EXPERIMENTATION, VALIDATION, AND UNCERTAINTY ANALYSIS FOR ENGINEERS
EXPERIMENTATION, VALIDATION, AND UNCERTAINTY ANALYSIS FOR ENGINEERS

FOURTH EDITION

HUGH W. COLEMAN and W. GLENN STEELE

WILEY
For our parents

Mable Morrison Coleman and Marion Templeton Coleman
and
Mary Evans Steele and Wilbur Glenn Steele, Sr.
who were reared and came of age during the Great Depression.
Their work ethic, nurturing, support, and respect for the value of education
have always been our foundation.

H.W.C
W.G.S
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td></td>
<td>xv</td>
</tr>
<tr>
<td>1</td>
<td>Experimentation, Errors, and Uncertainty</td>
<td>1</td>
</tr>
<tr>
<td>1-1</td>
<td>Experimentation, 2</td>
<td></td>
</tr>
<tr>
<td>1-1.1</td>
<td>Why Is Experimentation Necessary?, 2</td>
<td></td>
</tr>
<tr>
<td>1-1.2</td>
<td>Degree of Goodness and Uncertainty Analysis, 3</td>
<td></td>
</tr>
<tr>
<td>1-1.3</td>
<td>Experimentation and Validation of Simulations, 5</td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>Experimental Approach, 6</td>
<td></td>
</tr>
<tr>
<td>1-2.1</td>
<td>Questions to Be Considered, 7</td>
<td></td>
</tr>
<tr>
<td>1-2.2</td>
<td>Phases of Experimental Program, 8</td>
<td></td>
</tr>
<tr>
<td>1-3</td>
<td>Basic Concepts and Definitions, 8</td>
<td></td>
</tr>
<tr>
<td>1-3.1</td>
<td>Errors and Uncertainties, 9</td>
<td></td>
</tr>
<tr>
<td>1-3.2</td>
<td>Categorizing and Naming Errors and Uncertainties, 13</td>
<td></td>
</tr>
<tr>
<td>1-3.3</td>
<td>Estimating Standard Uncertainties, 15</td>
<td></td>
</tr>
<tr>
<td>1-3.4</td>
<td>Determining Combined Standard Uncertainties, 16</td>
<td></td>
</tr>
<tr>
<td>1-3.5</td>
<td>Elemental Systematic Errors and Effects of Calibration, 18</td>
<td></td>
</tr>
<tr>
<td>1-3.6</td>
<td>Expansion of Concept from “Measurement Uncertainty” to “Experimental Uncertainty”, 20</td>
<td></td>
</tr>
<tr>
<td>1-3.7</td>
<td>Repetition and Replication, 22</td>
<td></td>
</tr>
<tr>
<td>1-3.8</td>
<td>Associating a Percentage Coverage or Confidence with Uncertainty Estimates, 24</td>
<td></td>
</tr>
<tr>
<td>1-4</td>
<td>Experimental Results Determined from a Data Reduction Equation Combining Multiple Measured Variables, 25</td>
<td></td>
</tr>
</tbody>
</table>
## CONTENTS

1-5 Guides and Standards, 27
   1-5.1 Experimental Uncertainty Analysis, 27
   1-5.2 Validation of Simulations, 29

1-6 A Note on Nomenclature, 31

References, 31
Problems, 32

2 Coverage and Confidence Intervals for an Individual Measured Variable 33

2-1 Coverage Intervals from the Monte Carlo Method for a Single Measured Variable, 34

2-2 Confidence Intervals from the Taylor Series Method for a Single Measured Variable, Only Random Errors Considered, 35
   2-2.1 Statistical Distributions, 35
   2-2.2 The Gaussian Distribution, 36
   2-2.3 Confidence Intervals in Gaussian Parent Populations, 42
   2-2.4 Confidence Intervals in Samples from Gaussian Parent Populations, 43
   2-2.5 Tolerance and Prediction Intervals in Samples from Gaussian Parent Populations, 48
   2-2.6 Statistical Rejection of Outliers from a Sample Assumed from a Gaussian Parent Population, 51

2-3 Confidence Intervals from the Taylor Series Method for a Single Measured Variable: Random and Systematic Errors Considered, 55
   2-3.1 The Central Limit Theorem, 55
   2-3.2 Systematic Standard Uncertainty Estimation, 56
   2-3.3 The TSM Expanded Uncertainty of a Measured Variable, 58
   2-3.4 The TSM Large-Sample Expanded Uncertainty of a Measured Variable, 61

2-4 Uncertainty of Uncertainty Estimates and Confidence Interval Limits for a Measured Variable, 63
   2-4.1 Uncertainty of Uncertainty Estimates, 63
   2-4.2 Implications of the Uncertainty in Limits of High Confidence Uncertainty Intervals Used in Analysis and Design, 65

References, 67
Problems, 68
3 Uncertainty in a Result Determined from Multiple Variables

3-1 General Uncertainty Analysis vs. Detailed Uncertainty Analysis, 72
3-2 Monte Carlo Method for Propagation of Uncertainties, 73
  3-2.1 Using the MCM in General Uncertainty Analysis, 73
  3-2.2 Using the MCM in Detailed Uncertainty Analysis, 75
3-3 Taylor Series Method for Propagation of Uncertainties, 78
  3-3.1 General Uncertainty Analysis Using the Taylor Series
       Method (TSM), 79
  3-3.2 Detailed Uncertainty Analysis Using the Taylor Series
       Method (TSM), 80
3-4 Determining MCM Coverage Intervals and TSM Expanded
       Uncertainty, 82
  3-4.1 MCM Coverage Intervals for a Result, 82
  3-4.2 TSM Expanded Uncertainty of a Result, 85
3-5 General Uncertainty Analysis Using the TSM and MSM Approaches
   for a Rough-walled Pipe Flow Experiment, 87
  3-5.1 TSM General Uncertainty Analysis, 88
  3-5.2 MCM General Uncertainty Analysis, 89
  3-5.3 Implementation Using a Spreadsheet, 89
  3-5.4 Results of the Analysis, 92
3-6 Comments on Implementing Detailed Uncertainty Analysis
   Using a Spreadsheet, 95
   References, 96
   Problems, 97

4 General Uncertainty Analysis Using the Taylor Series
   Method (TSM)

4-1 TSM Application to Experiment Planning, 100
4-2 TSM Application to Experiment Planning: Special Functional
   Form, 103
4-3 Using TSM Uncertainty Analysis in Planning an Experiment, 107
4-4 Example: Analysis of Proposed Particulate Measuring System, 109
  4-4.1 The Problem, 109
  4-4.2 Proposed Measurement Technique and System, 109
  4-4.3 Analysis of Proposed Experiment, 110
  4-4.4 Implications of Uncertainty Analysis Results, 112
  4-4.5 Design Changes Indicated by Uncertainty Analysis, 113
CONTENTS

4-5 Example: Analysis of Proposed Heat Transfer Experiment, 114
  4-5.1 The Problem, 114
  4-5.2 Two Proposed Experimental Techniques, 115
  4-5.3 General Uncertainty Analysis: Steady-State Technique, 117
  4-5.4 General Uncertainty Analysis: Transient Technique, 121
  4-5.5 Implications of Uncertainty Analysis Results, 123

4-6 Examples of Presentation of Results from Actual Applications, 124
  4-6.1 Results from Analysis of a Turbine Test, 124
  4-6.2 Results from Analysis of a Solar Thermal Absorber/Thruster Test, 125

References, 126

Problems, 127

5 Detailed Uncertainty Analysis: Overview and Determining Random Uncertainties in Results 131
  5-1 Using Detailed Uncertainty Analysis, 131
  5-2 Detailed Uncertainty Analysis: Overview of Complete Methodology, 134
  5-3 Determining Random Uncertainty of Experimental Result, 137
    5-3.1 Example: Random Uncertainty Determination in Compressible Flow Venturi Meter Calibration Facility, 139
    5-3.2 Example: Random Uncertainty Determination in Laboratory-Scale Ambient Temperature Flow Test Facility, 141
    5-3.3 Example: Random Uncertainty Determination in Full-Scale Rocket Engine Ground Test Facility, 143
    5-3.4 Summary, 146

References, 146

6 Detailed Uncertainty Analysis: Determining Systematic Uncertainties in Results 147
  6-1 Estimating Systematic Uncertainties, 149
    6-1.1 Example: Estimating Uncertainty in Property Values, 152
    6-1.2 Example: Estimating Systematic Uncertainties in the Turbulent Heat Transfer Test Facility (THTTF), 153
    6-1.3 Example: An “Optimum” Calibration Approach Used in a Test to Determine Turbine Efficiency, 163
  6-2 Determining Systematic Uncertainty of Experimental Result Including Correlated Systematic Error Effects, 165
    6-2.1 Example: Correlated Systematic Error Effects with “% of Full Scale” (%FS) Systematic Uncertainties, 168
CONTENTS

6-2.2 Example: Correlated Systematic Error Effects with “% of Reading” Systematic Uncertainties, 170
6-2.3 Example: Correlated Systematic Error Effects with Systematic Uncertainties that Vary with Set Point, 171
6-2.4 Example: Correlated Systematic Error Effects When Only Some Elemental Sources Are Correlated, 172
6-2.5 Example: Correlated Systematic Error Effects When Determining Average Velocity of a Fluid Flow, 176

6-3 Comparative Testing, 177
6-3.1 Result Is a Difference of Test Results, 178
6-3.2 Result Is a Ratio of Test Results, 181

6-4 Some Additional Considerations in Experiment Execution, 183
6-4.1 Choice of Test Points: Rectification, 183
6-4.2 Choice of Test Sequence, 188
6-4.3 Relationship to Statistical Design of Experiments, 189
6-4.4 Use of a Jitter Program, 191
6-4.5 Comments on Transient Testing, 193
6-4.6 Comments on Digital Data Acquisition Errors, 193

References, 194

Problems, 195

7 Detailed Uncertainty Analysis: Comprehensive Examples 199

7-1 TSM Comprehensive Example: Sample-to-Sample Experiment, 199
7-1.1 The Problem, 199
7-1.2 Measurement System, 200
7-1.3 Zeroth-Order Replication-Level Analysis, 201
7-1.4 First-Order Replication-Level Analysis, 205
7-1.5 Nth-Order Replication-Level Analysis, 206

7-2 TSM Comprehensive Example: Use of Balance Checks, 207

7-3 Comprehensive Example: Debugging and Qualification of a Timewise Experiment, 210
7-3.1 Orders of Replication Level in Timewise Experiments, 211
7-3.2 Example, 212

7-4 Comprehensive Example: Heat Exchanger Test Facility for Single and Comparative Tests, 216
7-4.1 Determination of the Uncertainty in \( \Delta q \) for Two Core Designs Tested Sequentially Using the Same Facility and Instrumentation, 224
7-5 Case Study: Examples of Single and Comparative Tests of Nuclear Power Plant Residual Heat Removal Heat Exchanger, 230
7-5.1 Single Test Results for an RHR Heat Exchanger (HX1), 231
7-5.2 Comparative Test Approach for the Decrease in Fouling Resistance and Its Uncertainty, 234

References, 235
Problems, 235

8 The Uncertainty Associated with the Use of Regressions 239

8-1 Overview of Linear Regression Analysis and Its Uncertainty, 240
  8-1.1 Uncertainty in Coefficients, 241
  8-1.2 Uncertainty in \( Y \) from Regression Model, 241
  8-1.3 \((X_i, Y_i)\) Variables Are Functions, 243

8-2 Determining and Reporting Regression Uncertainty, 243
  8-2.1 MCM Regression Uncertainty Determination, 244
  8-2.2 TSM Regression Uncertainty Determination, 244
  8-2.3 Reporting Regression Uncertainties, 244

8-3 Method of Least Squares Regression, 246

8-4 First-Order Regression Example: MCM Approach to Determine Regression Uncertainty, 249

8-5 Regression Examples: TSM Approach to Determine Regression Uncertainty, 252
  8-5.1 Uncertainty in First-Order Coefficients, 252
  8-5.2 Uncertainty in \( Y \) from First-Order Regression, 253
  8-5.3 Uncertainty in \( Y \) from Higher-Order Regressions, 255
  8-5.4 Uncertainty in \( Y \) from Regressions in Which \( X \) and \( Y \) Are Functional Relations, 255
  8-5.5 Uncertainty Associated with Multivariate Linear Regression, 257

8-6 Comprehensive TSM Example: Regressions and Their Uncertainties in a Flow Test, 259
  8-6.1 Experimental Apparatus, 261
  8-6.2 Pressure Transducer Calibration and Uncertainty, 261
  8-6.3 Venturi Discharge Coefficient and Its Uncertainty, 265
  8-6.4 Flow Rate and Its Uncertainty in a Test, 269

References, 273
Problems, 273
9 Validation of Simulations 277

9-1 Introduction to Validation Methodology, 277
9-2 Errors and Uncertainties, 278
9-3 Validation Nomenclature, 279
9-4 Validation Approach, 280
9-5 Code and Solution Verification, 284
9-6 Interpretation of Validation Results Using $E$ and $u_{\text{val}}$, 284
  9-6.1 Interpretation with No Assumptions Made about Error Distributions, 285
  9-6.2 Interpretation with Assumptions Made about Error Distributions, 285
9-7 Estimation of Validation Uncertainty $u_{\text{val}}$, 286
  9-7.1 Case 1: Estimating $u_{\text{val}}$ When Experimental Value $D$ of Validation Variable Is Directly Measured, 287
  9-7.2 Cases 2 and 3: Estimating $u_{\text{val}}$ When Experimental Value $D$ of Validation Variable Is Determined from Data Reduction Equation, 290
  9-7.3 Case 4: Estimating $u_{\text{val}}$ When Experimental Value $D$ of Validation Variable Is Determined from Data Reduction Equation That Itself Is a Model, 295
9-8 Some Practical Points, 297
References, 299

Answers to Selected Problems 301

Appendix A Useful Statistics 305

Appendix B Taylor Series Method (TSM) for Uncertainty Propagation 311

B-1 Derivation of Uncertainty Propagation Equation, 312
B-2 Comparison with Previous Approaches, 316
  B-2.1 Abernethy et al. Approach, 316
  B-2.2 Coleman and Steele Approach, 317
  B-2.3 ISO Guide 1993 GUM Approach, 318
  B-2.4 AIAA Standard, AGARD, and ANSI/ASME Approach, 319
  B-2.5 NIST Approach, 319
B-3 Additional Assumptions for Engineering Applications, 319
  B-3.1 Approximating the Coverage Factor, 320
References, 322
Appendix C Comparison of Models for Calculation of Uncertainty 325
  C-1 Monte Carlo Simulations, 325
  C-2 Simulation Results, 328
  References, 336

Appendix D Shortest Coverage Interval for Monte Carlo Method 337
  Reference, 338

Appendix E Asymmetric Systematic Uncertainties 339
  E-1 Procedure for Asymmetric Systematic Uncertainties Using TSM Propagation, 340
  E-2 Procedure for Asymmetric Systematic Uncertainties Using MCM Propagation, 344
  E-3 Example: Biases in a Gas Temperature Measurement System, 344
  References, 351

Appendix F Dynamic Response of Instrument Systems 353
  F-1 General Instrument Response, 353
  F-2 Response of Zero-Order Instruments, 355
  F-3 Response of First-Order Instruments, 356
  F-4 Response of Second-Order Instruments, 359
  F-5 Summary, 362
  References, 362

Index 363
PREFACE

When we first agreed to produce a fourth edition, our plan was to update some material on the Monte Carlo Method (MCM) and to add a few examples. However, we quickly realized two things.

First, since we are both now retired from our professor day jobs, this edition is the first where we have the luxury of adequate time.

Second, since publication of the first edition we have taught our two-day short course based on the book to more than 125 classes containing people of almost every conceivable engineering and scientific background. Using what we learned from their questions and suggestions, we adjusted the sequence and logic of the presentation of the basic concepts in the course—and that sequence and logic no longer corresponded to the presentation in the book.

In this fourth edition, we have updated the sequence of presentation of basic ideas by introducing some topics earlier and by expanding the discussion of others. The chapter on uncertainty propagation has been rewritten to stress that the MCM has become the primary method for propagating uncertainties. This revised Chapter 3 also more clearly shows our concepts of “general uncertainty analysis” and “detailed uncertainty analysis” and shows how to use a spreadsheet to apply the MCM even in complex uncertainty propagation situations. The presentation of detailed uncertainty analysis has been divided into three chapters: determining random uncertainties in results, determining systematic uncertainties in results, and comprehensive examples. We have also added a large number of examples drawn from our personal experience, including a comprehensive example that covers all areas of uncertainty analysis.
Perhaps the thing that sets our book apart is its basis on the wide background of our experiences during the combined 100 years of our engineering careers. Our expertise in heat transfer, fluid mechanics, propulsion, energy systems, and uncertainty analysis has been used in applications for land-based, aviation, naval, and space projects, and we have personally worked on testing programs from laboratory-scale to full-scale trials. We have been active participants in professional societies’ committees developing experimental uncertainty and validation standards for use by practicing engineers. Our interactions in classes of our short course with participants from industry, laboratories, and academia have been stimulating and informative.

We would like to acknowledge the invaluable contributions of all our students to this work and also the contributions of our university and technical committee colleagues. We are always excited to present the material to classes and to observe how the students quickly see the myriad of applications of applied uncertainty analysis to their specific tests and analyses.

HUGH W. COLEMAN and W. GLENN STEELE
November 2017
When the word \textit{experimentation} is encountered, most of us immediately envision someone in a laboratory “taking data.” This idea has been fostered over many decades by portrayals in periodicals, television shows, and movies of an engineer or scientist in a white lab coat writing on a clipboard while surrounded by the piping and gauges in a refinery or by an impressive complexity of laboratory glassware. In recent years, the location is often a control room filled with computerized data acquisition equipment with lights blinking on the racks and panels. To some extent, the manner in which laboratory classes are typically implemented in university curricula also reinforces this idea. Students often encounter most instruction in experimentation as demonstration experiments that are already set up when the students walk into the laboratory. Data are often taken under the pressure of time, and much of the interpretation of the data and the reporting of results is spent on trying to rationalize what went wrong and what the results “would have shown if…”

\textit{Experimentation is not just data taking.} Any engineer or scientist who subscribes to the widely held but erroneous belief that experimentation is making measurements in the laboratory will be a failure as an experimentalist. The actual data-taking portion of a well-run experimental program generally constitutes a small percentage of the total time and effort expended. In this book we examine and discuss the steps and techniques involved in a logical, thorough approach to the subject of experimentation.
1-1 EXPERIMENTATION

1-1.1 Why Is Experimentation Necessary?

Why are experiments necessary? Why do we need to study the subject of experimentation? The experiments run in science and engineering courses demonstrate physical principles and processes, but once these demonstrations are made and their lessons taken to heart, why bother with experiments? With the laws of physics we know, with the sophisticated analytical solution methods we study, with the increasing knowledge of numerical solution techniques, and with the awesome computing power available, is there any longer a need for experimentation in the real world?

These are fair questions to ask. To address them, it is instructive to consider Figure 1.1, which illustrates a typical analytical approach to finding a solution to a physical problem. Experimental information is almost always required at one or more stages of the solution process, even when an analytical approach is used. Sometimes experimental results are necessary before realistic assumptions and idealizations can be made so that a mathematical model of the real-world process can be formulated using the basic laws of physics. In addition, experimentally determined information is generally present in the form of physical property values and the auxiliary equations (e.g., equations of state) necessary for obtaining a solution. So we see that even in situations in which the solution approach is analytical (or numerical), information from experiments is included in the solution process.

From a more general perspective, experimentation lies at the very foundations of science and engineering. Webster’s [1] defines science as “systematized knowledge derived from observation, study, and experimentation carried on in order to understand the natural world.”

![Figure 1.1 Analytical approach to solution of a problem.](image-url)
to determine the nature or principles of what is being studied.” In discussing the scientific method, Shortley and Williams [2] state: “The scientific method is the systematic attempt to construct theories that correlate wide groups of observed facts and are capable of predicting the results of future observations. Such theories are tested by controlled experimentation and are accepted only so long as they are consistent with all observed facts.”

In many systems and processes of scientific and engineering interest, the geometry, boundary conditions, and physical phenomena are so complex that it is beyond our present technical capability to formulate satisfactory analytical or numerical models and approaches. In these cases, experimentation is necessary to define the behavior of the systems and/or processes (i.e., to find a solution to the problem).

### 1-1.2 Degree of Goodness and Uncertainty Analysis

If we are using property data or other experimentally determined information in an analytical solution, we should certainly consider how “good” the experimental information is. Similarly, anyone comparing results of a mathematical model with experimental data (and perhaps also with the results of other mathematical models) should certainly consider the degree of goodness of the data when drawing conclusions based on the comparisons. This situation is illustrated in Figure 1.2.

In Figure 1.2a the results of two different mathematical models are compared with each other and with a set of experimental data. The authors of the two models might have a fine time arguing over which model compares better with the data. In Figure 1.2b, the same information is presented, but a range representing the uncertainty (likely amount of error) in the experimental value of $Y$ has been plotted for each data point. It is immediately obvious that once the degree of goodness of the $Y$ value is taken into consideration it is fruitless to argue for the validity of one model over another based only on how well the model results match the data. The “noise level” established by the data uncertainty effectively sets the resolution at which such comparisons can be made.

We will discuss such “validation” comparisons between simulation results and experimental results in considerable detail as we proceed. At this point, we will note that the experimental values of $X$ will also contain errors, and so an uncertainty should also be associated with $X$. In addition, the simulation result also has uncertainty arising from modeling errors, errors in the inputs to the model, and possibly errors from the algorithms used to numerically solve the simulation equations.

From this example, one might conclude that even a person with no ambition to become an experimentalist needs an appreciation of the experimental process and the factors that influence the degree of goodness of experimental data and results from simulations.

Whenever the experimental approach is to be used to answer a question or to find the solution to a problem, the question of how good the results will be should
be considered long before an experimental apparatus is constructed and data are taken. If the answer or solution must be known within, say, 5% for it to be useful to us, it would make no sense to spend the time and money to perform the experiment only to find that the probable amount of error in the results was considerably more than 5%.

In this book we use the concept of uncertainty to describe the degree of goodness of a measurement, experimental result, or analytical (simulation) result.
An uncertainty \( u \) defines an interval that is estimated to contain the actual value of an error of unknown sign and magnitude.

Schenck [3] quotes S. J. Kline as defining an experimental uncertainty as “what we think the error would be if we could and did measure it by calibration.”

An error \( \delta \) is a quantity that has a particular sign and magnitude, and a specific error \( \delta_i \) is the difference caused by error source \( i \) between a quantity (measured or simulated) and its true value. As we will discuss in detail later, it is generally assumed that each error whose sign and magnitude are known has been removed by correction. Any remaining error is thus of unknown sign and magnitude,\(^1\) and an uncertainty \( u \) is estimated with the idea that \( \pm u \) characterizes the range containing \( \delta \).

Uncertainty \( u \) is thus an estimate: a \( \pm u \) interval\(^2\) is an estimate of a range within which we believe the actual (but unknown) value of an error \( \delta \) lies. This is illustrated in Figure 1.3, which shows an uncertainty interval \( \pm u_d \) that contains the error \( \delta_d \) whose actual sign and magnitude are unknown.

Uncertainty analysis (the analysis of the uncertainties in experimental measurements and in experimental and simulation results) is a powerful tool. This is particularly true when it is used in the planning and design of experiments. As we will see in Chapter 4, there are realistic, practical cases in which all the measurements in an experiment can be made with 1% uncertainty yet the uncertainty in the final experimental result will be greater than 50%. Uncertainty analysis, when used in the initial planning phase of an experiment, can identify such situations and save the experimentalist much time, money, and embarrassment.

1-1.3 Experimentation and Validation of Simulations

Over the past several decades, advances in computing power, modeling approaches, and numerical solution algorithms have increased the ability of the

---

\(^1\)There are asymmetric errors that are more likely to (or are certain to) have one sign rather than the other. Treatment of these by either “zero centering” or estimating asymmetric uncertainties is discussed in Chapter 6 and Appendix E.

\(^2\)In our initial conceptual presentation, for simplicity we discuss the uncertainty interval as if it is symmetric. This is not a necessary assumption and will be discarded as we present more detailed discussion of the characteristics of errors and uncertainties.
EXPERIMENTATION, ERRORS, AND UNCERTAINTY

scientific and engineering community to simulate real-world processes to the point that it is realistic for predictions from surprisingly detailed simulations to be used to replace much of the experimentation that was previously necessary to develop designs for new systems and bring them to the market. The new systems to which we refer cover the gamut from simple mechanical and structural devices to rocket engine injectors to commercial aircraft to military weapons systems to nuclear power systems.

In the past, it was necessary to test (experimentally determine) subsystem and system performance at numerous set points covering the expected domain of operation of the system. For large, complex systems the required testing program can be prohibitively expensive, if not outright impossible, with available finite resources. The current approach seeks to replace some or much of the experimentation with (cheaper) simulation results that have been validated with experimental results at selected set points—but to do this with confidence one must know “how good” the predictions are at the selected set points. This has led to the emergence of the field called verification and validation (V&V) of simulations (e.g., models, codes).

The verification part refers to application of approaches to determine that the algorithms solve the equations in the model correctly and to estimate the numerical uncertainty if the equations are discretized as, for example, in the finite-difference, finite-element, and finite-volume approaches used in computational mechanics. Verification addresses the question of whether the equations are solved correctly but does not address the question of how well the equations represent the real world. Validation is the process of determining the degree to which a model is an accurate representation of the real world—it addresses the question of how good the predictions are.

Verification is a necessary component of the validation process and will be described briefly with references cited to guide the reader who desires more detail, but more than that is beyond the scope of what we want to cover in this book. Since experimentation and the uncertainties in experimental results and in simulation results are central issues in validation, the details of validation are covered in this book. Basic ideas and concepts are developed as they arise naturally in the discussion of experimental uncertainty analysis—for example, estimating the uncertainty in the simulation result due to the uncertainties in the simulation inputs. The application of the ideas and concepts in validation are covered in Chapter 9 with detailed discussion and examples.

1-2 EXPERIMENTAL APPROACH

In most experimental programs the experimental result (or question to be answered) is not directly measured but is determined by combining multiple measured variables using a data reduction equation (DRE). Examples are determination of the rate of heat transfer within a heat exchanger by measuring flow rates and temperatures and using tabulated fluid properties. Likewise, all
of the dimensionless groups such as drag coefficient, Nusselt number, Reynolds number, and Mach number that are used to present the results of a test are themselves DREs. In addition to determining the appropriate DRE(s) for the experimental program, other questions must be answered.

1-2.1 Questions to Be Considered

When an experimental approach is to be used to find a solution to a problem, many questions must be considered. Among these are the following:

1. What question are we trying to answer? (What is the problem?)
2. How accurately do we need to know the answer? (How is the answer to be used?)
3. What physical principles are involved? (What physical laws govern the situation?)
4. What experiment or set of experiments might provide the answer?
5. What variables must be controlled? How well?
6. What quantities must be measured? How accurately?
7. What instrumentation is to be used?
8. How are the data to be acquired, conditioned, and stored?
9. How many data points must be taken? In what order?
10. Can the requirements be satisfied within the budget and time constraints?
11. What techniques of data analysis should be used?
12. What is the most effective and revealing way to present the data?
13. What unanticipated questions are raised by the data?
14. In what manner should the data and results be reported?

Although by no means all-inclusive, this list does indicate the range of factors that must be considered by the experimentalist. This might seem to be a discouraging and somewhat overwhelming list, but it need not be. With the aid of uncertainty analysis and a logical, thorough approach in each phase of an experimental program, the apparent complexities often can be reduced and the chances of achieving a successful conclusion enhanced.

A key point is to avoid becoming so immersed in the many details that must be considered that the overall objective of the experiment is forgotten. This statement may sound trite, but it is true nonetheless. We perform an experiment to find the answer to a question. We need to know the answer within some uncertainty, the magnitude of which is usually determined by the intended use of the answer. Uncertainty analysis is a tool that we use to make decisions in each phase of an experiment, always keeping in mind the desired result and uncertainty. Properly applied, this approach will guide us past the pitfalls that are usually not at all obvious and will enable us to obtain an answer with an acceptable uncertainty.
1-2.2 Phases of Experimental Program

There are numerous ways that a general experimental program can be divided into different components or phases. For our discussions in this book, we consider the experimental phases to be planning, design, construction, debugging, execution, data analysis, and reporting of results. There are not sharp divisions between these phases—in fact, there is generally overlap and sometimes several phases will be ongoing simultaneously (as when something discovered during debugging leads to a design change and additional construction on the apparatus).

In the planning phase we consider and evaluate the various approaches that might be used to find an answer to the question being addressed. This is sometimes referred to as the preliminary design phase.

In the design phase we use the information found in the planning phase to specify the instrumentation needed and the details of the configuration of the experimental apparatus. The test plan is identified and decisions made on the ranges of conditions to be run, the data to be taken, the order in which the runs will be made, and so on.

During the construction phase, the individual components are assembled into the overall experimental apparatus, and necessary instrument calibrations are performed.

In the debugging phase, the initial runs using the apparatus are made and the unanticipated problems (which must always be expected!) are addressed. Often, results obtained in the debugging phase will lead to some redesign and changes in the construction and/or operation of the experimental apparatus. At the completion of the debugging phase, the experimentalist should be confident that the operation of the apparatus and the factors influencing the uncertainty in the results are well understood.

During the execution phase, the experimental runs are made and the data are acquired, recorded, and stored. Often, the operation of the apparatus is monitored using checks that were designed into the system to guard against unnoticed and unwanted changes in the apparatus or operating conditions.

During the data analysis phase, the data are analyzed to determine the answer to the original question or the solution to the problem being investigated. In the reporting phase, the data and conclusions should be presented in a form that will maximize the usefulness of the experimental results.

In the chapters that follow we discuss a logical approach for each of these phases. We will find that the use of uncertainty analysis and related techniques (e.g., balance checks) will help to ensure a maximum return for the time, effort, and financial resources invested.

1-3 BASIC CONCEPTS AND DEFINITIONS

There is no such thing as a perfect measurement. All measurements of a variable contain inaccuracies. Because it is important to have an understanding of these
inaccuracies if we are to perform experiments (use the experimental approach to answer a question) or if we are simply to use values that have been determined experimentally, we must carefully define the concepts involved. As stated in the previous section, generally a data reduction equation DRE is used to combine multiple measured variables into a test result, so it is necessary to consider errors and uncertainties within the context of a single measured variable and then to consider how those errors and uncertainties propagate through the DRE to produce the errors and uncertainties in the result. There are currently two approaches used to model the propagation.

The first is the Monte Carlo Method (MCM), which samples errors drawn from assumed distributions and simulates running the test many times with a different set of sampled errors each time. The authoritative international guide for this approach is the Joint Committee for Guides in Metrology (JCGM) [4], which we will henceforth refer to as the 2008 GUM and use as the standard reference for the MCM.

The second is the classic Taylor Series Method (TSM) with higher-order terms neglected (thus making it less “exact”). The authoritative international guide for this approach is from the International Organization for Standardization (ISO) [5], which we will henceforth refer to as the 1993 GUM and use as the standard reference for the TSM.

The two propagation methods are discussed in detail beginning in Chapter 3, but some differences in the way they model the errors and uncertainties will be discussed in the following sections and also in Chapter 2 as we consider the concepts of coverage intervals and confidence intervals in cases in which a directly measured variable is the desired experimental result and thus there is no DRE.

1-3.1 Errors and Uncertainties

Consider a variable $X$ in a process that is considered to be steady so that its true value ($X_{\text{true}}$) is constant. Measurements of the variable are influenced by a number of elemental error sources—such as the errors in the standard used for calibration and from an imperfect calibration process; errors caused by variations in ambient temperature, humidity, pressure, vibrations, electromagnetic influences; unsteadiness in the “steady-state” phenomenon being measured; errors due to undesired interactions of the transducer with the environment; errors due to imperfect installation of the transducer; and others.

As an example, suppose that a measurement system is used to make $N$ successive measurements of $X$ and that the measurements are influenced by five significant error sources, as shown in Figure 1.4.

The first three of those measurements are given by

$$X_1 = X_{\text{true}} + (\delta_1)_1 + (\delta_2)_1 + (\delta_3)_1 + (\delta_4)_1 + (\delta_5)_1$$

$$X_2 = X_{\text{true}} + (\delta_1)_2 + (\delta_2)_2 + (\delta_3)_2 + (\delta_4)_2 + (\delta_5)_2$$

$$X_3 = X_{\text{true}} + (\delta_1)_3 + (\delta_2)_3 + (\delta_3)_3 + (\delta_4)_3 + (\delta_5)_3$$

(1.1)
EXPERIMENTATION, ERRORS, AND UNCERTAINTY

Figure 1.4 Measurement of a variable influenced by five error sources.

where $\delta_1$ is the value of the error from the first source, $\delta_2$ the value of the error from the second source, and so on. Each of the measurements $X_1$, $X_2$, and $X_3$ has a different value since errors from some of the sources vary during the period when measurements are taken and so are different for each measurement while others do not vary and so are the same for each measurement. Using traditional nomenclature, we assign the symbol $\beta$ (beta) to designate an error that does not vary during the measurement period and the symbol $\epsilon$ (epsilon) to designate an error that does vary during the measurement period. For this example, we will assume that the errors from sources 1 and 2 do not vary and the errors from sources 3, 4, and 5 do vary, so that Eq. 1.1 can be written

$$
X_1 = X_{\text{true}} + \beta_1 + \beta_2 + ((\epsilon_3)_1) + ((\epsilon_4)_1) + ((\epsilon_5)_1)
$$

$$
X_2 = X_{\text{true}} + \beta_1 + \beta_2 + ((\epsilon_3)_2) + ((\epsilon_4)_2) + ((\epsilon_5)_2)
$$

$$
X_3 = X_{\text{true}} + \beta_1 + \beta_2 + ((\epsilon_3)_3) + ((\epsilon_4)_3) + ((\epsilon_5)_3)
$$

(1.2)

Since just by looking at the measured values we cannot distinguish between $\beta_1$ and $\beta_2$ or among $\epsilon_1$, $\epsilon_2$, and $\epsilon_3$, Eq. (1.3) describes what we actually have,

$$
X_1 = X_{\text{true}} + \beta + (\epsilon)_1 \quad X_2 = X_{\text{true}} + \beta + (\epsilon)_2 \quad X_3 = X_{\text{true}} + \beta + (\epsilon)_3
$$

(1.3)

where now

$$
\beta = \beta_1 + \beta_2 \quad \epsilon_1 = \epsilon_3 + \epsilon_4 + \epsilon_5
$$

(1.4)

This process of making successive measurements of $X$ is shown schematically in Figure 1.5. In Figure 1.5a, the first measurement $X_1$ is shown. The difference between the measured value and the true value is the total error $\delta_{X_1}$, which is the sum of the invariant error $\beta$ (the combination of all of the errors from the invariant elemental error sources) and the varying error $(\epsilon)_1$ (the combination at the time $X_1$ is measured of all of the errors from the error sources that vary during the period.
BASIC CONCEPTS AND DEFINITIONS

Figure 1.5  Effect of errors on successive measurements of a variable $X$.

that our $N$ measurements are taken). In Figure 1.5b the second measurement is also shown, and of course the total error $\delta X_2$ differs from $\delta X_1$ because the varying error $e$ is different for each measurement.

If we continued to acquire additional measurements, we could plot a histogram, which presents the fraction of the $N$ total measurements with values between $X$ and $X + \Delta X$, $X + \Delta X$ and $X + 2\Delta X$, $X + 2\Delta X$ and $X + 3\Delta X$, and so on, versus $X$, where $\Delta X$ is the bin width. This is shown schematically in Figure 1.5c and allows us to view the distribution of the total of $N$ measured values. This distribution of the sample population of $N$ measurements often tends to have a larger number of the measured values near the mean of the sample and a decreasing number of measured values as one moves away from the mean. A mean value $\bar{X}$ can be calculated, as can a standard deviation $s$, which is an indicator of the width of the distribution of the $X$ values (the amount of “scatter” of the measurements caused by the errors from the elemental sources that varied during the measurement period).

As the number of measurements approaches infinity, the parent population distribution would likely appear as shown in Figure 1.5d (although it would not necessarily be exactly symmetric), with the mean $\mu$ offset from $X_{\text{true}}$ by $\beta$, the combination of all of the invariant errors. Of course, we never have an infinite number of measurements, but conceptually the idea of the parent population distribution is very useful to us.

An example of this behavior exhibited by a real set of measurements is shown in Figure 1.6. A thermometer immersed in an insulated container of water was read
EXPERIMENTATION, ERRORS, AND UNCERTAINTY

Figure 1.6 Histogram of temperatures read from a thermometer by 24 students.

independently by 24 of our students to the nearest tenth of a degree Fahrenheit. Unknown to the students, the thermometer was biased (“read high”) by a little over a degree, and the “true” temperature of the water was about 96.0 F. The temperatures read by the students are distributed around an average value of about 97.2 F and are biased (offset) from the true value of 96.0 F.

With such a data sample, what we would like to do is use information from the sample to specify some range \( (X_{\text{best}} \pm u_X) \) within which we think \( X_{\text{true}} \) falls. Generally we take \( X_{\text{best}} \) to be equal to the average value of the \( N \) measurements (or to the single measured value \( X \) if \( N = 1 \)). The uncertainty \( u_X \) is an estimate of the interval \( (\pm u_X) \) that likely contains the magnitude of the combination of all of the errors affecting the measured value \( X \). Look back at the first measurement of \( X \) illustrated in Figure 1.5a and imagine that it is influenced by five significant error sources as in Figure 1.4. Then, recalling Eq. 1.2, the expression for the first measured value \( X_1 \) is given by

\[
X_1 = X_{\text{true}} + \beta_1 + \beta_2 + (\epsilon_3)_1 + (\epsilon_4)_1 + (\epsilon_5)_1
\]  

To associate an uncertainty with a measured \( X \) value, we need to have elemental uncertainty estimates for all of the elemental error sources. That is, \( u_1 \) is an uncertainty that defines an interval \( (\pm u_1) \) within which we think the value of \( \beta_1 \) falls, while \( u_3 \) is an uncertainty that defines an interval \( (\pm u_3) \) within which we think the value of \( \epsilon_3 \) falls.

Using the concepts and procedures in the 2008 GUM [4] and the 1993 GUM [5], a standard uncertainty \( (u) \) is defined as an estimate of the standard deviation of