William L. Oberkampf and Christopher J. Roy

# Verification and Validation in Scientific Computing

CAMBRIDGE

# VERIFICATION AND VALIDATION IN SCIENTIFIC COMPUTING

WILLIAM L. OBERKAMPF
CHRISTOPHER J. ROY



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521113601



First published 2010

Printed in the United Kingdom at the University Press, Cambridge

*A catalog record for this publication is available from the British Library*

*Library of Congress Cataloging in Publication data*
Oberkampf, William L., 1944–
Verification and validation in scientific computing / William L. Oberkampf, Christopher J. Roy.
p. cm.
Includes index.
ISBN 978-0-521-11360-1 (hardback)
1. Science – Data processing. 2. Numerical calculations – Verification. 3. Computer programs – Validation.
4. Decision making – Mathematical models. I. Roy, Christopher J. II. Title.
Q183.9.O24 2010
502.85 – dc22 2010021488

ISBN 978-0-521-11360-1 Hardback

# VERIFICATION AND VALIDATION IN SCIENTIFIC COMPUTING

Advances in scientific computing have made modeling and simulation an important part of the decision-making process in engineering, science, and public policy. This book provides a comprehensive and systematic development of the basic concepts, principles, and procedures for verification and validation of models and simulations. The emphasis is placed on models that are described by partial differential and integral equations and the simulations that result from their numerical solution. The methods described can be applied to a wide range of technical fields, such as the physical sciences, engineering, and technology, as well as to a wide range of applications in industry, environmental regulations and safety, product and plant safety, financial investing, and governmental regulations.

This book will be genuinely welcomed by researchers, practitioners, and decision-makers in a broad range of fields who seek to improve the credibility and reliability of simulation results. It will also be appropriate for either university courses or independent study.

WILLIAM L. OBERKAMPF has 39 years of experience in research and development in fluid dynamics, heat transfer, flight dynamics, and solid mechanics. He has worked in both computational and experimental areas, and taught 30 short courses in the field of verification and validation. He recently retired as a Distinguished Member of the Technical Staff at Sandia National Laboratories.

CHRISTOPHER J. ROY is an Associate Professor in the Aerospace and Ocean Engineering Department at Virginia Tech. After receiving his PhD from North Carolina State University in 1998, he spent five years working as a Senior Member of the Technical Staff at Sandia National Laboratories. He has published numerous articles on verification and validation in the area of computational fluid dynamics. In 2006, he received a Presidential Early Career Award for Scientists and Engineers for his work on verification and validation in computational science and engineering.

**To our wives, Sandra and Rachel**

# Preface

Modeling and simulation are used in a myriad of ways in business and government. The range covers science, engineering and technology, industry, environmental regulations and safety, product and plant safety, financial investing, design of military systems, governmental planning, and many more. In all of these activities models are built that are mental constructs of how we believe the activity functions and how it is influenced by events or surroundings. All models are abstractions of the real activity that are based on many different types of approximation. These models are then programmed for execution on a digital computer, and the computer produces a simulation result. The simulation result may have high fidelity to the actual activity of interest, or it may be complete nonsense. The question is: how can we tell which is which? This book deals with various technical and procedural tools that can be used to assess the fidelity of modeling and simulation aspects of scientific computing. Our focus is on physical processes and systems in a broad range of the natural sciences and engineering.

The tools discussed here are primarily focused on mathematical models that are represented by differential and/or integral equations. Many of these mathematical models occur in physics, chemistry, astronomy, Earth sciences, and engineering, but they also occur in other fields of modeling and simulation. The topics addressed in this book are all related to the principles involved in assessing the credibility of the models and the simulation results. We do not deal with the specific details of modeling the physical process or system of interest, but with assessment procedures relating to the fidelity of the models and simulations. These procedures are typically described by the terms *verification* and *validation*.

We present the state of the art in verification and validation of mathematical models and scientific computing simulations. Although we will discuss the terminology in detail, *verification* can simply be described as "solving the equations right" and *validation* as "solving the right equations." Verification and validation (V&V) are built on the concept of quantitative accuracy assessment. V&V do not answer the entire question of simulation credibility, but they are key contributors. V&V could be described as the processes that provide evidence of the correctness and/or accuracy of computational results. To measure correctness, one must have accurate benchmarks or reference values with which to compare. However, the majority of simulations of complex processes do not have a computable or measurable reference value. For these situations we must rely on numerical error estimation and estimation of the effects of all of the contributors to uncertainty in system responses. In verification, the primary benchmarks are highly accurate solutions to specific, although limited, mathematical models. In validation, the benchmarks are high-quality experimental measurements of system response quantities of interest. These experimental measurements, and the detailed information of the system being tested, should also have carefully estimated uncertainty in all of the quantities that are needed to perform a simulation of the experiment.

Mathematical models are built and programmed into software for the purpose of making predictions of system responses for cases where we do not have experimental data. We refer to this step as *prediction*. Since prediction is the usual goal of modeling and simulation, we discuss how accuracy assessment results from V&V activities enter into prediction uncertainty. We discuss methods for including the estimated numerical errors from the solution of the differential and/or integral equations into the prediction result. We review methods dealing with model input uncertainty and we present one approach for including estimated model uncertainty into the prediction result. The topic of how to incorporate the outcomes of V&V processes into prediction uncertainty is an active area of current research.

Because the field of V&V for models and simulations is in the early development stage, this book does not simply provide a prescriptive list of steps to be followed. The procedures and techniques presented will apply in the majority of cases, but there remain many open research issues. For example, there are times where we point out that some procedures may not be reliable, may simply not work, or may yield misleading results.

As the impact of modeling and simulation has rapidly increased during the last two decades, the interest in V&V has also increased. Although various techniques and procedures have been developed in V&V, the philosophical foundation of the field is *skepticism*. Stated differently, if the evidence for computer code correctness, numerical error estimation, and model accuracy assessment are not presented as part of a prediction, then the V&V perspective presumes these activities were not done and the results should be questioned. We feel this is the appropriate counter balance to commonly unsubstantiated claims of accuracy made by modeling and simulation. As humankind steadily moves from decision making primarily based on system testing to decision making based more heavily on modeling and simulation, increased prudence and caution are in order.

# Acknowledgments

Although only two names appear on the cover of this book, we recognize that if other people had not been there for us, and many others had not helped, this book would have never been written. These people provided training and guidance, created opportunities, gave advice and encouragement, corrected us when we were wrong, and showed the way to improved understanding of the subject. Although there were many pivotal individuals early in our lives, here we only mention those who have contributed during the last decade when the idea for this book first came to mind.

Timothy Trucano, Frederic Blottner, Patrick Roache, Dominique Pelletier, Daniel Aeschlimam, and Luís Eça have been critical in generously providing technical insights for many years. We have benefited from their deep knowledge of verification and validation, as well as a number of other fields. Jon Helton and Scott Ferson have guided our way to an understanding of uncertainty quantification and how it is used in risk-informed decision making. They have also provided key ideas concerning how to connect quantitative validation results with uncertainty estimates in model predictions. Without these people entering our technical lives, we would not be where we are today in our understanding of the field.

Martin Pilch created opportunities and provided long-term funding support at Sandia National Laboratories, without which we would not have been able to help advance the state of the art in V&V. He, along with Paul Hommert, Walter Rutledge, and Basil Hassan at Sandia, understood that V&V and uncertainty quantification were critical to building credibility and confidence in modeling and simulation. They all recognized that both technical advancements and changes in the culture of people and organizations were needed so that more reliable and understandable information could be provided to project managers and decision makers.

Many colleagues provided technical and conceptual ideas, as well as help in working through analyses. Although we cannot list them all, we must mention Mathew Barone, Robert Croll, Sharon DeLand, Kathleen Diegert, Ravi Duggirala, John Henfling, Harold Iuzzolino, Jay Johnson, Cliff Joslyn, David Larson, Mary McWherter-Payne, Brian Rutherford, Gary Don Seidel, Kari Sentz, James Stewart, Laura Swiler, and Roger Tate. We have benefited from the outstanding technical editing support through the years from Rhonda Reinert and Cynthia Gruver. Help from students Dylan Wood, S. Pavan Veluri, and John Janeski in computations for examples and/or presentation of graphical results was vital.

Reviewers of the manuscript have provided invaluable constructive criticism, corrections, and suggestions for improvements. Edward Allen, Ryan Bond, James Carpenter, Anthony Giunta, Matthew Hopkins, Edward Luke, Chris Nelson, Martin Pilch, William Rider, and William Wood reviewed one or more chapters and helped immeasurably in improving the quality and correctness of the material. Special recognition must be given to Tim Trucano, Rob Easterling, Luís Eça, Patrick Knupp, and Frederick Blottner for commenting on and correcting several draft chapters, or in some cases, the entire manuscript. We take full responsibility for any errors or misconceptions still remaining.

We were blessed with encouragement and patience from our wives, Sandra and Rachel. They tolerated our long hours of work on this book for longer than we deserved.

# Contents

# 1
# Introduction

This chapter briefly sketches the historical beginnings of modeling and simulation (M&S). Although claiming the beginning of anything is simply a matter of convenience, we will start with the stunning invention of calculus. We then discuss how the steadily increasing performance and decreasing costs of computing have been another critical driver in advancing M&S. Contributors to the credibility of M&S are discussed, and the preliminary concepts of verification and validation are mentioned. We close the chapter with an outline of the book and suggest how the book might be used by students and professionals.

## 1.1 Historical and modern role of modeling and simulation

### *1.1.1 Historical role of modeling and simulation*

For centuries, the primary method for designing an engineered system has been to improve the successful design of an existing system incrementally. During and after the system was built, it would be gradually tested in a number of ways. The first tests would usually be done during the building process in order to begin to understand the characteristics and responses of the new system. This new system was commonly a change in the old system's geometrical character, materials, fastening techniques, or assembly techniques, or a combination of all of these changes. If the system was intended to be used in some new environment such as a longer bridge span, a taller structure, or propelled at higher speeds, the system was always tested first in environments where the experience base already existed. Often, during the building and testing process, design or assembly weaknesses and flaws were discovered and modifications to the system were made. Sometimes a catastrophic failure of a monumental project would occur and the process would start over: occasionally after attending the funeral of the previous chief designer and his apprentices (DeCamp, 1995). In ancient times, chief designers understood the consequences of a major design failure; they had skin in the game.

After the invention of calculus by Newton and Leibniz around 1700, the mathematical modeling of physics slowly began to have an impact on concepts for the understanding of nature and the design of engineered systems. The second key ingredient to have an impact on mathematical physics was the invention of logarithms by John Napier about 1594 (Kirby *et al.*, 1956). A mathematical model is of little practical use until it is exercised, which today is referred to as obtaining a *simulation result*. Until the existence and use of logarithms, it was not practical to conduct simulations on a routine basis. Then, not long after the invention of logarithms, the slide rule was invented by William Oughtred. This device provided a mechanical method for adding and subtracting logarithms and enabling rapid multiplication and division of numbers. The slide rule and mechanical calculators revolutionized not only simulation, but also such fields as surveying, navigation, and astronomy. Even though by today's standards the combination of mathematical theory and computing machines would be called "Before Computers," it provided the opportunity for the beginning of massive changes in science, engineering, and technology.

Starting with the Industrial Revolution, roughly around 1800 in England, the impact of modeling and simulation on engineering and design began to grow rapidly. However, during the Industrial Revolution, M&S was always an adjunct to experimentation and testing of engineered systems, always playing a minor support role. The primary reason for this was that computations were typically done by hand on a slide rule or mechanical calculator. By the early 1960s, programmable digital computers began to appear in a wide number of industrial, academic, and governmental organizations. During this time period, the number of arithmetic calculations commonly done for a simulation grew from hundreds or thousands to millions of calculations. It would be reasonable to identify the 1960s as the beginning of widespread scientific computing. In this book, we restrict the term *scientific computing* to the numerical solution of models given by partial differential equations (PDEs) or integro-differential equations. During the 1960s, computer power reached the level where scientific computing began to have a significant effect on the design and decision making of engineered systems, particularly aerospace and military systems. It is appropriate to view scientific computing as a field within the broader topic of M&S, which today includes systems that would have, for example, fundamental involvement with human behavior, such as economic and investment modeling, and individual and social modeling.

There were a few important exceptions, such as nuclear weapons design in the US, where scientific computing began to significantly influence designs in the 1940s and 1950s. The initial impetus for building much faster computers was the Cold War between the US and the Soviet Union. (See Edwards, 1997 for a perspective of the early history of electronic computing and their influence.) M&S activities were primarily modeling activities in the sense that models were simplified until it was realistic to obtain simulation results in an acceptable time period so as to have an impact on the design of a system or research activity. Relative to today's standards, these were extremely simplified models because there was relatively minimal computing power. This in no way denigrates the M&S conducted during the 1940s or the century before. Indeed, one could convincingly argue that the M&S conducted before the 1960s was more creative and insightful than present day scientific computing because the modeler had to sort carefully through what was physically and mathematically important to decide what could be ignored. This took great understanding, skill, and experience regarding the physics involved in the system of interest.

One of the most stunning scientific computing articles to appear during the 1960s was "Computer Experiments in Fluid Dynamics" by Harlow and Fromm (1965). This article, probably more than any other, planted the seed that scientific computing should be thought of as the third pillar of science, along with theory and experiment. During the 1970s and 80s, many traditionalists strongly resisted this suggestion, but that resistance faded as the power of scientific computing became dominant in advancing science and engineering. It is now widely accepted that scientific computing does indeed provide the third pillar of science and engineering and that it has its own unique strengths and weaknesses.

From a historical perspective, it should be recognized that we are *only beginning* to build this third pillar. One could argue that the pillar of experiment and measurement has been built, tested, and continually refined since the beginning of the Italian Renaissance in the 1400s. One could also argue that this pillar has much earlier historical roots with the Mesopotamian, Egyptian, Babylonian, and Indus Valley civilizations. The pillar of theory, i.e., theoretical physics, has been built, tested, and refined since the late 1700s. Understanding the strengths and weaknesses of each of these pillars has not come without major controversies. For example, the importance of uncertainty estimation in experimental measurements, particularly the importance of using different measurement techniques, is well understood and documented. History has shown, even in modern times, the bitter and sometimes destructive debates that occur when there is a paradigm shift, e.g., the shift from Newtonian mechanics to relativistic mechanics. In a century or so, when present day human egos and organizational and national agendas have faded, science and engineering will admit that the pillar of scientific computing is just now beginning to be constructed. By this we mean that the weaknesses and failings of all the elements contributing to scientific computing are beginning to be better understood. More importantly, the weaknesses and failings are often simply ignored in the quest for publicity and grabbing media headlines. However, we must learn to balance this youthful enthusiasm and naiveté with the centuries of experience and errors encountered during the building of the pillars of experiment and theory.

### 1.1.2 Changing role of scientific computing in engineering

#### 1.1.2.1 Changing role of scientific computing in design, performance and safety of engineering systems

The capability and impact of scientific computing has increased at an astounding pace. For example, scientific simulations that were published in research journals in the 1990s are now given as homework problems in graduate courses. In a similar vein, what was at the competitive leading edge in scientific computing applied to engineering system design in the 1990s is now common design practice in industry. The impact of scientific computing has also increased with regard to helping designers and project managers improve their decision making, as well as in the assessment of the safety and reliability of manufactured products and public works projects. During most of this scientific computing revolution, system design and development were based primarily on testing and experience in the operating environment of the system, while scientific computing was commonly a secondary contributor in both preliminary and final design. For example, if there was some type of system failure, malfunction, or manufacturing issue that could not be solved quickly by testing, scientific computing was frequently called on for assistance and insight. Another common mode for the use of scientific computing was to reduce the number of design-then-test-then-redesign iterations that were needed for a product to perform better than competing products or to meet reliability or safety requirements. Specialized mathematical models for components or features of components were commonly constructed to better understand specific performance issues, flaws, or sensitivities of the components. For example, models were made to study the effect of joint stiffness and damping on structural response modes. Similarly, specialized mathematical models were built so that certain impractical, expensive, or restricted tests could be eliminated. Some examples were tests of high-speed entry of a space probe into the atmosphere of another planet or the structural failure of a full-scale containment vessel of a nuclear power plant.

As scientific computing steadily moves from a supporting role to a leading role in engineering system design and evaluation, new terminology has been introduced. Terminology such as *virtual prototyping* and *virtual testing* is now being used in engineering development to describe scientific computing used in the evaluation and "testing" of new components and subsystems, and even entire systems. As is common in the marketing of anything new, there is a modicum of truth to this terminology. For relatively simple components, manufacturing processes, or low-consequence systems, such as many consumer products, virtual prototyping can greatly reduce the time to market of new products. However, for complex, high-performance systems, such as gas turbine engines, commercial and military aircraft, and rocket engines, these systems continue to go through a long and careful development process based on testing, modification, and retesting. For these complex systems it would be fair to say that scientific computing plays a supporting role.

The trend toward using scientific computing more substantially in engineering systems is driven by increased competition in many markets, particularly aircraft, automobiles, propulsion systems, military systems, and systems for the exploration for oil and gas deposits. The need to decrease the time and cost of bringing products to market is intense. For example, scientific computing is relied on to reduce the high cost and time required to test components, subsystems, and complete systems. In addition, scientific computing is used in the highly industrialized nations of the world, e.g., the US, European Union, and Japan, to improve automated manufacturing processes. The industrialized nations increasingly rely on scientific computing to improve their competitiveness against nations that have much lower labor costs.

The safety aspects of products or systems also represent an important, sometimes dominant, element of both scientific computing and testing. The potential legal and liability costs of hardware failures can be staggering to a company, the environment, or the public.