

Standard

Assessment of Experimental Uncertainty With Application to Wind Tunnel Testing

AIAA standards are copyrighted by the American Institute of Aeronautics and Astronautics (AIAA), 1801 Alexander Bell Drive, Reston, VA 20191-4344 USA. All rights reserved.

AIAA grants you a license as follows: The right to download an electronic file of this AIAA standard for temporary storage on one computer for purposes of viewing, and/or printing one copy of the AIAA standard for individual use. Neither the electronic file nor the hard copy print may be reproduced in any way. In addition, the electronic file may not be distributed elsewhere over computer networks or otherwise. The hard copy print may only be distributed to other employees for their internal use within your organization.

Standard

Assessment of Experimental Uncertainty with Application to Wind Tunnel Testing

Sponsor

American Institute of Aeronautics and Astronautics

Abstract

This AIAA Standard provides a new methodology for assessment of experimental uncertainty and a technique for evaluating wind tunnel error sources. The methodology is then applied to a force and pressure test. The document is revision to the original document which was based on a Report of the NATO Advisory Group on Aerospace Research and Development (AGARD).

Library of Congress Cataloging-in-Publication Data

Assessment of experimental uncertainty with application to wind tunnel testing.

p. cm.

"AIAA S-071A-1999"

Includes bibliographical references (.p),

ISBN 1-56347-363-1 (softcover) – ISBN 1-56347-364-X (electronic)

1. Wind tunnels. 2. Airplanes—Models—Testing. I. American Institute of Aeronautics and Astronautics.

TL567.W5 A87 1999

629.134'52 21—dc21

99-041147

Published by

American Institute of Astronautics and Aeronautics
1801 Alexander Bell Drive, Suite 500, Reston, VA 20191

Copyright © 1999 American Institute of Aeronautics and Astronautics
All rights reserved.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without prior written permission of the publisher.

Printed in the United States of America.

Contents

Foreword	v
1. Introduction	1
References	3
2. Uncertainty Assessment Methodology	4
2.1 Overview	4
2.2 Estimating Uncertainty Components in Measured Variables	6
2.2.1 Definitions	6
2.2.2 Estimating Precision Limits	6
2.2.3 Estimating Bias Limits	8
2.3 Estimating Uncertainty Components for Experimental Results	9
2.3.1 Definitions	9
2.3.2 Propagation of Precision Limits into an Experimental Result	10
2.3.3 Propagation of Bias Limits into an Experimental Result	11
2.4 Summary of Methodology	12
2.5 Reporting Uncertainties	13
References	13
3. Wind Tunnel Error Sources	19
3.1 Introduction	19
3.2 Process Error Sources	19
3.2.1 Test Technique	20
3.2.2 Model Shape and Finish	22
3.2.3 Tunnel Flow Quality	22
3.2.4 Instrumentation	23
3.2.5 Math Models	24
3.3 Significance of Error Sources	25
3.4 Concluding Remarks	30
4. Application of the Uncertainty Methodology to a Force and Pressure Test	31
4.1 Introduction	31
4.2 Test Description	31
4.3 Uncertainty of a Measurement System	33
4.3.1 Uncertainty Evaluation Data	33
4.3.2 Outlier Detection	35
4.3.3 Uncertainty of an Instrumentation System Calibration	37
4.3.4 Discussion of Systems with Multiple Dependent Channels	39
4.3.5 Uncertainty Evaluation of Other Test Systems	39
4.3.6 Estimated Uncertainties of Various Parameters	41
4.4 Data Reduction and Estimated Uncertainty of the Forebody Drag Coefficient	41
4.4.1 Tunnel Conditions	42
4.4.2 Model Attitude	44
4.4.3 Measured Gross Axial and Normal Forces	46
4.4.4 Model Aerodynamic Axial and Normal	48
4.4.5 Model Base Axial Force	50
4.4.6 Wind Tunnel Aerodynamic Drag Coefficient	54
4.5 Adjustment to the Aerodynamic Reference Condition	55
4.6 Reporting Uncertainty	56
References	57
5. Summary and Recommendations	80
Nomenclature	82

Annexes

2-A A Comprehensive Uncertainty Analysis Methodology 14

2-B Identification and Elimination of Outliers in Samples 18

4-A Uncertainty Methodology for Multiple Channel Instrumentation Systems..... 58

4-B Determination and Evaluation of the Partial Derivatives Used in the Text..... 60

4-C The Effect of Determining the Partial Derivatives with Respect to Dependent Parameters 70

4-D Pressure Integration Example..... 73

4-E Uncertainty of an Incremental Value..... 78

Downloaded by Siavash Sohrab on April 21, 2015 | http://arc.aiaa.org | DOI: 10.2514/4.473647

Foreword

Measurement uncertainty has long been a topic of discussion and controversy within the aerospace community. The problem is not the lack of good methodology in this area, references on the subject exist and are readily available. The difficulty has been in the application of the methodology by researchers and engineers with consistency and regularity.

The AIAA Standards Technical Council approved an AIAA Standard on measurement uncertainty in 1995. The AIAA has adopted the standard set forth by the NATO Advisory Group for Aerospace Research and Development (AGARD) through AGARD publication AR-304. This standard is the result of an AGARD working group and represents methodology consistent with that adopted by the AIAA Ground Test Technical Committee which has provided support and guidance in the adoption and review of this publication.

At the AGARD Fluid Dynamics Panel symposium on Aerodynamic Data Accuracy and Quality: Requirements and Capabilities in Wind Tunnel Testing in October 1987, continuing important problems related to aerodynamic data quality were noted. The technical evaluator suggested that improved treatment of data uncertainty would help alleviate some problems. The panel approved an examination of data quality assessment methodologies with the intent of recommending specific improvements.

Measurement uncertainty is a complex subject involving both statistical techniques and engineering judgment. The method reported here was adapted from currently accepted practices by Working Group 15 of the AGARD Fluid Dynamics Panel and has been revised by the Standards Subcommittee of the AIAA Ground Testing Technical Committee. The objective of the document, is to provide a rational and practical framework for quantifying and reporting uncertainty in wind tunnel test data. The quantitative assessment method was to be compatible with existing methodologies within the technical community. Uncertainties that are difficult to quantify were to be identified and guidelines given on how to report these uncertainties. Although this document uses wind tunnel testing examples exclusively in its treatment of experimental uncertainty, the methodology is applicable all experimental test processes.

The members of the AGARD Working Group were (affiliation shown as of the publication date of the AGARD document):

Mr. Robin D. Galway (NRC, Institute for Aerospace Research, Ottawa, Canada)	Mr. Travis W. Binion (Micro Craft Technology, Arnold AFB, TN, USA)
Mr. Claude Armand (ONERA, Centre de Modane Avrieux, France)	Mr. David M. Cahill (Micro Craft Technology, Arnold AFB, TN, USA)
Mr. Claude Quemard (ONERA, Centre de Modane Avrieux, France)	Dr. Hugh W. Coleman (University of Alabama in Huntsville, AL, USA)
Dr. Gunter Viehweger (DLR, Köln, Germany)	Dr. Keith L. Kushman, Chairman (USAF-AEDC, Arnold AFB, TN, USA)
Mr. Jan H. A. te Boekhorst (NLR, Amsterdam, Netherlands)	Dr. Frank Steinle (NASA Ames Research Center, CA, USA)
Dr. David S. Woodward (DRA, Farnborough, UK)	
Mr. Keith Pallister (ARA, Bedford, UK)	

The members of the Standards Subcommittee of the AIAA Ground Test Technical Committee who have reviewed and revised this document have for some time been collaborating with the AGARD working group and deliberating themselves on how best to present this methodology as a consensus document. The AIAA, as an accredited non-government standards developer, was deemed to be ideal for this purpose. This standard will be complimented by a recommended practice on its use.

The AIAA Standards Procedures provide that all approved Standards, Recommended Practices, and Guides are advisory only. Their use by anyone engaged in industry or trade is entirely voluntary. There is no agreement to adhere to any AIAA standards publication and no commitment to conform to or be guided by a standards report. In formulating, revising, and approving standards publications, the Committees on Standards will not consider patents which may apply to the subject matter. Prospective users of the publications are responsible for protecting themselves against liability for infringement of patents or copyrights, or both.

The AIAA Ground Test Technical Committee (Dr. A. George Havener, Chairman) and Standards Subcommittee approved the original version of the document in June 1994. The AIAA Standards Technical Council (Ali M. Ghovanlou, Chairman) approved the original version of the document in May 1995.

The Standards Subcommittee of the AIAA Ground Test Technical Committee (Ms. Laura J. McGill, Chairperson) consisted of the following individuals at the time Revision A of the document was approved in June 1999:

Mr. David M. Cahill (Sverdrup Tech., AEDC)	Mr. Daniel E. Marren (AEDC/White Oak)
Mr. Daniel Cresci (GASL)	Mr. Mathew L. Rueger (Boeing)
Dr. Susan T. Hudson (NASA Marshall Space Flight Center)	Mr. William A. Straka (Applied Research Lab., Penn State University)
Mr. Wayne Kalliomaa (Air Force Research Lab.)	Dr. James C. Yu (NASA Langley Research Center)

The AIAA Standards Executive Council accepted the document for publication in August 1999. The Standards Subcommittee would like to thank the following individuals for their review of this document. Approval of the document was unanimous.

Dr. Hugh W Coleman (University of Alabama in Huntsville)
 Dr. Frank W. Steinle Jr. (Sverdrup Tech., AEDC)
 Mr. Robin D. Galway (NRC Institute for Aerospace Research, Canada, Retired)

The AIAA and the GTTC would like to thank Mr. David M. Cahill for his outstanding efforts to develop the revisions, edit the document, and coordinate the revision process to produce Revision A of the Standard.

Summary of the major revisions and corrections made in Revision A:

Chapter 1 – Included the rationale for retaining the Bias Limit and Precision Limit nomenclature

Chapter 2.5 – Included statement concerning the traceability of uncertainty to a primary standards organization such as the National Institute of Standards and Technology (NIST).

Chapter 4.3.3 – Clarification of transducer drift in final paragraph in the section.

Chapter 4.4.3 – Revision in the methodology used to estimate the correlated bias of W_A . Also, clarification of the method used to determine the measured gross forces and moments.

Chapter 4.4.5 – Correction for the way the bias and precision limits for the reference pressure, P_{REF} , were propagated and the calculation of the partial derivatives taken with respect to P_{REF} (thanks to Larry Meyn for pointing out this error). Revision to include a second correlated bias resulting from the bias limit associated with the drift in the transducers caused by a change in operating temperature.

1. Introduction

Wind tunnel data are often presented without reference to the quality of the results. When data uncertainty is considered, it is normally in the form of repeatability from a few supposedly identical tests. Only rarely are estimates of uncertainty based on professional calibrations of facilities and instrumentation, a thorough review of the process producing the data, and comprehensive accounting of significant biases inherent in the experiment.

The development of new and modified aircraft is frequently compromised by inadequate consideration of experimental error. References 1.1 through 1.4 are some of the AGARD publications that have reported important problems over the past 20+ years. An AGARD Symposium in 1987 (Ref. 1.5) entitled "Aerodynamic Data Accuracy and Quality: Requirements and Capabilities in Wind Tunnel Testing" highlighted continuing problems. Two important improvements in data quality assessment practices are clearly needed. The first is to adopt a consistent approach for integrating uncertainty analyses into all phases of a test. The second is to provide a complete professional analysis and documentation of uncertainty for each test. This report describes an engineering approach to wind tunnel data quality assessment that can alleviate many of the problems documented in Ref. 1.5. The method developed in this report is general. Aircraft aerodynamic testing was selected as a specific example application to provide a focus for describing and applying the method.

An important concern of an aircraft developer is the risk inherent in predicting the flight performance of a full-scale system. There are numerous contributors to the uncertainty of flight predictions, as shown in Fig. 1.1. Note that some of the contributions occur as the result of analyses that use wind tunnel data as the starting point. Model protuberance, propulsion temperature effects, and extrapolation from reference conditions are examples of such analyses. Other contributions to data uncertainty are directly related to the wind tunnel test. While the method

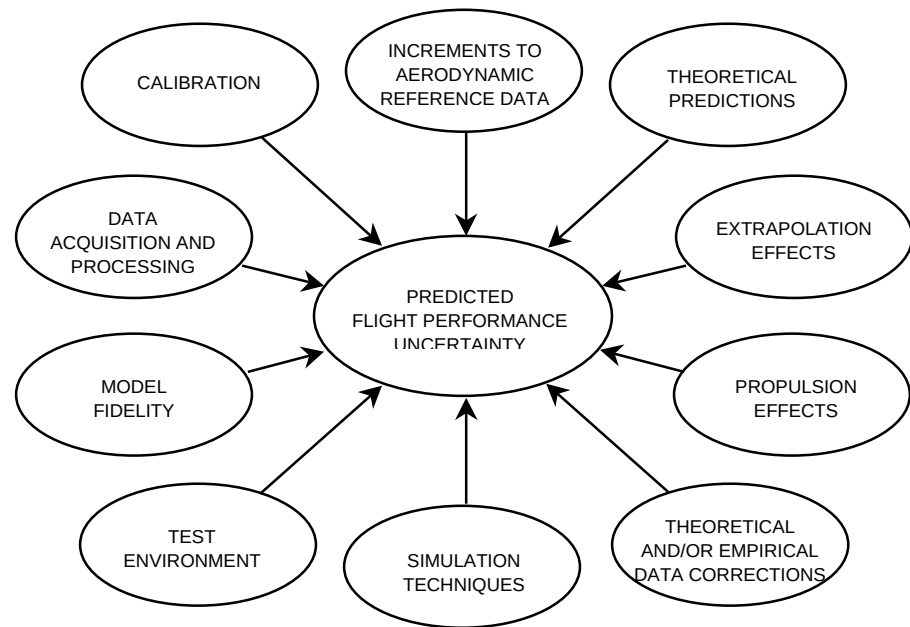


Figure 1.1 Contributions to predicted flight performance uncertainty

presented applies to any contribution, the scope of this report is limited to those associated with the wind tunnel test. The discussion is pointed toward providing the uncertainty of data for an unsupported rigid model in free air at the wind tunnel test conditions (Mach number, Reynolds number, boundary-layer state...), commonly known as the aerodynamic reference condition. Figure 1.2 illustrates typical contributors to uncertainty within the scope of this report. A well-defined, useful reference condition and related uncertainty analysis should be reported by wind tunnel facilities for all tests, regardless of type.

The terms "data quality" and "uncertainty" are used interchangeably throughout this report to reinforce the concept that intelligent design, execution, and documentation of a test adds great value to the results and must be done in a structured, consistent framework to gain the greatest benefit. It is clear that the risk involved in predicting flight performance is directly related to how well tests are designed to provide useful simulations of flight and suitably accurate data. Risk is managed by careful objective and subjective reasoning about the primary sources of error in the prediction processes. Experimentalists know that a

comprehensive uncertainty analysis uses quantitative estimates that are developed in a structured manner with as much rigor as is appropriate and possible. An assessment of the system development and test processes involves judgments about the “quality” of the results to be produced. A major component of such judgments must be an uncertainty estimate.

Data quality assessment should be a key part of the entire wind tunnel testing process. A simple schematic of the process (Fig. 1.3) shows considerations for uncertainty influencing the decision whether to test or not, the design of the experiment, and the conduct of the test. Figure 1.3 also shows the important step of proper analysis and documentation of the uncertainty of final results. This report addresses the entire process shown in Figure 1.3. Chapter Two focuses on the analysis and reporting method. Chapter Three overviews the test process for wind tunnels and provides insight about possible error sources and their significance. Chapter Four presents a specific example of the method applied to wind tunnel testing. Key points are summarized and recommendations are presented in Chapter Five.

A few considerations were prominent in development of this uncertainty method:

1. Simplification of the analyses to permit the greatest possible insight with the least possible effort is important. The method is general enough to apply to any test. The importance of focusing only on significant error sources is stressed, as is the use of “end to end” calibrations to avoid needlessly determining uncertainty contributions of every element in a process.
2. Recent technical contributions to the practice of uncertainty analysis have been incorporated. Therefore, correlated bias errors and an improved way to treat errors for small sample sizes are discussed in Chapter Two. The entire method is consistent with concepts developed and evolving in international standards committees (Ref. 2.2)

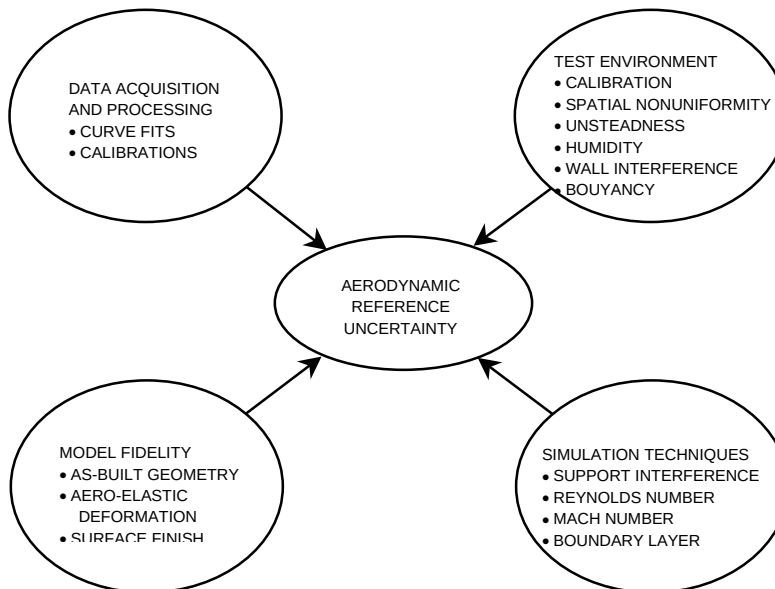


Figure 1.2 Contributions to aerodynamic reference uncertainty

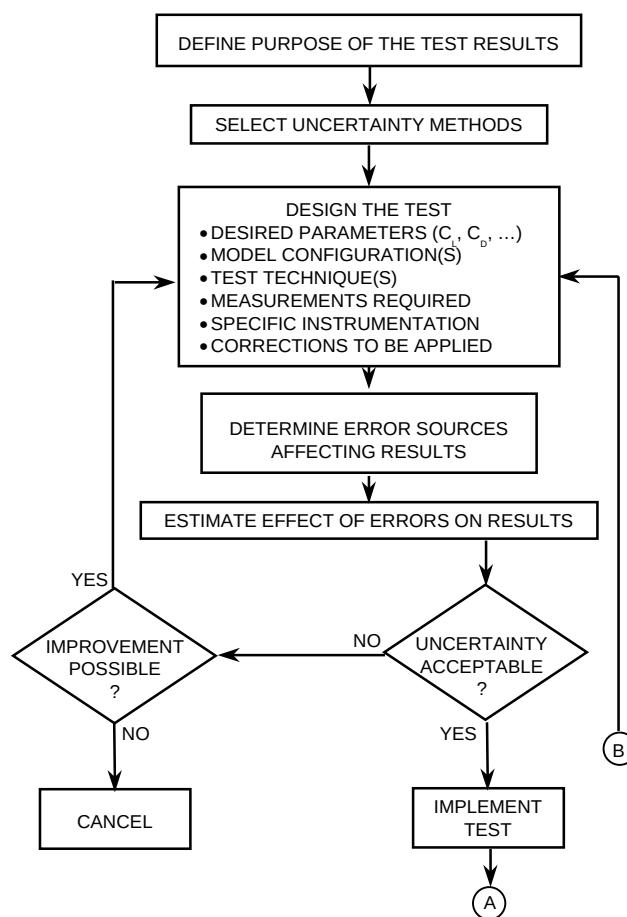


Figure 1.3 Integration of Uncertainty Considerations in an Experimental Process

3. The method should provide a basis for meaningful and efficient communication of quality assessments. A standard confidence level and process are recommended, and a consistent reporting procedure, including precision and bias limits and total uncertainty, is described. In the body of the report, the methodology for uncertainty estimates based on large sample sizes is discussed. This method can be used in most wind tunnel testing, and uncertainty estimates should be those appropriate over long sampling times.
4. Although the methodology contained in this document is in agreement with current international standards, a decision was made to retain the terms Bias Limit and Precision Limit rather than Type A and Type B uncertainties. The terms are familiar to the practicing engineer and retain the distinction between systematic and random sources of error regardless of whether they are determined from statistical (Type A) or other (Type B) methods.

This report assumes the reader is familiar with wind tunnel testing and desires an improved understanding of uncertainty analysis techniques. References describing techniques and testing experiences in detail are cited, but only information essential to the report is included in the text. Readers are encouraged to develop an appreciation for the importance of a professional treatment of data uncertainty and to accept the challenge of improving the quality of their experiments and data.

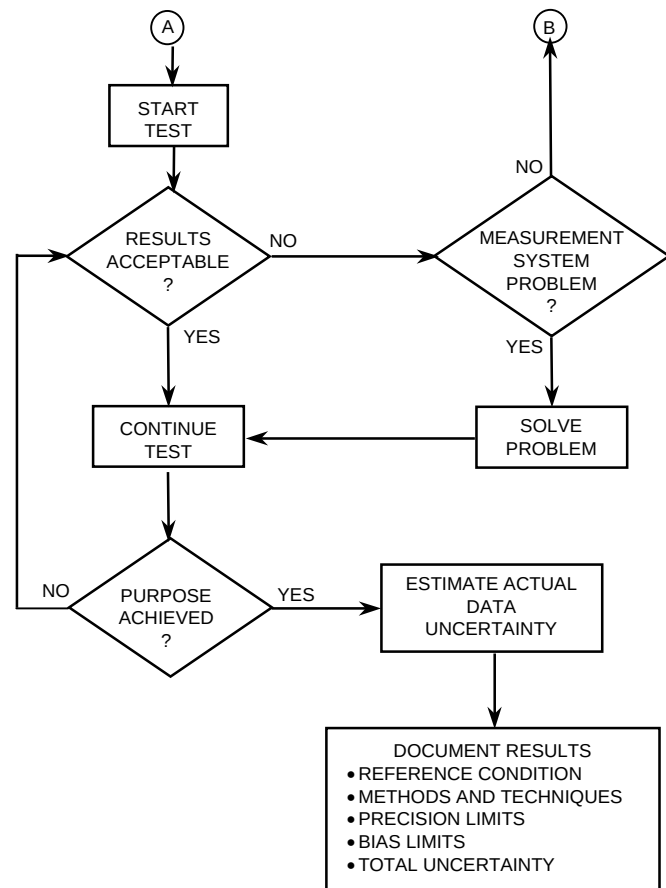


Figure 1.3 Concluded

The units used in this report are the SI units for force (N, Newtons), pressure (Pa, Pascals or Newtons per sq. meter), length (m, meters), and angles (degrees and radians). For simplicity, the units associated with a parameter are only shown for the result of a calculation and not in the equation.

References

- 1.1. Kuchemann, D., "Problems in Wind Tunnel Testing Techniques." AGARD-R-601, November 1972.
- 1.2. Carter, E.S. and Poisson-Quinton, Ph., "Flight/Ground Testing Facilities Correlation." AGARD-CP-187, April 1976.
- 1.3. Dietz, R.O. and Laster, M.L., "Wind Tunnel Corrections for High Angle of Attack Models." AGARD R-692, February 1981.
- 1.4. Ohman, L.H., "Wind Tunnel and Testing Techniques." AGARD-CP-348, February 1988.
- 1.5. Monnerie, B. and Ohman L., "Aerodynamic Data Accuracy and Quality: Requirements and Capabilities in Wind Tunnel Testing." AGARD-CP-429.

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 chapter, an uncertainty analysis methodology is presented, and its application in the different phases of an experimental program is discussed. The methodology is based primarily on material from Ref. 2.1 and is consistent with the most current drafts of international guidelines and standards [Refs. 2.2, and 2.3]. Definitions of specific terms are made as required in the following text, and the international vocabulary of metrological terms (VIM) is incorporated herein as Ref. 2.4.

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 experimentally determined value 95 times out of 100.

Errors can be considered to be composed of 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.

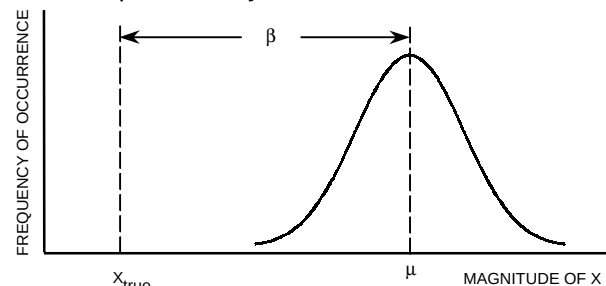


Figure 2.1 Errors in the measurement of a variable X

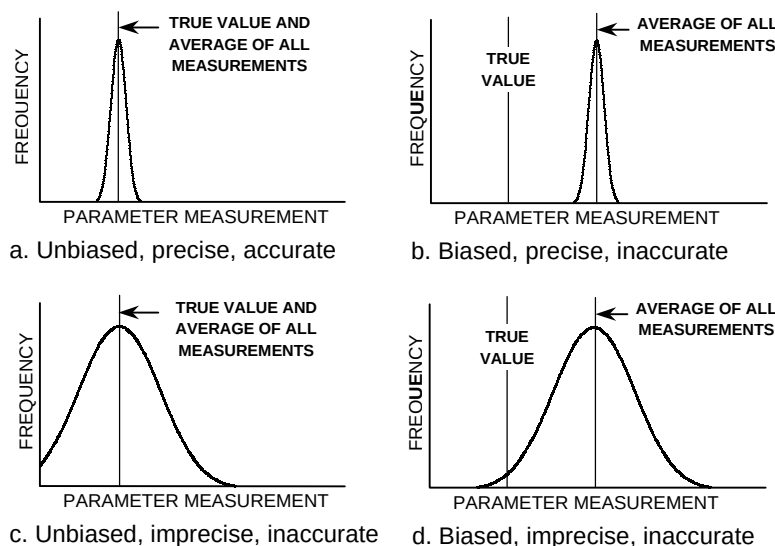


Figure 2.2 Measurement Error (bias, precision, and accuracy)

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 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 drag coefficient of a particular vehicle configuration in a wind tunnel test. Defining drag coefficient as

$$C_D = \frac{2F_D}{\rho V^2 A} \quad (2-1)$$

one can envision that errors in the values of the variables on the right-hand side of Eq. (2-1) will cause errors in the experimental result C_D .

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 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 pressure measurement could be the unsteady test conditions, orifice, tubing, transducer, transducer environment, signal amplifier, power supply, analog-to-digital converter, and recording device. In typical wind tunnel 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 pressure 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. For example, the bias caused by imperfect orifices 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

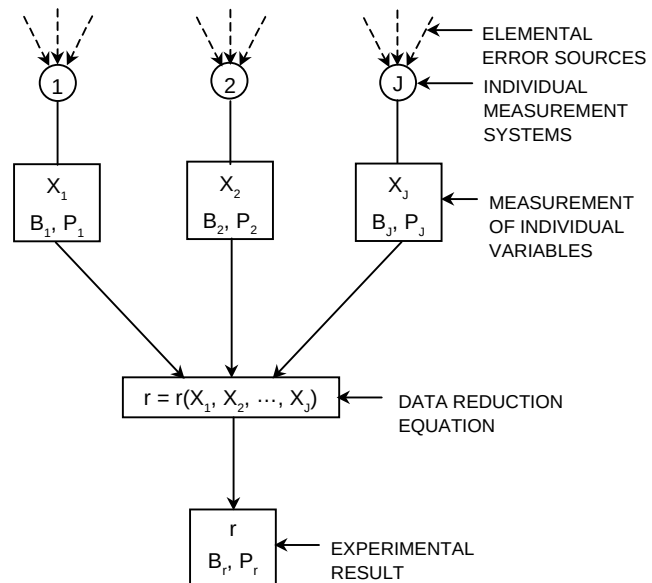


Figure 2.3 Propagation of Errors into an Experimental Result

systems. The examples in Chapter 4 describe ways to estimate the bias and precision of a measurement system.

In Section 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 2.3 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 chapter 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-A, a more comprehensive (and more complex) methodology (Ref. 2.2) that is valid for either small or large sample sizes and either Gaussian or non-Gaussian error distributions is discussed. In Annex 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 2.3.

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-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

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

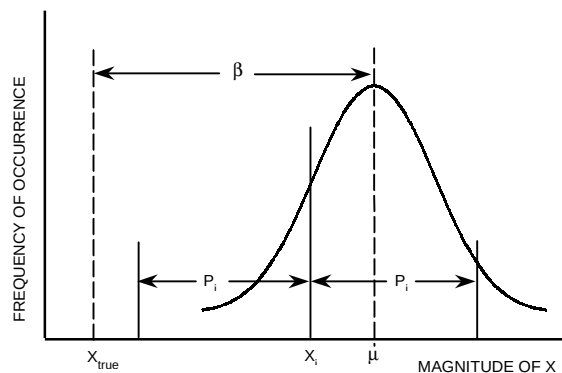


Figure 2.4 95-percent Confidence Precision Limit Interval Around a Single Reading of a Variable X_i

¹ A discussion of what constitutes "large" sample sizes is given in Annex 2-A. In most practical wind tunnel 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.)

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

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

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 (Ref. 2.5) by the U. S. National Institute of Standards and Technology (NIST): "To be consistent with current international practice, the value of K to be used at NIST for calculating U is, by convention, $K = 2$. 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-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 the following equation and shown graphically in Fig. 2.5

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

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

One method for identifying outliers is discussed in Annex 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. Aerodynamic folklore is full of stories about vehicles which "discovered" adverse events during flight tests only to find the same event in the ground test data which was erroneously considered to be an outlier. (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

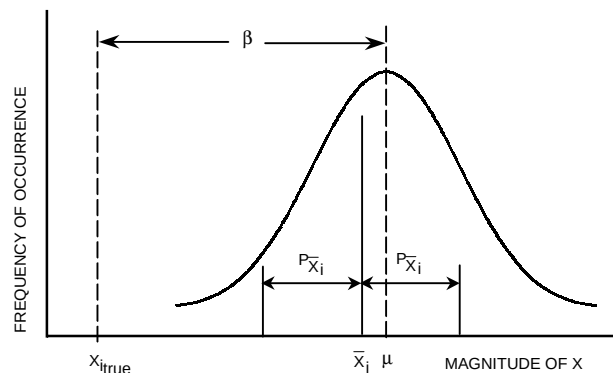


Figure 2.5 95-percent Confidence Precision Limit Interval Around the Mean Value of a sample of Readings of a Variable X_i

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.

A more typical situation in wind tunnel 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

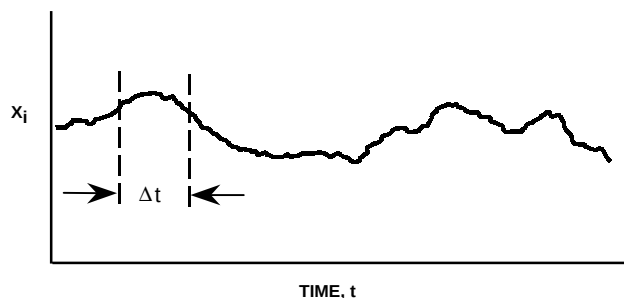


Figure 2.6 Variation of a Variable X_i with Time for a "Steady" Experimental Condition

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 even days and others may not be cyclic at all, but vary binarily. In most wind tunnel 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** (Ref. 2.6). 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 (Ref. 2.1).

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 .

More discussion of assumed bias error distributions is given in Annex 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

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

The elemental bias limits, $(B_J)_k$, 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 curve fit procedures, and perhaps analytical estimates of installation errors (such as wall interference effects, sting effects, etc).

As Moffat (Ref. 2.7) suggests, there can be additional conceptual bias errors resulting from not measuring the variable whose symbol appears in the data reduction equation. An example would be a free-stream velocity value measured at a particular axial position on the tunnel centerline and used as "the" free-stream velocity at that cross section in determining C_D , but there may be a cross-sectional gradient of velocity at that location 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 precision 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 reestablished 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 previously 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 as

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

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

Here the precision limits are assumed to be based on large sample sizes. Procedures for small sample estimates are discussed in Annex 2-A. Equation (2-13) is an approximate equation that can be derived (Ref. 2.1) 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

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

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 (Ref. 2.1), 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 (Ref. 2.1) 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

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

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

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 oc-

curs 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. Treatment of these situations is illustrated in Chapter 4.

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_r 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_r 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.

Reference 2.1 presents a derivation of the propagation equation for bias errors, including the effects of correlated elemental bias sources and discussions of the approximation of such terms in practical applications.

2.4 Summary of Methodology

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

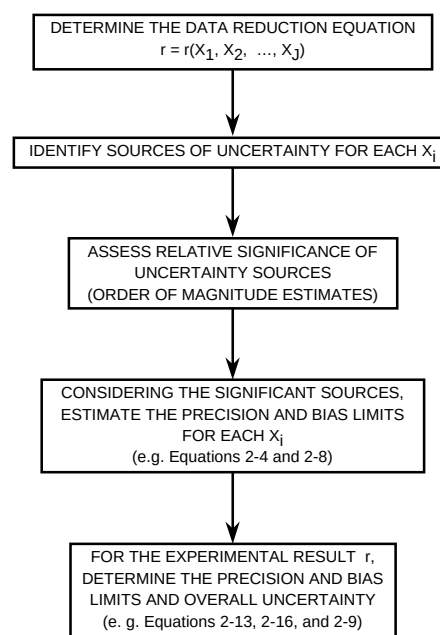


Figure 2.7 Summary of the Uncertainty Assessment Methodology

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.

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. The details should include explanations of error sources for the measurement processes (Sec. 3.2) and documentation of the instrumentation calibration hierarchy (traceability). The hierarchy should culminate in a primary standards organization such as the National Institute of Standards and Technology (NIST).

References

- 2.1 Coleman, H. W. and Steele, W. G., *Experimentation and Uncertainty Analysis for Engineers*, Second edition. John Wiley & Sons, Inc., New York, NY, 1999.
- 2.2 "Guide to the Expression of Uncertainty in Measurement," ISO, First edition, ISBN 92-67-10188-9, 1995.
- 2.3 "Accuracy (Trueness and Precision) of Measurement Methods and Results (6 parts)," ISO 5725, 1994 and 1998, prepared by TC69/SC6.
- 2.4 "International Vocabulary of Basic and General Terms in Metrology," ISO, Second edition, ISBN 92-67-01075-1, 1993.
- 2.5 Taylor, B. N. and Kuyatt, C. E., "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," NIST Technical Note 1297, January 1993.
- 2.6 Steele, W. G., Taylor, R. P., Burrell, R. E., and Coleman, H. W. "The Use of Data from Previous Experience to Estimate the Precision Uncertainty of Small Sample Experiments," *AIAA Journal*, Vol. 31, No. 10, 1993.
- 2.7 Moffat, R. J., "Describing the Uncertainties in Experimental Results," *Experimental Thermal and Fluid Science*, Vol. 1, 1988.

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 (Ref. 2.2) 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 Ref. 2.1. 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 Ref. 2.2 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

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

Table 2-A-1. The t Distribution*

$\begin{array}{c} C \\ v \end{array}$	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

The effective number of degrees of freedom v_r for determining t is given (approximately) by the so-called Welch-Satterthwaite formula (Ref. 2.2) 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 Ref. 2.2 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

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

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 (Ref. 2.2). 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 $\pm a$, 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 wind tunnel 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 wind tunnel 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 (Ref. 2.2). 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”.) 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 $v_r \approx v_{S_3} \geq 9$ for $N_3 \geq 10$ (recalling Eq. (2-A-7)). If, on the other hand, $\theta_3 b_3$ is dominant then $v_r \approx v_{b_3} \geq 9$ 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_i , 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

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

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 chapter.

The methodology discussed in the body of this chapter is recommended for use in practical wind tunnel 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 endeavors 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 disturbed 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, computerized routines are available to scan data sets and flag suspected outliers. The suspected outliers should then be analyzed with respect to the data set in order to make a judgment 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 (Ref. 2.1).

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

$$A_0 = 0.720185, \quad A_1 = 0.674947,$$

$$A_2 = -0.0771831, \quad A_3 = 0.00733435,$$

$$A_4 = -0.00040635, \text{ and } A_5 = 0.00000916028.$$

Table 2-B-1. Chauvenet's Criterion

Number of Readings (N)	τ
3	1.38
4	1.54
5	1.65
6	1.73
7	1.80
8	1.87
9	1.91
10	1.96
15	2.13
20	2.24
25	2.33
50	2.57
100	2.81
300	3.14
500	3.29

3. Wind Tunnel Error Sources

3.1 Introduction

The decision process for establishing a wind tunnel test is idealized in Fig. 1.3. The choice of whether or not to conduct a particular test in a certain facility should be governed by the ability of the expected test outcome to achieve the test objectives within the allowable uncertainties. **The outcome of any test will be dependent on the entire experimental process.** The process includes the design of the experiment, the techniques employed, the instrumentation selected, the flow quality of the facility, and both the reduction and presentation of data and data adjustments or corrections to convert the data to the appropriate reference condition. Under the premise that unacceptable results are worse than none at all, once the best experiment that can be conducted is found to be short of what is needed, the test should not be conducted. The usefulness in quantifying uncertainty for a test is thus to aid in designing the experiment and in formalizing the expected quality of the outcome so that the customer will know what can be expected from the test. Further evaluation and documentation of the test results, including uncertainty, will serve as a reference basis beneficial to both the subjective appraisal of test quality and the assessment of uncertainty for future tests.

As stated above, an experiment's outcome is process dependent. The purpose of this chapter is to discuss some aspects of the process of producing aerodynamic reference coefficients and the various factors that contribute to uncertainty at each step of the process. The chapter highlights considerations essential to the identification and assessment of significant error sources, both of which are key steps in the test process shown in Fig. 1.3. The definitions pertaining to uncertainty and the mathematical tools for estimating uncertainty are discussed in Chapter 2 of this report. Specific examples of the application of the methodology in Chapter 2 for the estimation of uncertainty for a variety of error sources are contained in Chapter 4.

In principle, the application of the tools of Chapter 2 is relatively simple (Fig. 2.7). But, because of the scope of the process from sensor selection to aerodynamic reference results, the application can be quite time consuming if carried out in rigorous detail. When the approach of Chapter 2 is utilized, rigor is required to assure that bias error sources are identified as being either correlated or uncorrelated to arrive at the estimate of uncertainty. The discussion in this chapter assumes that **the planned test techniques are employed faithfully and that no undetected measurement problems occur.** Subject to the considerations stated above, the approach described in Chapter 2 and Fig. 2.7 can be used to define the uncertainty.

3.2 Process Error Sources

Every contribution to the flow of the data stream from sensor to reported data is a source of uncertainty in the final product. There is a choice as to what is defined as the final product. For example, the final product could be defined as merely being the coefficients that are produced without any adjustments for flow quality, support interference, or tare loads. The uncertainty could also be defined in relation to an aerodynamic reference condition. In Fig. 3.1, the progression of corrections which leads to the final corrected out-

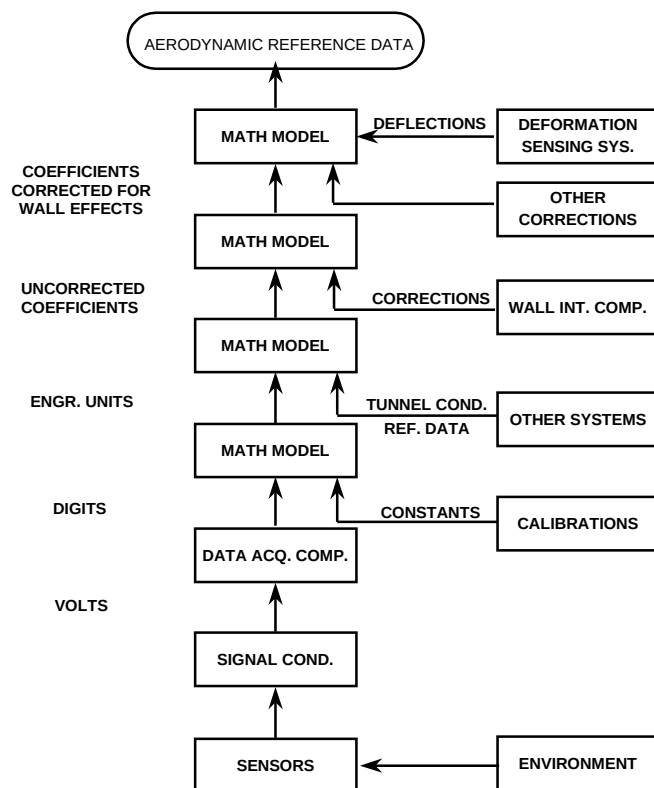


Figure 3.1 Simplified Data Flow Process

put from the test at the aero-reference condition is depicted. Here, the final output is assumed to be the derived parameters which can be used as the starting point in the process of extrapolating/predicting full-scale flight results. The parameters include coefficients, nondimensional quantities (Mach number, Reynolds number, etc.), fluid and stream properties (temperature, enthalpy, dynamic pressure, etc.), angles, and deflections. The process of extrapolating to flight is not considered here; this process has its own uncertainties. Depending on the test objectives, it may not be necessary to estimate the uncertainty at the aerodynamic reference condition. For example, if an increment is being sought between two conditions, corrections such as wall interference and buoyancy may, with good approximation, be the same for each condition and thus can be neglected in the assessment of uncertainty for the increment in question (see Section 2.3.3 for discussion of correlated bias error).

The first step, then, in performing an uncertainty analysis is to trace the entire data flow process from sensor output through all the conversions and adjustments to the data in arriving at the final result. Implicit in the data flow are all of the factors which serve as tributaries to the flow such as the calibrations (traceable to national standards) and all corrections applied (empirical, semi-empirical, and theoretical). Further, it is assumed that sufficient significant digits are presented such that any adjustments to or manipulations of the data flow will not create additional uncertainty.

In each case, uncertainty estimates must be made to decide whether a significant contribution to either bias or precision error is produced. Uncertainty should not automatically be estimated for each potential contributor in the entire chain making up the process. Instead, it is recommended that opportunities for end-to-end uncertainty assessments be identified. There are two reasons for this recommendation. First, an end-to-end assessment will often produce a more realistic (usually lower) estimate of process uncertainty. Second, long-term experience may be of more value in estimating overall uncertainty than knowledge derived from expending resources to determine individual uncertainties for a series of error sources. Regardless of the method employed, it is expected that the first time an assessment of overall uncertainty is developed, it will be more involved than subsequent ones since future assessments will benefit from knowledge gained from prior work.

In Fig. 3.1, the simplified data flow process is shown as starting with the environment affecting the sensors which then respond through the signal conditioning to initiate a data flow. This data flow is processed and modified by input from various sources as it progresses through the several levels to the aerodynamic reference data level. The aerodynamic reference condition is assumed to represent the reference aerodynamic model shape in a free-air environment at wind tunnel test conditions. Departures from any of these aerodynamic reference conditions lead to corrections to the data stream.

Figure 3.2 depicts a more detailed bottom-up data flow process which begins with the environment affecting the sensors (balance, transducers, etc.). The purpose of the increased detail is to illustrate the range of the contributors to the data stream, particularly the environment. The representation is not exhaustive. For example, model motion (inertial forces) is shown coming from the environment and feeding into the balance only and not other instrumentation such as angle sensors and transducers. In general, the sources shown in Fig. 3.2 can be classified as test technique, model, tunnel, instrumentation, and math model related, each of which will be discussed in the following sections.

3.2.1 Test Technique

Although a major potential source of uncertainty, test technique is not shown explicitly in Fig. 3.2. However, it is intimately related to the design of the experiment and should be considered along with the uncertainty sources implied in Fig. 3.2. **It is essential that each step of the test technique be examined to identify and estimate uncertainty contributions to the desired result.** In general, the effects of uncertainty due to test technique must be assessed on the basis of experience aided, whenever practical, by computation. Some examples of test technique-related error sources are presented below.

Model selection and mounting — The method of mounting a model and the choice of the model size are factors which lead to a variety of installation effects that may or may not be acceptable for the purpose of the test. The choice of the mounting system will compromise the test results particularly if done improperly. The choice of the size of a model may be influenced by wall interference considerations. The desire

to have the model large for any number of reasons (reduce uncertainty in extrapolating for Reynolds number effects, increase model fidelity, ease of installing instrumentation, increased strength) may be significantly offset by the increased uncertainty resulting from wall interference effects. The decision to apply, or not to apply wall interference corrections should be governed by the test objectives. How one assesses both support system effects and wall interference corrections and computes their influence on the data is assumed to be part of the math model that produces data adjusted for derived wall effects. Because of the lack of definitive experiments, the uncertainty associated with these adjustments is largely a subjective estimate.

Balance selection — The choice of the capacity of the balance is part of the test technique. Frequently, a compromise in the capacity of the balance is made because the primary purpose of the test is for stability and control over a wide range of conditions and not for drag. When cruise drag is also one of the primary objectives, the proper technique is to change balances and run a drag test with the correctly sized balance. A related common occurrence, driven by schedule and economics, is to try to simultaneously obtain a massive amount of pressure data and force data. This entails crossing the balance with a host of wires and flexible tubing which generally renders the drag measurement useless for accurate determination of cruise performance. The proper technique for drag is to remove all extraneous items bridging the balance and conduct the drag portion of the test separately.

Boundary-layer simulation — Simulation of proper boundary-layer development and verifying the result is an important test technique. In principle, the techniques for sizing and application of artificial roughness are widely known. There are many techniques for application that are all effective. However, there can be secondary effects that become important in situations where accurate aerodynamic reference drag data are required. To assess these effects properly, additional testing is needed. If not verified, the roughness strips can either be partially ineffective (particularly true if the strips are not maintained) or produce too much drag of their own.

Flow angle correction — Stream angle assessment methods are widely used test techniques. An average effect is generally obtained by testing the model both upright and inverted. The choice of how to do this (roll both model and balance, or roll only model) is a part of the technique. The latter has the advantage of eliminating errors in the original positioning of the balance in the process of transferring from the balance axis to the model axis while the former maintains the same load distribution among the balance components. This approach of averaging data from upright and inverted tests is accurate if there is no spanwise variation in stream angle. In that case, drag is affected differently from lift, and a more involved computational technique is required. Similar effects in the lateral direction can lead to large errors in directional static stability assessment if the model moves off centerline and the stream angle in the lateral direction is not constant. In this event, it isn't sufficient to assume that the model is symmetrical and apply a constant

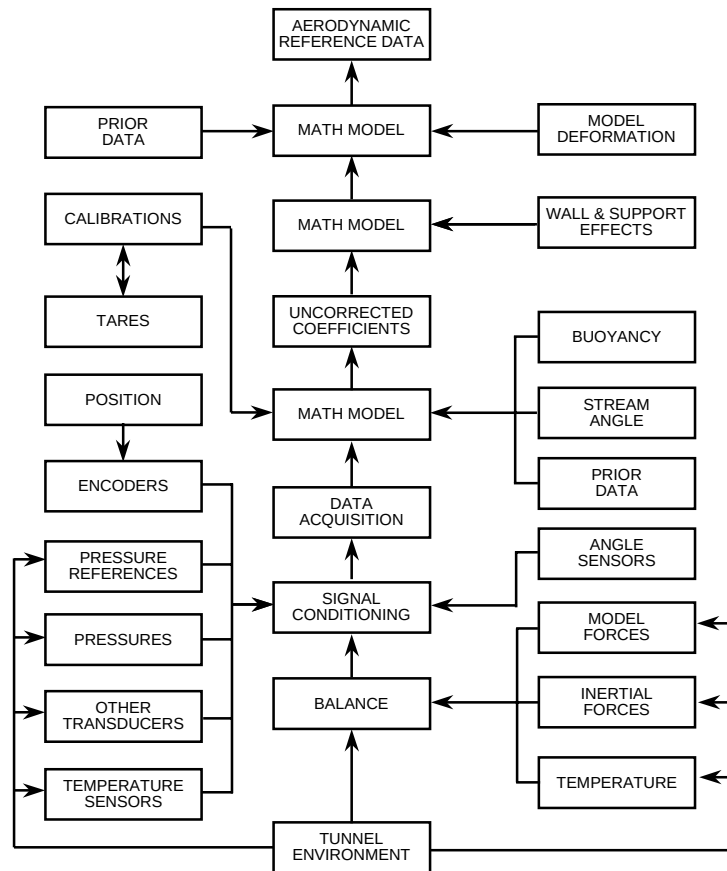


Figure 3.2 Detail Data Flow Process