
1

INTRODUCTION

In many fields of science and engineering, mathematical models are used to represent complex processes and results are used for system management and risk analysis. The methods commonly used to develop and apply such models often do not take full advantage of either the data available for model construction and calibration or the developed model. This book presents a set of methods and guidelines that, it is hoped, will improve how data and models are used.

This introductory chapter first describes the contributions of the book, including a description of what is on the associated web site. Sections 1.2 and 1.3 provide some context for the book by reviewing inverse modeling and considering the methods covered by the book relative to other paradigms for integrating data and models. After providing a few definitions, Chapter 1 concludes with a discussion of the expertise readers are expected to possess and some suggested readings and an overview of Chapters 2 through 15.

1.1 BOOK AND ASSOCIATED CONTRIBUTIONS: METHODS, GUIDELINES, EXERCISES, ANSWERS, SOFTWARE, AND POWERPOINT FILES

The methods presented in the book include (1) sensitivity analysis for evaluating the information content of data, (2) data assessment strategies for identifying (a) existing measurements that dominate model development and predictions

Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty. By Mary C. Hill and Claire R. Tiedeman
Published 2007 by John Wiley & Sons, Inc.

and (b) potential measurements likely to improve the reliability of predictions, (3) calibration techniques for developing models that are consistent with the data in some optimal manner, and (4) uncertainty evaluation for quantifying and communicating the potential error in simulated results (e.g., predictions) that often are used to make important societal decisions.

The fourteen guidelines presented in the book focus on practical application of the methods and are organized into four categories: (1) model development guidelines, (2) model testing guidelines, (3) potential new data guidelines, and (4) prediction uncertainty guidelines.

Most of the methods presented and referred to in the guidelines are based on linear or nonlinear regression theory. While this body of knowledge has its limits, it is very useful in many circumstances. The strengths and limitations of the methods presented are discussed throughout the book. In practice, linear and nonlinear regression are best thought of as imperfect, insightful tools. Whether regression methods prove to be beneficial in a given situation depends on how they are used. Here, the term beneficial refers to increasing the chance of achieving one or more useful models given the available data and a reasonable model development effort. The methods, guidelines, and related exercises presented in this book illustrate how to improve the chances of achieving useful models, and how to address problems that commonly are encountered along the way.

Besides the methods and guidelines, the book emphasizes the importance of how results are presented. To this end, the book can be thought of as emphasizing two criteria: valid statistical concepts and effective communication with resource managers. The most advanced, complex mathematics and statistics are worth very little if they cannot be used to address the societal needs related to the modeling objectives.

The methods and guidelines in this book have wide applicability for mathematical models of many types of systems and are presented in a general manner. The expertise of the authors is in the simulation of groundwater systems, and most of the examples are from this field. There are also some surface-water examples and a few references to other fields such as geophysics and biology. The fundamental aspects of systems most advantageously addressed by the methods and guidelines presented in this work are those typical of groundwater systems and shared by many other natural systems. Of relevance are that groundwater systems commonly involve (1) solutions in up to three spatial dimensions and time, (2) system characteristics that can vary dramatically in space and time, (3) knowledge about system variability in addition to the data used directly in regression methods, (4) available data sets that are typically sparse, and (5) nonlinearities that are often significant but not extreme.

Four important additional aspects of the book are the exercises, answers, software, and PowerPoint files available for teaching.

The exercises focus on a groundwater flow system and management problem to which students apply all the methods presented in the book. The system is simple, which allows basic principles to be clearly demonstrated, and is designed to have aspects that are directly relevant to typical systems. The exercises can be conducted

using the material provided in the book, or as hands-on computer exercises using instructions and files available on the web site http://water.usgs.gov/lookup/get?crresearch/hill_tiedeman_book.

The web site includes instructions for doing the exercises using files directly and/or using public-domain interface and visualization capabilities. It may also include instructions for using selected versions of commercial interfaces. The instructions are designed so that students can maximize the time spent understanding the ideas and the capabilities discussed in the book.

Answers to selected exercises are provided on the web site.

The software used for the exercises is freely available, open source, well documented, and widely used. The groundwater flow system is simulated using the Ground-Water Flow Process of MODFLOW-2000 (Harbaugh et al., 2000; Hill et al., 2000). The sensitivity analysis, calibration, and uncertainty aspects of the exercises can be accomplished using MODFLOW-2000's Observation, Sensitivity, and Parameter-Estimation Processes or UCODE_2005 (Poeter et al., 2005). Most of the sensitivity analysis, calibration, and uncertainty aspects of the exercises also can be conducted using PEST (Doherty, 1994, 2005). Relevant capabilities of MODFLOW-2000 and UCODE_2005 are noted as methods and guidelines are presented; relevant capabilities of PEST are noted in some cases. The public-domain programs for interface and visualization are MFI2K (Harbaugh, 2002), GWChart (Winston, 2000), and ModelViewer (Hsieh and Winston, 2002). The web sites from which these programs can be downloaded are listed with the references and on the book web site listed above.

The methods and guidelines presented in this book are broadly applicable. Throughout the book they are presented in the context of the capabilities of the computer codes mentioned above to provide concrete examples and encourage use.

PowerPoint files designed for teaching of the material in the book are provided on the web site. The authors invite those who use the PowerPoint files to share their additions and changes with others, in the same spirit with which we share these files with you.

The use of trade, firm, or product names in this book is for descriptive purposes only and does not imply endorsement by the U.S. Government.

The rest of this introductory chapter provides a brief overview of how regression methods fit into model calibration (Section 1.2), some perspective of how the ideas presented here relate to other ideas and past work (Section 1.3), some definitions (Section 1.4), a description of expertise that would assist readers and how to obtain that expertise (Section 1.5), and an overview of Chapters 2 through 15 (Section 1.6).

1.2 MODEL CALIBRATION WITH INVERSE MODELING

During calibration, model input such as system geometry and properties, initial and boundary conditions, and stresses are changed so that the model output matches related measured values. Many of the model inputs that are changed can be characterized using what are called "parameters" in this work. The measured values related

to model outputs often are called “observations” or “observed values,” which are equivalent terms and are used interchangeably in this book.

The basic steps of model calibration are shown in Figure 1.1. In the context of the entire modeling process, effectively using system information and observations to constrain the model is likely to produce a model that more accurately represents the simulated system and produces more accurate predictions, compared to a modeling procedure that uses these types of data less effectively. The ideas, methods, and guidelines presented in this book are aimed at helping to achieve more effective use of data.

The difficulties faced in simulating natural systems are demonstrated by the complex variability shown in Figure 1.2 as discussed by Zhang et al. (2006).

Four issues fundamental to model calibration are discussed in the next four sections. These include parameter definition or parameterization, which is the mechanism used to obtain a tractable and hopefully meaningful representation of

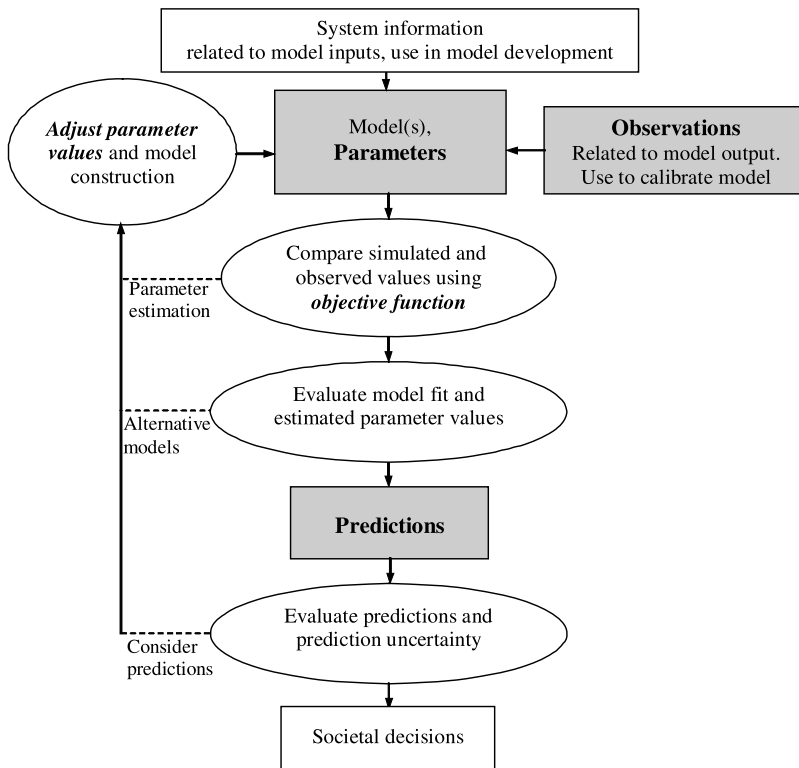


FIGURE 1.1 Flowchart showing the major steps of calibrating a model and using it to make predictions. Bold, italicized terms indicate the steps that are directly affected by nonlinear regression, including the use of an objective function to quantify the comparison between simulated and observed values. Predictions can be used during calibration as described in Chapter 8. (Adapted from Herb Buxton, U.S. Geological Survey, written communication, 1990.)



FIGURE 1.2 Experimental results from a subsiding tank, showing the kind of complexity characteristic of deltaic deposits in a subsiding basin. (Reproduced with permission from Paola et al. 2001.)

systems such as that shown in Figure 1.2; the objective function mentioned in Figure 1.1; the utility of inverse modeling, which is also called parameter estimation in this book; and using the model to quantitatively connect observations, parameters, and predictions.

1.2.1 Parameterization

The model inputs that need to be estimated are often distributed spatially and/or temporally, so that the number of parameter values could be infinite. The observations, however, generally are limited in number and support the estimation of relatively few parameters. Addressing this discrepancy is one of the greatest challenges faced by modelers in many fields. Typically, so-called parameterization is introduced that allows a limited number of parameter values to define model inputs throughout the spatial domain and time of interest. In this book, the term “parameter” is reserved for the values used to define model inputs. Consider the parameters defined in three groundwater model examples.

Example 1: One parameter represents the hydraulic conductivity of a hydrogeologic unit that occupies a prescribed volume of the model domain and is hydraulically distinctive and relatively uniform.

Example 2: One parameter represents a scalar multiplier of spatially varying recharge rates initially specified by the modeler for a given geographic area on the basis of precipitation, vegetation, elevation, and topography.

Example 3: One parameter represents the hydraulic head at a constant-head boundary that is used to simulate the water level in a lake.

This book focuses primarily on models for which a limited number of parameters are defined. Alternative methods are discussed in Section 1.3.2.

Historically, observed and simulated values, such as hydraulic heads, flows, and concentrations for groundwater systems, often were compared subjectively, so that it was difficult to determine how well one model was calibrated relative to another. In addition, in modeling of groundwater and other types of systems, adjustments of parameter values and other model characteristics were accomplished mostly by trial and error, which is time consuming, subjective, and inconclusive.

Formal methods have been developed that attempt to estimate parameter values given a mathematical model of system processes and a set of relevant observations. These are called inverse methods, and generally they are limited to the estimation of parameters as defined above. Thus, the terms “inverse modeling” and “parameter estimation” commonly are synonymous, as in this book. For some models, the inverse problem is linear, in that the observed quantities are linear functions of the parameters. In many circumstances of practical interest, however, the inverse problem is nonlinear, and its solution is not as straightforward as for linear problems. This book discusses methods for nonlinear inverse problems. One method of solving such problems is nonlinear regression, which is the primary solution method discussed in this book.

The complexity of many real systems and the scarcity of available data sets result in inversions that are often plagued by problems of insensitivity, nonuniqueness, and instability, regardless of how model calibration is achieved. Insensitivity occurs when the observations do not contain enough information to support estimation of the parameters. Nonuniqueness occurs when different combinations of parameter values match the observations equally well. Instability occurs when slight changes in, for example, parameter values or observations radically change simulated results. All these problems are exacerbated when the system is nonlinear. These problems are usually more easily detected when using formal inverse modeling and associated methods than when using trial-and-error methods for calibration. Detecting these problems is important to understanding the value of the resulting model.

1.2.2 Objective Function

In inverse modeling, the comparison of simulated and observed values is accomplished quantitatively using an objective function (Figure 1.1). The simulated and observed values include system-dependent variables (e.g., hydraulic head for the groundwater flow equation or concentration for the groundwater transport equation) and other system characteristics as represented by prior information on parameters. Parameter values that produce the “best fit” are defined as those that produce the smallest value of the objective function.

1.2.3 Utility of Inverse Modeling and Associated Methods

Recent work has clearly demonstrated that inverse modeling and associated sensitivity analysis, data needs assessment, and uncertainty evaluation methods provide

capabilities that help modelers take greater advantage of their models and data, even for simulated systems that are very complex (i.e., Poeter and Hill, 1997; Faunt et al., 2004). The benefits include

1. Clear determination of parameter values that produce the best possible fit to the available observations.
2. Graphical analyses and diagnostic statistics that quantify the quality of calibration and data shortcomings and needs, including analyses of model fit, model bias, parameter estimates, and model predictions.
3. Inferential statistics that quantify the reliability of parameter estimates and predictions.
4. Other evaluations of uncertainty, including deterministic and Monte Carlo methods.
5. Identification of issues that are easily overlooked when calibration is conducted using trial and error methods alone.

Quantifying the quality of calibration, data shortcomings and needs, and uncertainty of parameter estimates and predictions is important to model defensibility and transparency and to communicating the results of modeling studies to managers, regulators, lawyers, concerned citizens, and to the modelers themselves.

Despite its apparent utility, in many fields, such as groundwater hydrology, the methods described in this book are not routinely used, and calibration using only trial-and-error methods is more common. This, in part, is due to lack of familiarity with the methods and the perception that they require more time than trial-and-error methods. It is also because inverse modeling and related sensitivity analysis methods clearly reveal problems such as insensitivity and nonuniqueness, and thereby reveal inconvenient model weaknesses. Yet if they are revealed, such weaknesses often can be reduced or eliminated. This occurs because knowledge of the weaknesses can be used to determine data collection and model development effort needed to strengthen the model. We hope this text will encourage modelers to use, and resource managers to demand, the more transparent and defensible models that result from using the types of methods and ideas described in this book.

1.2.4 Using the Model to Quantitatively Connect Parameters, Observations, and Predictions

The model quantitatively connects the system information and the observations to the predictions and their uncertainty. The entities Parameters, Observations, and Predictions are in bold type in Figure 1.1 because these entities are directly used by or produced by the model, whereas the system information often is indirectly used to create model input. Many of the methods presented in this book take advantage of the quantitative links the model provides between what is referred to in this book as the triad of the observations, parameters, and predictions.

The depiction of model calibration shown in Figure 1.1 is unusual in that it suggests simulating predictions and prediction uncertainty as model calibration proceeds. When execution times allow, it is often useful to include predictive analyses during model calibration so that the dynamics affecting model predictions can be better understood. Care must be taken, of course, not to use such simulations to bias model predictions.

1.3 RELATION OF THIS BOOK TO OTHER IDEAS AND PREVIOUS WORKS

This section relates the ideas of this book to predictive models and other literature.

1.3.1 Predictive Versus Calibrated Models

When simulating natural systems, the objective is often to produce a model that can predict, accurately enough to be useful, for assessing the consequences of introducing something new in the system. In groundwater systems, this may entail new pumpage or transport of recently introduced or potential contamination.

Ideally, model inputs would be determined accurately and completely enough from directly related field data to produce useful model results. This is advantageous because the resulting model is likely to be able to predict results in a wide range of circumstances, and for this reason such models are called *predictive models* (e.g., see Wilcock and Iverson, 2003; National Research Council, 2002). However, commonly quantities simulated by the model can be more readily measured than model inputs. The best possible determination of model inputs based on directly related field data can produce model outputs that match the measured equivalents poorly. If the fit is poor enough that the utility of model predictions is questionable, then a decision needs to be made about how to proceed. The choices are to use the predictive model, which has been shown to perform poorly in the circumstances for which testing is possible, or to modify the model so that, at the very least, it matches the available measured equivalents of model results. A model modified in this way is called a *calibrated model*.

There is significant and important debate about the utility of predictive and calibrated models, and it is our hope that the debate will lead to better methods of measuring quantities directly related to model inputs. We would rejoice with all others in the natural sciences to be able to always use predictive models. Until then, however, it is our opinion that methods and guidelines that promote the best possible use of models and data in the development of calibrated models are critical. It is also our belief that such methods and guidelines can play a role in informing and focusing the efforts of developing field methods that may ultimately allow predictive models to be used in more circumstances.

1.3.2 Previous Work

For the most part, comments in this introductory chapter are limited to the history, evolution, and status of nonlinear regression and modeling as related to groundwater systems. Comments about how specific methods or ideas relate to previous

publications appear elsewhere in the book. This section contains the broadest discussion of parameterization methods presented in the book.

The topics covered by this book have been addressed by others using a variety of different methods, and have been developed for and applied to many different fields of science and engineering. We do not attempt to provide a full review of all work on these topics. Selected textbooks are as follows. Parker (1994), Sun (1994), Lebbe (1999), and Aster et al. (2005) discuss nonlinear regression in the field of geophysics. More general references for nonlinear regression and associated analyses include Bard (1974), Beck and Arnold (1977), Belsley et al. (1980), Seber and Wild (1989), Dennis and Schnabel (1996), and Tarantola (2005). Saltelli et al. (2000, 2004) provide comprehensive overviews of sensitivity-analysis methods. This book focuses on what Saltelli et al. describe as local sensitivity methods, and includes new sensitivity-analysis methods not included in the previous books.

The pioneers of using regression methods in groundwater modeling were Cooley (1977) and Yeh and Yoon (1981). Some of the material in this book was first published in U.S. Geological Survey reports (Cooley and Naff, 1990; Hill, 1992; Hill, 1994; Hill, 1998). Cooley and Naff (1990) presented a modified Gauss–Newton method of nonlinear regression that with some modification is used in Chapter 5, and residual analysis ideas derived from early editions of Draper and Smith (1998) that are used in Chapter 6. Hill (1992) presents sensitivity-analysis and residual-analysis methods used in Chapters 4 and 6. Cooley and Naff (1990), and Hill (1992), and Hill (1994) present methods of residual analysis and linear uncertainty analysis that are used in Chapters 6 and 8. Hill (1998) enhanced the methods presented in the previous works and presents the first version of the guidelines that are described in Chapters 10 through 14. Various aspects of the guidelines have a long history, and relevant references are cited in later chapters. To the authors' knowledge, these guidelines provide a more comprehensive foundation for the calibration and use of models of complex systems than any similar set of published guidelines. In general, the book expands the previously presented material, presents some new methods, and includes an extensive set of exercises.

Achieving Tractable Problems Regression is a powerful tool for using data to test hypothesized physical relations and to calibrate models in many fields (Seber and Wild, 1989; Draper and Smith, 1998). Despite its introduction into the groundwater literature in the 1970s (reviewed by McLaughlin and Townley, 1996), regression is only starting to be used with any regularity to develop numerical models of complicated groundwater systems. The scarcity of data, nonlinearity of the regression, and complexity of the physical systems cause substantial difficulties. Obtaining tractable models that represent the true system well enough to yield useful results is arguably the most important problem in the field. The only options are (1) improving the data, (2) ignoring the nonlinearity, and/or (3) carefully ignoring some of the system complexity. Scarcity of data is a perpetual problem not likely to be alleviated at most field sites despite recent impressive advances in geophysical data collection and analysis (e.g., Eppstein and Dougherty, 1996; Hyndman and Gorelick, 1996; Lebbe, 1999; Dam and Christensen, 2003). Methods that ignore nonlinearity are presented by, for example, Kitanidis (1997) and Sun (1994, p. 182). The large

changes in parameter values that occur in most nonlinear regressions of many problems after the first iteration, however, indicate that linearized methods are unlikely to produce satisfactory results in many circumstances. This leaves option 3, which is discussed in the following paragraphs.

Defining a tractable and useful level of parameterization for groundwater inverse problems has been an intensely sought goal, focused mostly on the representation of hydraulic conductivity or transmissivity. Suggested approaches vary considerably. The most complex parameterizations are cell- or pixel-based methods in which hydraulic conductivity or transmissivity parameters are defined for each model cell, element, or other basic model entity, and prior information or regularization is used to stabilize the solution (e.g., see Tikhonov and Arsenin, 1977; Clifton and Neuman, 1982; Backus, 1988; McLaughlin and Townley, 1996). The simplest parameterizations require homogeneity, such that, at the extreme, one parameter specifies hydraulic conductivity throughout the model.

As more parameters are defined and the information contained in the observations is overwhelmed, prior information on parameters and/or regularization on observations and/or parameters become necessary to attain a tractable problem. In this book, we use definitions of prior information and regularization derived from Backus (1988). When applied to parameters, prior information and regularization produce similar penalty-function terms in the objective function. For prior information, the weighting used approximates the reliability of the prior information based on either classical or Bayesian statistical arguments. Essentially, classical statistical arguments are based on sampling methods; Bayesian statistical arguments are, at least in part, based on belief (Bolstad, 2004). In contrast, for regularization the weighting generally is determined as required to produce a tractable problem, as represented by a unique set of estimated parameter values. The resulting weights generally are much larger than can be justified based on what could possibly be known or theorized about the parameter values and distribution. For both prior information and regularization, the values used in the penalty function need to be unbiased (see the definition in Section 1.4.2).

Between the two extreme parameterizations mentioned previously, there is a wide array of designs ranging from interpolation methods such as pilot points (RamaRoa et al., 1995; Doherty, 2003; Moore and Doherty, 2005, 2006) to zones of constant value designed using geologic information (see Chapter 15 for examples). For example, the Regularization Capability of the computer code PEST (Doherty, 1994, 2005) typically allows many parameters to be estimated. Indeed, the number of parameters may exceed the number of observations. Parameter estimation is made possible by requiring that the parameter values satisfy additional considerations. Most commonly, the parameter distribution is required to be smooth. This and other considerations are discussed by Tikhonov and Arsenin (1977) and Menke (1989). More recent approaches include the superparameters of Tonkin and Doherty (2006) and the representer method of Valstar et al. (2004). The former uses singular value decomposition to identify a few major eigenvectors from sensitivity matrices; only the “superparameters” defined by the eigenvectors are estimated by regression.

Parameterizations with many parameters are advantageous in that they minimize user-imposed simplifications, but they have the following problems: (1) they do not eliminate the scale problem if heterogeneities smaller than the grid or parameter scale are important, as they often are in transport problems, for example; (2) they generally require more and better hydraulic-conductivity or transmissivity data than are available in most circumstances or unsupportable assumptions about smoothness; and (3) they can easily lead to overfitting the observations and a resulting decline in predictive accuracy. Historically, parameterization methods that resulted in many parameters also were unable to accommodate easily knowledge about geologic structure. Gradually, the ability to apply geologic constraints within the context of many defined parameters is being developed and provides exciting possibilities.

Simpler parameterizations (simpler in that there are fewer defined parameters) can be achieved using zonation, interpolation, or eigenvectors of the variance–covariance matrix of grid-scale parameters (e.g., Jacobson, 1985; Sun and Yeh, 1985; Cooley et al., 1986; RamaRao et al., 1995; Eppstein and Dougherty, 1996; Reid, 1996; D’Agnese et al., 1999; Tonkin and Doherty, 2006). Stochastic methods (e.g., Gelhar, 1993; Kitanidis, 1995; Yeh et al., 1995; Carle et al., 1998) also generally fall into this category, although they share some of the characteristics of the grid-based methods. These simpler parameterizations produce a more tractable problem, but it is not clear what level of simplicity diminishes utility.

The principle of parsimony (Box et al., 1994; Parker, 1994) suggests that simple models should be considered, but the perception remains that many complex systems cannot be adequately represented using parsimonious models. For example, Gelhar (1993, p. 341) claims that for groundwater systems “there is no clear evidence that [nonlinear regression] methods [using simple parameterizations] actually work under field conditions.” Indeed, Beven and Binley (1992) even suggest that for some problems it may be best to abandon the concept of parameterizations simple enough to produce an optimal set of parameter values.

A concept as useful as parsimony should not be given up lightly, yet there have been few conclusive evaluations of the parameter complexity needed to produce useful results for groundwater models (Hill et al., 1998). In this book we proceed from the point of view that it is best to introduce complexity slowly and carefully, which is taken to mean increase the number of parameters slowly and carefully. One reason for this approach is that models with a few parameters can be used to learn things about a system that are true for all parameterizations but are more difficult to determine when many parameters are defined. As related to the famous quote by George E. P. Box, “All models are wrong, but some are useful,” the idea is that parsimony is likely to play an important role in achieving useful models. We suggest that simpler parameterizations are useful for many models and for the initial phases of development of all models.

Direct and Indirect Inverse Modeling In groundwater inverse modeling, methods have been classified as indirect and direct (Neuman, 1973; Yeh, 1986; Sun, 1994). This book considers indirect inverse modeling, which uses available observation data and optimization techniques to estimate model input values.

Direct inverse modeling is dramatically different: available, usually sparse observations are interpolated or extrapolated everywhere in the model domain to create “observations” throughout the system. Using these “observations,” the differential equations describing the simulated processes (such as groundwater flow or transport) are used to calculate the model input values (parameters) directly. The direct inverse modeling methods have been in existence longer than the indirect methods but have been shown consistently to be unstable in the presence of common measurement errors (Yeh, 1986). The direct methods do not use sensitivities and rarely calculate them, so these methods cannot be used to compute many of the statistics used for model evaluation that are presented in this book.

1.4 A FEW DEFINITIONS

This section defines what is meant by a linear and a nonlinear model in the context of parameter estimation. It also defines four terms that are often confusing and states how the terms are used in this book.

1.4.1 Linear and Nonlinear

As discussed in Section 1.1, this book focuses on models for which parameter estimation is nonlinear. In this context, nonlinearity results when simulated equivalents to observations are nonlinearly related to parameters. For example, consider groundwater flow.

In a confined groundwater flow system, hydraulic head is a linear function of space and time, which is why superposition can be used (Reilly et al., 1987). In contrast, for the same circumstances, head is a nonlinear function of many parameter values of interest, such as hydraulic conductivity. The simplest form of the groundwater flow equation, Darcy’s Law, can be used to demonstrate both linearity with respect to the spatial dimension and nonlinearity with respect to hydraulic conductivity. This was shown by Hill et al. (2000, pp. 16–18) and is presented here in a modified form.

Darcy’s Law relates the hydraulic head along the length of a cylinder packed with saturated porous media and flow through the cylinder. Darcy’s Law can be expressed as

$$Q = -KA \frac{\partial h}{\partial X} \quad (1.1)$$

where Q = flow produced by imposing different hydraulic heads at opposite ends of a cylinder containing homogenous, saturated, porous media [L^3/T];

K = hydraulic conductivity of the saturated porous media [L/T];

A = cross-sectional area of the cylinder [L^2];

X = distance along an axis parallel to the length of the cylinder and, therefore, parallel to the direction of flow [L];

h = hydraulic head at any distance X along the cylinder [L].

Equation (1.1) can be solved for the hydraulic head at any distance, X , to achieve

$$h = h_0 - \frac{Q}{KA}X \quad (1.2)$$

where h_0 is the hydraulic head at $X = 0$.

The derivatives $\partial h/\partial Q$ and $\partial h/\partial K$ are sensitivities in a parameter-estimation problem in which Q and K are estimated. By using partial derivative notation, the derivatives of Eq. (1.2) with respect to X , Q , and K are

$$\frac{\partial h}{\partial X} = -\frac{Q}{KA} \quad (1.3)$$

$$\frac{\partial h}{\partial Q} = -\frac{1}{KA}X \quad (1.4)$$

$$\frac{\partial h}{\partial K} = -\frac{Q}{K^2A}X \quad (1.5)$$

The hydraulic head is considered to be a linear function of X because $\partial h/\partial X$ is independent of X . Hydraulic head also is a linear function of Q , because $\partial h/\partial Q$ is independent of Q . However, hydraulic head is a nonlinear function of K because $\partial h/\partial K$ is a function of K . As in this simple example, sensitivities with respect to flows, such as Q , are nearly always functions of aquifer properties; sensitivities with respect to aquifer properties, such as K , are nearly always functions of the aquifer properties and the flows. If Q and K are both estimated, both situations make the regression nonlinear.

1.4.2 Precision, Accuracy, Reliability, and Uncertainty

The terms precision, accuracy, reliability, and uncertainty are used in this book and by many others and can cause confusion. Formal definitions of these terms as related to estimated parameters and predictions are described here using an archery analogy and by relating them to the statistical terms bias and variance or standard error of the regression. (The archery analogy was suggested by Richard L. Cooley, retired from the U.S. Geological Survey, oral communication, 1988).

Precision: In archery, a set of shots is precise if the shots fall within a narrow range, regardless of whether they are near the bull's eye. A parameter estimate or prediction is more precise if associated coefficients of variation or confidence intervals are smaller. A model fits the observations more closely if the objective function is smaller, and this may indicate a more precise model depending on the measure used (see Chapter 6). More precise estimates or predictions are said to have lower variance. A precise parameter estimate results when the observations provide abundant information about the parameter, given the model construction. A precise prediction results when the parameters important to the prediction are precisely estimated.

Accuracy: In archery, a set of shots is accurate if the shots are distributed evenly about the bull's eye, though they may fall within a large radius around the bull's eye. Accurate estimates and predictions are, on average, close to the true, unknown value, but the range of values may be large. An accurate parameter estimate results when (1) the model is accurate and (2) the observations are unbiased. The observations may or may not provide abundant information about the parameter; abundant information would result in a parameter estimate that is both accurate and precise if points 1 and 2 were satisfied. An accurate prediction results when (1) the model is accurate and (2) parameter values important to the prediction are accurate. The observations may or may not provide much information about the parameters important to predictions. Accurate estimates and predictions are sometimes referred to as unbiased; inaccurate estimates and predictions are biased.

Reliability: In archery, a set of shots is reliable if the shots are distributed in a narrow range about the bull's eye. Reliable parameter estimates and predictions are both accurate and precisely determined. Reliable parameter estimates and predictions result when (1) the model accurately represents processes of importance to the observations and the predictions, and (2) the observations contain much information relevant to the predictions, so that the parameters important to the predictions are reliably estimated. From a probabilistic perspective, reliability is often defined as 1.0 minus the probability of failure.

Uncertainty: The direct inverse of reliability, so often defined as the probability of failure.

While these terms have distinct meanings, in practice, “accurate” often is used when “precise” is more applicable. In this book, we had to choose between always using these terms as defined here, or recognizing that many readers would proceed without having these definitions firmly in mind and would possibly be confused by proper usage. In some circumstances we chose more common usage to create what we thought would be an easier learning experience.

1.5 ADVANTAGEOUS EXPERTISE AND SUGGESTED READINGS

Most of this book requires little expertise in statistics and mathematics. Familiarity with basic statistics is useful, including definitions of the following terms: samples and populations; mean, standard deviation, variance, and coefficient of variation of samples and populations; normal probability distribution; log-normal probability distribution; confidence interval; and significance level. Familiarity with simple linear regression also is helpful. Good elementary references for these topics include Benjamin and Cornell (1970), Ott (1993), Davis (2002), and Helsel and Hirsch (2002). Useful advanced texts include Cook and Weisberg (1982), Seber and Wild (1989), and Draper and Smith (1998).

To use the exercises to learn the principles of sensitivity analysis, nonlinear regression, and associated evaluation of the regression, students will benefit from understanding groundwater flow problems well enough to follow the discussions of the physical problem considered. To perform the optional simulations of the groundwater model used in many of the exercises that accompany the methods, students will benefit from familiarity with the computer program MODFLOW-2000 (McDonald and Harbaugh, 1988; Harbaugh et al., 2000; Hill et al., 2000; Anderman and Hill, 2001).

When this book is used to teach a semester- or quarter-long academic course, it may be desirable to start with two to four weeks of instruction on statistics and linear regression. Recommended topics include graphical data analysis, hypothesis testing, simple linear regression, and multiple linear regression. If, for example, Helsel and Hirsch (2002) is used, the readings and exercises in Table 1.1 address the suggested material.

If Davis (2002) is used to learn basic statistics, the topics in Table 1.2 are suggested.

TABLE 1.1 Suggested Reading Assignments and Exercises in Helsel and Hirsch (2002)

Chapter	Topic	Reading Assignment	Exercise
2	Graphical data analysis	Introduction; Section 2.1.5	None
3	Uncertainty	Sections 3.1, 3.2, 3.4	3.1 (parametric interval)
4	Hypothesis testing	Introduction; Sections 4.1, 4.2, and 4.4	4.1 (for untransformed data)
5	<i>t</i> -Tests	Introduction; Section 5.2	5.2
8	Correlation coefficients	Introduction; Sections 8.1 and 8.4	None
9	Simple linear regression	All except Section 9.6	9.1. Use data subsets to show the effect of small data sets.
11	Multiple regression	All except Section 11.8	11.1

TABLE 1.2 Suggested Reading Assignments and Exercises in Davis (2002)

Chapter	Topic	Reading Assignment
2	Summary statistics	pp. 34–39
2	Joint variation of two variables	pp. 40–46
2	Comparing normal populations	pp. 55–58
2	Testing the mean, <i>P</i> -values, significance	pp. 60–66
2	Confidence limits, <i>t</i> -distribution	pp. 66–75
4	Runs tests	pp. 185–191
4	Simple linear regression	pp. 191–204, 227–228
6	Multiple regression	pp. 462–470

1.6 OVERVIEW OF CHAPTERS 2 THROUGH 15

The primary topics of this book are (1) methods for sensitivity analysis, data assessment, model calibration, and uncertainty analysis developed on the basis of inverse modeling theory; and (2) guidelines for the effective application of these methods. The methods are presented in Chapters 3 to 9 and the guidelines are presented in Chapters 10 to 14. Field applications and tests of the methods and guidelines are presented in Chapter 15. Chapter 2 presents an overview of the exercises and the computer programs used in this work. Three appendixes go into greater depth concerning several aspects of the nonlinear regression method used and one appendix presents selected statistical tables. Chapters 2 through 15 are described in more detail in the following paragraphs.

Chapter 2 presents an overview of (1) three computer codes for inverse modeling that are used throughout the book, (2) a hypothetical groundwater management problem to which the methods are applied, and (3) exercises that use this groundwater management problem to clearly demonstrate the methods.

Chapters 3 to 5 present methods for measuring model fit, initial model sensitivity analysis, and parameter estimation. Chapter 3 discusses how observations of the simulated system are compared to equivalent simulated values using objective functions. Terms of the objective functions are defined, and least-squares objective-function surfaces are introduced. Chapter 4 discusses sensitivity analysis methods for evaluating the information that the observations provide toward estimating a set of parameters and using such an analysis to design parameterizations and decide what parameters to estimate. Several statistics are presented that are independent of model fit and thus can be applied prior to having achieved a successful inversion. These are called fit-independent statistics. Chapter 5 presents the modified Gauss–Newton gradient method for estimating parameter values that produce the best fit to the observations by minimizing the least-squares objective function.

Chapters 6 to 8 present methods for evaluating model fit, parameter estimates, data needs, and prediction sensitivity and uncertainty. Most of these methods involve calculating and evaluating diagnostic and inferential statistics and conducting graphical analyses. Chapter 6 discusses methods for evaluating model fit, including using residuals (differences between observed and simulated values) and weighted residuals to calculate statistical measures of fit, and graphs that can be used to help detect model error and assess normality of weighted residuals. Chapter 7 presents methods for evaluating estimated parameters and their uncertainty, including confidence intervals and measures of the support that the observations provide for the estimated parameter values. Methods for assessing model linearity are also discussed. Chapter 8 discusses evaluation of model predictions and their sensitivity and uncertainty, and methods for identifying data that would improve model predictions. Topics include measures for assessing the importance to predictions and to confidence intervals on predictions of observations and prior information on parameters. Monte Carlo methods of evaluating uncertainty are discussed briefly.

Chapter 9 presents methods for calibrating transient and transport models, and for recalibrating and reevaluating existing models when new data become available.

TABLE 1.3 Guidelines for Effective Model Calibration

<i>Model Development (Chapter 11)</i>
<ol style="list-style-type: none"> 1. Apply the principle of parsimony (start very simple; build complexity slowly) 2. Use a broad range of system information (soft data) to constrain the problem 3. Maintain a well-posed, comprehensive regression problem 4. Include many kinds of observations (hard data) in the regression 5. Use prior information carefully 6. Assign weights that reflect errors 7. Encourage convergence by making the model more accurate and by evaluating the observations 8. Consider alternative models
<i>Model Testing (Chapter 12)</i>
<ol style="list-style-type: none"> 9. Evaluate model fit 10. Evaluate optimized parameter values
<i>Potential New Data (Chapter 13)</i>
<ol style="list-style-type: none"> 11. Identify new data to improve simulated processes, features, and properties 12. Identify new data to improve predictions
<i>Prediction Uncertainty (Chapter 14)</i>
<ol style="list-style-type: none"> 13. Evaluate prediction uncertainty and accuracy using deterministic methods 14. Quantify prediction uncertainty using statistical methods

Exercises at the ends of Chapters 3 to 9 demonstrate the methods. Most of the exercises involve the simple hypothetical groundwater management problem mentioned in the beginning of this chapter.

Chapters 10 to 14 present fourteen guidelines that address using the methods presented in Chapters 3 to 9 to analyze, simulate, calibrate, and evaluate models of complex systems. The guidelines are grouped into four topics: (1) model development, (2) model testing, (3) potential new data, and (4) prediction uncertainty. Chapter 10 introduces the guidelines and Chapters 11 to 14 each focus on the guidelines that address one of the four topics.

Table 1.3 lists the guidelines to introduce the reader to the basic ideas they promote. For example, a fundamental aspect of the approach is to start simple and to build complexity slowly.

Chapter 15 addresses the use and testing of the methods and guidelines. First, issues of computer execution time, which are nearly always of concern when calibrating models, are discussed. Then, selected publications describing tests of the guidelines using synthetic test cases and use of the guidelines in field applications are listed. The remainder of Chapter 15 discusses a few aspects of two field cases to illustrate some of the methods and guidelines presented in the book.