

# Preface

A mathematical model constructed for a real system must be calibrated by data and its uncertainty must be assessed before used for prediction, decision making, and management purposes. After more than a half century of study, however, the construction of a reliable model for complex systems is still a challenging task.

In mathematical modeling, the “prediction problem” or the “forward problem” (inputs  $\rightarrow$  outputs) uses as model inputs a fixed model structure, known model parameters, given system controls, and other necessary information to find the system states (as model outputs). Unless all properties of the modeled system can be measured directly, model inputs tend to always contain unknowns or uncertainties that have to be determined indirectly. Model calibration (outputs  $\rightarrow$  inputs) uses the measured system states and other available information to identify or estimate the unknown model inputs. Thus, in a certain sense it is the “inverse problem” of model prediction. The history of studying model calibration is probably as long as the history of forward modeling, but the progress of study had been slow due to the very nature of inversion: identifying the causes from results is always more difficult than predicting the results on the basis of known causes. Various optimization-based data-fitting methods developed for solving the classical inverse problem in mathematics have been proven to be successful for model calibration, only if a system’s structure is simple and well-defined, and both the number and dimensions of unknown parameters are low. When these assumptions do not hold, the use of data-fitting may produce an unacceptable model.

With the advances in computing, instrumentation, and information technologies, more and more sophisticated numerical models have been developed for simulating complicated physical, chemical, and biological processes observed in environment, energy, water resources, and other scientific and engineering fields. The advent of highly sophisticated software packages has made solution to the “forward problem” much easier, but, at the same time, calibrating the resulting model becomes more difficult due to the increase of model complexity. Modelers gradually realize that: (1) the requirement of model uniqueness has to be given up; (2) the classical concept of model inversion must be extended to include model structure identification, model reduction, and model error quantification; and (3) five fundamental problems of modeling technology (i.e., the model selection problem, model calibration problem, model reliability problem, model application problem, and data collection

problem) must be considered systematically rather than separately or sequentially during the model construction stage. The model calibration problem has thus become more challenging but, at the same time, the process of solving the problem becomes more interesting and rewarding than the simple data-fitting exercise. Empowered by these methodological understandings and newly developed tools in mathematics and statistics, the research on model construction has made significant progress in recent years. As a result, existing methods have been improved and new and promising approaches have emerged.

This book provides a comprehensive introduction on all aspects of constructing useful models: from the deterministic framework to the statistical frameworks; from the classical inverse problem of parameter estimation to the extended inverse problem of system structure identification; from physical-based models to data-driven models; from model reduction to model uncertainty quantification; from data sufficiency assessment to optimal experimental design; and from basic concepts, theory, and methods to the state-of-the-art approaches developed for model construction. A central problem to be considered in this book is how to find surrogate models for predetermined model applications.

Chapter 1 is a general description of the modeling technology. Models that are often seen in environmental and water resources fields are introduced here and are used to exemplify different methods throughout the book. Based on different criteria of model calibration, three kinds of inverse problem are defined: the classical inverse problem (CIP) for parameter estimation, the extended inverse problem (EIP) for system identification, and the goal-oriented inverse problem (GIP) for model application. After reading this chapter, readers will be able to get a holistic picture of mathematical modeling, know the difficulties and problems of model construction, and learn how this book is organized.

Part I of this three-part book (Chapter 2–5) is contributed to the solution of CIP. Basic concepts and methods on linear model inversion and single-state nonlinear model inversion are given in Chapter 2; singular value decomposition and various nonlinear optimization algorithms are introduced briefly. In Chapter 3, the multi-state model inversion is cast into a multi-objective optimization problem and solved by the evolutionary algorithms. Regularization is also introduced in this chapter from the point of view of multi-criterion inversion. The inverse problem is reformulated and resolved in the statistical framework in Chapter 4. Monte Carlo based sampling methods, including the Markov Chain Monte Carlo method, are introduced for finding the posterior distribution. Various methods of model differentiation are given in Chapter 5. Model differentiation is a necessary tool for almost all topics covered in this book.

Part II of this book (Chapter 6–8) is dedicated to the solution of EIP. In Chapter 6, various methods for parameterizing deterministic functions or random fields are introduced. Principal component analysis and other linear and nonlinear dimension reduction methods, as well as their applications to inverse solution, are covered. Model structure identification and hyperparameter estimation are the main topics of Chapter 7, in which various adaptive parameterization approaches, the level set method, multiscale inversion, and geostatistical inversion are introduced. Methods for constructing data-driven models are given in Chapter 8, including linear regres-

sion and various machine learning methods such as artificial neural networks, support vector machine, and Gaussian process regression.

Part III of the book (Chapter 9–12) is contributed to the topic of model reliability. Chapter 9 introduces various data assimilation methods for inverse solution that allow us to update a model continuously to improve its reliability whenever new data become available. Methods used for uncertainty quantification, including Monte Carlo simulation, global sensitivity analysis, stochastic response surface, are introduced systematically in Chapter 10. The effects of model parameter uncertainty and model structure uncertainty on model outputs are assessed. To construct a more reliable model, more data are needed. Design of informative and cost-effective data collection strategies is the subject of Chapter 11, in which, optimal experimental design is formulated into a multi-objective optimization problem. The criteria of optimal design for linear model inversion are derived. For nonlinear model inversion, Bayesian and robust design methods, especially, the interval-identifiability-based robust design, are introduced. In Chapter 12, after the goal-oriented forward problem is described, the GIP is formulated and solved in both the deterministic and statistical frameworks. When the existing data are insufficient, a cost-effective experimental design method is given. Finally, the goal-oriented pilot-point method is described.

Preliminary mathematics required for reading this book is reviewed in details in three Appendices. To help readers better understand the text, review questions are given at the end of each chapter. All major methods introduced in this book are illustrated with numerical examples created by the authors, including the information on available toolboxes. Alex Sun authored Chapters 6, 8, 9, and 10 and edited the whole book. Other chapters are authored by Ne-Zheng Sun. This book can be used as a textbook for graduate and upper-level undergraduate students majoring in environmental engineering, hydrology, or geosciences. It also serves as an essential reference book for petroleum engineers, mining engineers, chemists, mechanical engineers, biologists, medical engineers, applied mathematicians, and others who perform mathematical modeling. Much of the research conducted by the authors over the years has been made possible by the support from U.S. National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), Department of Energy (DOE), Environmental Protection Agency (EPA), and Nuclear Regulatory Commission (NRC). We are grateful to all of our current and past collaborators for insightful discussions. Ne-Zheng Sun would like to give special thanks to Drs. Jacob Bear and William Yeh for their guidance, support, collaboration, and long-term friendship. Alex Sun would like to thank Drs. Yoram Rubin and Dongxiao Zhang, and his colleagues at the University of Texas at Austin for their advice and collaboration.

We are grateful to the editors Achi Dosanjh, Donna Chernyk, and Danielle Walker at Springer for their support and guidance in every step of the process. Finally, we would like to thank Fang and Zhenzhen for their endless love, sacrifice, and patience during this multiyear-long book project.

Santa Monica, CA  
Austin, TX

Ne-Zheng Sun  
Alex Sun

<http://www.springer.com/978-1-4939-2322-9>

Model Calibration and Parameter Estimation  
For Environmental and Water Resource Systems

Sun, N.-Z.; Sun, A.

2015, XXVIII, 621 p. 123 illus., 107 illus. in color.,  
Hardcover

ISBN: 978-1-4939-2322-9