	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 1 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

CONTENTS

1 PURPOSE OF PROCEDURE

2 UNCERTAINTY ASSESSMENT METHODOLOGY

2.1 Overview

2.2 Estimating Uncertainty Components in Measured Variables

2.2.1 Definitions

2.2.2 Estimating Precision Limits

2.2.3 Estimating Bias Limits

2.3 Estimating Uncertainty Components for Experimental Results

2.3.1 Definitions

2.3.2 Propagation of Precision Limits into an Experimental Result

2.3.2.1 Multiple Tests

2.3.2.2 Single Test with Single Readings

2.3.2.3 Single Test with Averaged Readings

2.3.3 Propagation of Bias Limits into an Experimental Result

2.4 Summary of the Methodology


2.5 Reporting Uncertainties

ANNEX 2-A: A Comprehensive Uncertainty Analysis Methodology

ANNEX 2-B: Identification and Elimination of Outliers in Samples

3. REFERENCES

Prepared	Approved
Resistance Committee of 22 nd ITTC	22 nd ITTC 1999
Date	Date

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 2 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

Uncertainty Analysis in EFD, Uncertainty Assessment Methodology

1 PURPOSE OF PROCEDURE

To provide a methodology for estimating the uncertainty in an experimental result at a 95% confidence level.

2 UNCERTAINTY ASSESSMENT METHODOLOGY

The methodology for estimating the uncertainties in measurements, and in the experimental results calculated from them, must be structured to combine statistical and engineering concepts in a manner that can be systematically applied to each step in the data uncertainty assessment determination. In this section, an uncertainty analysis methodology is presented, and its application in the different phases of an experimental program is discussed. The methodology is based on Chapter 2 of the AIAA (1995) standard with minor modifications for terminology and figure, table and equation numbering. The AIAA methodology is based on the material from Coleman & Steele (1989, 1st edition and 1999, 2nd edition) and is consistent with the most current drafts of international guidelines and standards (ISO, 1992 and 1993a). Definitions of specific terms are made as required in the following text, and the international vocabulary of metrological terms (VIM) is incorporated herein (ISO, 1993b).

2.1 Overview

The word **accuracy** is generally used to indicate the closeness of the agreement between an experimentally determined value of a quantity and its true value. **Error** is the difference between the experimentally determined value and the truth. Accuracy is said to increase as error approaches zero. The true values of standard measurement quantities (e.g., mass, length, time, volts, etc.) generally only reside in national standards laboratories. Only in rare instances is the true value of a quantity known. Thus, one is forced to estimate error, and that estimate is called an uncertainty, U . In general, the uncertainty of a quantity is a function of the value of that quantity. However, it is common practice to quote the same value of uncertainty for a range of values of the quantity, e.g., percent of full scale of an instrument. In this document, all estimates are assumed to be made at a 95-percent confidence level, meaning that the true value of the quantity is expected to be within the $\pm U$ interval about the

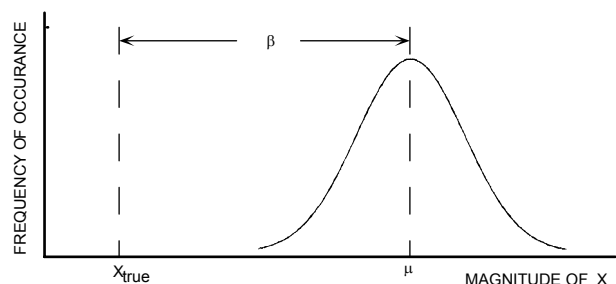

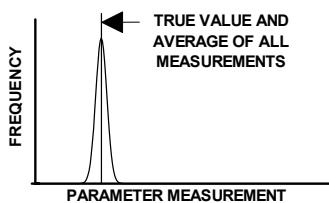


Figure 2.1 Errors in the measurement of a variable X .

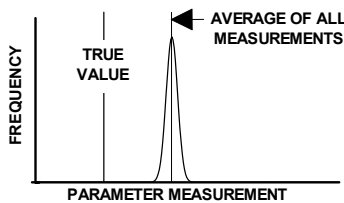
	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 3 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

experimentally determined value 95 times out of 100.

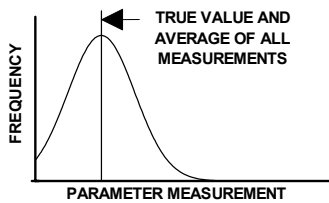
Errors can be considered to be composed of



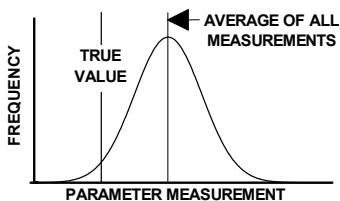
a. Unbiased, precise, accurate.



b. Biased, precise, inaccurate.



c. Unbiased, imprecise, inaccurate.



d. Biased, imprecise, inaccurate.


Figure 2.2 Measurement error (bias, precision, and accuracy).

two components: a **precision** (random) component and a **bias** (systematic) component. An error is classified as precision if it contributes to the scatter of the data; otherwise, it is a bias error. It is assumed that corrections have been made for all systematic errors whose values are known. The remaining bias errors are thus equally as likely to be positive as negative.

The effects of such errors on multiple readings of a variable X are illustrated in Fig. 2.1, where the bias error is denoted by β . The qualitative influence of various combinations of large and small precision and bias errors on accuracy is depicted in Fig. 2.2. For example, an accurate value is one with small bias and precision errors (Fig. 2.2a), whereas one may have small precision errors but inaccurate values (Fig. 2.2b)

Estimates of error are meaningful only when considered in the context of the process leading to the value of the quantity under consideration. In order to identify and quantify error sources, two factors must be considered: (1) the steps used in the processes to obtain the measurement of the quantity, and (2) the environment in which the steps were accomplished. Each factor influences the outcome.

In nearly all experiments, the measured values of different quantities are combined using a data reduction equation to form some desired result. A good example is the experimental determination of total resistance coefficient of a particular ship model configuration in a

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 4 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

in a towing tank test. Defining total resistance coefficient as

$$C_T = \frac{R_x}{0.5\rho U^2 S} \quad (2-1)$$

variables on the right-hand side of Eq. (2-1) will cause errors in the experimental result C_T .

A more general representation of a data reduction equation is

$$r = r(X_1, X_2, \dots, X_J) \quad (2-2)$$

where r is the experimental result determined from J measured variables X_i . If B and P are taken as estimates of the magnitude of bias and precision errors, respectively, the experimental situation is represented schematically in Fig. 2.3.

Each of the measurement systems used to measure the value of an individual variable X_i is influenced by a large number of elemental error sources. The effects of these elemental errors are manifested as a bias error (estimated by B_i) and a precision error (estimated by P_i) in the measured values of the variable. These errors in the measured values then propagate through the data reduction equation, thereby generating the bias and precision errors in the experimental result, r .

In performing an uncertainty analysis, it is convenient to consider the things which could produce errors in a measurement as **elements**. For example, the elements associated with a resistance measurement could be the unsteady test conditions, transducer, transducer environment, signal amplifier, power supply, analog-to-digital converter, and recording device.

one can envision that errors in the values of the

In typical towing tank experimental programs, it is generally not cost effective to try to estimate the precision errors of each elemental error source. It is usually far more effective to estimate the precision of a group of elements (such as the output of the entire measurement system for X_J – the P_J level in Fig. 2.3). This way, the measurement system precision is considered an element contributing to the total uncertainty of resistance measurements. Better yet would be to compute directly the precision of the result (P_r in Fig. 2.3) if multiple results at the same set point are available.

Bias errors, on the other hand, are generally easiest to estimate at a smaller elemental level.

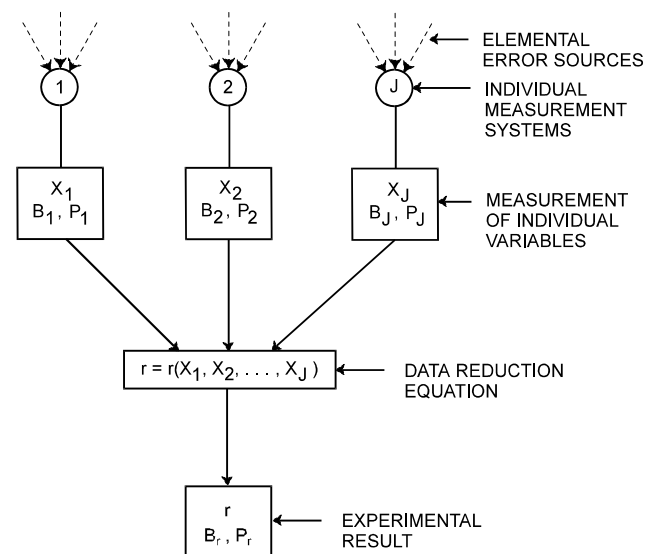



Figure 2.3 Propagation of errors into an experimental result.

	ITTC – Recommended Procedures		7.5 – 02 01 – 01 Page 5 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology		Effective Date 1999	Revision 00

For example, the bias caused by transducer set-up would not be identified in any single set of experimental data and thus must be estimated. However, this should not be taken to imply that bias estimates must be made for each component of the measurement system. Each measurement system should be calibrated in as large a piece as possible (ideally, an end-to-end calibration under operating conditions).

In most situations, such an approach removes the need to estimate the bias errors of individual components of measurement systems. The example in Section 2.4 describes ways to estimate the bias and precision of a measurement system.

In Section 2.2.2, the methodology for obtaining estimates of the precision errors and bias errors in the measured variables X_i is presented, and in Section 2.2.3 the methodology for obtaining estimates of the precision errors and bias errors in the experimental results determined from the X_i is presented. The methodology discussed in the body of this section assumes that error distributions are well-approximated by the Gaussian distribution, that uncertainty estimates are made at a 95-percent confidence level using large sample size techniques¹, and that all precision errors are uncorrelated. In Annex 2.2-A, a more comprehensive (and more complex) methodology (ISO, 1993a) that is valid for either small or large sample

¹ A discussion of what constitutes “large” sample sizes is given in Annex 2.2-A. In most practical towing tank test situations, if the dominant uncertainties are estimated based on 10 or more readings, then use of large sample size methodology is justified. (Of course, it is always desirable to have as many readings as possible so that a better estimate can be made of the true variance of the distribution from which the sample readings are taken.)

sizes and either Gaussian or non-Gaussian error distributions is discussed. In Annex 2.2-B, a method for identification of outliers in samples is presented.


2.2 Estimating Uncertainty Components in Measured Variables

In this section, the methodology for obtaining estimates of the precision errors and bias errors **in the measured variables** X_i is presented. The methodology for obtaining estimates of the precision errors and bias errors **in the experimental results** r , computed using the measured variables in data reduction equations of the form of Eq. (2-2), is discussed in Section 2.2.3.

2.2.1 Definitions

To estimate the magnitude of the precision errors in measurements of a variable X_i , a **precision limit** P_i is defined. As illustrated in Fig. 2.4, the $\pm P_i$ interval about a measurement of X_i is the band within which the (biased) mean value, μ , of the variable would fall 95 times out of 100 if the experiment were repeated many times under the same conditions using the same equipment. The precision limit is thus an estimate of the lack of measurement repeatability caused by random errors, unsteadiness, inability to reset experimental conditions exactly, etc.

To estimate the magnitude of the bias errors in measurements of a variable X_i , a **bias limit** B_i is defined. The bias limit is estimated with the understanding that the experimenter is 95-

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 6 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

percent confident that the true value of the bias error, if known, would be less than $|B_i|$.

The $\pm U_i$ **uncertainty** interval about the measured value of X_i is the band within which the experimenter is 95-percent confident the true value of the variable lies. The 95-percent confidence uncertainty is given by

$$U_i = (B_i^2 + P_i^2)^{\frac{1}{2}} \quad (2-3)$$

2.2.2 Estimating Precision Limits

The precision limit for a measured variable X_i is given by

$$P_i = K S_i \quad (2-4)$$

where K is the coverage factor and equals 2 for a 95-percent confidence level, S_i is the standard deviation of the sample of N_i readings of the variable X_i and is defined as

$$S_i = \left(\frac{\sum_{k=1}^{N_i} [(X_i)_k - \bar{X}_i]^2}{N_i - 1} \right)^{\frac{1}{2}} \quad (2-5)$$

and the mean value is defined as

$$\bar{X}_i = \frac{1}{N_i} \sum_{k=1}^{N_i} (X_i)_k \quad (2-6)$$

An interpretation of the $\pm P_i$ interval is shown in Fig. 2.4.

The use of $K = 2$ assumes a large sample size and Gaussian error distribution. It is instructive to note, however, a 1993 policy statement (Taylor & Kuyatt, 1993) by the U. S. National Institute of Standards and Technology (NIST): “To be consistent with current interna-

tional practice, the value of K to be used at NIST for calculating U_i is, by convention, $K =$

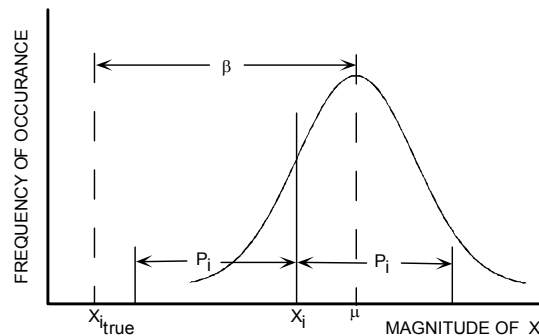


Figure 2.4 95-percent confidence precision limit interval around a single reading of a variable X_i .

2. Values of K other than 2 are only to be used for specific applications dictated by established and documented requirements.” A discussion of estimating the coverage factor K for “small” sample sizes is presented in Annex 2.2-A. Also discussed in that annex is the method for combining precision limits estimated at the elemental error source level (Fig. 2.3).

When a mean (averaged) value of X_i is to be used in Eq. (2-2) to determine the result r , the appropriate precision limit is the precision limit of the mean defined by

$$P_{\bar{X}_i} = \frac{P_i}{\sqrt{N_i}} \quad (2-7)$$

An interpretation of this precision limit is shown in Fig. 2.5.

Two questions that often arise in evaluating a precision limit from a sample of N_i readings are

1. What should be done with those data points (outliers) that are far from the majority of

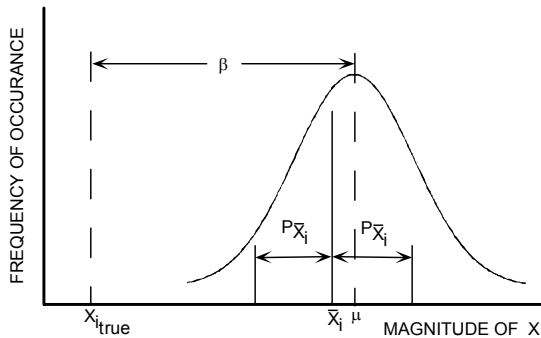


Figure 2.5 95-percent confidence precision limit interval around the mean value of a sample of readings of variable X_i

the points in the sample?

2. How should data variations that occur because of system unsteadiness rather than from random error sources be evaluated?

Procedures for identifying outliers are discussed in Annex 2.2-B. Note that apparent outliers can be due to two basic causes - truly spurious events not connected with the test, or phenomenologically relevant data variations that, though improbable, occur during the first N_i readings taken. Hydrodynamic folklore is full of stories about ships which “discovered” adverse events during sea trials only to find the same event in the model test data which was erroneously considered to be an outliers. (Note that outliers can only be identified in relation to a mean value computed from a number of samples taken at the same test conditions.) To avoid such occurrences, all outliers should be examined for relevance to the phenomena being investigated.

Consideration of the appropriate time interval for collection of the N_i readings is critical if appropriate precision limits are to be estimated. Consider, for example, an experiment in which some of the test variables have a time variation such as that shown in Fig. 2.6. If the question in the experiment is “what is the result for time interval Δt ?”, then M multiple sets of readings of the (X_1, \dots, X_J) taken over that interval can be used in the data reduction equation [Eq. (2-2)] to determine M values of the result r , and the mean result and appropriate precision limit can be computed using the techniques discussed below in Section 2.3.2.1.

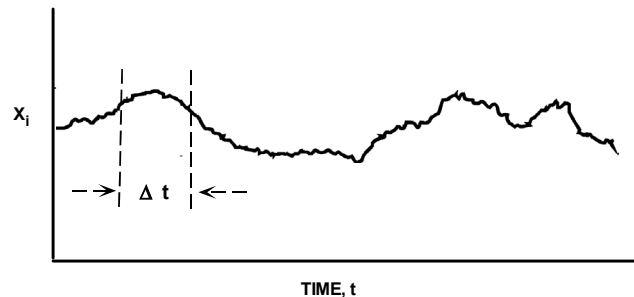



Figure 2.6 Variation of a variable X_i with time for a “steady” experimental condition.

A more typical situation in towing tank testing occurs when the test data are taken at “steady” conditions, but the actual variation of the X_i 's with time is similar to that shown in Fig. 2.6. In this case, one typically desires the result determined using the data reduction equation to be indicative of the value of the result over the interval during which several complete variations in the variables occur. However, it is usually not possible to take measurements over that entire interval, as some of the variations may have periods of hours or

	ITTC – Recommended Procedures		7.5 – 02 01 – 01 Page 8 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology		Effective Date 1999	Revision 00

even days and others may not be cyclic at all, but vary binarily. In most towing tank tests, measurements are taken over a short period with the full understanding that the interval for variation of some of the variables is much longer than the measurement time. In such a case, a value of X_i determined over such a relatively short Δt should be considered as a **single reading** and the appropriate precision limit is estimated by Eq. (2-4), not by Eq. (2-7). Note that this interpretation holds whether the value of X_i is the average of 10 , 10^3 or 10^6 readings taken during Δt .

One may obtain an appropriate estimate of the sample standard deviation [Eq. (2-5)] during the testing process by taking repeat data provided that all of the error sources contributing to the total precision are allowed to vary during the repeat process. For example, taking multiple-samples of data as a function of only time while holding all other conditions constant merely identifies the precision associated with the measurement system and the unsteadiness of the test conditions. The precision associated with other precision error sources, e.g., repeating test conditions, model positions, configuration variables, etc, must also be included to determine the proper precision limit for the variable of interest.


In a given test, the value for the precision limit to be assigned to a single reading would have to be based on **previous information** about that measurement **obtained over the appropriate time interval** (Steele et al., 1993). If such previous information consists of N_i repeated readings for each of the X_i variables, the precision limit for each variable can

be determined from the N_i previous readings using Eqs. (2-4) and (2-5). If previous readings of a variable over an appropriate interval are not available, then the experimenter must estimate a value for P_i using the best information available at that time.

The concept of a precision limit is very useful in all phases of an experimental program. For example, in the design phase of an experiment, a 95-percent confidence estimate of the “scatter” expected for a given measurement based on past experience with the measurement technique may be all that is available. In the debugging phase of the experiment, the comparison of the precision limits estimated in the design phase and the precision limits actually calculated from multiple samples in the debugging phase allows the experimenter to verify that all the factors that influence the precision of the measured variables have been properly taken into account (Coleman & Steele, 1999).

2.2.3 Estimating Bias Limits

A useful approach to estimating the magnitude of a bias error is to assume that it belongs to some assumed statistical distribution. For example, if a thermistor manufacturer specifies that 95-percent of samples of a given model are within $\pm 0.5^\circ\text{C}$ of the reference resistance-temperature (R-T) curve supplied, then one might assume that the systematic errors (the difference between the actual R-T curves of various thermistors and the reference curve) belong to a normal distribution with a standard deviation b_T equal to $(0.5^\circ\text{C})/2$, corresponding to a bias limit estimate $B_T = 2b_T$ (analogous to Eq. (2-4)) or 0.5°C .

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 9 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

More discussion of assumed bias error distributions is given in Annex 2.2-A. In the following, all bias errors are assumed to be normally distributed and the coverage factor $K = 2$, as before.

One might separate the bias errors which influence the measurement of a variable into different categories: calibration errors, data acquisition errors, data reduction errors, test technique errors, etc. Within each category, there may be several elemental sources of bias, as indicated schematically in Fig. 2.3. For instance, if for the J th variable, X_J , there are M elemental bias errors identified as significant and whose bias limits are estimated as

$$(B_J)_1, (B_J)_2, \dots, (B_J)_M$$

then the bias limit for the measurement of X_J is calculated as the root-sum-square (RSS) combination of the elemental limits

$$B_J = \left[\sum_{k=1}^M (B_J)_k^2 \right]^{1/2} \quad (2-8)$$

The elemental bias limits, $(B_i)_k$, must be estimated for each variable X_i using the best information one has available at the time. In the design phase of an experimental program, manufacturer's specifications, analytical estimates, and previous experience will typically provide the basis for most of the estimates. As the experimental program progresses, equipment is assembled, and calibrations are conducted, these estimates can be updated using the additional information gained about the accuracy of the calibration standards, errors associated with calibration process and curvefit procedures, and perhaps analytical estimates of

installation errors (such as wall interference effects, sting effects, etc).


As Moffat (1988) suggests, there can be additional conceptual bias errors resulting from not measuring the variable whose symbol appears in the data reduction equation. An example from a towing tank test would be a velocity value measured by the carriage and used as “the” ship model velocity in determining C_T , but there may be cross sectional gradient of the velocity (blockage effects) causing the “average” value to be different.

2.3 Estimating Uncertainty Components for Experimental Results

In the previous section, the methodology for obtaining estimates of the precision errors and bias errors in the measured variables X_i was discussed. In this section, the methodology is presented for obtaining estimates of the precision errors and bias errors in the experimental results r computed using those measured variables in data reduction equations of the form of Eq. (2-2).

2.3.1 Definitions

To estimate the magnitude of the precision component of uncertainty in an experimental result, the **precision limit of a result** P_r is defined. The $\pm P_r$ interval about a result is the band within which the (biased) mean result, μ_r , would fall 95 percent of the time if the experiment were repeated many times under the same conditions using the same equipment. The pre-

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 10 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

recision limit is representative of the scatter (or lack of repeatability) caused by random errors, unsteadiness, inability to reset experimental conditions exactly, etc.

To estimate the magnitude of the bias component of uncertainty in an experimental result, the **bias limit of a result** B_r is defined. The bias limit is estimated with the understanding that the experimenter is 95-percent confident that the true value of the bias error, if known, would be less than $|B_r|$.

The $\pm U_r$ **uncertainty interval** about the result is the band within which the experimenter is 95-percent confident the true value of the result lies. The 95-percent confidence uncertainty is defined as

$$U_r = (B_r^2 + P_r^2)^{\frac{1}{2}} \quad (2-9)$$

2.3.2 Propagation of Precision Limits into an Experimental Result

2.3.2.1 Multiple Tests

If a test is repeated a number of times so that multiple results at the same set point are available, then the best estimate of the result r would be \bar{r} where

$$\bar{r} = \frac{1}{M} \sum_{k=1}^M r_k \quad (2-10)$$

and where M is the number of separate test results. The precision limit for this result would be $P_r = P_{\bar{r}}$ calculated as

$$P_{\bar{r}} = \frac{K S_r}{\sqrt{M}} \quad (2-11)$$

where K is the coverage factor and is taken as 2, as before. S_r is the standard deviation of the sample of M results and is defined as


$$S_r = \left[\sum_{k=1}^M \frac{(r_k - \bar{r})^2}{M-1} \right]^{\frac{1}{2}} \quad (2-12)$$

Obviously, this cannot be computed until multiple results are obtained.

Also note that the precision limit computed is only applicable for those random error sources that were “active” during the repeat measurements. For example, if the model was not disassembled and reassembled between the multiple results, then the precision limit calculated would not account for the fact that the model may not be assembled exactly the same way every time to represent the full-scale article. Further, if the test conditions were not changed and then re-established between the multiple results, the variability due to resetting to a given test condition would not be accounted for.

2.3.2.2 Single Test with Single Readings

The often-encountered situation, discussed in Section 2.2.2, is when measurements of the variables are averaged over a period that is small compared to the periods of the factors causing variability in the experiment. A proper precision limit cannot be calculated from readings taken over such a small time interval. For such data, the measurement(s) of a variable X_i should be considered a single reading, and the precision limit must be estimated based on pre-

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 11 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

viously determined information (calibration data, previous testing in the same facility, previous testing using similar equipment, etc). Once estimates are obtained for the precision limits of all of the measured variables, the precision limit for the result is calculated using

$$P_r = \left[\sum_{i=1}^J (\theta_i P_i)^2 \right]^{\frac{1}{2}} \quad (2-13)$$

$$\text{where } \theta_i = \frac{\partial r}{\partial X_i} \quad (2-14)$$

Here the precision limits are assumed to be based on large sample sizes. Procedures for small sample estimates are discussed in Annex 2.2-A. Equation (2-13) is an approximate equation that can be derived (Coleman & Steele, 1999) using a Taylor series expansion and neglecting all terms higher than first order.

2.3.2.3 Single Test with Averaged Readings

If a test is performed in such a manner that some, but not all, of the X_i 's in Eq. (2-2) are determined as averages over appropriate time periods, then Eq. (2-13) should be used with the precision limits for the averaged variables being computed from Eq. (2-7). If a test is run such that all of the X_i 's could be determined as averages over appropriate time periods, then multiple individual test results can be determined, and the method of Section 2.3.2.1 should be used.

2.3.3 Propagation of Bias Limits into an Experimental Result

When a result is given by

$$r = r(X_1, X_2, \dots, X_J) \quad (2-15)$$

the bias limit of that result is related to the bias limits B_i of the measurements of the separate variables X_i by

$$B_r^2 = \left(\sum_{i=1}^J \theta_i^2 B_i^2 \right) + 2\theta_m \theta_n B'_m B'_n \quad (2-16)$$

where the quantities B'_m and B'_n are the portions of the bias limits for measurements of variables X_m and X_n that arise from the same sources and are presumed to be perfectly correlated (Coleman & Steele, 1999), and the bias limits B_i are estimates at 95-percent confidence of the magnitude of the bias errors in the measurements of the separate variables X_i as previously discussed. Equation (2-16) is an approximate equation that can be derived (Coleman & Steele, 1999) using a Taylor series expansion and neglecting all terms higher than first order. There is a term similar to the final term in Eq. (2-16) for each (m,n) pair of measured variables whose bias errors are correlated.

For example, if


$$r = r(X_1, X_2, X_3) \quad (2-17)$$

and it is possible for portions of the bias limits B_1 , B_2 , and B_3 to arise from the same source(s), then

$$B_r^2 = \theta_1^2 B_1^2 + \theta_2^2 B_2^2 + \theta_3^2 B_3^2 + 2\theta_1 \theta_2 B'_1 B'_2 + 2\theta_1 \theta_3 B'_1 B'_3 + 2\theta_2 \theta_3 B'_2 B'_3 \quad (2-18)$$

If, for instance, the measurements of X_1 and X_2 are each influenced by 4 elemental error sources and sources 2 and 3 are the same for both X_1 and X_2 , then

$$B_1^2 = (B_1)_1^2 + (B_1)_2^2 + (B_1)_3^2 + (B_1)_4^2 \quad (2-19)$$

	ITTC – Recommended Procedures		7.5 – 02 01 – 01 Page 12 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology		Effective Date 1999	Revision 00

$$B_2^2 = (B_2)_1^2 + (B_2)_2^2 + (B_2)_3^2 + (B_2)_4^2 \quad (2-20)$$

and

$$B_1' B_2' = (B_1)_2 (B_2)_2 + (B_1)_3 (B_2)_3 \quad (2-21)$$


Correlated bias errors are those that are not independent of each other. It is not unusual for the uncertainties in the results of experimental programs in the fluid and thermal sciences to be influenced by the effects of correlated bias errors in the measurements of several of the variables. A typical example occurs when different variables are measured using the same transducer, such as multiple pressures sequentially ported to and measured with the same transducer, or temperatures at different positions in a flow measured with a single probe that is traversed across the flow field. Obviously, the bias errors in the variables measured with the same transducer are not independent of one another. Another common example occurs when different variables are measured using different transducers, all of which have been calibrated against the same standard, a situation typical of electronically scanned pressure (ESP) measurement systems. In such a case, at least a part of the bias error arising from the calibration procedure will be the same for each transducer, and thus some of the elemental bias error contributions in the measurements of the variables will be correlated.

A comparative test program is another obvious instance where correlated bias error effects are of great importance. If a test article is tested sequentially at the same free-stream conditions and orientation with and without a configuration change, and the difference in lift coefficients is the experimental result, then most (if not all) of the elemental errors in the

measurement of an individual variable will arise from the same source in the two tests. **Note that the axiom “bias errors subtract out in comparative tests” is not generally correct, even though that is commonly accepted as a truism.** The partial derivatives in Eq. (2-16) are evaluated at the particular values of the measured variables, some of which are different in the two tests in a comparative program. Also, the bias limits can be functions of the measured value of a variable — this occurs when bias limits are of the “% of reading” type rather than the “% of full scale” type, for instance.

Depending on the particular experimental approach, the effect of correlated bias errors in the measurements of different variables can lead either to increased or to decreased uncertainty in the final experimental result as compared to the same approach with no correlated bias errors. Consider the final term in Eq. (2-16) - if some bias errors are correlated (B'_m , B'_n not equal to zero) and the partial derivatives (θ_m and θ_n) are of the same sign, the term is positive and B_T is increased. On the other hand, if some bias errors are correlated and the partial derivatives are of opposite signs, the term is negative and B_T is decreased. This observation suggests that the effect of correlated bias errors can sometimes be used to advantage if the proper strategies are applied in planning and designing the experiment - sometimes one would want to force correlation of bias errors using appropriate calibration approaches, sometimes not.

Coleman & Steele (1999) presents a derivation of the propagation equation for bias errors,

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 13 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

including the effects of correlated elemental bias sources and discussions of the approximation of such terms in practical applications.

2.4 Summary of the Methodology

The uncertainty assessment methodology is summarised in Fig. 2.7. For each experimental result, the data reduction equation [Eq. (2-2)] must be determined. Once this has been done,

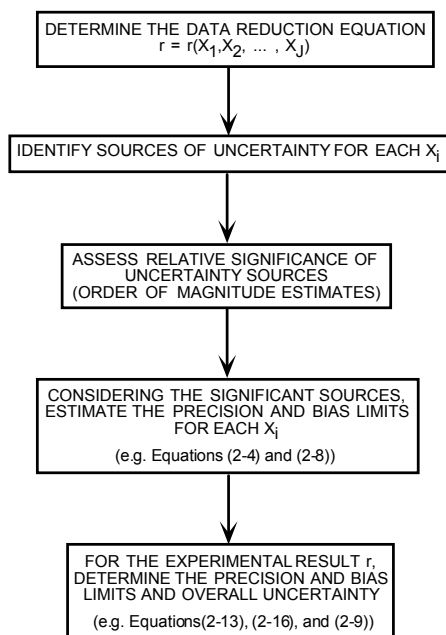



Fig. 2.7 Summary of the uncertainty assessment methodology.

the X_i 's that must be considered are known, and the sources of uncertainty for each X_i should be identified. (Note that a math model for a correction, such as for blockage or wall interference effects, is an X_i whose uncertainty must also be considered.)

Once the sources of uncertainty have been identified, their relative significance should be established. This is often done using order of magnitude estimates of the sources. As a “rule of thumb” for a given X_i , those uncertainty sources that are smaller than 1/4 or 1/5 of the largest sources are usually considered negligible. Resources can then be concentrated on obtaining estimates of those uncertainties of most importance.

For each X_i , estimates of the precision limit and the bias limit are then made. In most towing tank tests, it is generally not cost effective or necessary to try to estimate precision limits at the elemental error source level. It is far more effective to estimate the precision of the measurement systems (at the P_i level in Fig. 2.3 and as defined by Eq. (2-4)) or, even better, the precision of the mean result as given by Eq. (2-11) if multiple results at the same set point are available. Of course, if one encounters unacceptably large P 's, the elemental sources' contributions must be examined to see which need to be (or can be) improved. It is generally easiest to obtain an estimate of the bias limit for X_i by estimating the bias limits of the significant elemental sources and using Eq. (2-8).

The precision limit, bias limit, and overall uncertainty for the experimental result, r , are then found using Eqs. (2-13) [or (2-11)], (2-16) and (2-9). Note that the partial derivatives can be numerically approximated (using finite difference techniques, for example) if one prefers that to finding them analytically.

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 14 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

2.5 Reporting Uncertainties

For each experimental result, the bias limit, precision limit, and overall uncertainty should be reported. For situations in which the large sample assumption is not applicable, the small sample methodology and coverage factor used should be reported and discussed. If outliers are rejected, the circumstances and rationale used in rejecting them should be reported.

Details of the uncertainty assessments (as outlined in Fig. 2.7) should be documented either in an appendix to the primary test report or in a separate document that can be referenced in the primary test report.

ANNEX 2-A: A Comprehensive Uncertainty Analysis Methodology

In this Annex, a comprehensive uncertainty analysis methodology is presented and discussed. This methodology is applicable for either large or small sample sizes and for either Gaussian or non-Gaussian error distributions.

Consider the situation in which the experimental result is determined from

$$r = r(X_1, X_2, \dots, X_J) \quad (2-A-1)$$

where the X_i 's are the values of the measured variables. Then the **combined standard uncertainty** u_c (ISO, 1993a) is given by

$$u_c^2 = \sum_{i=1}^J \left(\theta_i^2 b_i^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{b_{ik}} b_i b_k (1 - \delta_{ik}) \right) + \sum_{i=1}^J \left(\theta_i^2 S_i^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{S_{ik}} S_i S_k (1 - \delta_{ik}) \right) \quad (2-A-2)$$

In Eq. (2-A-2), the S_i^2 are the variances of the precision error distributions of the X_i , the b_i^2 are the variances of the (assumed) bias error distributions of the X_i , the $\rho_{S_{ik}}$ are the correlation coefficients appropriate for the precision errors in variables X_i and X_k , the $\rho_{b_{ik}}$ are the correlation coefficients appropriate for the bias errors in variables X_i and X_k , δ_{ik} is the Kronecker delta defined to equal 1 when $i = k$ and 0 when $i \neq k$, and

$$\theta_i = \frac{\partial r}{\partial X_i} \quad (2-A-3)$$

Eq. (2-A-2) is an approximate equation obtained using a Taylor series expansion and neglecting all terms higher than first order. A derivation is given in Appendix B of Coleman & Steele (1999). No assumptions about type(s) of error distributions are made to obtain Eq. (2-A-2). To obtain an uncertainty U_r at some specified confidence level (such as the 95 percent chosen for use in this document) the combined standard uncertainty u_c must be multiplied by a coverage factor, K ,

$$U_r = K u_c \quad (2-A-4)$$

It is in choosing K that assumptions about the type(s) of the error distributions must be made.

An argument is presented in ISO (1993a) that the error distribution of the result, r , in Eq. (2-A-1), may often be considered Gaussian because of the Central Limit Theorem, even if the error distributions of the X_i are not normal. In fact, the same argument can be made for approximate normality of the error distributions of the X_i since the errors typically are

composed of a combination of errors from a number of elemental sources. If it is assumed that the error distribution of the result, r , is normal, then the value of K for 95-percent coverage corresponds to the 95-percent confidence level value (Table 2-A-1) from the t distribution so that

$$U_r^2 = t^2 \sum_{i=1}^J \left(\theta_i^2 b_i^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{b_{ik}} b_i b_k (1 - \delta_{ik}) \right) + t^2 \sum_{i=1}^J \left(\theta_i^2 S_i^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{S_{ik}} S_i S_k (1 - \delta_{ik}) \right) \quad (2-A-5)$$

The effective number of degrees of freedom v_r for determining t is given (approximately) by the so-called Welch-Satterthwaite formula (ISO, 1993a) as

$$v_r = \frac{u_c^4}{\sum_{i=1}^J \left(\left[(\theta_i S_i)^4 / v_{S_i} \right] + \left[(\theta_i b_i)^4 / v_{b_i} \right] \right)} \quad (2-A-6)$$

where u_c is given by Eq. (2-A-2) with all correlation coefficients set equal to zero and with

$$v_{S_i} = N_i - 1 \quad (2-A-7)$$

for the number of degrees of freedom associated with the S_i . For the number of degrees of freedom v_{b_i} to associate with a non-statistical estimate of b_i , it is suggested in ISO (1993a) that one might use the approximation


$$v_{b_i} \approx \frac{1}{2} \left(\frac{\Delta b_i}{b_i} \right)^{-2} \quad (2-A-8)$$

where the quantity in parenthesis is the relative uncertainty of b_i . For example, if one thought that the estimate of b_i was reliable to within ± 25 percent, then

Table 2-A-1 The t Distribution*.

$v \backslash C$	0.900	0.950	0.990	0.995	0.999
1	6.314	12.706	63.657	127.321	636.619
2	2.920	4.303	9.925	14.080	31.598
3	2.353	3.182	5.841	7.453	12.924
4	2.132	2.776	4.604	5.598	8.610
5	2.015	2.571	4.032	4.773	6.869
6	1.963	2.447	3.707	4.317	5.959
7	1.895	2.365	3.499	4.029	5.408
8	1.860	2.306	3.355	3.833	5.041
9	1.833	2.262	3.250	3.690	4.781
10	1.812	2.228	3.169	3.581	4.587
11	1.796	2.201	3.106	3.497	4.436
12	1.782	2.179	3.055	3.428	4.318
13	1.771	2.160	3.012	3.372	4.221
14	1.761	2.145	2.977	3.326	4.140
15	1.753	2.131	2.947	3.286	4.073
16	1.746	2.120	2.921	3.252	4.015
17	1.740	2.110	2.898	3.223	3.965
18	1.734	2.101	2.878	3.197	3.922
19	1.729	2.093	2.861	3.174	3.883
20	1.725	2.086	2.845	3.153	3.850
21	1.721	2.080	2.831	3.135	3.819
22	1.717	2.074	2.819	3.119	3.792
23	1.714	2.069	2.807	3.104	3.768
24	1.711	2.064	2.797	3.090	3.745
25	1.708	2.060	2.787	3.078	3.725
26	1.706	2.056	2.779	3.067	3.707
27	1.703	2.052	2.771	3.057	3.690
28	1.701	2.048	2.763	3.047	3.674
29	1.699	2.045	2.756	3.038	3.659
30	1.697	2.042	2.750	3.030	3.646
40	1.684	2.021	2.704	2.971	3.551
60	1.671	2.000	2.660	2.915	3.460
120	1.658	1.980	2.617	2.860	3.373
∞	1.645	1.960	2.576	2.807	3.291

* Shown are the values of the two-tailed t distribution for confidence level C and degrees of freedom v .

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 16 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

$$v_{b_i} \approx \frac{1}{2}(0.25)^{-2} \approx 8 \quad (2-A-9)$$

If b_i results from the influence of M elemental error sources $(b_i)_k$, then

$$b_i^2 = \sum_{k=1}^M (b_i)_k^2 \quad (2-A-10)$$

(An analogous equation holds for S_i if precision uncertainties $(S_i)_k$ are estimated for elemental error sources.) There are several distributions - Gaussian, rectangular and triangular, for instance - that might logically be assumed for bias errors (ISO, 1993a). For an assumed Gaussian distribution, one might estimate the 95-percent confidence bias limit $(B_i)_k$, make the large sample assumption so that $t = 2$, and then

$$(b_i)_k = \frac{(B_i)_k}{2} \quad (2-A-11)$$

If one estimates that it is equally probable for $(b_i)_k$ to lie anywhere within an interval $\pm a$ and highly unlikely that it would lie outside that range, then a rectangular error distribution of width $2a$ might be assumed and

$$(b_i)_k = \frac{a}{\sqrt{3}} \quad (2-A-12)$$

If one estimates that it is highly unlikely that $(b_i)_k$ would lie outside a range a , but that values near the midpoint are more likely than near the bounds, then a distribution shaped like an isosceles triangle of base $2a$ might be assumed and


$$(b_i)_k = \frac{a}{\sqrt{6}} \quad (2-A-13)$$

In most practical towing tank tests, it seems (from an engineering perspective) that the use of the preceding equations [(2-A-5) and (2-A-

6)] in this Annex would be excessively and unnecessarily complex and would tend to give a false sense of the degree of significance of the numbers computed using them. In determining what additional simplifying approximations can reasonably be made, the following factors should be considered.

The propagation equation [(Eq. (2-A-5))] is approximate - it is not an exact equation. Unavoidable uncertainties are always present in estimating the bias uncertainties b_i and in estimating their associated degrees of freedom, v_{b_i} . The S_i are usually estimated based on previously determined information (since in most towing tank tests it is not possible to obtain multiple readings of an X_i over an appropriate time interval), and the uncertainties associated with these S_i can be surprisingly large (ISO, 1993a). For samples of a Gaussian parent population with standard deviation σ , 95 out of 100 determinations of the S_i will scatter within an interval of $\pm 0.45\sigma$ if the S_i are determined from $N=10$ readings and within an interval of $\pm 0.25\sigma$ if the S_i are determined from $N=30$ readings (which has traditionally been considered a “large” sample).

Considering the 95-percent confidence t table (Table 2-A-1), one can see that for $v_r \geq 9$ the values of t are within about 13 percent of the large sample t -value of 2. This difference is relatively insignificant compared with the effects discussed in the preceding paragraph. For most engineering applications, it is proposed that Gaussian error distributions and $v_r \geq 9$ be assumed so that $t = 2$ always. (This could be called the “large sample-size assumption”.)

	ITTC – Recommended Procedures	7.5 – 02 01 – 01 Page 17 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology	Effective Date 1999	Revision 00

This eliminates the need for evaluation of v_r using Eq. (2-A-6) and the need to evaluate all of the v_{S_i} and v_{b_i} .

Consideration of Eq. (2-A-6) shows that, because of the exponent of 4 in each term, v_r is most influenced by the number of degrees of freedom of the largest of the $\theta_i S_i$ or $\theta_i b_i$ terms. If, for example, $\theta_3 S_3$ is dominant then $v_r \approx 3 v_{S_3} \geq 9$ for $N_3 \geq 10$ (recalling Eq. (2-A-7)). If, on the other hand, $\theta_3 b_3$ is dominant then $v_r \approx 3 v_{b_3} \geq 9$ when the relative uncertainty in b_i is about 24 percent or less (recalling Eq. (2-A-8)). Therefore, invoking the “large sample-size assumption” essentially means that if a $\theta_i S_i$ is dominant then its $N_i \geq 10$ or if a $\theta_i b_i$ is dominant then the relative uncertainty in that b_i is about 24 percent or less. If there is no single dominant term, but there are M different $\theta_i S_i$ and $\theta_i b_i$ that all have the same magnitude and of degrees of freedom v_a , then $v_r = M v_a$. If $M=3$, for example, v_a would only have to be 3 or greater for v_r to be equal to or greater than 9. Therefore, t can often legitimately be taken as 2 for estimating the uncertainty in a result determined from several measured variables even when the degrees of freedom associated with the measured variables are very small.

If the “large sample-size assumption” is made so that $t = 2$, then from Eq. (2-A-5) the 95-percent confidence expression for U_r becomes

$$U_r^2 = \sum_{i=1}^J \left(\theta_i^2 \left(2 \frac{B_i}{2} \right)^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{b_{ik}} \left(2 \frac{B_i}{2} \right) \left(2 \frac{B_k}{2} \right) (1 - \delta_{ik}) \right) + \sum_{i=1}^J \left(\theta_i^2 (2S_i)^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{S_{ik}} (2S_i)(2S_k)(1 - \delta_{ik}) \right) \quad (2-A-14)$$

Remembering the definition of the precision limit P_r [Eq. (2-4)], this equation can be written


$$U_r^2 = \sum_{i=1}^J \left(\theta_i^2 B_i^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{b_{ik}} B_i B_k (1 - \delta_{ik}) \right) + \sum_{i=1}^J \left(\theta_i^2 P_i^2 + \sum_{k=1}^J \theta_i \theta_k \rho_{S_{ik}} P_i P_k (1 - \delta_{ik}) \right) \quad (2-A-15)$$

If it is additionally assumed that precision errors in different variables are uncorrelated and if the correlated bias term is approximated as discussed in Section 2.3.3, then Eq. (2-A-15) reduces to those equations presented in the body of this section.

The methodology discussed in the body of this section is recommended for use in practical towing tank testing situations unless there are other overriding considerations which require the application of (the still approximate) Equations (2-A-5) and (2-A-6).

ANNEX 2-B: Identification and Elimination of Outliers in Samples

All experimental endeavours can produce data points which appear to be spurious. Such points (outliers) may be caused by intermittent malfunctions of the instrumentation or a physical perturbation not connected with the experiment. For example, a calibration of a pressure measurement system was recently dis-

	ITTC – Recommended Procedures		7.5 – 02 01 – 01 Page 18 of 19
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology		Effective Date 1999

turbed by random spikes in the data caused by a crane being operated in an adjacent room. Obviously, errors of this type should not be included in the uncertainty estimates, assuming that crane operation would be prohibited during a test. When such points occur, they should be removed from the data if the “best” estimate of the sample standard deviation is desired. Thus, all data should be inspected for spurious data points. Identification criteria should be based on engineering analysis of the instrumentation, the physics of the phenomena, theoretical predictions, and/or the history of similar experiments. To ease the burden of examining large amounts of data, computerised routines are available to scan data sets and flag suspected outliers. The suspected outliers should then be analysed with respect to the data set in order to make a judgement about their quality.

The effect of outliers (if they are not rejected) is to increase the estimate of the precision limit of the variable. One of the several techniques in common usage for determining if spurious data points are outliers is Chauvenet's criterion (Coleman & Steele, 1999).

Consider a sample of N measurements of a variable X with a sample standard deviation of S_X . The outlier tests are performed as follows.

Compute

$$\delta_k = |X_k - \bar{X}| \quad (2-B-1)$$

Determine τ from Table 2-B-1. If

$$\delta_k \geq \tau S_X \quad (2-B-2)$$

then X_k meets the criterion and is identified as an outlier.

In general, removing an outlier from the data sample will have a relatively small effect on the mean value, but can have a large effect on the sample standard deviation. There is a continuing controversy over whether the criterion should be applied only once, or more than once, to a given data set. Rejection of outliers should be documented and reported.

A curve-fit equation for τ using Chauvenet's criterion for $N < 833,333$ is


$$\tau = \sum_{i=1}^5 A_i [\ln(N)]^i \quad (2-B-3)$$

where

$$\begin{aligned} A_0 &= 0.720185, & A_1 &= 0.674947, \\ A_2 &= -0.0771831, & A_3 &= 0.00733435, \\ A_4 &= -0.00040635, \text{ and } & A_5 &= 0.00000916028. \end{aligned}$$

3. REFERENCES

- AIAA, 1995, “Assessment of Wind Tunnel Data Uncertainty,” AIAA S-071-1995.
- ANSI/ASME, 1985, “Measurement Uncertainty: Part 1, Instrument and Apparatus,” ANSI/ASME PTC 19.1-1985.
- Coleman, H.W. and Steele, W.G., 1999, “Experimentation and Uncertainty Analysis for Engineers,” 2nd Edition, John Wiley & Sons, Inc., New York, NY.
- ISO, 1992, “Measurement Uncertainty,” ISO/TC 69/SC 6.

	ITTC – Recommended Procedures		7.5 – 02 01 – 01 Page 19 of 19	
	Testing and Extrapolation Methods, General Uncertainty Analysis in EFD Uncertainty Assessment Methodology		Effective Date 1999	Revision 00

ISO, 1993a, “Guide to the Expression of Uncertainty in Measurement,” ISO, First edition, ISBN 92-67-10188-9.

ISO, 1993b, “International Vocabulary of Basic and General Terms in Metrology,” ISO, Second edition, ISBN 92-67-01075-1.

ITTC, 1990, “Report of the Panel of Validation Procedures”, 19th International Towing Tank Conference, Madrid, Spain, Proc. Vol. 1, pp. 577-603.

Moffat, R.J., 1988, “Describing the Uncertainties in Experimental Results,” Experimental Thermal and Fluid Science, Vol. 1.

Steele, W.G., Taylor, R.P., Burrell, R.E., and Coleman, H.W., 1993, “The Use of Data from Previous Experience to Estimate the Precision Uncertainty of Small Sample Experiments,” AIAA Journal, Vol. 31, No. 10.

Taylor, B.N. and Kuyatt, C.E., 1993, “Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results,” NIST Technical Note 1297.