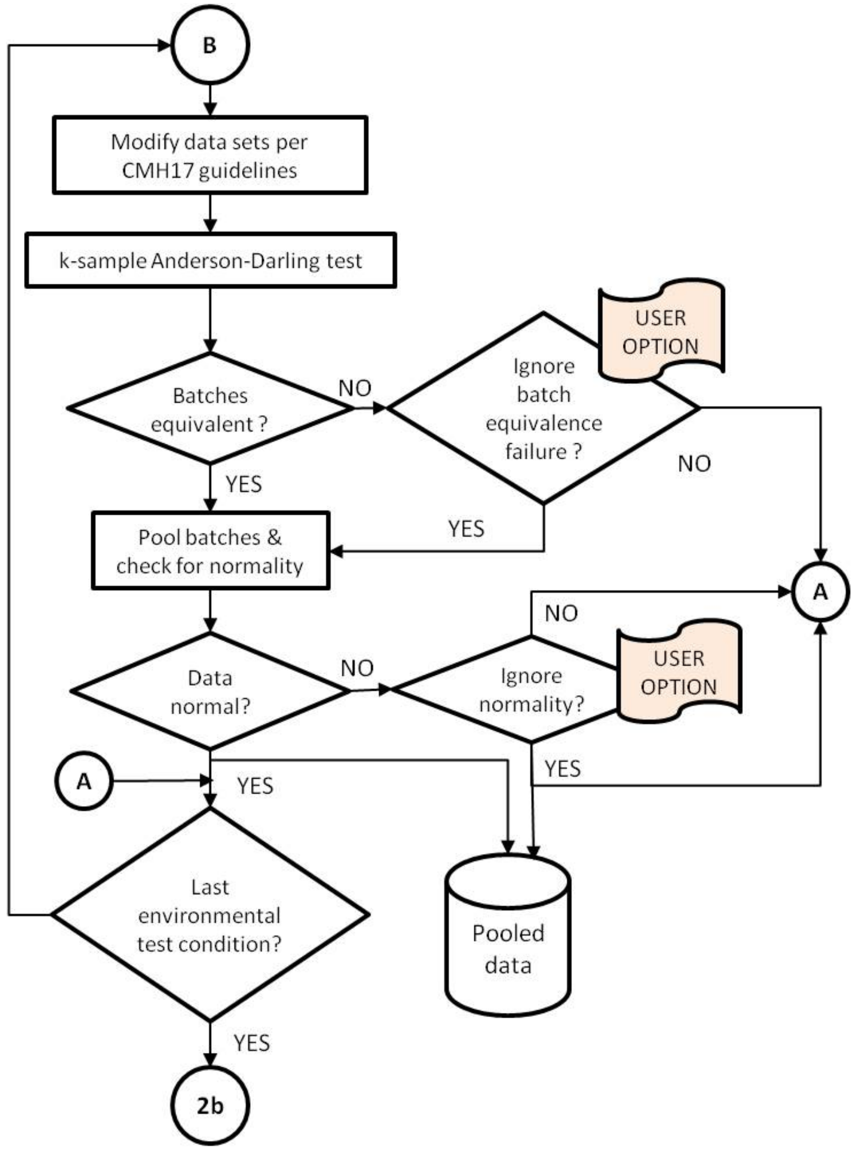


Figure 9. Flow Chart Showing the Analysis of CV Modified Data

4. USER GUIDE.

In this section, a step-by-step procedure for using the CMHSTATS program and its capabilities is presented. Hypothetical data sets (appendix A) are used to serve as examples for the exercise.

4.1 INPUT DATA.

The data input for the program may be provided through the use of template files or by manually keying in the values in the INPUT DATA sheet. The data input for the program must follow a specific format. Each data point must be accompanied by a batch identifier (alphanumeric) and an optional specimen number or identifier. The batches of data need not be contiguous and blank rows are acceptable. In addition, to accommodate grouping checks for Pooling analysis, the built-in alphanumeric identifiers must be used for each test condition. It is the responsibility of the user to correctly assign these identifiers based on the temperature and humidity combinations used during the generation of experimental data. A sample data template and acceptable data formats are shown in figure 10. The first six rows of Column A may be used for providing information about the material system, property, etc. On the eighth row, the alphanumeric for the test condition must be provided per the table at the top of the worksheet. Optional descriptions of the test conditions may also be provided. A single workbook may contain multiple worksheets of data sets, which may be analyzed individually or batch processed.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	Example Data 1										<test cond>	: use the following alphanumerics for identi				
2	material 1														Temperature	
3	property 1										Condition	CT	RT	ET1	ET2	ET3
4	ASTM Dxx										Ambient	CTA	RTA	ETA1	ETA2	
5	M& P Group										Wet	CTW	RTW	ETW1	ETW2	
6	Initial Qualification										Note: The temperatures associated with CT < RT < ET1 <ET2 < .					
7																
8	CTA	65F		RTA	75 F		ETA1	150F		ETW1	150F		ETW2	180F		<te
9	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Ba
10	1	1	118.37	1	1	84.96	1	1	83.74	1		106.36	1		99.02	
11	1	2	123.6	1	2	92.49	1	2	84.38	1		105.89	1		103.34	
12	1	3	115.22	1	3	96.82	1	3	94.8	1		88.46	1		100.3	
13	1	4	112.63	1	4	109.03	1	4	94.39	1		103.9	1		98.46	
14	1	5	116.56	1	5	97.89	1	5	101.7	1		80.21	1		92.26	
15	1	6	123.16	1	6	100.92	1	6	86.54	1		109.2	1		103.49	
16	2	1	128.59	1	7	103.69	1	7	92.38	1		61.01	1		113.73	
17	2	2	113.14	2	1	93.79	2	1	89.21	2		99.32	2		108.17	
18	2	3	121.42	2	2	107.53	2	2	100.69	2		115.86	2		108.42	
19	3	1	115.45	2	3	94.57	2	3	81.04	2		82.61	2		116.26	
20	3	2	120.03	2	4	93.88	2	4	91.34	2		85.37	2		121.05	
21	3	3	117.16	2	5	98.23	2	5	93.14	2		115.8	2		111.22	
22	3	4	112.93	2	6	111.35	2	6	85.82	2		44.32	2		104.57	
23	3	5	117.91	2	7	100.82	3	1	94.89	2		117.32	2		103.22	
24	3	6	120.19	3	1	100.38	3	2	95.81	2		88.67	3		99.39	
25	3	7	110.73	3	2	91.5	3	3	86.78	3		107.68	3		87.34	
26	2	4	134.32	3	3	100.08	3	4	94.4	3		108.96	3		102.73	
27	2	5	129.64	3	4	95.63	3	5	96.72	3		116.12	3		96.37	
28	2	6	117.98	3	5	109.3	3	6	89.9	3		80.23	3		99.59	
29				3	6	99.12	3	7	89.37	3		106.15	3		97.07	
30				3	7	100.07				3		104.67				
31										3		104.23				

Figure 10. Template for Input Data

4.2 ANALYSIS OF SINGLE DATA SET.

Example data sets are provided in the workbook titled EXAMPLE DATA SETS.xlsx (.xls for 97 through 2003 versions). The analysis of single data sets consisting of data at different environmental test conditions is shown in the following steps:

1. On the CMH-17 MAIN worksheet (figure 11) enter the name of the organization or department, material system, property (e.g., tension, compression, etc.), test method (e.g., ASTM D3039, etc.) and other pertinent information to identify the data set to be analyzed. The information may contain numbers, letters, and special characters (/ \ ? % * : | “ ”< > .). The program uses the information to label the report files. The special characters will be removed while forming file names.

2. Select the analysis options appropriate for the data sets. The options for the different statistical tests are chosen by using the drop-down lists as shown in figure 12. To use the default values (CMH-17 recommended), click on the RESET OPTIONS button as shown in the figure. Whenever the workbook is opened (after closing), the options are automatically set to the default values. For more information on the options, refer to CMH-17 handbook [1].

3. Set the “Batch process multiple data sets?” option to “no” using the drop-down list as shown in figure 12. This feature is not needed for analyzing single data sets.

4. Select a destination path for the report file by clicking the SELECT DESTINATION DIRECTORY button. The report file for the analysis will be saved under this directory. If a path is not selected, the program will prompt the user for a destination path at the end of the analysis.

5. Enter the data for analysis on the DATA INPUT worksheet. This may be done by manually entering the data or by reading the data from a template file. To read the data from a template file, use the IMPORT DATA FROM .XLS FILE button to select the data file. The program will prompt the user to select the appropriate worksheet in the data file as shown in figure 13. On selection of the worksheet, the program reads the data and writes it into the DATA INPUT worksheet. The user may now assign (if not assigned in the input file) the alphanumeric for identifying the test conditions by using the drop-down menus and space assigned for entering a description of the test condition (as shown in figure 14). The alphanumerics[*] for test conditions must be correctly assigned by the user and will dictate the correctness of grouping checks made by the program. The computer program identifies the test conditions based solely on the alphanumeric; any numerical/alphanumeric descriptions are for user reference only. In the presence of

[*] CTA, CTW ~ Cold Temperature Ambient and Cold Temperature Wet
RTA, RTW ~ Room Temperature Ambient and Room Temperature Wet
ETAx, ETWx ~ Elevated Temperature Ambient and Elevated Temperature Wet, x=1, 2, 3, 4, 5

duplicate identifiers, the program will skip the grouping check and will be indicated in the error messages.

6. (Optional) The user may conduct some diagnostic tests on the data sets entered on the DATA INPUT worksheet. The diagnostic tests include tests for outliers, batch equivalence, and normality. The significance levels selected under the analysis options will be used for these tests. As an example, to conduct the test on batch equivalence at the RTA condition, select the condition identification using the drop-down list next to the CHECK FOR BATCH EQUIVALENCE AT CONDITION button. The test may now be run by clicking on the same button. Upon running the statistical test, the program will display one of the message windows shown in figure 15, indicating the results of the test.

7. To begin the analysis, click on the COMPUTE BASIS VALUES button. The program begins the analysis of the test data and a status window opens, as shown in figure 16. The status window provides information about the progress of the analysis and a summary of the statistical tests at each test condition for which data were entered.

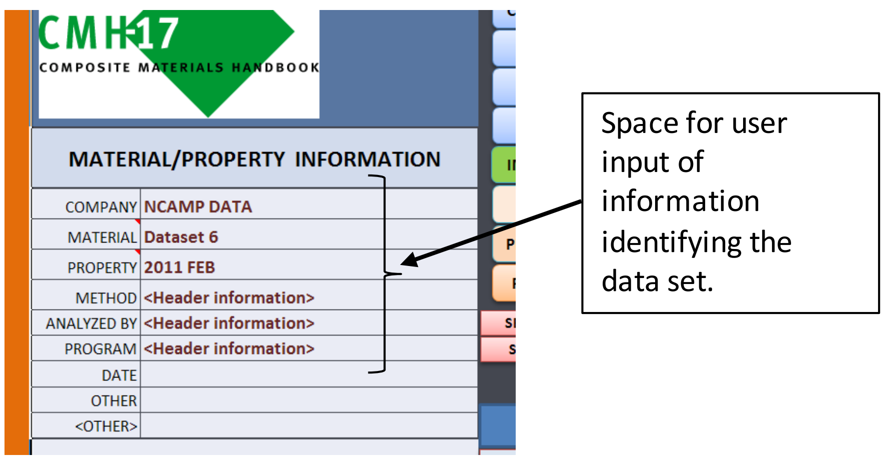


Figure 11. User Input of Information to Identify Data Set

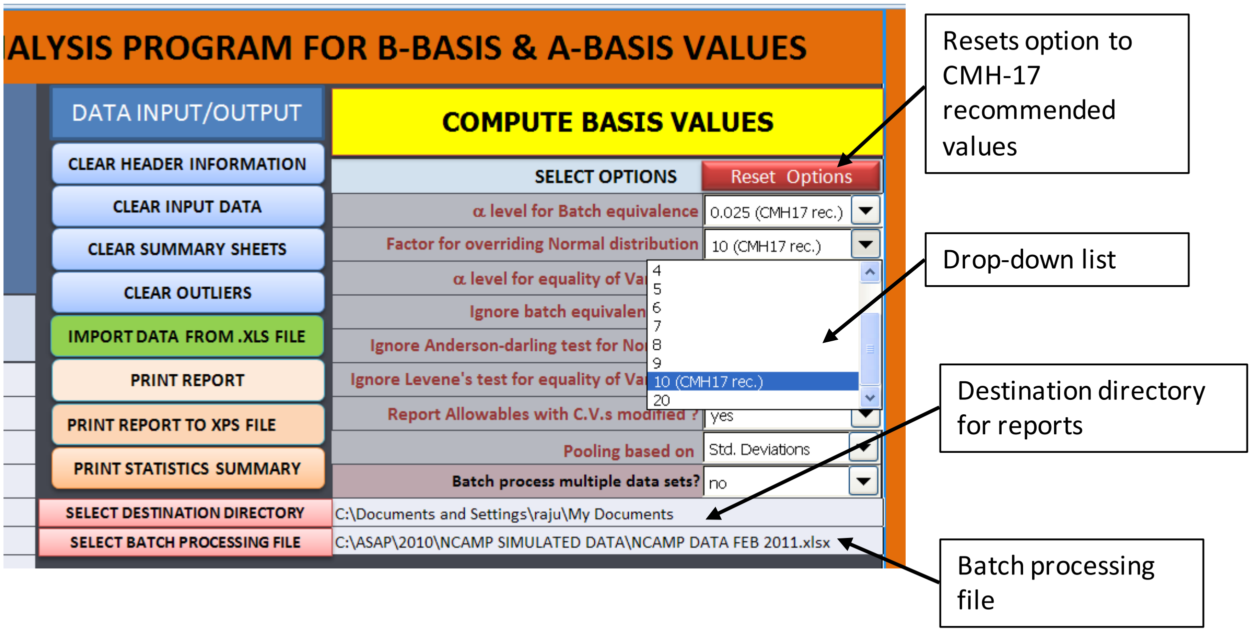


Figure 12. Selection of Options on CMH-17 MAIN Worksheet for Statistical Analysis

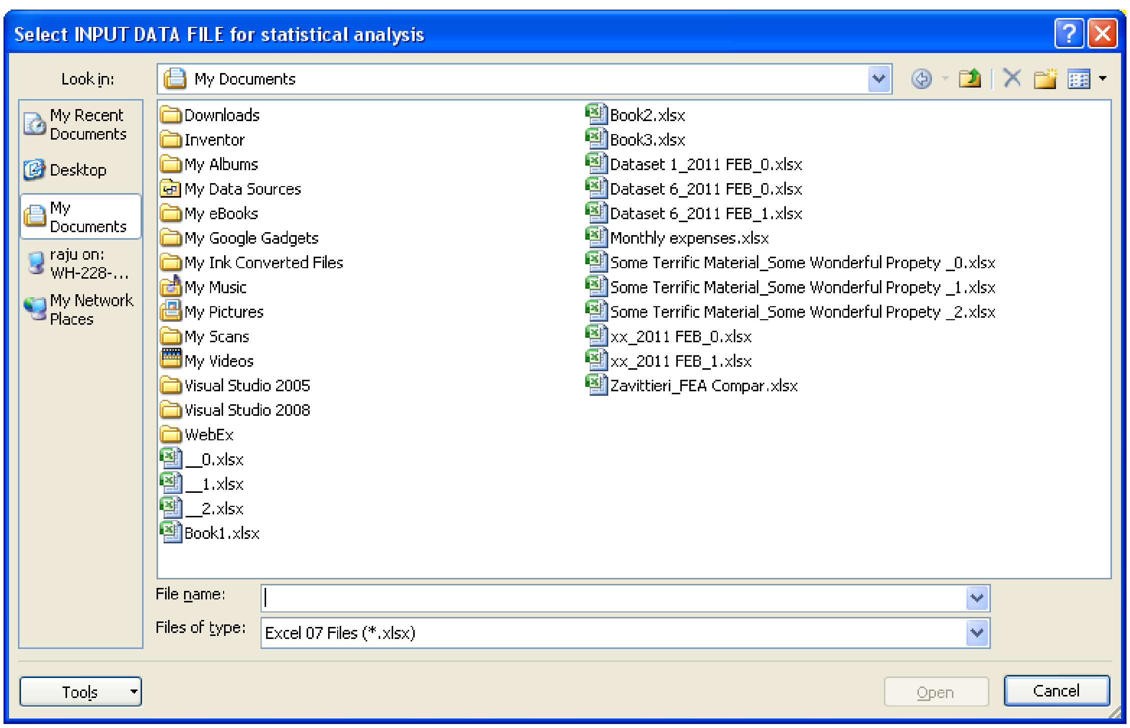


Figure 13. Program Window to Select Input Data File

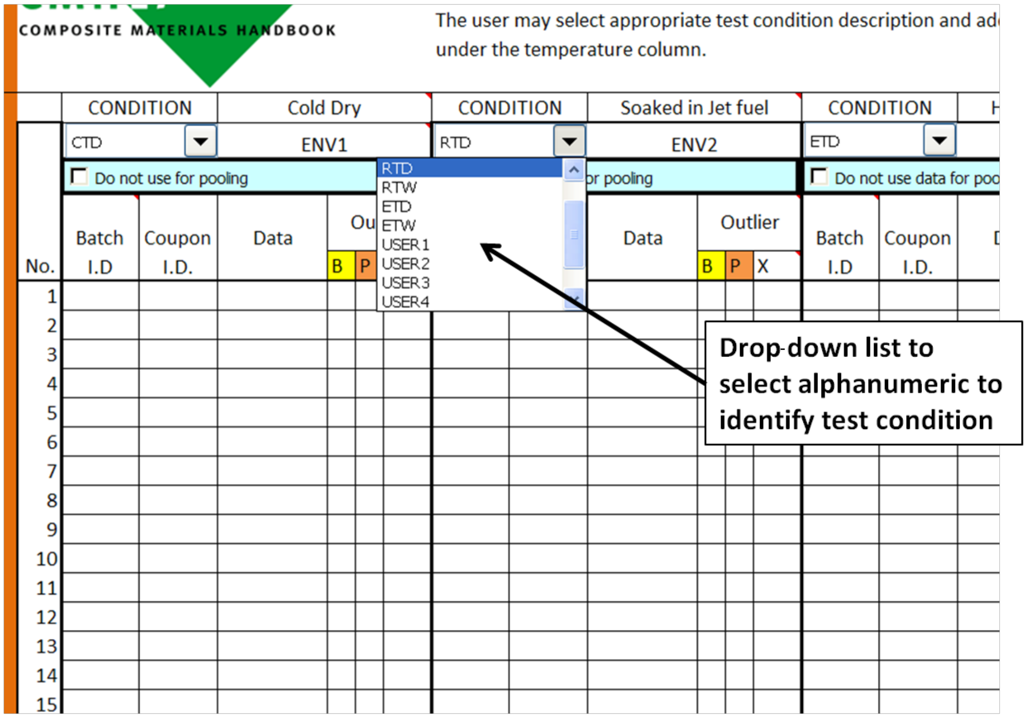


Figure 14. Drop-Down Menu for Assigning Alphanumeric Identifier for Test Condition

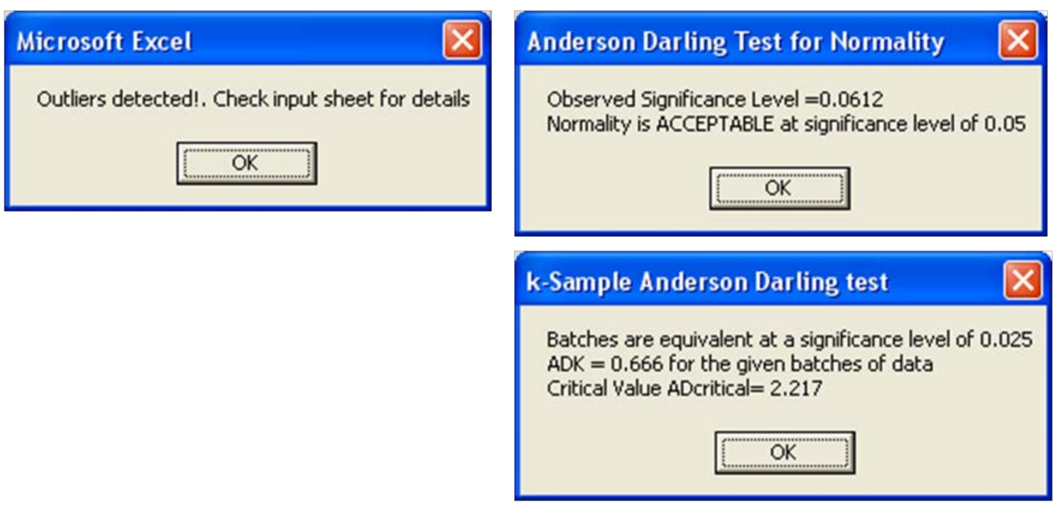


Figure 15. Message Windows Generated by Program After Conducting Diagnostic Tests

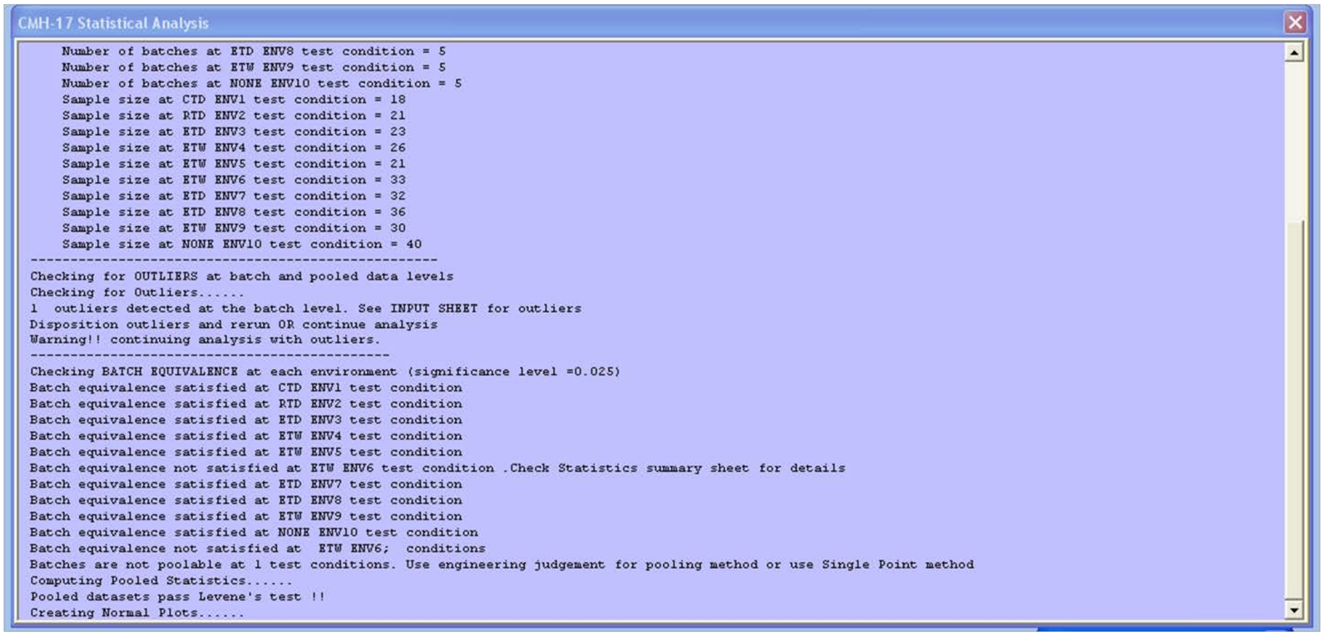


Figure 16. Status Window Indicating the Progress of the Analysis

4.3 ANALYSIS OF MULTIPLE DATA SETS (BATCH PROCESSING).

The analysis of multiple data sets is similar to single data sets, with the exception that the user interactions during the program execution are disabled. The following steps are involved in running a batch process:

1. Generate a batch processing file using the Excel data input file template. The multiple data sets should be included on separate worksheets of the same workbook. The number of worksheets in the workbook is limited by the version of Excel being used.

2. Select the analysis options appropriate for the data sets using the drop-down lists. To select the default values (CMH-17 recommended), click on the RESET OPTIONS button, as shown in figure 17. The same combination of options will be used for analyzing all data sets in the batch processing file.

3. Set the *Mulitple data refg?* option to “Yes” using the drop-down list.

4. Select the batch processing file by using the SELECT BATCH PROCESSING FILE button. If a batch processing file is not selected, the user will be repeatedly prompted for the file during runtime.

5. Select a destination for the report file by clicking the SELECT DESTINATION DIRECTORY button. The report files from the analysis will be saved under this directory.

6. To begin the analysis, click on the COMPUTE BASIS VALUES button. The program begins the analysis of the test data and a status screen opens as shown in figure 16. The status screen provides information about the progress of the analysis and a summary of the statistical tests for each test condition for which data was entered.

5. EXAMPLE PROBLEMS.

5.1 EXAMPLE PROBLEM I.

In this example (see appendix B, table B1), test data at environmental test conditions 1/m CTA, RTA, ETA1, ETW1, and ETW2 1/M are provided. The data may be entered manually or read from the EXAMPLE DATA SETS.xlsx workbook. Note that the specimens IDs are not necessary, but batch IDs are required. In this data set, the specimen IDs are provided for data at CTA, RTA, and ETA1 conditions, while only the batch IDs are provided for data at ETW1 and ETW2 test conditions. The alphanumeric identifiers for test conditions must be assigned correctly. If an input file is being used, the alphanumeric identifiers must be one of the values listed in the table at the top of the template worksheets. Note that the program is insensitive to the case of letters, spaces, or hyphens between characters. Therefore, CTA could be specified as any of the following: *cta*, *Cta*, *CT-A*, or *C T A*.

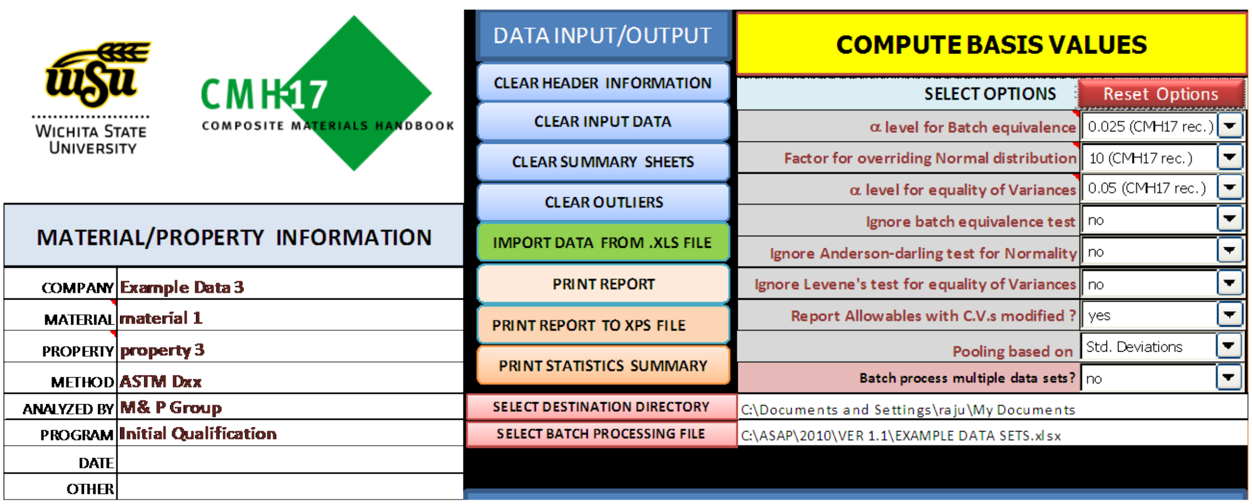


Figure 17. Options Chosen for Example Problem I

The analysis is started by clicking the COMPUTE BASIS VALUES button. The program analyzes the data set for sizes, invalid data points, and outliers. For the given data set, the program identifies two outliers in the ETW1 data. One of the data points is an outlier at the batch level, while the other is an outlier after pooling the batches as shown in figure 18. The analysis may be rerun after appropriately dispositioning the outliers. The user may retain the outliers in the data set but ignore them for analysis by placing any character in the column labeled “X” next to the outlier columns.

The summary of the results for the data set is shown in figure 19, which is a portion of the RESULTS SUMMARY worksheet. The accompanying comments generated by the program are listed in figure 20. Since the default options were used for the analysis, the program enforces the CMH-17 flow chart rigorously. As a result, the data at ETW1 and ETW2 test conditions are not included in the Pooling analysis as they fail the batch equivalence test (ETW2) and normality test (ETW1), as shown in the summary of diagnostic tests in figure 21. This summary is available for the user in the Statistics Summary worksheet. The three remaining test conditions are pooled based on equality of variances and the basis values are reported. The program further indicates (see figures 20 and 21) that the grouping of CTA, RTA, and ETA1 test conditions are acceptable per the CMH-17 guidelines.

For the data at ETW1 and ETW2, the SP analysis indicates values based on the nonparametric method and ANOVA. Further, the ANOVA values are highlighted to indicate lack of sufficient number of batches. A summary of diagnostic tests for SP analysis is shown in figure 22, which is again a portion of the Statistics Summary worksheet.

To aid engineering judgment per the CMH-17 guidelines, graphical representation of the data are provided in terms of batch plots (figure 23), quantile box plots, and normal plots. The normality of data at the ETW1 condition is clearly questionable as indicated by a dominant tail in the quantile box plot (figure 24) and an unsatisfactory fit of the normal curve to the data in the normal plots (figure 25).

A report file labeled C:\..path..\material1_property3_0.xlsx is created. This file contains copies of the input data sheet, summary, and plot sheets. If the data are reanalyzed without changing the material and property information on the main sheet, the subsequent report files are named by incrementing the last character of the filename (.._property3_1.xls).

CONDITION		Hot- Ambient			
ETA1		150F			
Do not use data for pooling					
Batch I.D	Coupon I.D.	Data	Outlier B	Outlier P	Outlier X
1	1	83.7			
1	2	84.4			
1	3	94.8			
1	4	94.4			
1	5	101.7			
1	6	86.5			
1	7	92.4			
2	1	89.2			
2	2	100.7			
2	3	81.0			
2	4	91.3			
2	5	93.1			
2	6	85.8			
3	1	94.9			
3	2	95.8			
3	3	86.8			
3	4	94.4			
3	5	96.7			
3	6	89.9			
3	7	89.4			

CONDITION		Hot-wet 150F			
ETW1		150F			
Do not use data for pooling					
Batch I.D	Coupon I.D.	Data	Outlier B	Outlier P	Outlier X
1		106.4			
1		105.9			
1		88.5			
1		103.9			
1		80.2			
1		109.2			
1		61.0			
2		99.3			
2		115.9			
2		82.6			
2		85.4			
2		115.8			
2		44.3		x	
2		117.3			
2		88.7			
3		107.7			
3		109.0			
3		116.1			
3		80.2	x		
3		106.2			
3		104.7			
3		104.2			

CONDITION		45 days 85% RH			
ETW2		180F			
Do not use data for pooling					
Batch I.D	Coupon I.D.	Data	Outlier B	Outlier P	Outlier X
1		99.0			
1		103.3			
1		100.3			
1		98.5			
1		92.3			
1		103.5			
1		113.7			
2		108.2			
2		108.4			
2		116.3			
2		121.1			
2		111.2			
2		104.6			
2		103.2			
3		99.4			
3		87.3			
3		102.7			
3		96.4			
3		99.6			
3		97.1			

Figure 18. Input Data Sheet Indicating the Outliers at ETW1 Test Condition

DATA SUMMARY										
STATISTIC	ENVIRONMENTAL TEST CONDITION									
	CTA 65F	RTA 75 F	ETA1 150F	ETW1 150F	ETW2 180F					
Sample Size	19	21	20	22	20					
No. of Batches	3	3	3	3	3					
Mean	119.42	99.15	91.35	96.92	103.30					
Std.Dev	6.25	6.52	5.56	18.80	8.11					
% Co.V	5.23	6.58	6.09	19.40	7.85					
Minimum	110.73	84.96	81.04	44.32	87.34					
Maximum	134.32	111.35	101.70	117.32	121.05					
SUMMARY OF BASIS VALUES FOR DATA POOLED ACROSS ENVIRONMENTAL TEST CONDITIONS										
Pooled % Co.V	6.02									
K_b	1.7507	1.7342	1.7421							
K_a	2.9208	2.9073	2.9138							
Pooled St.Dev. Basis Values (Original data)										
B-Basis Value	108.69	88.51	80.67							
A-Basis Value	101.52	81.32	73.49							
Pooled St.Dev. Basis Values (Modified C.V data)										
Modified %C.V	6.62	7.29	7.05	19.40	7.93					
Modified Pooled % Co.V	7.00									
B-Basis Value	106.82	86.66	78.81							
A-Basis Value	98.39	78.21	70.37							
SUMMARY OF BASIS VALUES USING SINGLE POINT APPROACH (Original Data)										
Distribution	Normal	Normal	Normal	Non-Parm.	ANOVA**					
B-Basis Value	107.25	86.72	80.64	37.89	63.20					
A-Basis Value	98.61	77.85	73.01	13.00	34.58					

Figure 19. Results Summary Sheet Showing the Statistics and Basis Values Based on SP and Pooling Analysis

COMMENTS
(1) Test data has been provided at 5 test condition(s)
(2) 1 outliers detected at the batch level. See INPUT SHEET for outliers
(3) 1 outliers detected at the pooled data level. See INPUT SHEET for outliers
(4) Disposition outliers and rerun analysis
(5) Batch equivalence not satisfied at ETW2 180F; conditions
(6) Batches are not poolable at 1 test conditions. Use engineering judgement for pooling method or use Single Point method
(7) Normality of test data at ETW1 150F test condition is questionable
(8) Pooled Conditions are : CTA 65F; RTA 75 F; ETA1 150F;
(9) -----------Levene's test based on equality of STANDARD DEVIATIONS-----------
(10) Pooled datasets pass Levene's test !!
(11) Pooled datasets form acceptable gouping. Pooling is VALID
(12) Single Point Analysis : Normal Distribution is indicated for CTA 65F Note : Normal Distbn Overriding factor selected = 10
(13) Single Point Analysis : Normal Distribution is indicated for RTA 75 F Note : Normal Distbn Overriding factor selected = 10
(14) Single Point Analysis : Normal Distribution is indicated for ETA1 150F Note : Normal Distbn Overriding factor selected = 10
(15) Single Point Analysis : Non-Paramteric method is indicated for ETW1 150F
(16) **Single Point Analysis : ANOVA Indicated for ETW2 180F. Exercise caution as number of batches is less than 5
(17) ------- C.V. Modified Data ---------
(18) Pooled Conditions are : CTA 65F; RTA 75 F; ETA1 150F;
(19) -----------Levene's test based on equality of STANDARD DEVIATIONS-----------
(20) Modified Pooled datasets pass Levene's test !!
(21) Pooled datasets form acceptable gouping. Pooling is VALID

Figure 20. Comments and Warning Messages Generated by the Program

SUMMARY OF DIAGNOSTIC TESTS										
Test Condition	1	2	3	4	5	6	7	8	9	10
Condition I.D.	CTA 65F	RTA 75 F	ETA1 150F	ETW1 150F	ETW2 180F					
Maximum Normed Residual Test for Outliers							*Significance Level α =		0.05	
Batch Level	0	0	0	1	0					
Pooled Data	0	0	0	1	0					
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						*Significance Level α =		0.025	
ADK	1.427	0.452	0.732	0.793	3.024					
ADC (α = 0.05)	1.924	1.935	1.930	1.940	1.930					
ADC (α = 0.025)	2.225	2.240	2.233	2.246	2.233					
ADC (α = 0.01)	2.624	2.644	2.634	2.652	2.634					
Same Population ?*	YES	YES	YES	YES	NO					
Modified CV Data - for pooling method										
ADK	1.067	0.452	0.607	0.793	2.854					
Same Population ?(α=0.025)	YES	YES	YES	YES	NO					

Summary of Diagnostic Statistics for Pooling Method

Anderson-Darling Test for Normality										
O.S.L. (original data)	0.184	0.395	0.646	0.006	0.429					
Normality acceptable ?	Yes	Yes	Yes	No	Yes					
O.S.L. (Modified data)	0.815	0.386	0.593	0.006	0.394					
Normality acceptable ?	YES	YES	YES	NO	YES					
Check for Normality Based on Graphical Method										
Pearson Coefficient r	0.965	0.981	0.992	0.926	0.982					
Normality acceptable ?	Yes	Yes	Yes	No	Yes					

Anderson-Darling Test for Normality of Pooled Data	
O.S.L. for Original Data	0.3047
O.S.L. for Modified Data	0.6347

Levene's Test for Equality of St.Devs.(FCALCULATED < FCRITICAL for equivalence)				
α =	0.100	0.050	0.025	0.010
FCRITICAL	2.398	3.159	3.938	4.998
FCALCULATED	0.057			
FCALCULATED - Mod. CV	0.276			

Check for Acceptable Grouping (Original Data) X ~ Data available ~ Acceptable for pooling

Environment	CT	RT	ET1	ET2	ET3	ET4	ET5	ET6	ET7	ET8
Ambient	X	X	X							
Wet			X	X						
Check for Acceptable Grouping (Modified Data)										
Ambient	X	X	X							
Wet			X	X						

USER COMMENTS:

Figure 21. Summary of Statistical Diagnostic Tests for Pooling Analysis

SUMMARY OF DIAGNOSTIC TESTS

Summary of Diagnostic Statistics for Single-Point Method

Distribution	Statistic	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Normal	Observed Significance Level (OSL)	0.184	0.395	0.646	0.006	0.429					
Log Normal		0.273	0.457	0.625	0.000	0.527					
Weibull		0.016	0.090	0.540	0.022	0.101					
Normal	Mean	119.42	99.15	91.35	96.92	103.30					
	Stdev	6.25	6.52	5.56	18.80	8.11					
	%C.V.	5.23	6.58	6.09	19.40	7.85					
Log Normal	Log Mean	4.78	4.59	4.51	4.55	4.63					
	Log Stdev	0.051	0.066	0.061	0.235	0.078					
Weibull	Scale parameter	122.49	102.16	93.91	103.84	107.05					
	Shape Parameter	18.15	16.41	18.20	7.29	13.14					
NON-PARAMETRIC STATISTICS	H-K ~ Hanson-Koopmans										
	B-Basis Method	H-K	H-K	H-K	H-K	H-K					
	A-Basis Method	H-K	H-K	H-K	H-K	H-K					
	B-Basis Rank	9	10	10	10	10					
	A-Basis Rank	n/a	n/a	n/a	n/a	n/a					
	B-Basis Hans-Koop k Factor	1.311	1.218	1.253	1.184	1.253					
	A-Basis Hans-Koop k Factor	2.428	2.311	2.367	2.260	2.367					
LEVENE'S EQUALITY OF VARIANCES TEST								*Significance Level α =		0.05	
	Fcalculated	3.869	0.535	0.727	1.505	0.124					
	Fcritical	3.634	3.555	3.592	3.522	3.592					
	Variances Equal ?	no	yes	yes	yes	yes					
ANALYSIS OF VARIANCE (ANOVA) STATISTICS											
	Sample Between-batch Mean Sq. (MSB)	105.22	7.83	9.15	257.41	304.38					
	Error Mean Square (MSE)	30.74	46.43	33.53	363.75	37.70					
	Estimate of Pop. Std. Deviation (S)	6.52	6.40	5.46	18.69	8.82					
	B-Basis Tolerance Limit Factor (TB)	3.54	1.90	1.93	1.89	4.55					
	A-Basis Tolerance Limit Factor (TA)	6.07	3.26	3.30	3.23	7.79					
	B-Basis Value	96.32	86.96	80.83	61.68	63.20					
	A-Basis Value	79.86	78.27	73.34	36.51	34.58					
SUMMARY OF BASIS VALUES											
B-Basis	NORMAL	107.25	86.72	80.64	61.46	87.68					
	LOGNORMAL	107.92	87.26	81.07	60.84	88.63					
	WEIBULL	100.88	82.89	77.59	65.20	82.19					
	NON-PARAMETRIC	108.59	82.31	78.62	37.89	83.83					
	ANALYSIS OF VARIANCE	96.32	86.96	80.83	61.68	63.20					
A-Basis	NORMAL	98.61	77.85	73.01	36.13	76.57					
	LOGNORMAL	100.52	79.79	74.56	44.34	79.64					
	WEIBULL	83.62	67.68	64.49	41.49	63.62					
	NON-PARAMETRIC	84.04	59.59	59.42	13.00	55.91					
	ANALYSIS OF VARIANCE	79.86	78.27	73.34	36.51	34.58					

Figure 22. Summary of Statistical Diagnostic Tests for SP Analysis

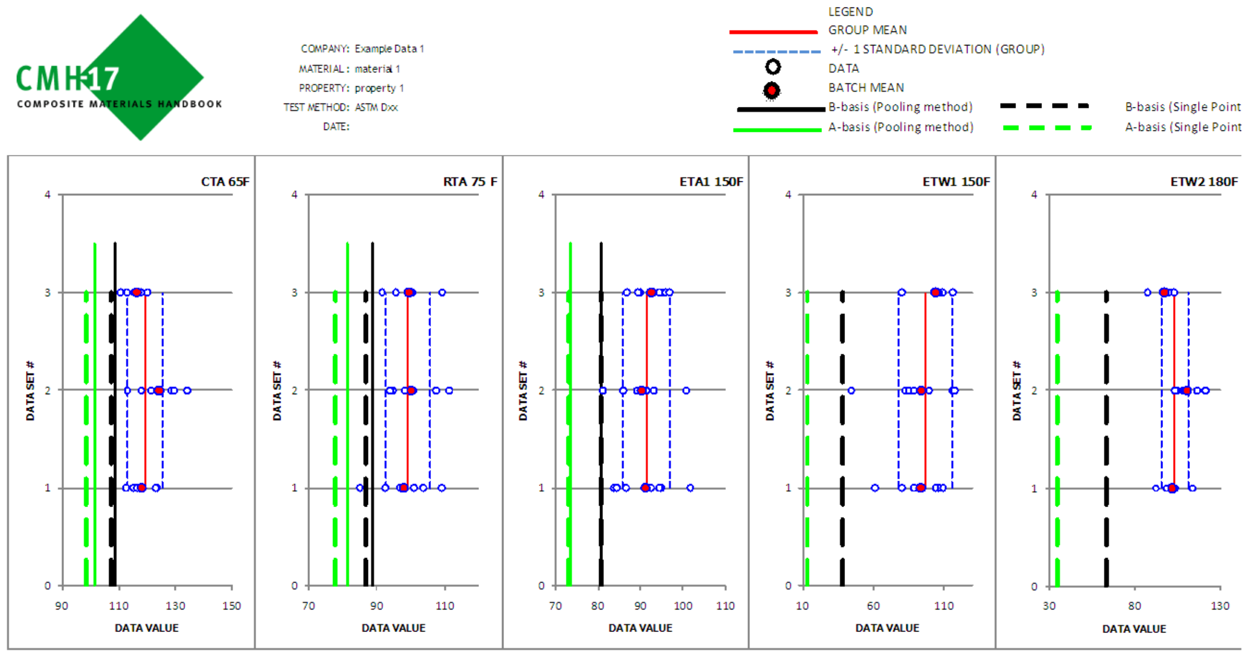


Figure 23. Batch Plots for the Test Data and Corresponding Average and Basis Value Lines

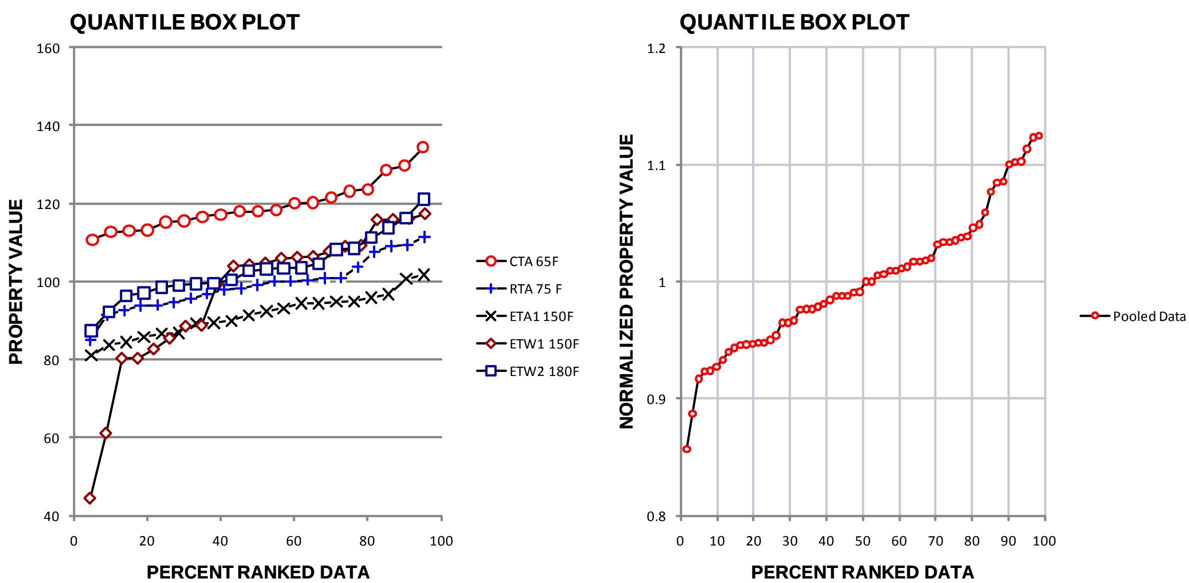


Figure 24. Quantile Box Plots for Raw Data and Normalized Pooled Data in Example Problem I

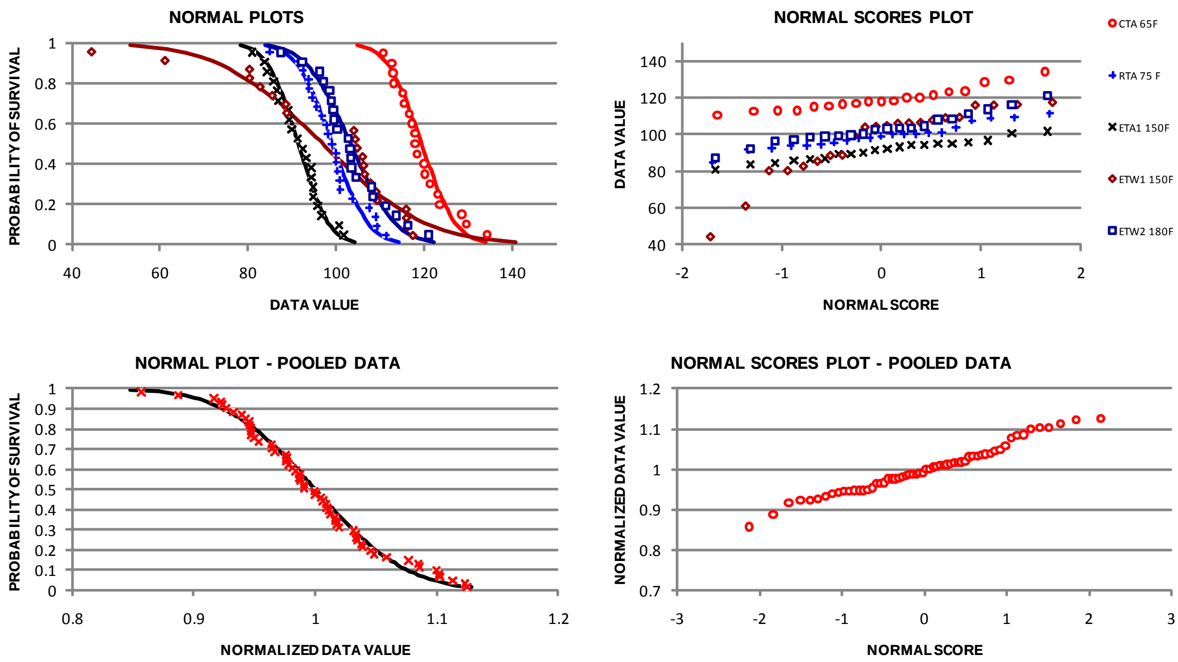


Figure 25. Normal Plots and Normal Scores Plots for Data in Example Problem I

5.2 EXAMPLE PROBLEM II.

In this example, the three data sets contained in the worksheets of the workbook EXAMPLE DATA SETS.xlsx are subjected to batch processing. The batch processing option is chosen as "yes" and the corresponding batch processing file is selected using the SELECT BATCH PROCESSING FILE button. In addition, unlike the previous example, the statistical tests for batch equivalence, normality, and Levene's test are bypassed using the option buttons (set them to "yes"). The batch processing of the three data sets can be started by clicking the COMPUTE BASIS VALUES button. The program analyzes the individual data sets successively. The error messages and user interaction is disabled during batch processing. However, the error messages are written into the results summary files created for individual data sets.

6. SUMMARY.

A Visual Basic for Applications with the Microsoft® Excel® user interface has been assembled to implement the CMH-17 statistical analyses for computing A- and B-basis allowables for material properties. The data input and selection of analysis options consistent with the CMH-17 guidelines are accomplished using the Microsoft Excel worksheets. The program accepts data through an input sheet by manually entering the data or reading it from files based on a template. The program is capable of analyzing up to 1000 data points each at ten individual test conditions. There are no limitations on the number of batches as long as the batch sizes are consistent with the CMH-17 guidelines. The statistical analysis is conducted using the Visual Basic application macros in the background. The secure macro program analyzes the data and interacts with the user (except during batch processing) in the presence of outliers, errors in data, etc. The program conducts both Single Point and Pooling analyses on the given data sets and reports the basis values. The program may be run in a compliant mode with default options with the CMH-17 guidelines strictly enforced, or part of the statistical test may be overridden using the options provided in the program. With each run of the analysis, the program creates an Excel workbook report file consisting of the input data, results and summary sheets, and plot sheets, but without the VBA macros attached. The program has been verified and validated against STAT17 and ASAP programs using the example data sets reported in this document.

7. REFERENCES.

1. SAE International, The Composites Materials Handbook, CMH-17 (Rev G), Vol. 1, 2012.

2. 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, Pennsylvania, 1989, pp. 111-135.

3. Tomblin, John S., Ng, Y. C., and Raju, K. S., "Material Qualification and Equivalency for Polymer Matrix Composite Material Systems," DOT/FAA/AR-00/47, April 2000.

4. Johnson, R.A. Miller, I., and Freund, J., *Probability and Statistics for Engineers*, 5th Ed., Prentice Hall, Englewood Cliffs, New Jersey, 1993.

APPENDIX A— SAMPLES OF WORKSHEETS

Figures A-1 through A-7 show the worksheets in the Microsoft® Excel® workbook report file.

COMPANY: Example Data 1
MATERIAL: material 1
PROPERTY: property 1
TEST METHOD: ASTM Dxx
DATE: May 25, 2011

CMH-17 COMPOSITE MATERIALS HANDBOOK

DATA SUMMARY

STATISTIC	ENVIRONMENTAL TEST CONDITION				
	CTA 65F	RTA 75 F	ETA1 150F	ETW1 150F	ETW2 180F
Sample Size	19	21	20	22	20
No. of Batches	3	3	3	3	3
Mean	119.42	99.15	91.35	96.92	103.30
Std.Dev	6.25	6.52	5.56	18.80	8.11
% Co.V	5.23	6.58	6.09	19.40	7.85
Minimum	110.73	84.96	81.04	44.32	87.34
Maximum	134.32	111.35	101.70	117.32	121.05

SUMMARY OF BASIS VALUES FOR DATA POOLED ACROSS ENVIRONMENTAL TEST CONDITIONS

	CTA 65F	RTA 75 F	ETA1 150F	ETW1 150F	ETW2 180F
Pooled % Co.V	10.72				
K_b	1.7126	1.6954	1.7036	1.6877	1.7036
K_a	2.8350	2.8201	2.8273	2.8135	2.8273
Pooled St.Dev. Basis Values (Original data)					
B-Basis Value	101.34	81.24	73.36	79.10	85.31
A-Basis Value	89.48	69.36	61.49	67.21	73.44
Pooled St.Dev. Basis Values (Modified C.V data)					
Modified %C.V	6.62	7.29	7.05	19.40	7.93
Modified Pooled % Co.V	11.07				
B-Basis Value	100.65	80.56	72.68	78.42	84.63
A-Basis Value	88.35	68.23	60.36	66.08	72.31

SUMMARY OF BASIS VALUES USING SINGLE POINT APPROACH (Original Data)

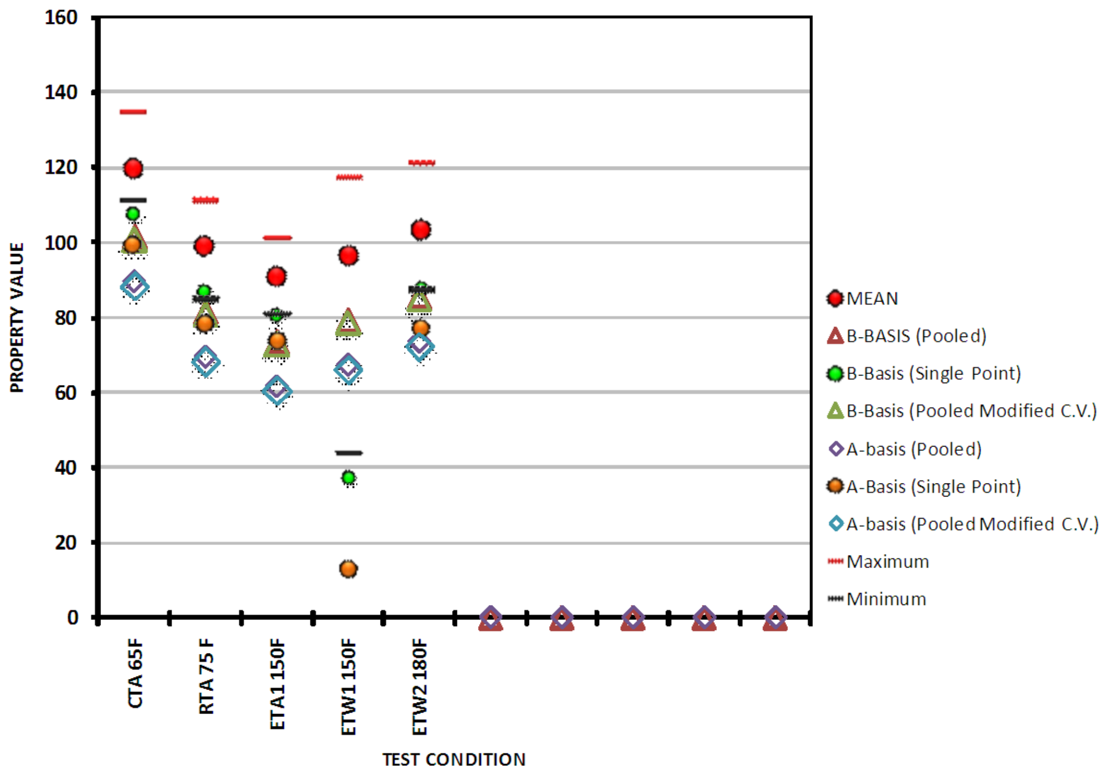
	CTA 65F	RTA 75 F	ETA1 150F	ETW1 150F	ETW2 180F
Distribution	Normal	Normal	Normal	Non-Parm.	Normal
B-Basis Value	107.25	86.72	80.64	37.89	87.68
A-Basis Value	98.61	77.85	73.01	13.00	76.57

COMMENTS
(1) Test data has been provided at 5 test condition(s)
(2) 1 outliers detected at the batch level. See INPUT SHEET for outliers
(3) 1 outliers detected at the pooled data level. See INPUT SHEET for outliers
(4) Disposition outliers and rerun analysis
(5) Batch equivalence not satisfied at ETW2 180F; conditions
(6) Batches are not poolable at 1 test conditions. Use engineering judgement for pooling method or use Single Point method
(7) Normality of test data at ETW1 150F test condition is questionable
(8) Normality of Pooled datasets is questionable at significance level of 0.05 Use engineering judgement or Single Point Method
(9) Pooled Conditions are : CTA 65F; RTA 75 F; ETA1 150F; ETW1 150F; ETW2 180F;
(10) -----------Levene's test based on equality of STANDARD DEVIATIONS----------
(11) Pooled datasets fail Levene's test!!, pooling is questionable. Use Engineering Judgement or Single Point Method
(12) Pooled datasets form acceptable gouping. Pooling is VALID
(13) Single Point Analysis : Normal Distribution is indicated for CTA 65F Note : Normal Distbn Overriding factor selected = 10
(14) Single Point Analysis : Normal Distribution is indicated for RTA 75 F Note : Normal Distbn Overriding factor selected = 10
(15) Single Point Analysis : Normal Distribution is indicated for ETA1 150F Note : Normal Distbn Overriding factor selected = 10

Figure A-1. Left Side of the Results Summary Worksheet

COMPANY: Example Data 1
MATERIAL: material 1
PROPERTY: property 1
TEST METHOD: ASTM Dxx
DATE:

CMH-17 COMPOSITE MATERIALS HANDBOOK



(16) Single Point Analysis : Non-Paramteric method is indicated for ETW1 150F

(17) Single Point Analysis : Normal Distribution is indicated for ETW2 180F Note : Normal Distbn Overriding factor selected = 10

(18) ------- C.V. Modified Data ---------

(19) Normality of Modified Pooled datasets is questionable at significance level of 0.05 Use engineering judgement or Single Point Metho

(20) Pooled Conditions are : CTA 65F; RTA 75 F; ETA1 150F; ETW1 150F; ETW2 180F;

(21) -----------Levene's test based on equality of STANDARD DEVIATIONS----------

(22) Modified Pooled datasets fail Levene's test!!, pooling is questionable. Use Engineering Judgement or Single Point Method

(23) Pooled datasets form acceptable gouping. Pooling is VALID

Figure A-2. Right Side of the Results Summary Worksheet

SUMMARY OF DIAGNOSTIC TESTS

Test Condition	1	2	3	4	5	6	7	8	9	10
Condition I.D.	CTA 65F	RTA 75 F	ETA1 150F	ETW1 150F	ETW2 180F					

Maximum Normed Residual Test for Outliers — *Significance Level α = 0.05

Batch Level	0	0	0	1	0					
Pooled Data	0	0	0	1	0					

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 — *Significance Level α = 0.025

ADK	1.427	0.452	0.732	0.793	3.024					
ADC (α = 0.05)	1.924	1.935	1.930	1.940	1.930					
ADC (α = 0.025)	2.225	2.240	2.233	2.246	2.233					
ADC (α = 0.01)	2.624	2.644	2.634	2.652	2.634					
Same Population ?*	YES	YES	YES	YES	NO					
Modified CV Data - for pooling method										
ADK	1.067	0.452	0.607	0.793	2.854					
Same Population ?(α=0.025)	YES	YES	YES	YES	NO					

Summary of Diagnostic Statistics for Pooling Method

Anderson-Darling Test for Normality

O.S.L. (original data)	0.184	0.395	0.646	0.006	0.429					
Normality acceptable ?	Yes	Yes	Yes	No	Yes					
O.S.L. (Modified data)	0.815	0.386	0.593	0.006	0.394					
Normality acceptable ?	YES	YES	YES	NO	YES					

Check for Normality Based on Graphical Method

Pearson Coefficient r	0.965	0.981	0.992	0.926	0.982					
Normality acceptable ?	Yes	Yes	Yes	No	Yes					

Anderson-Darling Test for Normality of Pooled Data

O.S.L. for Original Data	0.0005
O.S.L. for Modified Data	0.0101

Levene's Test for Equality of St.Devs.(FCALCULATED < FCRITICAL for equivalence)

α =	0.100	0.050	0.025	0.010
FCRITICAL	2.004	2.465	2.921	3.519
FCALCULATED	4.979			
FCALCULATED - Mod. CV	3.938			

Check for Acceptable Grouping (Original Data) — X ~ Data available; ~ Acceptable for pooling

Environment	CT	RT	ET1	ET2	ET3	ET4	ET5	ET6	ET7	ET8
Ambient	X	X	X							
Wet			X	X						

Check for Acceptable Grouping (Modified Data)

Environment	CT	RT	ET1	ET2	ET3	ET4	ET5	ET6	ET7	ET8
Ambient	X	X	X							
Wet			X	X						

USER COMMENTS:

Figure A-3. Left Side of the Statistics Summary Worksheet

SUMMARY OF DIAGNOSTIC TESTS

Summary of Diagnostic Statistics for Single-Point Method

Distribution	Statistic	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Normal	Observed Significance Level (OSL)	0.184	0.395	0.646	0.006	0.429					
Log Normal		0.273	0.457	0.625	0.000	0.527					
Weibull		0.016	0.090	0.540	0.022	0.101					
Normal	Mean	119.42	99.15	91.35	96.92	103.30					
	Stdev	6.25	6.52	5.56	18.80	8.11					
	%C.V.	5.23	6.58	6.09	19.40	7.85					
Log Normal	Log Mean	4.78	4.59	4.51	4.55	4.63					
	Log Stdev	0.051	0.066	0.061	0.235	0.078					
Weibull	Scale parameter	122.49	102.16	93.91	103.84	107.05					
	Shape Parameter	18.15	16.41	18.20	7.29	13.14					

NON-PARAMETRIC STATISTICS H-K ~ Hanson-Koopmans

B-Basis Method	H-K	H-K	H-K	H-K	H-K					
A-Basis Method	H-K	H-K	H-K	H-K	H-K					
B-Basis Rank	9	10	10	10	10					
A-Basis Rank	n/a	n/a	n/a	n/a	n/a					
B-Basis Hans-Koop k Factor	1.311	1.218	1.253	1.184	1.253					
A-Basis Hans-Koop k Factor	2.428	2.311	2.367	2.260	2.367					

LEVENE'S EQUALITY OF VARIANCES TEST *Significance Level α = 0.05

Fcalculated	3.869	0.535	0.727	1.505	0.124					
Fcritical	3.634	3.555	3.592	3.522	3.592					
Variances Equal ?	no	yes	yes	yes	yes					

ANALYSIS OF VARIANCE (ANOVA) STATISTICS

Sample Between-batch Mean Sq. (MSB)	105.22	7.83	9.15	257.41	304.38					
Error Mean Square (MSE)	30.74	46.43	33.53	363.75	37.70					
Estimate of Pop. Std. Deviation (S)	6.52	6.40	5.46	18.69	8.82					
B-Basis Tolerance Limit Factor (TB)	3.54	1.90	1.93	1.89	4.55					
A-Basis Tolerance Limit Factor (TA)	6.07	3.26	3.30	3.23	7.79					
B-Basis Value	96.32	86.96	80.83	61.68	63.20					
A-Basis Value	79.86	78.27	73.34	36.51	34.58					

SUMMARY OF BASIS VALUES

B-Basis	NORMAL	107.25	86.72	80.64	61.46	87.68					
	LOGNORMAL	107.92	87.26	81.07	60.84	88.63					
	WEIBULL	100.88	82.89	77.59	65.20	82.19					
	NON-PARAMETRIC	108.59	82.31	78.62	37.89	83.83					
	ANALYSIS OF VARIANCE	96.32	86.96	80.83	61.68	63.20					
A-Basis	NORMAL	98.61	77.85	73.01	36.13	76.57					
	LOGNORMAL	100.52	79.79	74.56	44.34	79.64					
	WEIBULL	83.62	67.68	64.49	41.49	63.62					
	NON-PARAMETRIC	84.04	59.59	59.42	13.00	55.91					
	ANALYSIS OF VARIANCE	79.86	78.27	73.34	36.51	34.58					

Figure A-4. Right Side of the Statistics Summary Worksheet

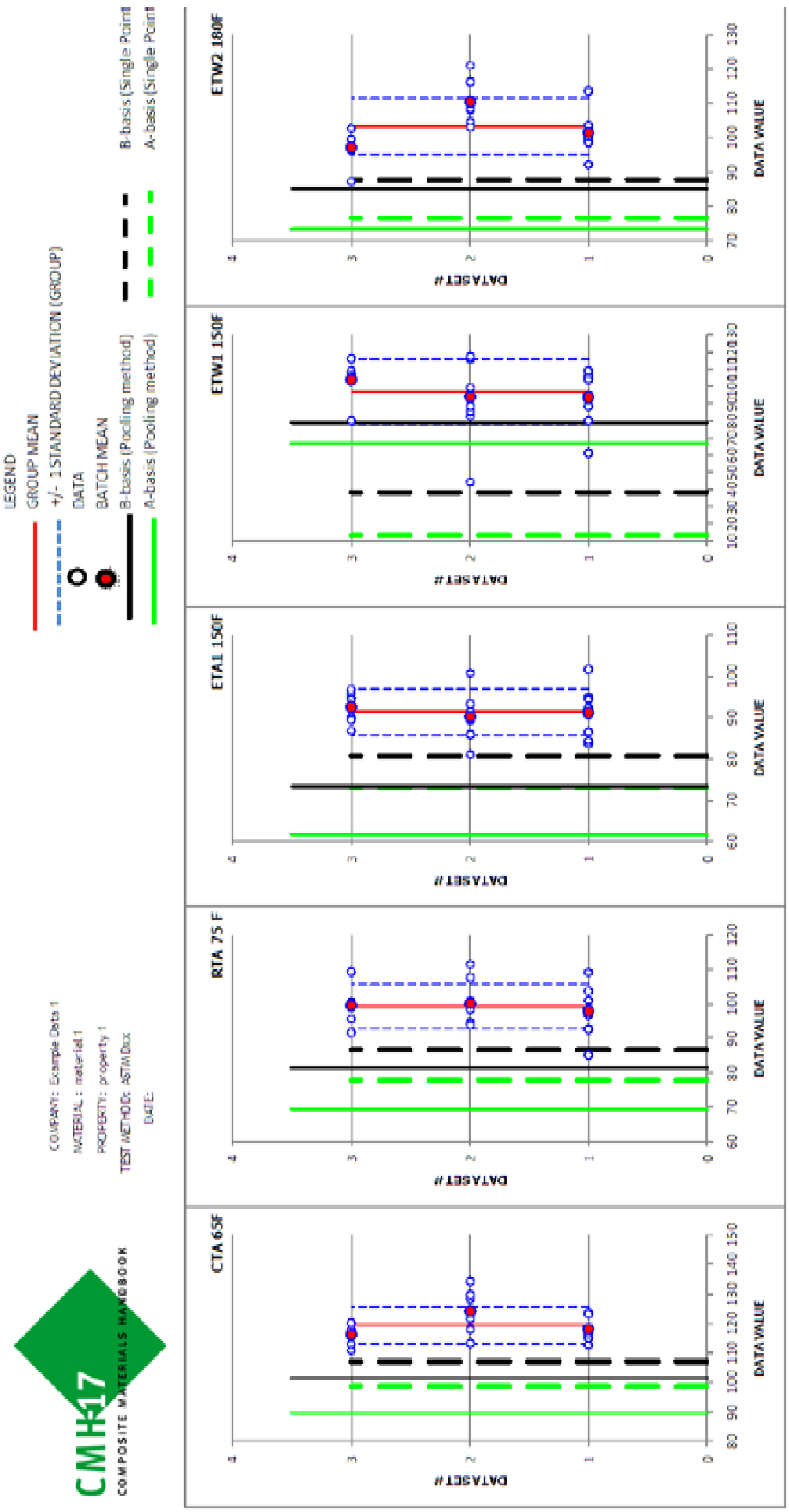


Figure A-5. Batch Plots Worksheet

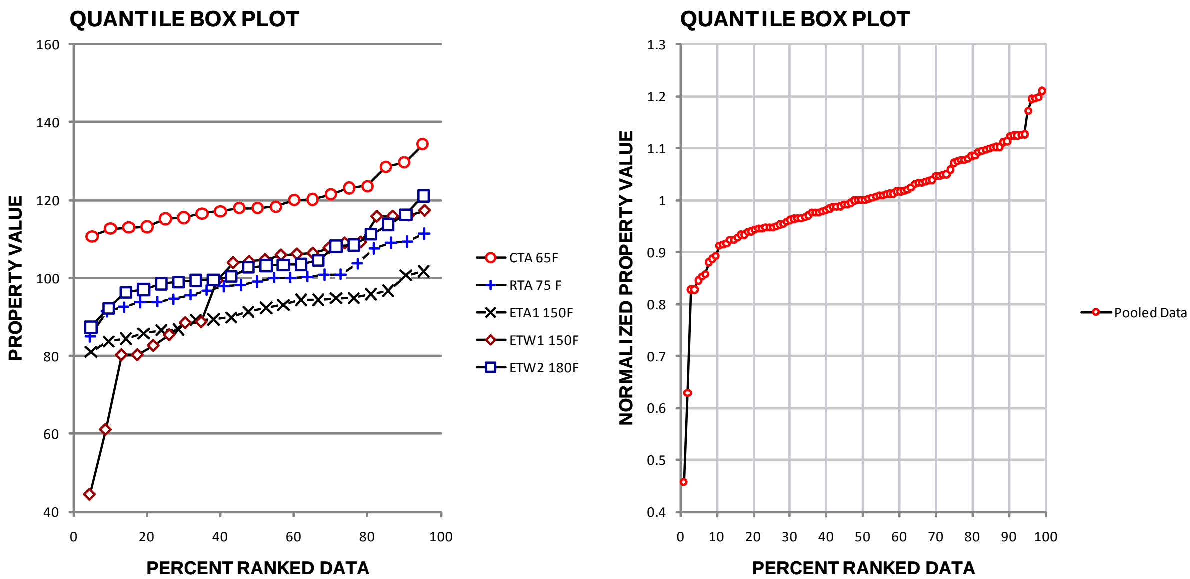


Figure A-6. Quantile Box Plots

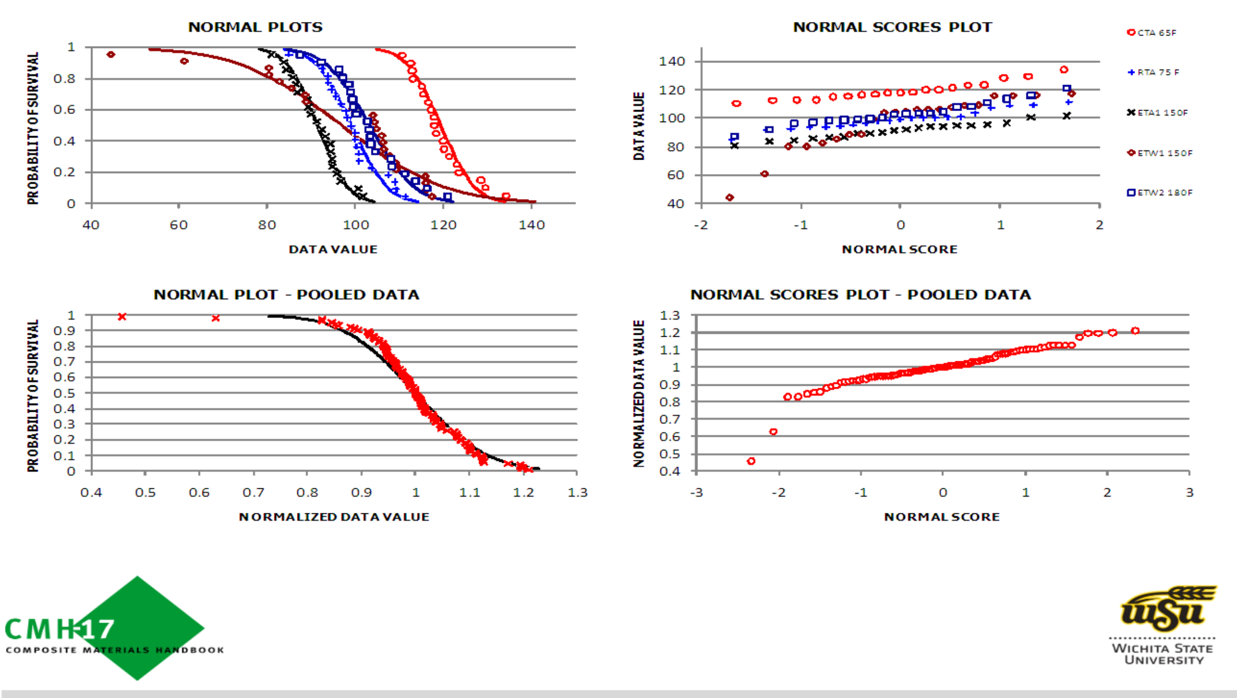


Figure A-7. Normal Plots Worksheets

APPENDIX B—EXAMPLE DATA SETS

Table B-1. Data Set 1 Used in Example Problem I

CTA	65F		RTA	75 F		ETA1	150F		ETW1	150F		ETW2	180F	
Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point
1	1	118.37	1	1	84.96	1	1	83.74	1		106.36	1		99.02
1	2	123.6	1	2	92.49	1	2	84.38	1		105.89	1		103.34
1	3	115.22	1	3	96.82	1	3	94.8	1		88.46	1		100.3
1	4	112.63	1	4	109.03	1	4	94.39	1		103.9	1		98.46
1	5	116.56	1	5	97.89	1	5	101.7	1		80.21	1		92.26
1	6	123.16	1	6	100.92	1	6	86.54	1		109.2	1		103.49
2	1	128.59	1	7	103.69	1	7	92.38	1		61.01	1		113.73
2	2	113.14	2	1	93.79	2	1	89.21	2		99.32	2		108.17
2	3	121.42	2	2	107.53	2	2	100.69	2		115.86	2		108.42
3	1	115.45	2	3	94.57	2	3	81.04	2		82.61	2		116.26
3	2	120.03	2	4	93.88	2	4	91.34	2		85.37	2		121.05
3	3	117.16	2	5	98.23	2	5	93.14	2		115.8	2		111.22
3	4	112.93	2	6	111.35	2	6	85.82	2		44.32	2		104.57
3	5	117.91	2	7	100.82	3	1	94.89	2		117.32	2		103.22
3	6	120.19	3	1	100.38	3	2	95.81	2		88.67	3		99.39
3	7	110.73	3	2	91.5	3	3	86.78	3		107.68	3		87.34
2	4	134.32	3	3	100.08	3	4	94.4	3		108.96	3		102.73
2	5	129.64	3	4	95.63	3	5	96.72	3		116.12	3		96.37
2	6	117.98	3	5	109.3	3	6	89.9	3		80.23	3		99.59
			3	6	99.12	3	7	89.37	3		106.15	3		97.07
			3	7	100.07				3		104.67			
									3		104.23			

Table B-2. Data Set 3 Used in Example Problem II

ETA5	250F		ETA4	230F		ETA3	220F		ETW2	200F		ETA2	200F	
Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point
1		69.91263	1		78.12067	1		97.1369	1		92.78339	1		96.77356
1		81.96462	1		82.55485	1		87.48967	1		88.37217	1		99.31894
1		70.68759	1		80.79616	1		91.61715	1		94.17975	1		99.07103
1		68.50212	1		74.80195	1		89.11911	1		80.19713	1		99.39722
1		74.79227	1		74.62671	1		82.16539	1		94.79738	1		94.54631
1		79.70756	1		79.88364	1		80.56426	1		97.31397	1		95.69947
2		81.37306	1		73.84307	1		81.58813	1		92.23522	1		105.1792
2		72.04628	2		82.07122	2		84.03589	1		93.58871	2		99.33323
2		78.69356	2		82.78033	2		84.67326	2		87.86631	2		101.9243
2		70.12193	2		89.0001	2		73.61682	2		90.76745	2		94.6774
2		79.80434	2		89.20018	2		82.68241	2		77.12314	2		98.07296
2		81.25645	2		77.68341	2		95.45282	2		79.06915	2		106.5309
3		70.61777	2		73.72375	2		84.18347	2		92.93094	2		100.9778
3		83.94673	2		75.33043	2		84.20755	2		86.09726	2		89.72538
3		68.5897	3		78.53682	2		82.44415	2		89.94332	3		97.14054
3		76.08194	3		84.74324	3		87.54882	2		95.40902	3		90.63035
3		77.13565	3		89.36263	3		79.75705	2		91.46176	3		100.4394
3		71.22007	3		83.66531	3		77.0137	3		105.2241	3		102.0987
			3		82.03035	3		83.08062	3		85.1773	3		109.2772
			3		80.26102	3		89.98098	3		100.2122	3		111.7935
			3		78.08158	3		85.67083	3		84.02894	3		99.12726
						3		81.41129	3		96.73305			
						3		86.70081	3		87.86386			
									3		94.44057			
									3		83.94017			
									3		96.18809			

Table B-2. Data Set 3 Used in Example Problem II (continued)

ETW2	180F		ETW1	150F		ETA1	150 F		RTA	75F		CTA	-65F	
Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point	Batch ID	Specimen ID	Data Point
1		98.93112	1		106.4655	1		124.4215	1		105.5764	1		122.382
1		100.5546	1		104.7068	1		119.1794	1		114.337	1		124.7999
1		93.6334	1		105.392	1		111.4722	1		115.6431	1		110.2058
1		98.72667	1		104.812	1		116.2045	1		103.991	1		127.4361
1		94.16671	1		100.6858	1		113.8614	1		104.1946	1		112.4427
1		94.55921	1		110.2591	1		111.334	1		123.711	1		127.2213
2		110.5548	1		116.9146	1		111.3942	2		107.0624	1		128.8502
2		104.3434	2		101.4666	1		102.8418	2		126.6586	1		115.253
2		100.78	2		96.33357	2		117.5634	2		99.00781	2		126.2763
2		95.55943	2		109.2691	2		102.6082	2		125.4528	2		123.7442
2		102.6547	2		103.9903	2		109.6383	2		111.2525	2		118.4421
2		102.5664	2		110.7485	2		116.8967	2		113.5103	2		127.7775
3		115.4784	3		95.75347	2		103.3405	3		116.4868	2		113.3414
3		104.2226	3		106.7166	2		111.9566	3		113.1856	2		135.5563
3		101.9968	3		99.63288	2		118.6275	3		121.8975	2		126.721
3		104.2745	3		103.876	2		108.7956	3		114.8993	2		104.6341
3		108.634	3		103.7815	3		102.1615	3		124.7786	3		118.9097
3		107.2194	3		92.77661	3		102.8791	3		112.2747	3		126.3145
3		104.774	3		109.8475	3		115.1475	4		114.2511	3		118.8232
4		105.8724	4		101.7854	3		106.8156	4		112.9998	3		127.5909
4		101.824	4		100.6064	3		105.3757	4		104.842	3		116.4338
4		98.53446	4		111.9736	3		110.8045	4		111.1233	3		114.7959
4		97.97111	4		107.2377	3		115.677	4		119.475	3		118.2892
4		106.8033	4		105.5569	4		120.3053	4		117.0532	3		113.7031
4		102.0766	4		111.2557	4		117.4213	5		119.0843	4		125.05
4		101.8957	5		108.3833	4		101.4644	5		106.1484	4		122.5022
5		116.6167	5		112.4104	4		104.528	5		110.9653	4		116.4338
5		105.5596	5		92.74354	4		109.992	5		112.6651	4		117.6799
5		93.02456	5		101.4151	4		107.367	5		115.0392	4		117.4215
5		100.8295	5		99.86447	4		120.3456	5		107.2034	4		118.5496
5		104.6101	5		95.29519	5		114.5475				4		123.3763
5		99.96993	5		97.42337	5		110.6279				4		129.868
5		105.1448				5		103.0632				5		111.2706
						5		111.7697				5		123.8456
						5		106.6505				5		120.4661
						5		113.6863				5		121.4476
												5		113.715
												5		120.9595
												5		122.6993
												5		109.5929