Chapter 2
Principles of Environmental Modeling

Everything should be made as simple as possible, but not simpler.
Albert Einstein

We have three primary scientific tools at our disposal to evaluate transformation and transport processes in the environment or to find solutions to environmental pollution problems and make decisions based on these solutions. These are, in no particular order: (i) direct field observations; (ii) laboratory scale tests and physical modeling studies; and, (iii) mathematical modeling. We recognize that transformation and transport processes that may occur in the environment and the accurate characterization of these processes both in the physical and also the mathematical domain are extremely complex. Thus, each of these tools has its appropriate place and mutually supporting role, as well as advantages and disadvantages of its use in understanding and solving environmental pollution problems.

It is well established in the literature that field observations tend to be costly but necessary. They are commonly used after the primary symptoms of the problems emerge at a contamination site. In this sense, they are extremely useful in characterizing the extent of the environmental problem, identifying its bounds or in evaluating whether the proposed remedial strategies are contributing to the solution of the environmental problem at a specific site. Laboratory studies, on the other hand, may be only useful in understanding the basic principles governing the problem at a micro or molecular scale. Findings and knowledge gained at this scale may experience significant problems in up-scaling the results to the field-scale analysis. Nevertheless, laboratory studies are extremely useful for both solving problems and for understanding micro scale issues at various stages of environmental pollution investigations and remediation.

In this book, among other topics, we will focus our attention on the use of mathematical modeling techniques in evaluating environmental transformation and transport processes. Thus it is important that we discuss problems we may encounter during model building and application, and the expectations we may have from a modeling study in an environmental application. First we should agree that

mathematical models cannot help us in the problem recognition stage of an environmental pollution problem. However, they are very useful tools in the "gaining control" and "finding solutions" stages of our problem solution spectrum. They are cost effective and can be easily set up to test "what if" scenarios associated with a remedial application or a contamination problem. This cannot be easily studied with the other two scientific tools. The downside is the approximate nature of these tools which should always be kept in mind when their outcome is utilized. The level of contribution of each of these three tools to an analysis throughout the environmental problem solving spectrum is shown in Fig. 2.1.

Mathematical models are an abstraction of the environmental system and they are based on our understanding of the physical principles that govern the system. Since models are always going to be an abstraction of a system or a physical process, their outcome should always go through a careful and detailed interpretation stage before the results obtained from a model are determined to be representative of the behavior of the process or the system modeled (Fig. 2.2).

The purpose of mathematical model building and modeling is to simulate the behavior of the environmental system being modeled. Models are built to represent the system behavior in a controlled and cost effective computational environment. In this sense, modeling has become a common building block of most scientific applications. Using this tool we may observe, analyze, synthesize and rationalize the behavior of these systems under controlled conditions, and also we may evaluate the performance of the proposed solutions to an environmental problem. A common feature of all models is that they are all based on the "concept" of
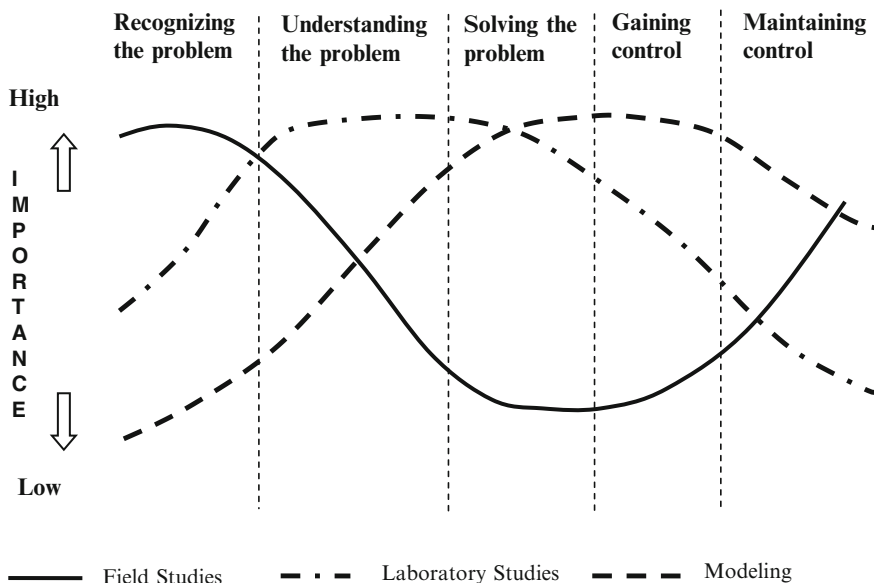

Fig. 2.1 Importance of scientific tools in assessing, evaluating and solving environmental problems
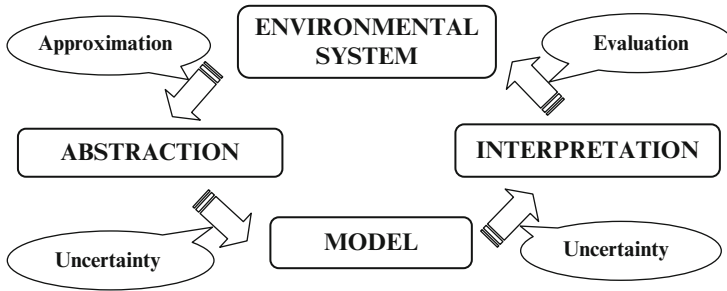
Fig. 2.2 Principles of modeling philosophy

simplification of the environmental system they are built to represent. This simplification may be achieved either through reducing the dimensionality of the system, elimination of less important processes that govern or affect the system, or through the introduction of simplified definitions for the parameters and variables that are used to describe the system. All of these or a selected subset of these simplifications are always observed in models built to represent an environmental process or an environmental system. Before we describe and make use of the models that are included in this text and also in the ACTS and RISK computational platforms, it is important that we review the modeling terminology from this perspective since it is necessary and extremely important for the reader to understand the limitations of models and modeling procedures in general. Otherwise, models or modeling may end up becoming a dangerous tool if their output is interpreted as the absolute truth without regard to the inherent simplifications and limitations they may have, or used as if they represent the environmental system under all circumstances. As a rule of thumb, modeling should always be considered to be a cost effective, efficient but approximate substitute for observing the modeled system behavior in its natural environment. Since observation of a process cannot always be achieved in a timely and cost effective manner, the models are here to stay among our scientific arsenal of tools as an important and alternative method.

The three evaluation tools identified above also differ from one another in the instruments that they may use to perform the analysis. In this sense, field study tools and laboratory tools are more closely related. Both of these methods may use electronic instrumentation to record and measure macro scale or micro scale processes. To provide a systematic procedure, these instruments may be linked to a computer or the observations can be done manually. On the other hand, computers are an essential component of all mathematical modeling studies. The language used in this analysis is primarily the language of mathematics. The interpretation of mathematics in the computer is done through coded systems, which nowadays can take the form of object or class oriented computer programming languages. As a simple definition one can say that a computer program written in any language to solve a mathematical problem is an orderly collection of coded instructions to the computer which perform certain mathematical tasks described in the mathematical model. A collection of coded programs for an application are commonly identified

as software. The ACTS and RISK computational platforms, in this sense, can be identified as software that can be used in the modeling of multimedia environmental transformation and transport problems and health risk analysis.

Finally, the analysis tools described above should always be used in coordination with one another. Field studies should support the laboratory studies or vice versa, and mathematical modeling should support both of these efforts and vice versa. The advantages of any one tool should be exploited to the utmost for the benefit of finding a satisfactory solution to the problem. The outcome of each tool should be checked and verified with the outcome of the other tool. In this sense, these tools should be viewed as complementing, rather than competing scientific methods.

2.1   Modeling Principles

Principal steps involved in modeling and the uncertainty and approximations introduced at each step are summarized in Fig. 2.2 in their simplest form. As a preliminary definition, one can say that to model is to abstract from the natural system a description which addresses a question we have posed for the system. All models are developed to answer a specific question about the system outcome. The use of models in a specific application cannot and should not go beyond the question posed during the model development stage. This is an inherent approximation and limitation that is involved in all models. After this stage several other uncertainties are introduced in model coding and analysis. Some of these uncertainties are associated with mathematical representations used in modeling and others are related to the choice of model parameter characterization during implementation. When the model is used in the simulation phase it may produce a significant amount of output. The evaluation of this output is identified as the interpretation stage. Thus the overarching goal of mathematical modeling is first to come up with an abstract representation of an environmental system and to characterize this abstraction in a mathematically consistent manner such that it yields easy to use and understandable representations of the outcome, and second to use the outcome to interpret the behavior of the modeled system within the bounds of the model. Within this sequence, approximations and uncertainties are introduced to the analysis at each stage as shown in Fig. 2.2.

A common aspect of all mathematical models is that there is an input and an output component. Outputs are tied to inputs in some mathematical sense which describes the behavior of the abstracted physical problem. Since all models are approximate representations of a natural system, they are commonly designed to accept only a subset of all possible inputs an environmental system may have. Consequently models can only generate a subset of outputs that is expected from an environmental system. In other words we can never see the complete output or picture of the modeled system. To the extend that the inputs are limited the outputs will be limited as well.

When completed, models are used in simulation. Simulations are done to provide the data necessary in decision making or in evaluating the behavior of

the system that is modeled. Decision making is based on simulation results and simulations themselves should not be interpreted as decision making. Human interaction or other heuristic mathematical models are always necessary in decision making which will be based on the outputs obtained from a model. Simulation results generated by models only provide us with the pieces of the puzzle that will help us make the appropriate decision. Evaluating the behavior of the modeled system should also be interpreted the same way. Simulation output only gives us the pieces of the puzzle needed to evaluate the system behavior.

Developing abstracted conceptual systems and a computational code for the conceptual systems is the scientific part of the modeling effort which may introduce scientific uncertainties (Lemons 1996). Simulation can be identified as the labor intensive part. Interpretation of the outcome and decision making can be considered to be the artistic part of the overall modeling effort (Fig. 2.2).

Fallout in modeling is the tendency to model in too much detail rather than modeling a finite manageable abstraction. The key to avoid this pitfall is to model around a question that needs to be answered rather than shooting for a universal representation. A simple model can always be fine tuned (calibrated) to overcome the approximations introduced through simplification. As a rule of thumb the following are key elements of a successful modeling effort:

    i. Understand the problem and clearly state the question that needs to be addressed.
    ii. Evaluate existing models first, do not re-invent the wheel.
    iii. Create a conceptual model that is logical and represents the conceptual model in consistent mathematical terms.
    iv. In developing the model involve the user or think like a user.
    v. Simplify the conceptual model, its mathematical interpretation and its user interface. This may lead to a trial and error process. Don't be shy of remodeling.
    vi. When complete make sure that the model satisfies the objective and mission of the effort (see item 1).
    vii. Design the simulations such that they provide answers to the question posed. Do not expect answers beyond the questions posed.
    viii. Always remember that the purpose of modeling is the knowledge gained from a model and not the models themselves.

2.2   Model Building and Model Types

In model building the starting point should always be the identification of the goals of the modeling study. In this context, the following alternative goals can be cited:

    i. The modeling study is going to be a scientific study in which different hypotheses regarding the governing principles of the study will be tested,

     dominant processes of the problem will be identified, bounds of the parame-
     ter ranges that define these processes will be quantified.
  ii. The modeling study will be used to characterize a study area, i.e. to deter-
     mine the site specific parameters that are associated with the processes
     included in the model.
 iii. The model will be based on well established basic principles and will be
     used as a predictor either to reconstruct a past event or simulate the future
     behavior of an environmental process at a site.
 iv. The model will be used as an imbedded predictor (slave application) within a
     master application and will be used repeatedly to supply data to the master
     application. Simulators used in optimization models or statistical applica-
     tions (Monte Carlo analysis) fall into this category and may include the goals
     identified in item 3.
  v. The model will be used to support engineered decisions that will be made at
     a site and the purpose of modeling is the evaluation of the performance of
     these decisions.

Given the list of goals stated above, we should expect the following character-
istics to be the dominant features of the model built. In case 1 the model should be
considered to be modular. The construction and solution method of the model
should allow for inclusion or exclusion of certain sub-processes to the model with
relative ease. Complexity of the model is of no concern in these applications. The
purpose is to include all possible and important sub-processes into the model. In
case 2 the model will be used in the inverse modeling sense. In these applications,
independent parameters of the model are treated as unknowns and dependent
variables are treated as known variables and the solution process is based on the
intrinsic relation between the independent and dependent variables. These models
are not expected to include many independent parameters; otherwise, the solution
becomes impossible. These models rely heavily on accurate field data on dependent
variables. In case 3 the model will be used as a predictor. In this case the model
should include all the dominant sub-processes of the problem studied, independent
of the availability of accurate definitions of the parameters that are necessary to
define these sub-processes. During simulation these parameters will be varied
anyway, and the model output sensitivity with respect to these parameters will be
documented. In case 4 the model should yield results efficiently with minimal
computation time. For this to happen one may either resort to closed form solutions
(analytical) or simplified models that may not include complex sub-processes
which may exist in the overall system. In this case, as another simplification
alternative, one may choose to represent complex processes in their simplest
approximate forms. For example, in contaminant transformation and transport
analysis one may either choose not to include chemical reactions, that is only
simulate transport of a conservative chemical behavior, or represent this chemical
reaction as a first order reaction for a single species application. These are all
acceptable simplifications for a class of applications. For case 5 the model will be
used to test the "what if" scenarios with respect to an environmental decision that

will be made at a site. In this sense, the model should definitely include the best and most accurate definition of the sub-process that is being evaluated at the site. Secondary sub-processes that may not influence the main process may be given lesser importance in the construction of the model. In all of these cases the dimensionality of the model is determined based on the available data and the complexity desired by the model builder. Whatever the goal of the modeling study is, one always has to recognize that the tool at hand is an approximate representation of the process that is being modeled.

From the perspective of inclusion of some mathematical reasoning into the analysis of system behavior, as a general rule, the three procedures discussed above are available: (i) physical modeling (laboratory); (ii) empirical modeling (laboratory and field scale); and, (iii) computational modeling. In physical modeling the natural system being modeled is duplicated by a scaled model which is geometrically and dynamically similar to the large scale system. In this case the mathematical processes are used to arrive at similarity laws that are based on the similarity of the force ratios which govern the behavior of the natural system. Observations are conducted on the scaled model and the results are projected to the large scale system, again using the same similarity laws. Mathematical reasoning behind empirical models is based on induction supported by the data collected in field or laboratory studies. In a sense, the empirical approach represents our declaration that the system modeled is very complex, or not fully understood, and that the only alternative left for us is to represent the system by the use of a black box approach. In some cases, the empirical equations that are developed may even end up being dimensionally non-homogeneous, such as the case of the well known Manning's equation in open channel flow analysis. This is a further an indication that the natural process modeled is not well understood. Sometimes modelers get around the issue of dimensional non-homogeneity by attributing dimensions to the proportionality constants that are used in the empirical model. This of course may lead to a dimensionally homogeneous equation but does not resolve the issue of how well we understand the process that is modeled. Some of these models are so well established in the technical literature that we do not question their validity, such as the Manning's equation used in open channel flow analysis, which is sometimes inhibiting. In other cases statistical methods are used to verify the predictions made from these models.

Finally, computational models (mechanistic modeling) are based on deductive reasoning. Derivation of these models is tied to fundamental principles that govern the system. In these models, more often than not, it is impossible to include all sub-processes affecting the behavior of a complex system. Thus, as stated earlier, these models commonly include simplifying assumptions which should be accounted for when they are put to use. In this sense, although these models are generic models, i.e. can be used in any large or small scale modeling study, we use calibration methods to overcome this deficiency and adjust the model response to a site or an application to represent a specific behavior. A classification of mathematical models is given in Fig. 2.3.
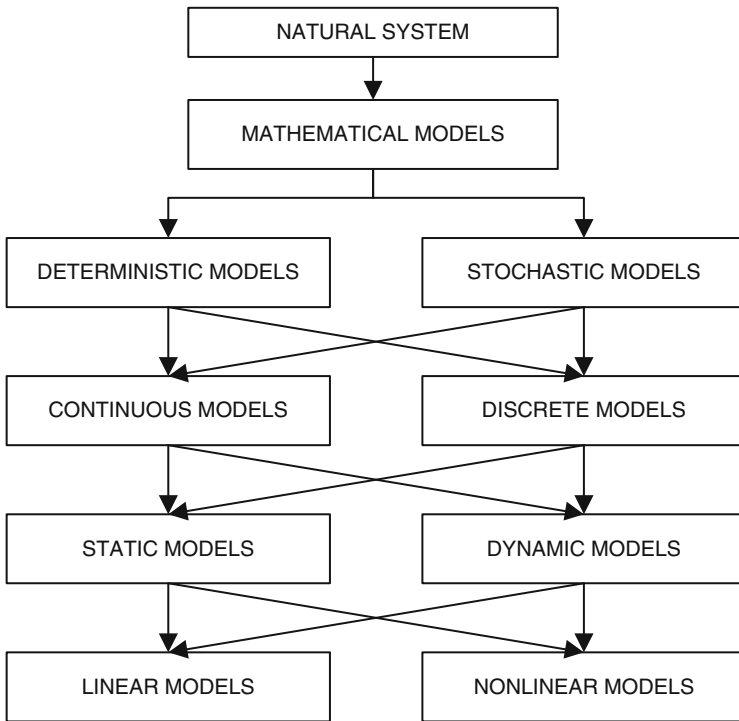
Fig. 2.3 Classification of mathematical models

The distinction in this classification is that deterministic models always produce the same output for a given input. On the other hand stochastic, a word of Greek origin which is synonymous with "randomness" and means "pertaining to chance," describes models in which a random set of inputs producing set of outputs that are interpreted statistically. Thus, stochastic is often used as the counterpart of the modeling exercise which is "deterministic," which means that random phenomena are not involved. Continuous models are based on the general mathematical property obeyed by mathematical objects and imply expressions in which all elements of the objects are within a neighborhood of nearby points. The continuity principle applies to dependent as well as independent variables of a mathematical model and implies smoothly varying properties, i.e. at least continuous first derivatives. Their counterpart is discrete models in which mathematical objects are not continuous and abrupt variation of parameters is expected. Static and dynamic refer to the dependence of the model on the independent variable "time". Static models are time independent and dynamic models are time dependent. Mathematical models that satisfy both the principles of additivity and homogeneity are considered to be linear models. These two rules, the additivity and homogeneity – taken together, lead to the possibility of the use of the principle of superposition. Nonlinear models are mathematical systems in which the behavior of the system is not expressible as

a linear operation of its descriptors. Nonlinear models may exhibit behavior and results which are extremely hard (or impossible) to predict under current knowledge or technology.

Mathematical model building is a complex process. However, a systematic path to successful model building can be defined and this path should be followed to avoid common mistakes that may render the overall effort fruitless. Following the commonly accepted principles, a model building path is given in Fig. 2.4. The modeling framework, as identified in Fig. 2.4, includes standard checks and balances that should be used in model building, no matter what the purpose of the model may be. Remodeling is always an integral path of this process to improve on what is being built.

2.3 Model Calibration, Validation, Verification and Sensitivity Analysis

Since all models are simplifications of a complex system they need to be calibrated and verified before they are used in simulation. Validation and sensitivity analysis of models is also another concept that needs to be addressed and clarified. The literature on the definition and use of these concepts is abundant and sometimes confusing. Most of the confusion is associated with the concept of validation of models (Gentil and Blake 1981; Tsang 1991; Mayer and Butler 1993; Power 1993; Oreskes et al. 1994a, b; Rykiel 1996). For example validation is sometimes considered essential (Power 1993) and sometimes validation of models is considered impossible (Starfield and Bleloch 1986; Oreskes et al. 1994a, b), and some technicians of this field indicate that models can only be invalidated (Holling 1978; McCarl 1984). Due to this confusion and conflicting definitions it is appropriate to review the meaning of these terms as well as the interpretation of the very important terms "calibration" and "sensitivity analysis" from a mathematical modeling perspective.

Model Calibration: Models include parameters and constants that need to be associated with values. These parameters are used as input to the mathematical models to produce numerical output. Ideally, these parameters should have a good definition and a physical basis for the environmental system studied. Usually these parameters either are calculated using the mathematical representation of this physical basis, or they are measured in field or laboratory studies. More often than not, however, the values of these parameters are unknown or only known approximately. Thus a range of these parameters can be input to a model to yield the best outcome when compared to an observation made in a field or laboratory study. Thus, appropriate values of the parameters are needed in the model to achieve the appropriate output that is observed at a site. Calibration of a model can then be identified as the stage where we adjust the parameters of the mathematical model such that the model agreement is maximized with respect to the observation data we have on the modeled system. In this sense, model calibration is fine tuning the
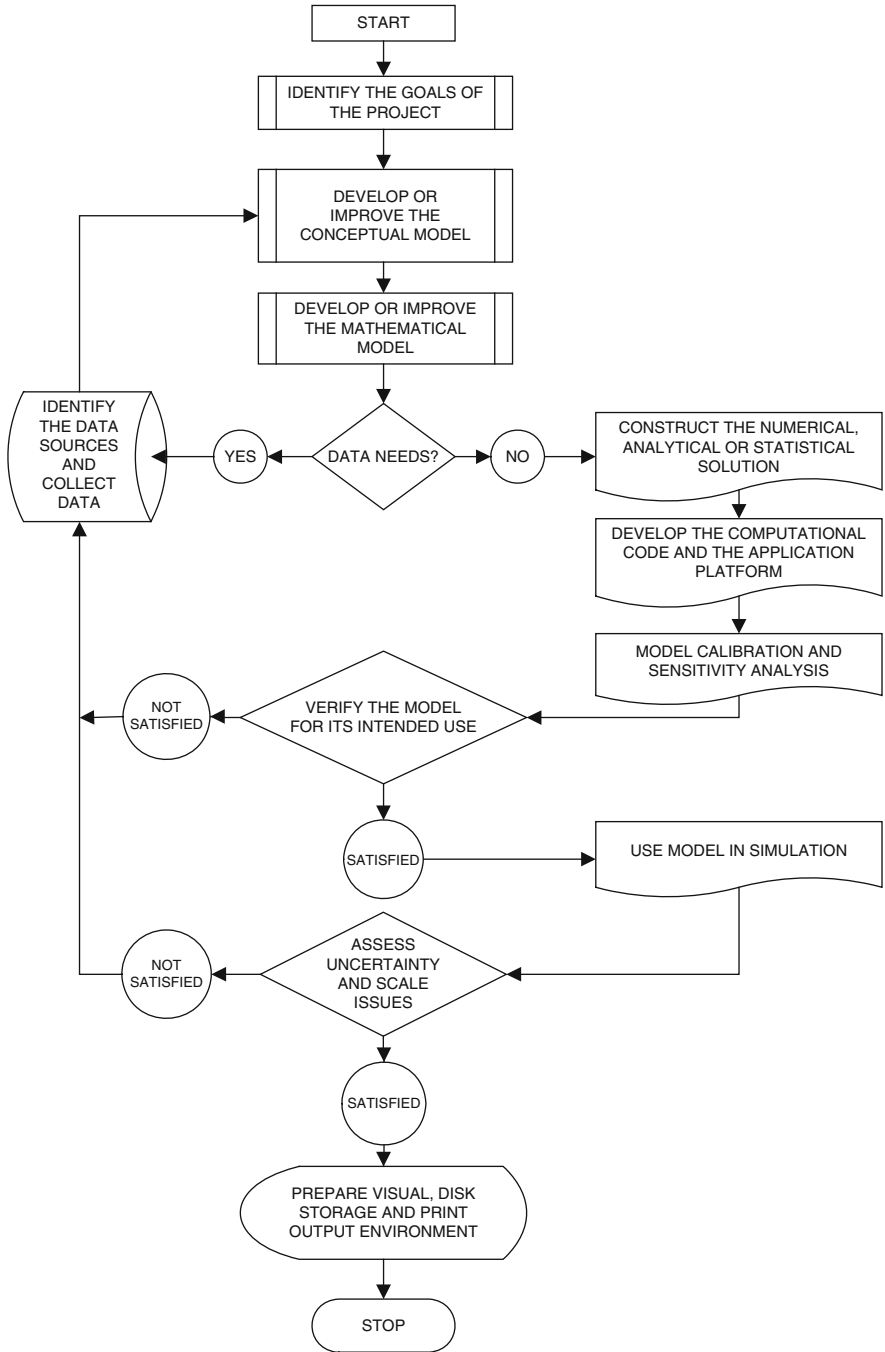
Fig. 2.4 Model building framework

model to a set of data on the natural system. Calibration of a model can be done manually, i.e. by trial and error adjustment of model parameters or it can be automated using stochastic procedures. Success in calibration, or lack of it, may yield information on how reasonable the modeler was in conceptualizing the natural system and mathematical representation of the conceptualized system. If a model fails to calibrate, it may mean that the conceptualization and mathematical representation stages need to be revisited. This also emphasizes the importance of remodeling in model development (Fig. 2.4). Calibration should not be interpreted as an inverse modeling technique which is used in parameter identification problems. Calibration procedure basically readies a model for its further use in simulation.

Model Verification: The confusion pointed out earlier may originate from the way we use the words 'verify' and 'validate'. In ordinary language, they are synonymous. From the perspective of modeling terminology these two words are used to describe two distinct concepts. Verification is a demonstration that the modeling formalism is correct. There are two types of verification avenues in modeling: (i) mechanical; and, (ii) logical. The former is associated with the debugging process of a computer program and in mathematical models, which shows that the mathematics and their numeric calculations are mechanically correct. A more important and difficult verification issue is the latter: showing that the program logic is correct. Some logical errors in a model may only appear under special circumstances that may not routinely occur in an application. Thus, these errors may not be recognized in routine applications of the model. Verification is thus a technical matter that identifies how faithfully and accurately ideas are translated into a computer code or mathematical formalisms (Law and Kelton 1991). In the case of large (complex) models, it is extremely difficult to verify that the model is entirely accurate and error free under all circumstances. Models are thus generally verified for the normal circumstances in which they are expected to be applied, and such verification is presumed inapplicable if the model is run outside this range. It is important to distinguish verification logic which relates to program operation from conceptual model logic which refers to the ecological logic used in structuring the model. Verification of models is needed in both aspects.

In summary, verification of a model is the stage at which we quantify the predictive capability of a mathematical model. This may be accomplished through a comparison of the output obtained from a model, which is based on input data, or with a set of observation data we have on a natural system which is based on the same input data. It is important to note that the observation data used in the calibration stage should be distinctly different from the data set used in the verification stage. That is, the data used for verification should be such that the calibration parameters should be fully independent of the verification data. The verified model can then be used for forecasting.

Model Validation: The absolute validity of a model can never be determined (NRC 1990). This statement is a strong reference to the impossibility of validation of a model. This reference to the impossibility of validation of models is somewhat

relaxed in a statement in which Hoover and Perry state that: "The computer model is verified by showing that the computer program is a correct implementation of the logic of the model. Verifying the computer model is quite different from showing that the computer model is a valid representation of the real system and that verified model does not guarantee a valid model" (Hoover and Perry 1989), which implies that "validity" of a model is a possibility. To clear this confusion we need to expand on these definitions.

The term model uncertainty which is linked to model validation is used to represent lack of confidence that the mathematical model is a "correct" formulation of the problem solved. Model uncertainty exists if the model produces an incorrect result even if we input the exact values for all of the model parameters. The best method for assessing model uncertainties is through model validation (Hoffman and Hammonds 1994), a process in which the model predictions are compared to numerous independent data sets obtained. Thus, as is the case with verification, validation is better understood as a process that results in an explicit statement about the behavior of a model. A common definition of validation can be the demonstration that a model, within its domain of applicability, possesses satisfactory accuracy consistent with the intended application of the model (Sargent 1984; Curry et al. 1989). This demonstration indicates that the model is acceptable for use. But that does not imply that it represents the absolute truth for the system modeled, nor even that it is the best model available. For operational validation, this demonstration involves a comparison of simulated data with data obtained by observation and measurement of the real system. Such a test cannot demonstrate the logical validity of the model's scientific content (Oreskes et al. 1994b). Validation only demonstrates that a model meets some specified performance standard under specified conditions. It is often overlooked that the "specified conditions" include all implicit and explicit assumptions about the real system the model represents as well as the environmental context it covers. That is, that a model is declared validated only within a specific context, is an integral part of the certification. If the context changes, the model must be re-validated; however, that does not invalidate the model for the context in which it was originally validated (Rykiel 1996). Validation is a "yes" or "no" proposition in the sense that a model does or does not meet the specified validation criteria. These criteria may include requirements for statistical properties (goodness-of-fit) of the data generated by the model, and thus are not necessarily deterministic. Ambiguous situations may develop when the model meets some but not all of the criteria. The criteria may need to be prioritized, and the model may be validated with respect to these priorities. Because modeling is an iterative process, validation criteria may evolve along with the model. This is more typically the case with scientific research models than with engineering models. From a technical perspective, a valid model is the one whose scientific or conceptual content is acceptable for its purpose.

Sensitivity Analysis: Sensitivity analysis, on the other hand, can be considered to be a component of simulation through which the modeler evaluates the response of the model to changes in input parameters or boundary conditions of the model.

Sensitivity of model response to the input data and parameters of the model and the model output obtained is critical and must be quantified both during calibration and verification stages. Through this process, discrepancies between the model output and observation must be minimized to the extent that is possible by identifying and minimizing sources of error. These error sources include measurement errors, conceptual error in model development and approximation errors that may exist in mathematical representations. The goal of sensitivity analysis is to estimate the rate of change in the output of a model with respect to changes in model inputs or parameters. This knowledge is important for:

  i. Evaluating the applicability range of the model developed;
 ii. Determining parameters for which it is important to have more accurate values; and,
iii. Understanding the behavior of the system being modeled at critical points of solution – possibly at singular points.

The choice of the method of sensitivity analysis depends on:

  i. The sensitivity measure employed;
 ii. The desired accuracy in the estimates of the sensitivity measure; and
iii. The computational cost involved in calculating the error.

Consider a contaminant transport model in which several parameters P_i characterize the contaminant concentration C as a continuous function in a linear mathematical function, $C = f(P_1, P_2, P_3, ..., P_n)$ from which some reference value of C can be calculated, $C_o = f(P_1^o, P_2^o, P_3^o, ..., P_n^o)$. For this case some of the more common sensitivity measures S_{ij}, which can be used, are:

Local gradient measure:

$$S_{ij} = \frac{\partial C_i}{\partial P_j^i}$$

Normalized gradient measure:

$$S_{ij} = \frac{\partial C_i}{\partial P_j^i} \frac{P_j^i}{C_i}$$

Normalized variance measure:

$$S_{ij} = \frac{\partial C_i}{\partial P_j^i} \frac{std\{P_j\}}{std\{C_i\}}$$

$$(2.1)$$

Expected value measure:

$$S_{ij} = C_i[\mathrm{E}(P_i)]$$

Extreme value measure:

$$S_{ij} = \left\{ \max C_i\left(P_j^i\right), \min C_i\left(P_j^i\right) \right\}$$

Normalized response measure:

$$S_{ij} = \left(C_o - C_i\left(P_j^i\right) \right) \Big/ C_i(P)$$

Average response measure:

$$S_{ij} = \sum_j C_i\left(P_j^i\right) \Big/ \sum_j P_j^i$$

where $\mathbf{E}$ is the expected value measure and the expected value of P_i is the mean value of parameters P_i.

Based on the choice of the sensitivity measure and the variation in the model parameters, methods of sensitivity analysis can be broadly classified into one of the following categories:

i. Variation in parameters or model formulation: In this approach, the model is run for a set of sample points (different combinations of parameters of concern) or with straightforward changes in model structure (e.g., in model resolution). Sensitivity measures that are appropriate for this type of analysis include the response from arbitrary parameter variation, normalized response and extreme value measure. Of these measures, the extreme values are often of critical importance in environmental applications.

ii. Sensitivity analysis over the solution domain: In this case the sensitivity involves the study of the system behavior over the entire range of parameter variation, often taking the uncertainty in the parameter estimates into account.

iii. Local sensitivity analysis: In this case the model sensitivity to input and parameter variation in the vicinity of a sample point(s) is evaluated. This sensitivity is often characterized through gradient measures.

The discussion of the terms calibration, verification, validation and sensitivity analysis given above outlines the basic principles involved in any modeling and model development effort. There are numerous models that are available in the scientific literature which may be used to analyze a multitude of physical processes. These models are sometimes identified as off-the-shelf models from which the users may download a code and implement it in a specific application that is of interest to the user. Here, it is important to note that the user must be fully aware of the limitations and the application range of the model used for the intended purpose. In certain cases some of these models have become so common in the literature that we no longer truly check the application rage of the model downloaded and we do not verify if the model truly fits the physical problem being modeled. In certain cases there are model applications in which the physical system modeled is restricted just to fit the system into a readily available off-the-shelf model. This practice can be characterized as fitting a physical system to a model rather than fitting a model to a physical system. This approach in modeling should be avoided at all times, at all cost. One should never try to define a physical system based on the limitations of the model that may be readily available. One should always remember the hierarchical steps involved in modeling. The description of the physical system always comes first, while the development of the model to describe the system follows behind.

2.4 Model Scales, Error and Uncertainty

The term "scale" refers to the characteristic spatial or temporal dimensions at which entities, patterns, and processes can be observed and characterized to capture the important features of an environmental process. Borrowing from cartography

concepts, as environmental modelers we define scale as having two components: grain and extent. The former corresponds to the smallest spatial and temporal sampling units used to gather a series of observations or perform a computation. Extent is the total area or time frame over which observations or computations related to a particular grain are made (O'Neill and King 1998). For example, this may be defined for an observation of a hydrologic process, or it may be defined for a modeled environment (Klemes 1983; Bloschl and Sivapalan 1995; Singh 1995). All environmental processes, large-scale or small-scale, have their own characteristic scales of reference, which are necessary to capture details of the processes modeled or observed. Independent of the size of the model used, all environmental models, as covered in this book, are based on some mathematical representation of a physical process which is scale dependent (Gupta et al. 1986). When analysts use large-scale models to predict small-scale events, or when small-scale models are used to predict large-scale events, problems may arise (Fig. 2.1).

From groundwater flow and contaminant transport models to flow and transport in river channel networks to overland flow in a watershed or air shed models, the environmental processes occur over a wide range of scales and may span about ten orders of magnitude in space and time. When we attempt to model an integrated system the first question one should ask is: "if it is necessary to link all components of the environmental cycle into one system model?" The answer to that question should not be based on whether these components are separable or not. In a global sense they are not. However, the answer to that question should be based on whether one wants to separate them or not depending on the goals of the project and the importance of the contribution of the sub-processes to the understanding and evaluation of that goal. For example, if one is not interested in observing or reflecting the effect of one subcomponent on the other, then one can easily isolate an environmental process and analyze that subcomponent alone. For example, there are numerous groundwater flow and contaminant transport models which are extensively used in the literature just to study groundwater systems (McDonald and Harbaugh 1988; Aral 1990a, b). In their analysis, groundwater would receive input from surface water, but the reverse influence cannot be considered. On the other hand, if the simulation of multipathway interaction of an environmental process is the goal, than an integrated systems modeling approach is a must, and therein one encounters the difficulties of integration over scales (Gunduz and Aral 2005).

The transfer of data or information across scales, or linking sub-process models through a unified scale, is referred to in the literature as "scaling." Up-scaling consists of taking information from smaller scales to derive processes at larger scales, while downscaling consists of decomposing information at one scale into its constituents at smaller scales (Jarvis 1995). In the context of absolute space and time, scaling primarily involves a change in the geometric and temporal structure of the data and their corresponding attributes. In using the term "absolute scale" here we are referring to the definitions used in an Eulerian coordinate system in which distances between points in time and space are well defined geometric and differential entities. Thus, linking sub-process parameters within the well defined rules can be considered to be objective and to be independent of one's viewpoint or

frame of reference in solving a problem. From a relative perspective, scaling becomes a more complex task than it would be in an absolute framework. In a relative scale framework one focuses on the sub-environmental processes and defines space and time as a measure of the relationship between these sub-processes. In a way one can interpret this definition as a Lagrangian frame of reference.

The relative scales concept represents the transcending concepts that link processes at different levels of space and time. It entails a change in scale that identifies major factors operational on a given scale of observation, their congruency with those on lower and higher scales, and the constraints and feedbacks on those factors (Caldwell et al. 1993). With this definition, one can observe that two processes that occur in close proximity by the definition of an absolute scale may be very distant from one another in terms of a relative scale sense. An example could be the case of the two hydrologic processes, overland flow and saturated groundwater flow, that normally are separated by an unsaturated zone. These two hydrologic processes could be close to each other in an absolute sense, but in terms of their interaction with one another, they could be very distant in a relative space and time frame of reference, due to limiting transfer rates that may exist in the unsaturated zone. In such cases, when scaling is considered the relative frame of reference should take precedence.

As expressed by Jarvis (1995), what makes scaling a real challenge is the nonlinearity between processes and variables scaled, and the heterogeneity in the properties that determine the rates of processes in a relative frame of reference. Therefore, it is important to realize that scaling requires an understanding of the complex hierarchical organization of the geographic and temporal worlds in which different patterns and processes are linked to specific scales of observation, and in which transitions across scales are based on geographically and temporarily meaningful rules (Marceau 1999).

Scaling and its effects on environmental modeling are commonly linked to the heterogeneity of the system modeled. However, this link should also include the refinement necessary to resolve the mathematical nonlinearities incorporated into an environmental process. Scale differences necessary to resolve nonlinearities, such as the nonlinearities introduced by the dependence of the higher order chemical reaction terms on rate constants as opposed to the easily solved differential equation that accompanies the first order reaction rates can be given as an example. Thus nonlinearity and heterogeneity are the two important factors that need to be considered in scaling. The greater the degree of heterogeneity and nonlinearity, the smaller the scale one would have to use to represent such variability or resolve such nonlinearity.

The other component of scaling effect arises in the interpretation of field data. Integrated environmental models use a variety of parameters to represent the characteristics of an application domain. However, data on large scale domain parameters are often limited. The task is then to transform this spatially limited data to a scale which can be used as an input in large scale applications. The question to answer here is what scale one should use to represent this data without losing accuracy during the extrapolation process. As the spatial scale of the model increases from a small area to a large area, the extrapolation of limited spatial data

to a large scale system would introduce errors in the analysis from the start, which should be avoided.

An optimum scale of an integrated model should then reflect the "functional scale" (Aral and Gunduz 2003), that provides a compromise between the resolution of nonlinearities of the mathematical model, availability and extrapolation of data and the heterogeneity of the system. Thus, in environmental modeling, in order to resolve scale and scaling problems, one should first attempt to answer the following fundamental questions:

    i. What is the appropriate scale of study for a particular hydrologic sub-process in the study?
    ii. How close these sub-processes are in a relative frame of reference?
    iii. How can one accurately transfer the necessary information from one process scale to another for closure?

When answering these questions we end up with a so called compromised scale which we identify as the functional scale (Aral and Gunduz 2003).

Scales of Sub-processes: Different scales of space and time govern the flow and transport phenomena in the environmental cycle. For an integrated environmental model these scales vary by several orders of magnitude in terms of the idealization of the solution domain, the computational step size and the simulation extent that is necessary to capture the important aspects of the process modeled as well as the proper scales that are necessary to interpret the input data.

One important aspect of integrating various sub-processes is the selection of the method applied to solve the equations that define the system. In this regard, coupling via iterative solution and coupling via simultaneous solution are the most advanced levels of solving the sub-processes in an integrated fashion. In iterative solutions, each sub-process model is solved separately and integrated sequentially by using the contributions from the other sub-processes. When each sub-model is solved, the common parameters linking these systems are checked for convergence (i.e., deviation from the previous solution). If the solutions of these common parameters are not sufficiently close, the solution procedure is repeated until the differences between subsequent solutions are below a pre-determined convergence criteria. This iterative coupling approach is slow, especially when more than two sub-processes are linked together. On the other hand this approach would be less restrictive from the perspective of scaling concerns since each sub-process can be analyzed within its own scale.

In the simultaneous solution approach, all sub-process models are solved together using a common idealization scale and a common time step. In this approach all sub-model solution matrices are grouped in a single matrix structure and solved at once. Hence, this method requires the use of the smallest idealizations and smallest time step of all sub-models, which may be impractical for the coupling processes requiring idealization and time steps from the two extremes. For example linking the two processes such of saturated groundwater flow and transport and the unsaturated groundwater flow and transport falls into category. Attempting to solve such a system simultaneously results in small idealization scales and time steps and

creates incompatibility between systems. For example, unsaturated flow requires small time steps in the order of seconds to describe the vertical movement of moisture in the unsaturated domain whereas the groundwater flow can be run with time steps in the order of days. If a simultaneous solution technique is used to couple these two systems, then the entire system would need to be run with the time step of the unsaturated zone. This condition is computationally costly and inefficient for the groundwater flow and contaminant transport simulations. On the other hand, this approach is more accurate than the iterative method since it does not involve improvement of the solution by iterating on the common parameters of the two sub-models (Gunduz and Aral 2003a, b, c, d). Thus the wide array of time scales required to simulate efficiently the flow and transport processes in the environment is the most important problem of environmental modeling. The incompatibility of the sub-process time scales makes the overall coupling of the system difficult and sometimes impractical.

Suggested Solutions to Scaling Problems in Integrated Environmental Modeling: In large scale environmental modeling, the scale issues and up-scaling or down-scaling difficulties outlined above must be resolved if we are to develop an integrated representation of these processes. Technicians in the field of modeling believe that these problems can be resolved through some compromises. In order to develop an order of importance list of compromises that can be considered, the modeler has to introduce concepts such as:

   i. Order of importance;
  ii. Domain of importance;
 iii. Functional scales; and,
 iv. Hybrid modeling concepts.

In an integrated modeling effort, the order of importance ranking of different sub-processes can be achieved by the analysis of the data associated with the environment under study. For example in an environment where the groundwater table is high and the unsaturated flow zone thickness is very small, it may not be a significant loss of accuracy if the unsaturated zone is not modeled as a distributed model but instead is represented in terms of lump parameter models. Similar order of importance analysis evaluation can be made for overland flow as well as for the contaminant transport modeling. In arid regions or for rainfall events which are not significant, the contribution of this component may also be represented in terms of lumped parameter models rather than distributed parameter models. However, in all cases the groundwater flow zone and the river channel flow zone will play an important role in the overall watershed hydrology and should be included in the analysis in terms of distributed models for improved accuracy of representation of these sub-processes in the integrated environmental model.

The domain of importance concept arises from the analysis of the type of the problem solved. For example, if the concern is the transport of a certain contaminant source in the watershed, and if this source is not located in the unsaturated zone, then modeling the hydrologic processes in the unsaturated zone in detail with the use of distributed models may not be necessary. Similarly, if it is known that the

flux of water between the unsaturated and the saturated zones is negligible, there is no need to complicate the analysis by including the unsaturated zone. On the contrary, there may not be any need to model the saturated groundwater flow when the top few meters of the soil column are of concern to the modeler and the groundwater table is at a much deeper elevation. Such simplifying judgments are a direct consequence of the available data for the domain modeled and are essential components of engineering evaluations to be made in a modeling study.

The functional scales concept is associated with the limitations of the integrated domain scales. If all sub-processes are important in an integrated environmental modeling effort and the use of distributed models is the goal, then one has to analyze the final time and space scales that are necessary to combine these models in an integrated system. At that point one may clearly see that this is not possible given the computational difficulties or long computation times required to solve the system. In such cases a compromise, as described earlier, is again the only solution.

Data availability is another limiting aspect of the integrated large scale environmental modeling studies. More often than not, field data is not available to justify the use of a distributed model at a large scale. This may be observed at a sub-process scale, in which case there is no reason to force a distributed model application for that sub-process as well. Otherwise, unforeseen errors will be introduced to the modeling effort. The availability of the alternative models, which range from simplified to more detailed system representations, or from small scale to large scale models, aids in evaluating the applicability of the low resolution models. If the results of the low resolution models (either in detail or in scale complexity) agree closely with those of the high resolution models, then the low resolution models are preferable, since they typically require lower computational resources and lesser input data.

Given the limitations on computational resources, computational methods and data limitations, the outcome of the integrated modeling compromises, as discussed above, is clearly to direct the modeler towards the use of hybrid models in integrated environmental modeling. In these models, lumped parameter models are used along with distributed parameter models to develop an integrated system.

Uncertainty and Error: The discussion above leads to uncertainty and error associated with environmental models and modeling (Figs. 2.2 and 2.5). Uncertainty in transformation and transport models arises in the following two stages of modeling: (i) model conceptualization or model building; and, (ii) model application. As mentioned above, model building uncertainty arises under several conditions, including the following:

i. When alternative sets of scientific or technical assumptions for developing a model exist (model structure);
ii. When models are simplified for purposes of tractability (model detail – inclusion or exclusion of sub-processes); and,
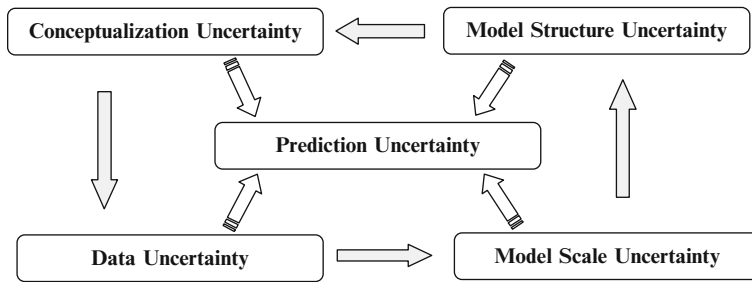
Fig. 2.5 Uncertainty sources in modeling

  iii. When a coarse discretization and of data is used to reduce the computation
       demands of the model (model resolution – scale issues and statistical
       uncertainty).

The uncertainties and errors in simulation may arise from uncertainty in model
inputs or parameters (i.e., parametric or data uncertainty). When a model applica-
tion involves both model and data uncertainties, it is important to identify the
relative magnitudes of the uncertainties associated with data and model formula-
tion. Such a comparison is useful for focusing resources where they are most
appropriate (e.g., data gaps versus model refinement).

Uncertainties in model parameter estimates may stem from a variety of sources.
Even though many parameters could be measured or calculated up to some degree
of precision, there are often significant uncertainties associated with their estimates.
Some uncertainties and errors can be identified as:

   i. Random errors in analytic devices used in field and laboratory measurements;
  ii. Systematic biases that occur due to imprecise calibration;
 iii. Extrapolation of data from one scale to another; and,
  iv. Inaccuracy in the assumptions used to infer the actual quantity of interest
      from observations of a "surrogate" parameter or estimation of parameters
      based on mildly representative samples.

Uncertainty analysis should not be confused with sensitivity analysis. In uncer-
tainty analysis one attempts to describe the entire set of possible outcomes of a
model together with their associated probabilities of occurrence. In sensitivity
analysis one determines the relative change in model output given changes in
model input values.

Model errors can be evaluated by analyzing the variation in dependent variables
in the model based on the variation of the independent variables of the model, i.e.
the parameters of the model. Taylor series analysis is commonly used in this
analysis. Since Taylor series will be used in several different contexts in this
book it is appropriate to introduce a review of this topic.

A Taylor series is the sum of functions composed of continually increasing
derivatives. For a dependent variable such as contaminant concentration $C(P)$,

which depends on only one independent parameter P, the value of the function $C(P)$ at points near P_o can be approximated by the following Taylor series,

$$C(P_o + \Delta P) = C(P_o) + \frac{\Delta P}{1!}\left(\frac{dC}{dP}\right)\bigg|_{P_o} + \frac{(\Delta P)^2}{2!}\left(\frac{d^2C}{dP^2}\right)\bigg|_{P_o} + \frac{(\Delta P)^3}{3!}\left(\frac{d^3C}{dP^3}\right)\bigg|_{P_o}$$
$$+ \cdots + \frac{(\Delta P)^n}{n!}\left(\frac{d^nC}{dP^n}\right)\bigg|_{P_o} + R_{n+1}$$

(2.2)

in which P_o is some reference value of the parameter P, ΔP is the increment in the parameter P and $(P_o + \Delta P)$ identifies the point where the concentration C is to be evaluated $C(P_o + \Delta P)$ and R_{n+1} represents the remainder terms of a Taylor series expansion. In Eq. (2.2) the derivatives of $C(P)$ are evaluated at P_o. Using the definition above a first order approximation can be defined by keeping the terms of the Taylor series up to and including the first derivative as follows,

$$C(P_o + \Delta P) \approx C(P_o) + \Delta P\left(\frac{dC}{dP}\right)\bigg|_{P_o}$$

(2.3)

Similarly, the second and third order approximations to Taylor series are given by

$$C(P_o + \Delta P) \approx C(P_o) + \Delta P\left(\frac{dC}{dP}\right)\bigg|_{P_o} + \frac{(\Delta P)^2}{2!}\left(\frac{d^2C}{dP^2}\right)\bigg|_{P_o}$$

(2.4)

and

$$C(P_o + \Delta P) \approx C(P_o) + \Delta P\left(\frac{dC}{dP}\right)\bigg|_{P_o} + \frac{(\Delta P)^2}{2!}\left(\frac{d^2C}{dP^2}\right)\bigg|_{P_o}$$
$$+ \frac{(\Delta P)^3}{3!}\left(\frac{d^3C}{dP^3}\right)\bigg|_{P_o}$$

(2.5)

respectively. The accuracy of a Taylor series approximation improves as the order of the Taylor series increasesas shown in Eqs. (2.3) through (2.5). In these equations an approximate relationship is implied since the remainder terms of the Taylor series are omitted. Referring back to Eq. (2.3), we can associate the point P_o with the mean value of the parameter distribution P. Accordingly, the Eq. (2.3) will represent the value of C(a space is needed here such as) C around the mean value of P. We can now write an equation for the variance of the concentration C, using the definition of variance of $C(P)$ about the mean P_o, $S^2(C(P_o))$,

$$S^2(C(P)) = S^2(P)\left(\frac{dC}{dP}\right)^2\bigg|_{P_o}$$

(2.6)

where $S(P)$ is the sample standard deviation, and $S^2(P)$ is the sample variance around the mean P_o. Eq. (2.6) implies that the variance in the dependent variable (uncertainty) is a function of the variance (uncertainty) in the parameter P, the sensitivity of the dependent variable to the changes in the parameter P around its mean, $\left(\frac{dC}{dP}\right)^2\big|_{P_o}$ and the variance in the parameters $S^2(P)$.

For a multivariate relationship, $C(P^i)$, $i = 1, 2, 3, ..., n$ the first order Taylor series expansion, Eq. (2.3), can be written as,

$$C\left(P_o^1 + \Delta P^1, P_o^2 + \Delta P^2, P_o^3 + \Delta P^3, ..., P_o^n + \Delta P^n\right)$$
$$\approx C\left(P_o^1, P_o^2, P_o^3, ..., P_o^n\right) + \sum_{i=1}^{n} \Delta P^i \left(\frac{\partial C}{\partial P^i}\right)\bigg|_{P_o^i} \qquad (2.7)$$

which yields the variance relation,

$$S^2\left(C\left(P_o^1, ..., P_o^n\right)\right) \approx \sum_{i=1}^{n} S^2\left(P_o^i\right)\left(\frac{\partial C}{\partial P_o^i}\right)^2$$
$$+ 2\sum_{j=1}^{n-1}\sum_{i=j+1}^{n} \left(\frac{\partial C}{\partial P_o^i}\right)\left(\frac{\partial C}{\partial P_o^j}\right) S\left(P_o^i\right) S\left(P_o^j\right) \Phi\left(P_o^i, P_o^j\right) \qquad (2.8)$$

where P_o^i is the mean of the ith parameter, $S\left(P_o^i\right)$ and $S^2\left(P_o^i\right)$ are the standard deviation and the variance of the ith parameter around its mean respectively, $S^2\left(C\left(P_o^i\right)\right)$ is the variance of $C(P^i)$ around the means P_o^i, $\Phi\left(P_o^i, P_o^j\right)$ is the correlation coefficient in a linear least squares regression between the parameters P^i and P^j (Crow et al. 1960; Reckhow and Chapra 1983; Bogen and Spear 1987; Ayyub and McCuen 1997; Conover 1999).

Monte Carlo analysis is another method used to evaluate parameter sensitivity to solution. Since this approach is used extensively in the ACTS and RISK software we will review this topic in more detail in Chapter 7.

2.5   Methods of Solution

Some mathematical models are relatively simple and their solution can be achieved using analytical methods, sometimes referred to as a closed form solution. Numerical calculation based on an analytical solution can be exact or approximate. Its accuracy depends on the complexity of the analytical solution. More complex models may require numerical solution which are all inherently approximate solutions to the problem. Both solutions will require computer based calculations to relate the model inputs to model outputs.

As indicated above statistical models and statistical calculations are also a necessary component of a modeling exercise. If not explicitly used in the modeling

itself, statistical methods will become an important component in the sensitivity, calibration and verification phases of the modeling exercise.

In the case of the ACTS and RISK software analytical solutions will commonly be employed, since the models included in these software platforms are considered as screening models and in that sense are simpler representations of the modeled system. To perform sensitivity analysis the ACTS and RISK software also includes a Monte Carlo module in all models where the models can be run in a stochastic mode.

2.6 Modeling Terminology

The modeling field is quite a diverse field of science. It is important for the professionals working in the environmental health field to familiarize themselves with various concepts and methods employed in this field to be able to understand the outcomes and limitations of environmental modeling and use them in environmental health analysis appropriately. For this purpose a review of the following references are recommended, (Gentil and Blake 1981; USEPA 1984; Starfield and Bleloch 1986; Hoover and Perry 1989; Law and Kelton 1991; Tsang 1991; Mayer and Butler 1993; Oreskes et al. 1994b; Lemons 1996; Schnoor 1996; Abdel-Magid et al. 1997; Saltelli et al. 2000; Anderson and Bates 2001; Nirmalakhandan 2002; Aral and Gunduz 2003). The acronyms used in this field are given in Appendix A of this book. The list of terms and their definitions given in Appendix B are also included in this book to familiarize the reader with the terminology used in the environmental modeling field as a starting point.

References

Abdel-Magid IS, Mohammed A-WH et al (1997) Modeling methods for environmental engineers. CRC Lewis, Boca Raton, FL

Anderson MG, Bates PD (2001) Model validation: perspectives in hydrological science. Wiley, West Sussex, England

Aral MM (1990a) Groundwater modeling in multilayer aquifers: steady flow. Lewis, Chelsea, MI

Aral MM (1990b) Groundwater modeling in multilayer aquifers: unsteady flow. Lewis, Chelsea, MI

Aral MM, Gunduz O (2003) Scale effects in large scale watershed modeling. In: Singh V, Yadava RN (eds) Advances in hydrology. Proceedings of the international conference on water and environment (WE-2003), Bhopal, India, pp 37–51

Ayyub BM, McCuen RH (1997) Probability, statistics, and reliability for engineers. CRC Press, Boca Raton, FL

Bloschl G, Sivapalan M (1995) Scale issues in hydrological modeling: a review. In: Kalma JD, Sivapalan MC (eds) Scale issues in hydrological modeling. Wiley, Chichester, UK, pp 9–48

Bogen KT, Spear RC (1987) Integrating uncertainty and interindividual variability in environmental risk assessments. Risk Anal 7:427–436

Caldwell MM, Matson PA et al (1993) Prospects for scaling. In: Ehleringer JR, Field CB (eds) Scaling physiological processes: leaf to globe. Academic, San Diego, CA, pp 223–230

Conover WJ (1999) Practical nonparametric statistics. Wiley, New York

Crow EL, Davis FA et al (1960) Statistics manual. Dover Publications Inc, New York

Curry GL, Deuermeyer BL et al (1989) Discrete simulation. Holden-Day, Oakland, CA

Gentil S, Blake G (1981) Validation of complex ecosystems models. Ecol Model 14(1–2):21–38

Gunduz O, Aral MM (2003a) Simultaneous solution of coupled surface water/groundwater flow systems. In: Brebbia CA (ed) River basin management II. WIT Press, Southampton, UK, pp 25–34

Gunduz O, Aral MM (2003b) An integrated model of surface and subsurface flows. In: Brebbia CA (ed) Water resources management II. WIT Press, Southampton, UK, pp 367–376

Gunduz O, Aral MM (2003c) A simultaneous solution approach for coupled surface and subsurface flow modeling, multimedia environmental simulations laboratory. School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, 98 pp

Gunduz O, Aral MM (2003d) Hydrologic modeling of the lower Altamaha River Basin. In: 2003 Georgia water resources conference, University of Georgia, Athens, GA

Gunduz O, Aral MM (2005) River networks and groundwater flow: a simultaneous solution of a coupled system. J Hydrol 301(1–4):216–234

Gupta VK, Rodriguez-Iturbe I et al (1986) Scale problems in hydrology. D. Reidel, Dordrecht, The Netherlands

Hoffman FO, Hammonds JS (1994) Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 14(5):707–712

Holling CS (1978) Adaptive environmental assessment and management. Wiley, New York

Hoover SV, Perry RF (1989) Simulation. Addison–Wesley, Reading, MA

Jarvis PG (1995) Scaling processes and problems. Plant Cell Environ 18:1079–1089

Klemes V (1983) Conceptualization and scale in hydrology. J Hydrol 65:1–23

Law AW, Kelton WD (1991) Simulation modeling and analysis. McGraw-Hill, New York

Lemons J (ed) (1996) Scientific uncertainty and environmental problem solving. Blackwell Science, Cambridge, MA

Marceau D (1999) The scale issue in social and natural sciences. Can J Remote Sens 25(4):347–356

Mayer DG, Butler DG (1993) Statistical validation. Ecol Model 68(1–2):21–32

McCarl BA (1984) An overview: with some emphasis on risk models. Rev Market Agric Econ 52:153–173

McDonald MG, Harbaugh AW (1988) A modular three-dimensional finite difference groundwater flow model. U.S. Geological Survey Techniques of Water Resources Investigations, Book 6, Chapter A1, 586 pp

Nirmalakhandan N (2002) Modeling tools for environmental engineers and scientists. CRC, Boca Raton, FL

NRC (1990) Groundwater models; scientific and regulatory applications. National Research Council, Washington, DC, p 303

O'Neill RV, King AW (1998) Homage to St. Michael: or why are there so many books on scale? In: Peterson DL, Parker VT (eds) Ecological scale, theory and applications. Columbia University Press, New York, pp 3–15

Oreskes N, Belitz K et al (1994a) The meaning of models – response. Science 264(5157): 331–341

Oreskes N, Shraderfrechette K et al (1994b) Verification, validation, and confirmation of numerical-models in the earth-sciences. Science 263(5147):641–646

Power M (1993) The predictive validation of ecological and environmental-models. Ecol Model 68(1–2):33–50

Reckhow KH, Chapra SC (1983) Engineering approaches for lake management. Butterworth, Boston, MA

Rykiel EJ (1996) Testing ecological models: the meaning of validation. Ecol Model 90(3): 229–244

Saltelli A, Chan K et al (2000) Sensitivity analysis. Wiley, West Sussex, England

Sargent RG (1984) A tutorial on verification and validation of simulation models. In: Proceedings of the 1984 Winter Simulation Conference. IEEE 84CH2098-2, Dallas, TX

Schnoor JL (1996) Environmental modeling: fate and transport of pollutants in water, air, and soil. Wiley, New York

Singh VP (1995) Watershed modeling. In: Singh VP (ed) Computer models of watershed hydrology. Water Resources Publications, Highlands Ranch, CO, pp 1–22

Starfield AM, Bleloch AL (1986) Building models for conservation and wildlife management. Macmillan, New York

Tsang CF (1991) The modeling process and model validation. Ground Water 29(6):825–831

USEPA (1984) Evaluation and selection of models for estimating air emissions from hazardous waste treatment, storage and disposal facilities. Office of Air Quality Planning and Standards, Research Triangle Park, NC