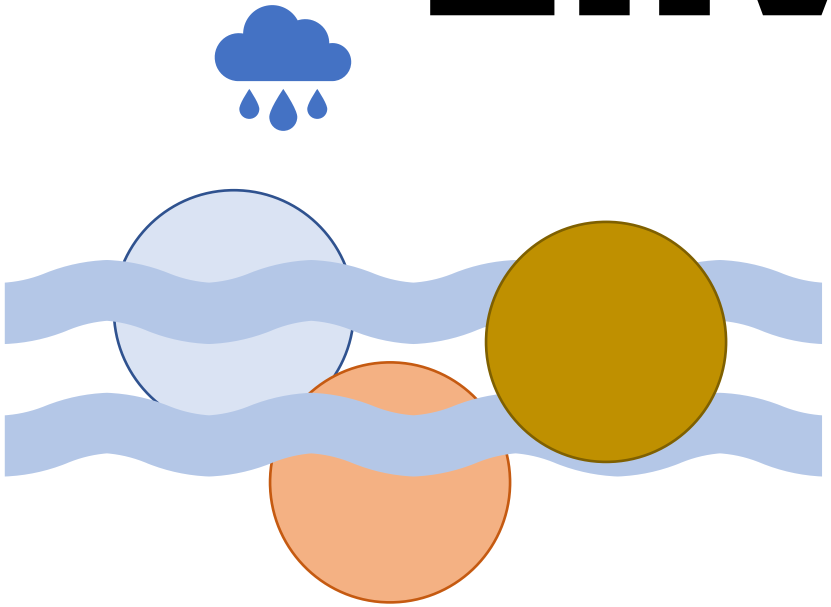
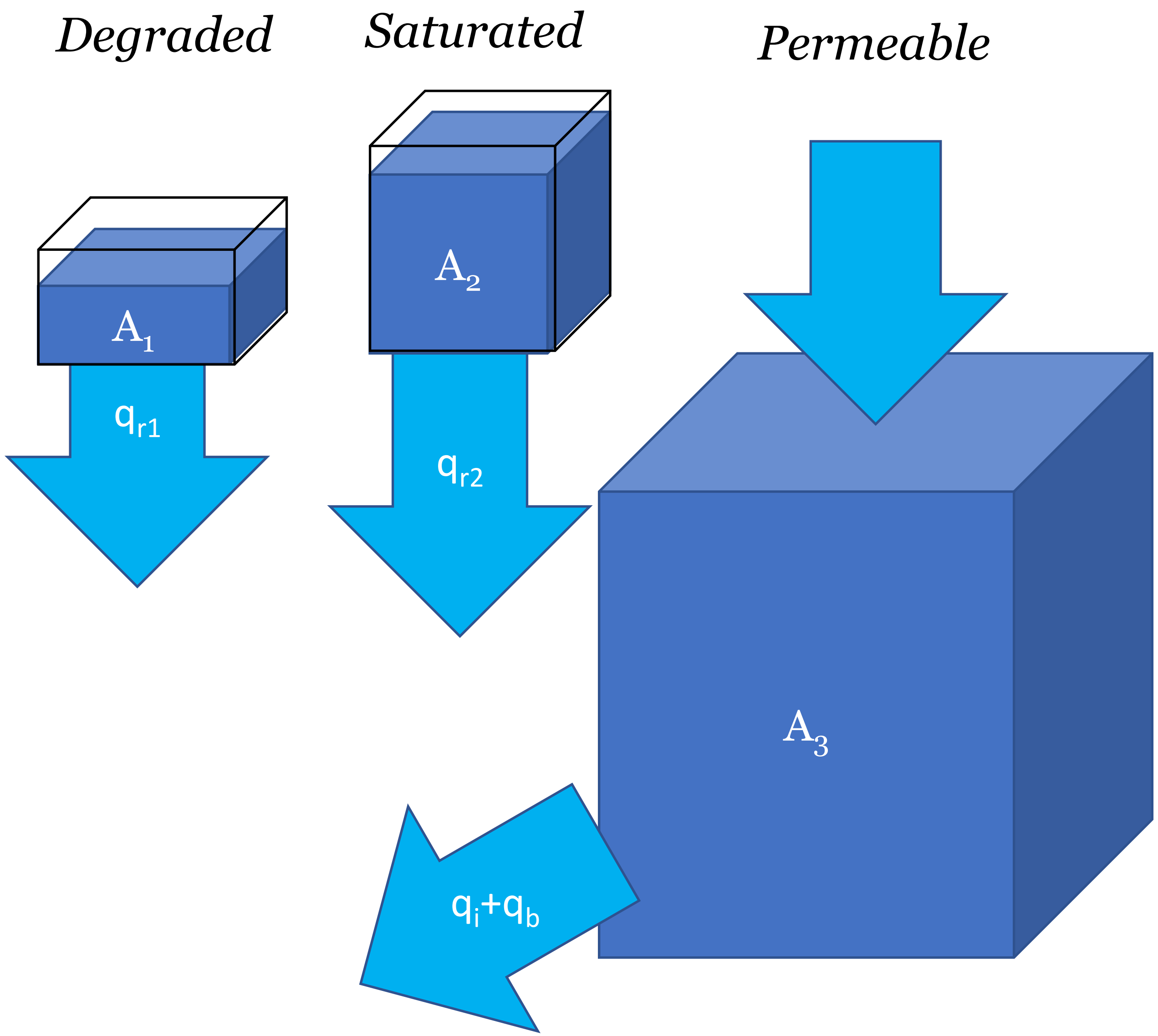
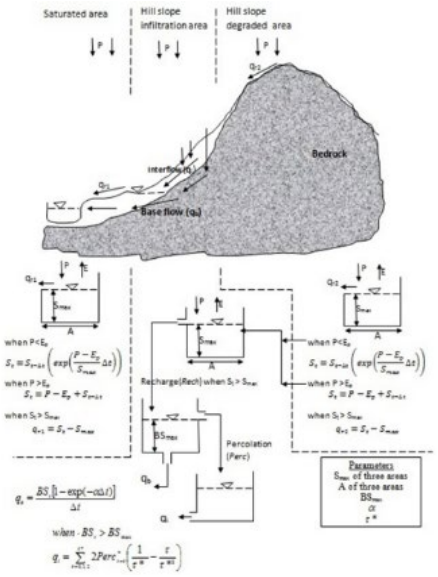
The “Modeling” Environment



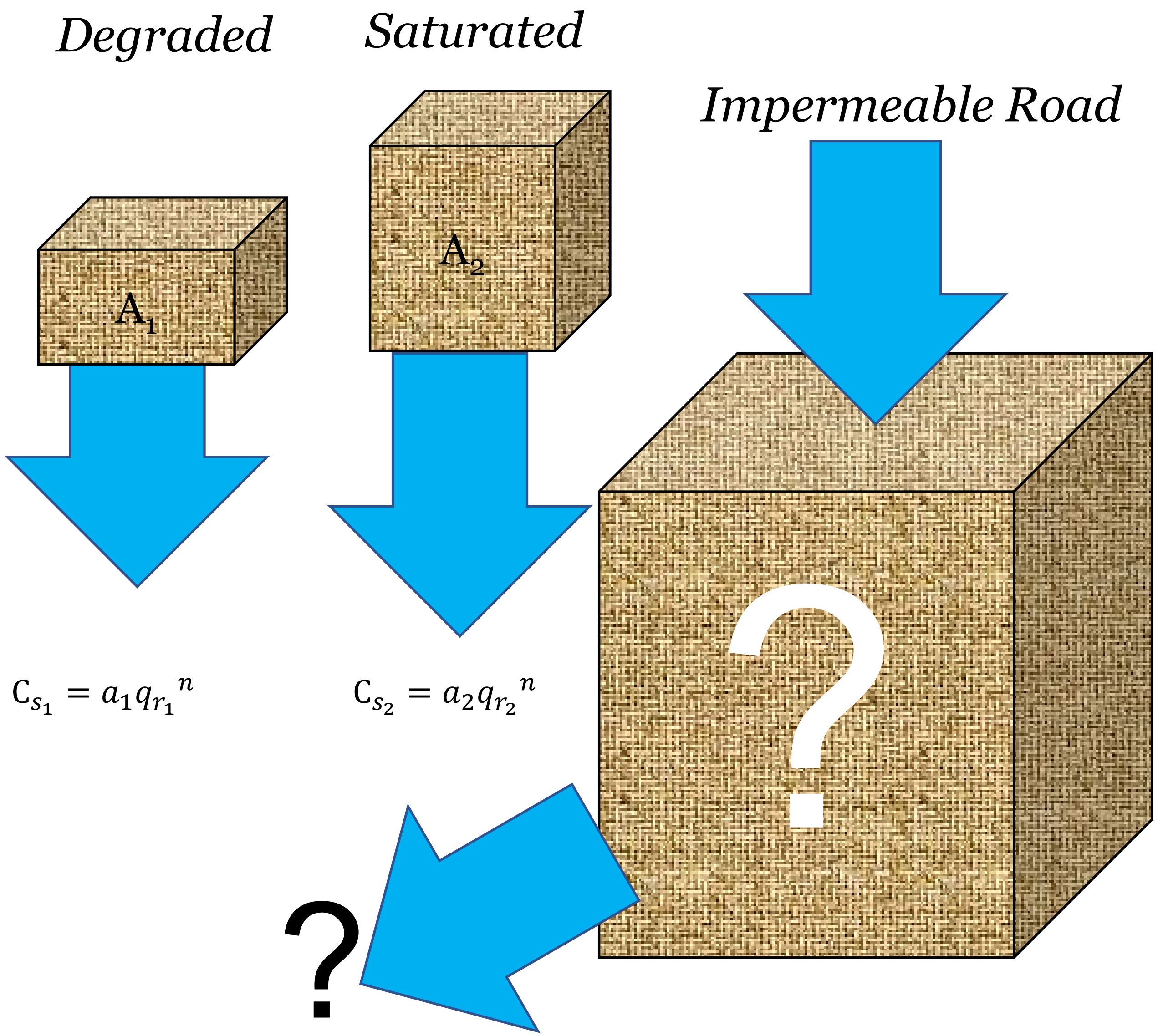


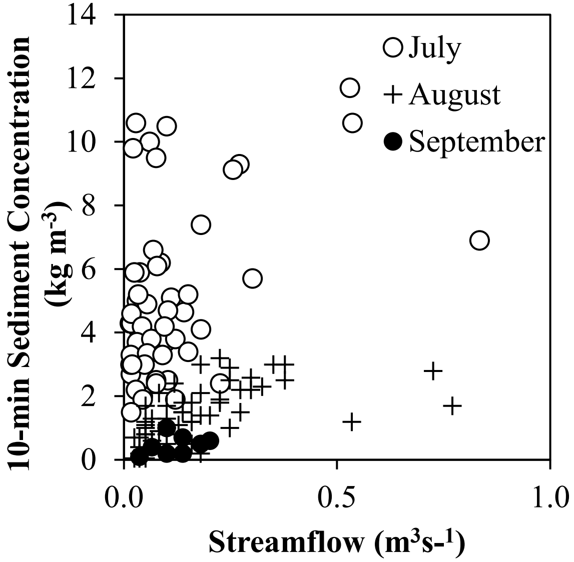
$$Q = A_1 q_{r_1} + A_2 q_{r_2} + A_3(q_i + q_b)$$

The Parameter Efficient Distributed (PED) model (Tilahun et al., 2013,2015) uses a water-balance saturation excess framework.



Erosion is simulated where runoff is occurring.



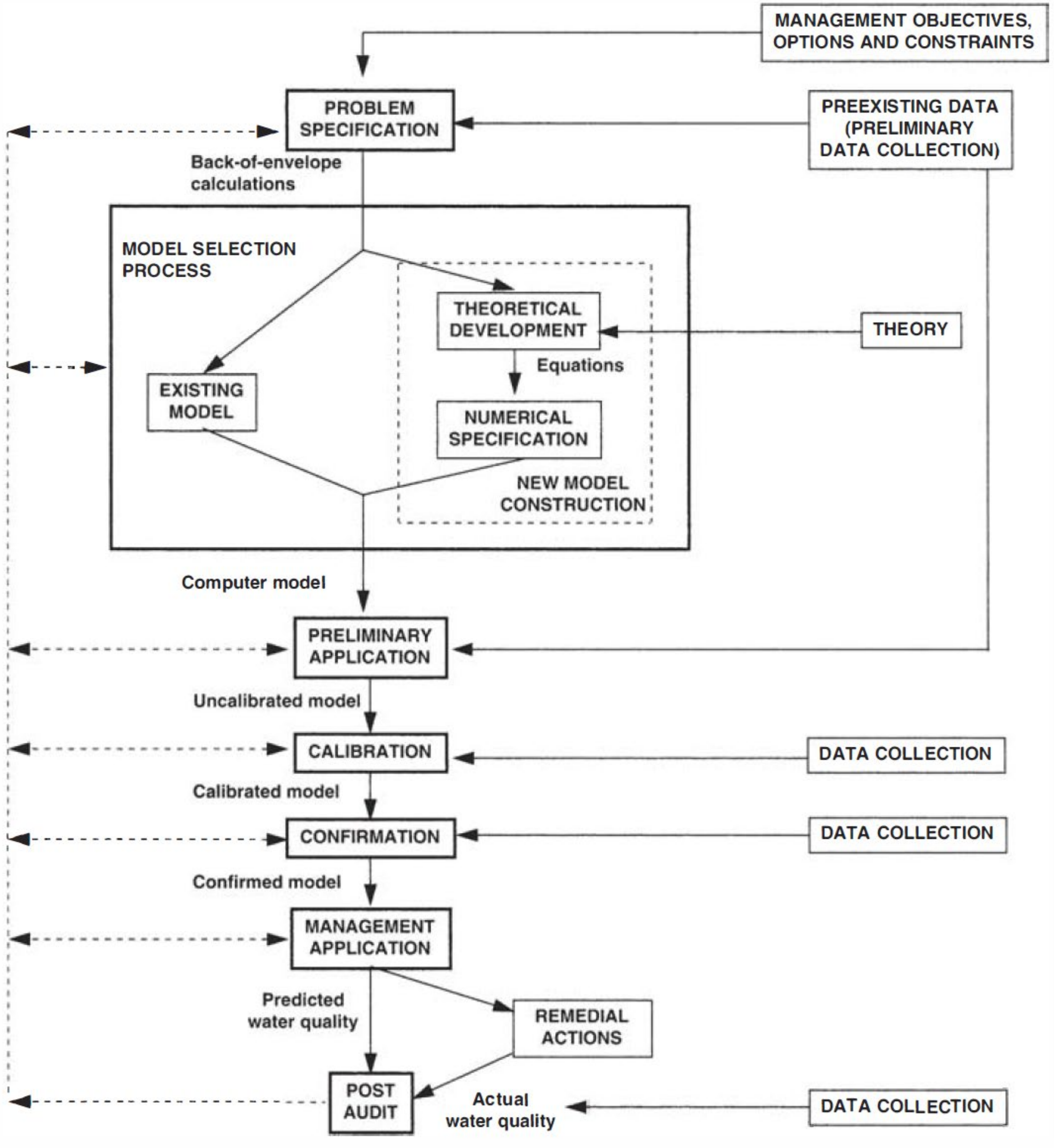


$$C = \frac{\left(A_1 {q_{r_1}}^{1.4}[a_{s1} + H(a_{t1} - a_{s1})] + A_2 {q_{r_2}}^{1.4}[a_{s2} + H(a_{t2} - a_{s2})]\right) + ?}{A_1 q_{r_1} + A_2 q_{r_2} + A_3(q_i + q_b) + ?}$$

Roads create an uncertain amount of runoff and sediment supply and prevented the simple use of the PED model.

Water-quality modeling seems to be prone to oversimplification since (in order to develop realistic models) you must first assimilate a considerable body of technical information dealing with theory, mathematics, and computers.

However, there is a much larger picture for the modeling process…



(*a*) Process (*b*) Information

The Water-Quality-Modeling Process

- Prbm Specification: Getting Started
- Model Selection
- Preliminary Application
- Calibration
- Confirmation and Robustness
- Management Applications
- Post-Audit

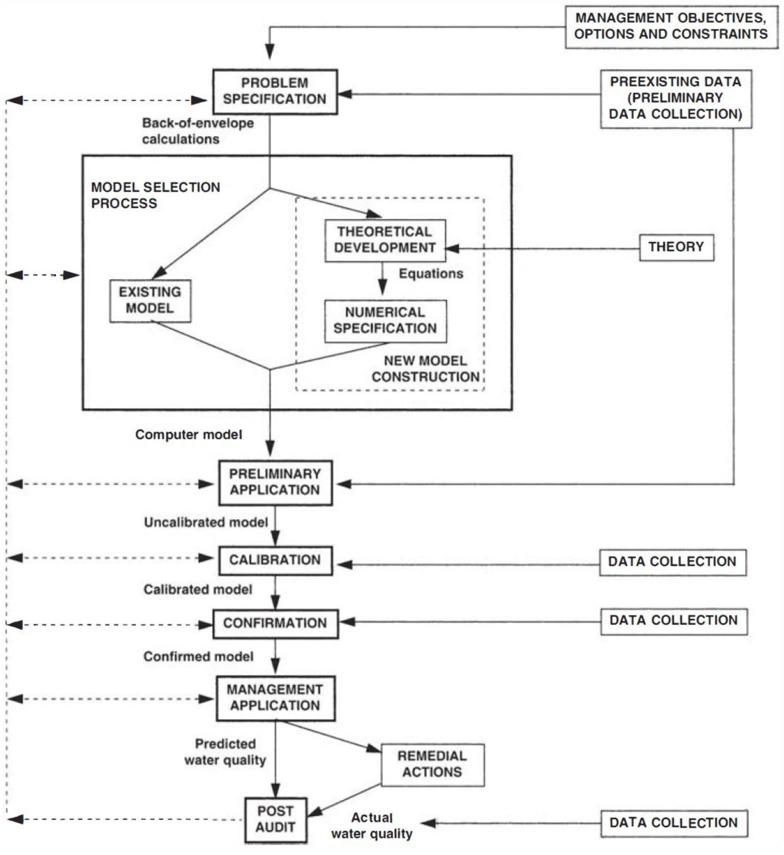
Model Sensitivity

- Parameter Perturbation
- First-Order Sensitivity Analysis
- Monte Carlo Analysis

Assessing Model Performance

Segmentation and Model Resolution

Water-Quality-Modeling Process: Problem Specification; Getting Started



(*a*) Process (*b*) Information

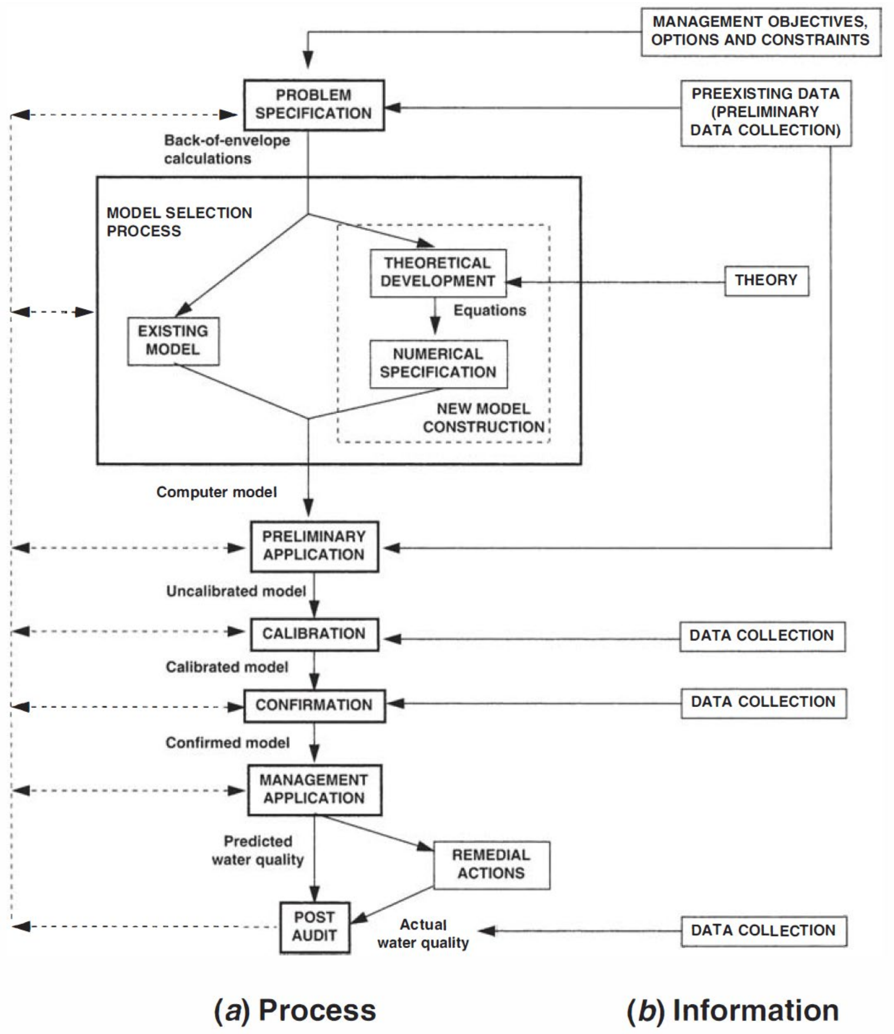
Problem Specification: the water-quality engineer must be provided with a clear delineation of the objectives of the customer (individual, corporation, municipality, regulatory agency).

Two information sources feed this phase:

Management objectives, options, constraints

Physics, chemistry, biology of water body/drainage basin

Water-Quality-Modeling Process: Model Selection



Next step is to obtain a model (usually existing software)…why:

-work already done

-widely used (legal, regulatory)

-credibility

Sometimes water-quality problem is not yet available as a modeling process.

This necessitates theoretical development & numerical development/validation

Water-Quality-Modeling Process: Theoretical Development

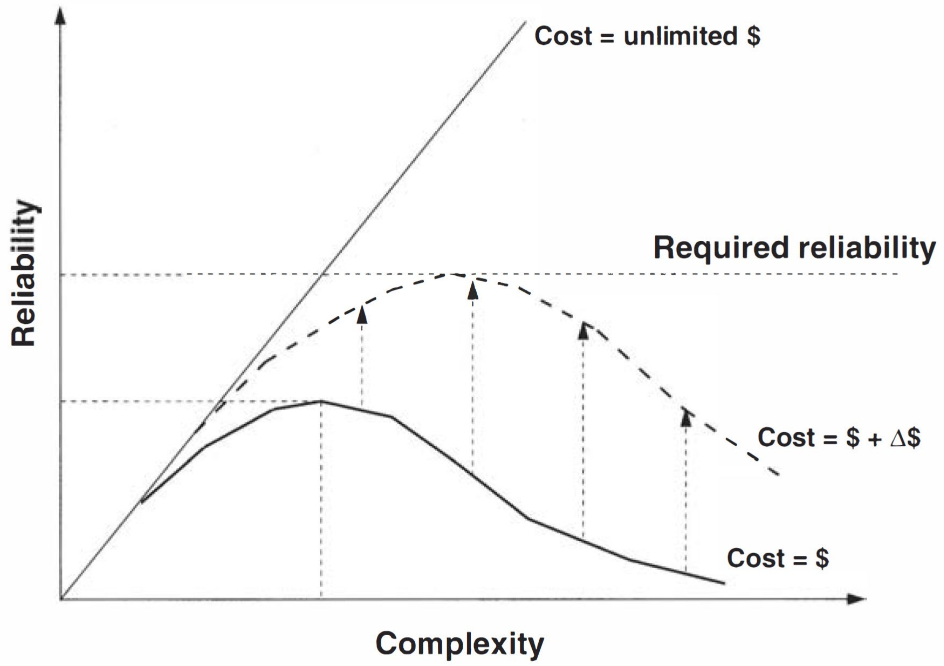
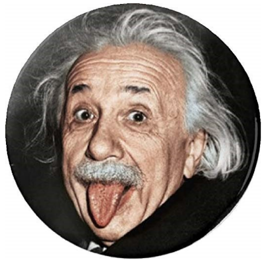


FIGURE 18.2
The trade-off between model reliability and complexity.



Requires variables, parameters, and associated continuity equations.

Continuity equations: mass and/or energy

Momentum balances: for hydrodynamics

Also requires decision on model complexity/cost tradeoffs

Strive to develop the simplest model that is consistent with data and problem requirements (Ockham's razor).

Numerical specification and validation

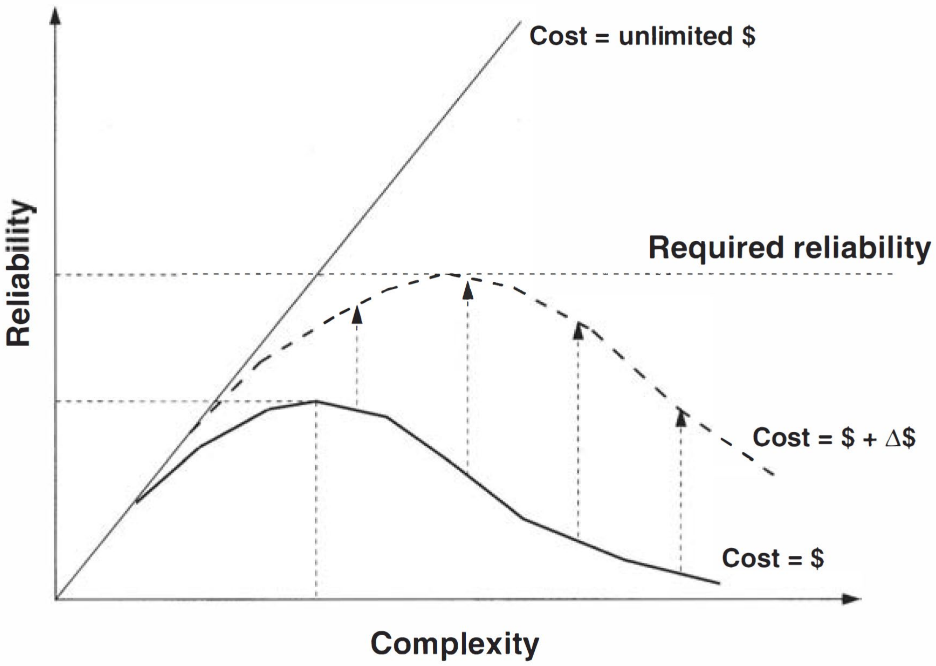
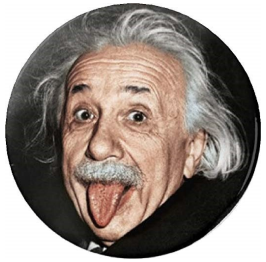


FIGURE 18.2
The trade-off between model reliability and complexity.



Now equations must be implemented on the computer: algorithm design (data structure and numerical solution techniques), coding in computer language, debugging, testing, and documentation.

Testing (validation, confirmation)

- Mass balances
- Simplified solutions
- Range of conditions
- Graphical results
- Benchmarking

Preliminary Application

This stage is useful for identifying data deficiencies and theoretical gaps. It can also provide important context for designing the field and laboratory studies required to fill the gaps.

Here we furthermore investigate which model parameters have the greatest impact on model predictions with a ***model sensitivity analysis*** (method of varying each parameters by a set percent and observing how the predictions vary).

Calibration

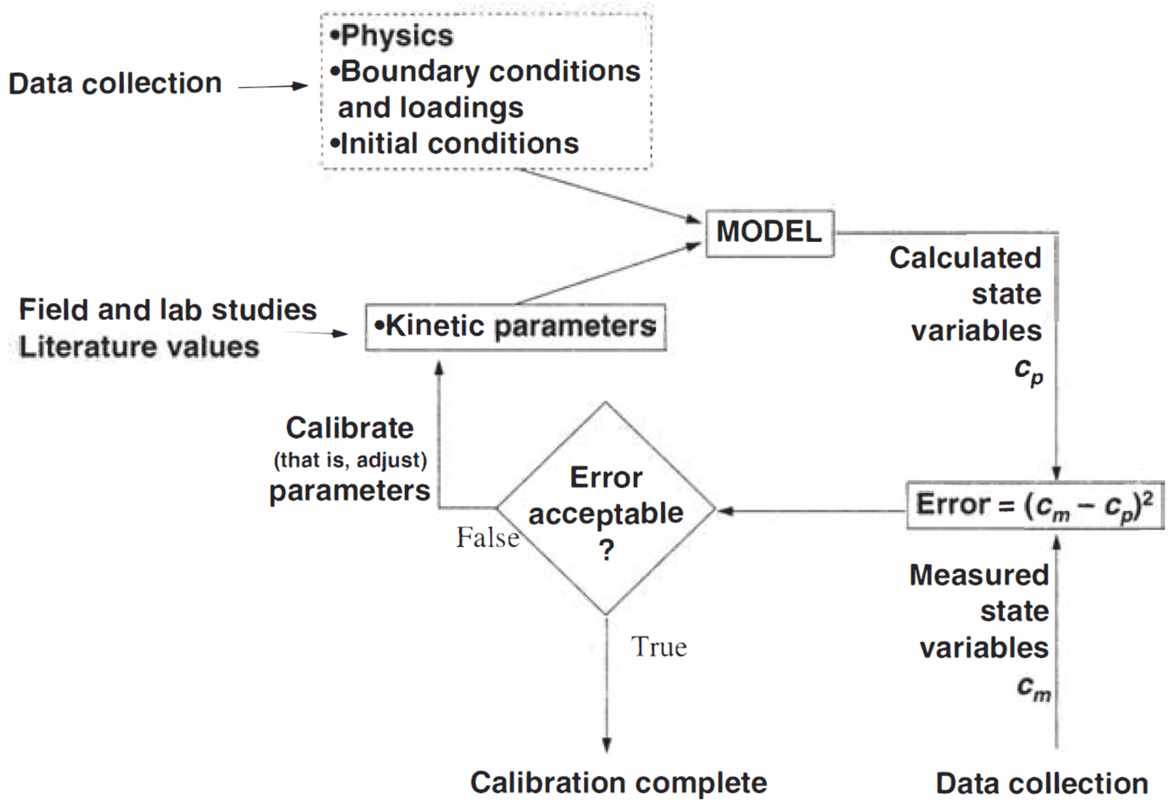


FIGURE 18.3
Schematic diagram of the model calibration process.

Calibration: "Tuning" the model to fit a data set. Fitting means to vary model parameters to obtain an optimal agreement between model calculations and the data set.

E.g. calibrating on low flows (7Q10) during summer for wastewater discharge.

Several types of information must be fed into a model to systematically obtain a best fit:

Calibration

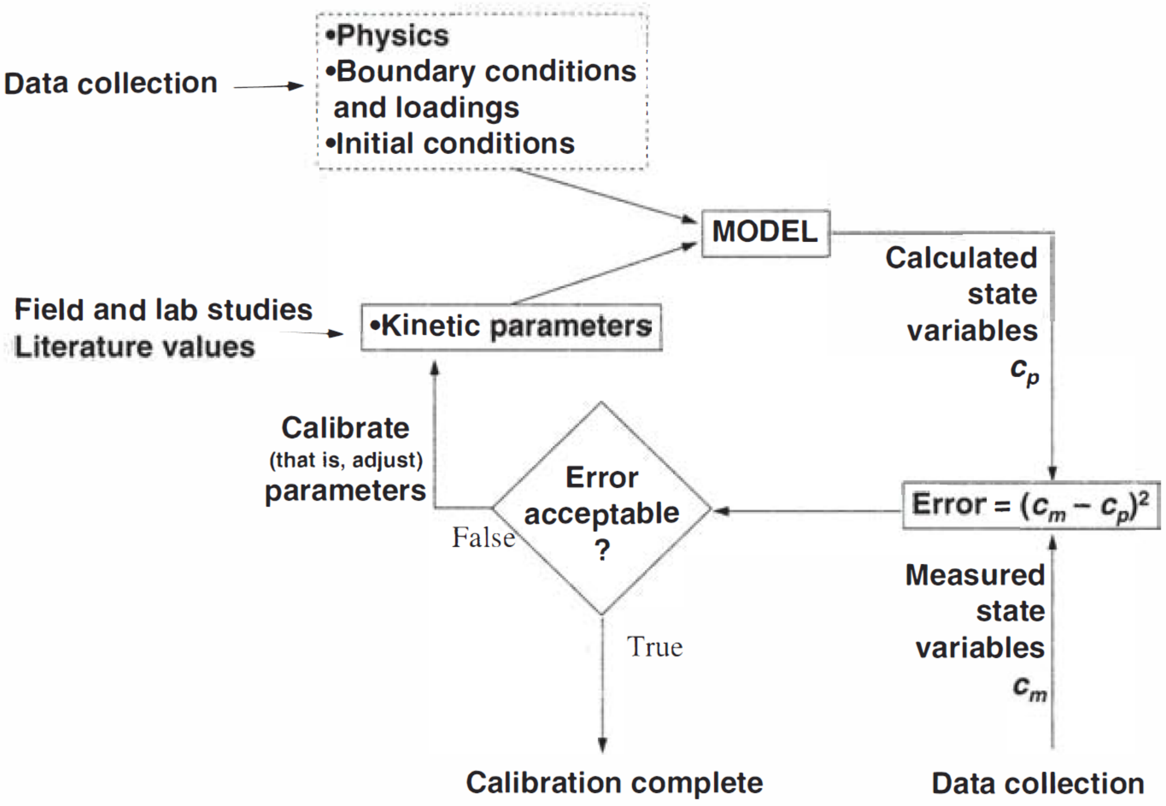


FIGURE 18.3
Schematic diagram of the model calibration process.

Several types of information must be fed into a model to systematically obtain a best fit:

-Forcing functions and physical parameters:

Boundary conditions & loads

Initial conditions

Physics

-Calibration parameters: kinetics

Adjustments can be made by trail and error or through automated techniques.

Confirmation and Robustness

At this point a calibrated model fits a single data set, now it must be confirmed.

-On a new data set (or several data sets)

-with new physical parameters

-and updated forcings reflecting new conditions

If it matches the new conditions (confirmation!), if not, discrepancies must be investigated…. Recall "Oreskes et al 1994: *all that can be concluded is that our testing has not proved the model wrong*". The actual goal then is to establish the model's robustness (via a large number and diversity of confirming observations)

Management Application

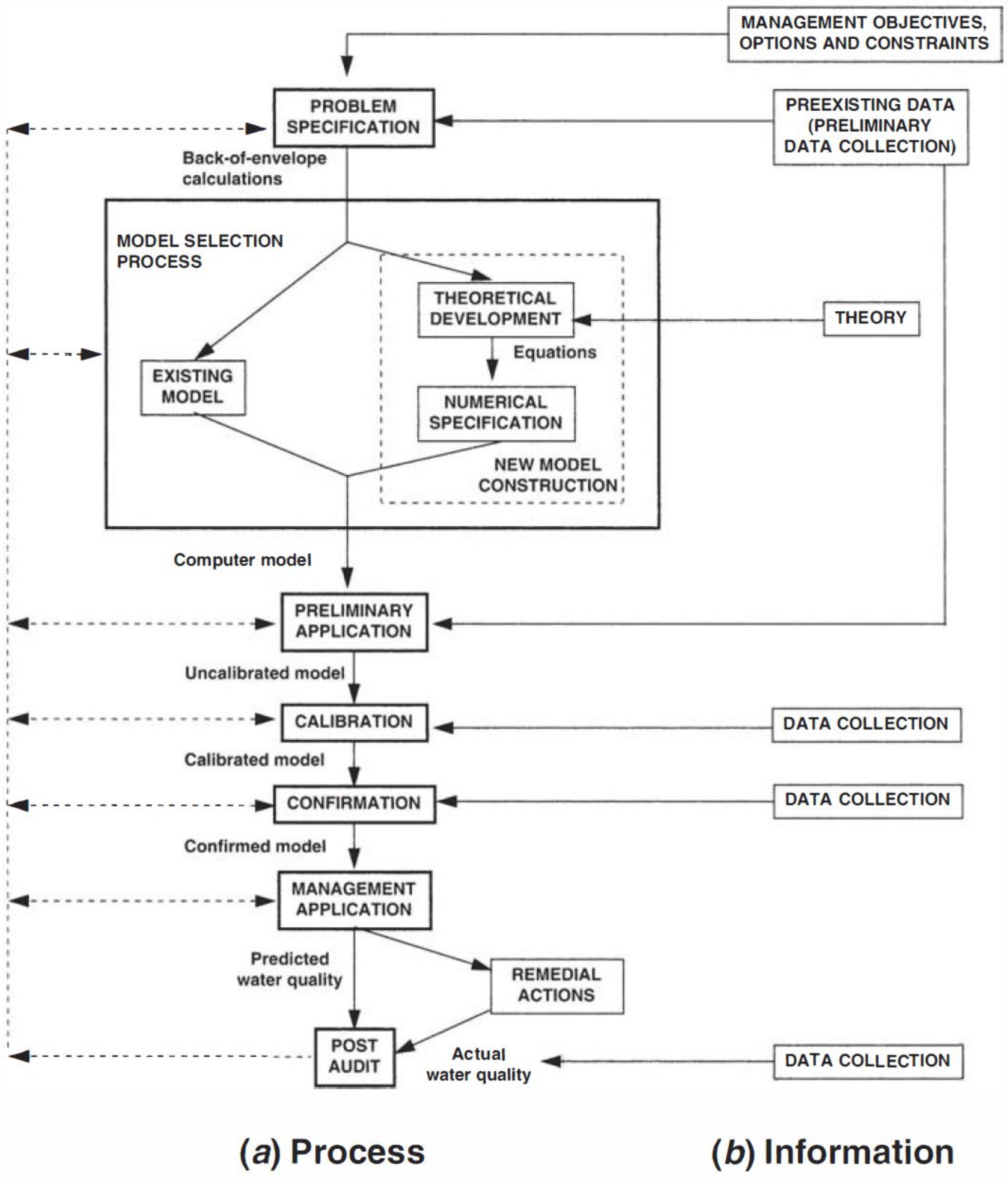
Many modeling studies result in remedial actions (wastewater treatment plant construction/upgrade, aeration, dredging implemented).

We can model the effectiveness of these actions by modifying the parameters and forcing functions to view effects on state variables.

Post-Audit

After remedial action has been implemented, a check can be made to view the accuracy of predictions.

Differences between predictions and resulting water quality have often occurred. These cases can be useful for discovering missing mechanisms and information for the improvement of the robustness of the frameworks.



The Water-Quality-Modeling Process

Prbm Specification: Getting Started

Model Selection

Preliminary Application

Calibration

Confirmation and Robustness

Management Applications

Post-Audit

Model Sensitivity

Parameter Perturbation

First-Order Sensitivity Analysis

Monte Carlo Analysis

Assessing Model Performance

Segmentation and Model Resolution

Model Sensitivity: Parameter Perturbation & 1st Order Analysis

One way to get a general understanding of the behavior of a water-quality model is to conduct a sensitivity analysis.

Common implementations are:

-Simple parameter perturbation

-First-order sensitivity analysis

-Monte Carlo approaches

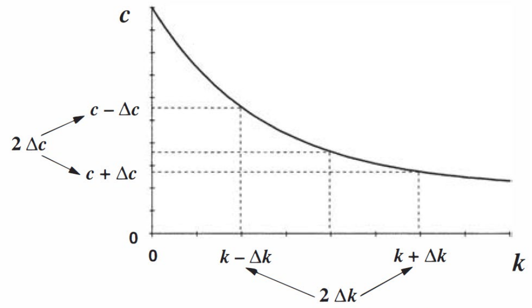
Model Sensitivity: Parameter Perturbation

For a simple mass-balance equation for a well-mixed lake, we illustrate both (parameter perturbation, first-order sensitivity analysis).

$$V\frac{dc}{dt} = Q\,c_{in} - Qc - kVc$$

which at steady state is solved for as:

$$c = \frac{Q}{Q + kV}c_{in}$$



(a) Parameter perturbations

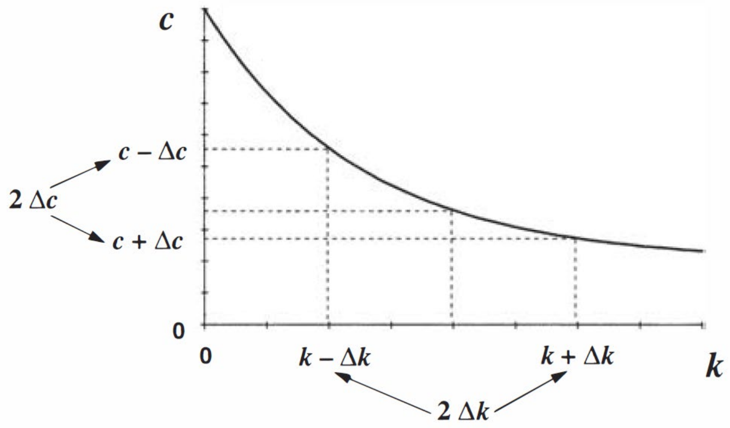
Recognize that c is a function of the parameters and forcing functions: i.e. $c = f(Q, k, V, c_{in})$. Hence once way to visualize the dependence of the solution on one of the parameters is to plot c versus the parameter (e.g. k)

Model Sensitivity: Parameter Perturbation

Parameter perturbation consists of varying each of the model parameters (e.g. raised or lowered a fixed percent) while holding all other terms constant.

The corresponding variations of the state variables reflect the sensitivity of the solution of the varied parameter.

$$\Delta c = \frac{c(k + \Delta k) - c(k - \Delta k)}{2}$$



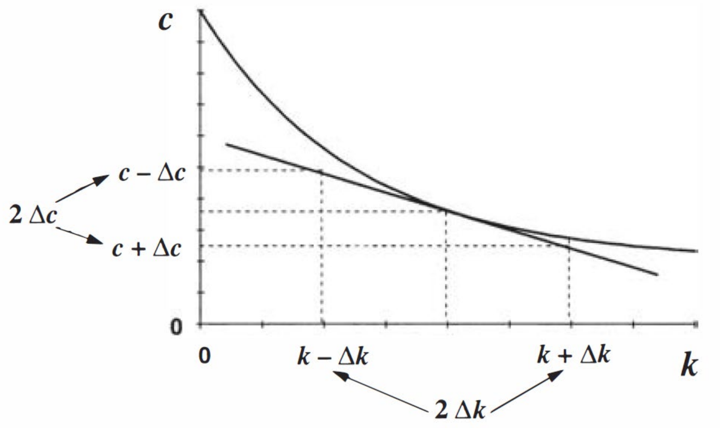
(a) Parameter perturbations

First-order Sensitivity Analysis

An alternative technique that yields similar information is based on a first-order sensitivity analysis. This approach uses the derivative of the function with respect to the parameter as an estimate of the sensitivity. One way is to employ first-order Taylor-series expansions of the model around the value of the parameter. E.g.

$$c(k + \Delta k) = c(k) + \frac{\partial c(k)}{\partial k} \Delta k$$

$$c(k - \Delta k) = c(k) - \frac{\partial c(k)}{\partial k} \Delta k$$



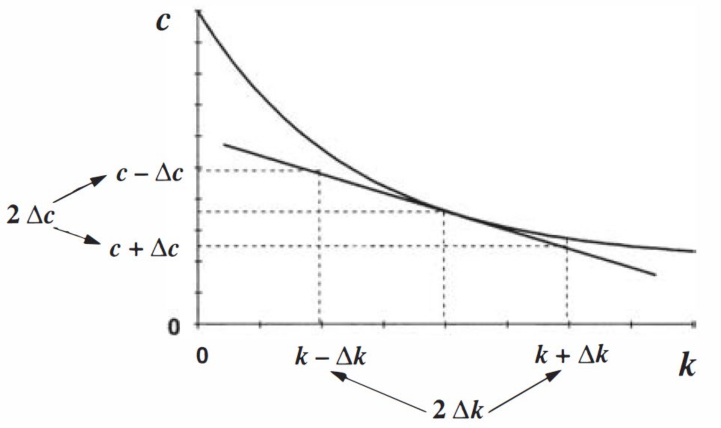
(b) First-order analysis

First-order Sensitivity Analysis

These equations can be subtracted from each other

$$c(k + \Delta k) = c(k) + \frac{\partial c(k)}{\partial k}\Delta k$$

$$c(k - \Delta k) = c(k) - \frac{\partial c(k)}{\partial k}\Delta k$$



(b) First-order analysis

To yield:

$$\Delta c = \frac{c(k + \Delta k) - c(k - \Delta k)}{2} = \frac{\partial c(k)}{\partial k}\Delta k$$

(the sign indicates variation of the prediction, see graph's negative slope)

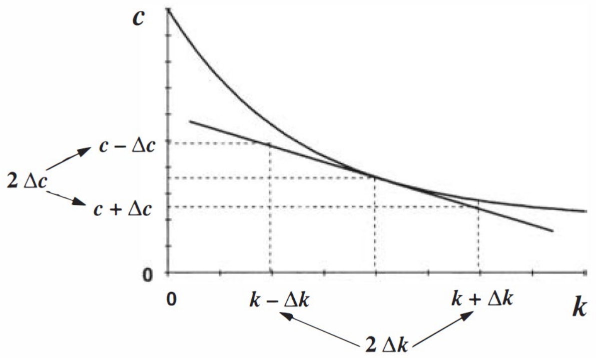
Condition numbers

A refinement on these previous sensitivity analyses is to express the results as condition numbers: For 1st order ($\Delta c = \frac{\partial c(k)}{\partial k} \Delta k$) , dividing both sides by c and multiplying right hand side by k/k.

$$\frac{\Delta c}{c} = CN_k \frac{\Delta k}{k}$$

where CN_k is the condition number for the parameter k.

$$CN_k = \frac{k}{c} \frac{\partial c}{\partial k}$$



(b) First-order analysis

For perturbation analysis the discrete form would be used:

$$CN_k = \frac{k}{c} \frac{\Delta c}{\Delta k}$$

EXAMPLE 18.1. SENSITIVITY ANALYSIS. As depicted below, two chemical species react within a lake. Mass balances for the two reactants can be written as

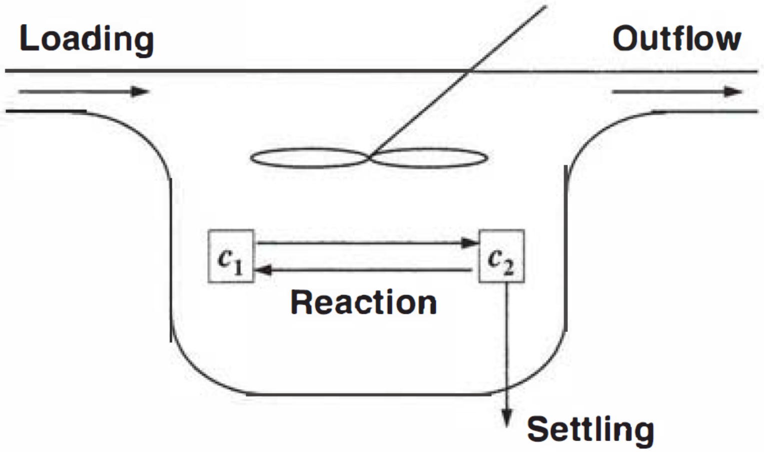


FIGURE E18.1

$$V\frac{dc_1}{dt} = Qc_{\text{in}} - Qc_1 - k_{12}Vc_1 + k_{21}Vc_2$$

$$V\frac{dc_2}{dt} = -Qc_2 + k_{12}Vc_1 - k_{21}Vc_2 - v_sA_sc_2$$

where V = volume of lake = 50,000 m^3
Q = inflow = outflow = 50,000 m^3 d^{-1}
A_s = surface area of lake sediments = 16,667 m^2
k_{12} = first-order conversion rate of c_1 to c_2 = 1 to 4 d^{-1}
k_{21} = first-order conversion rate of c_2 to c_1 = 0.5 to 0.7 d^{-1}
v_s = settling velocity of c_2 = 0 to 1 m d^{-1}

The ranges for the parameters connote literature values. The term c_{in} represents the concentration of the inflowing stream (mg L^{-1}), which during the period of study has a mean value of approximately 10 mg L^{-1}. Use this information to estimate the sensitivity of the model to the three parameters k_{12}, k_{21}, and v_s. Employ first-order sensitivity analysis and express your results as a table of condition numbers.

Model Sensitivity: Parameter Perturbation & 1st Order Analysis

One problem with the previous example is that the uncertainty of the parameter estimates were not considered. Some parameters are known more definitively than others. A solution is to propagate the range for the parameters through either the perturbation or the first-order approach.

This can be done by using $\frac{\Delta c}{c} = CN_k \frac{\Delta k}{k}$ to propagate the relative error in each parameter into the resulting error in the prediction.

EXAMPLE 18.2. UNCERTAINTY ANALYSIS. For the case from Example 18.1, determine the expected uncertainties of the concentrations due to the uncertainties of the parameters.

Monte Carlo Analysis

In MCA, the distribution of the parameters is characterized rather than given a prescribed range.

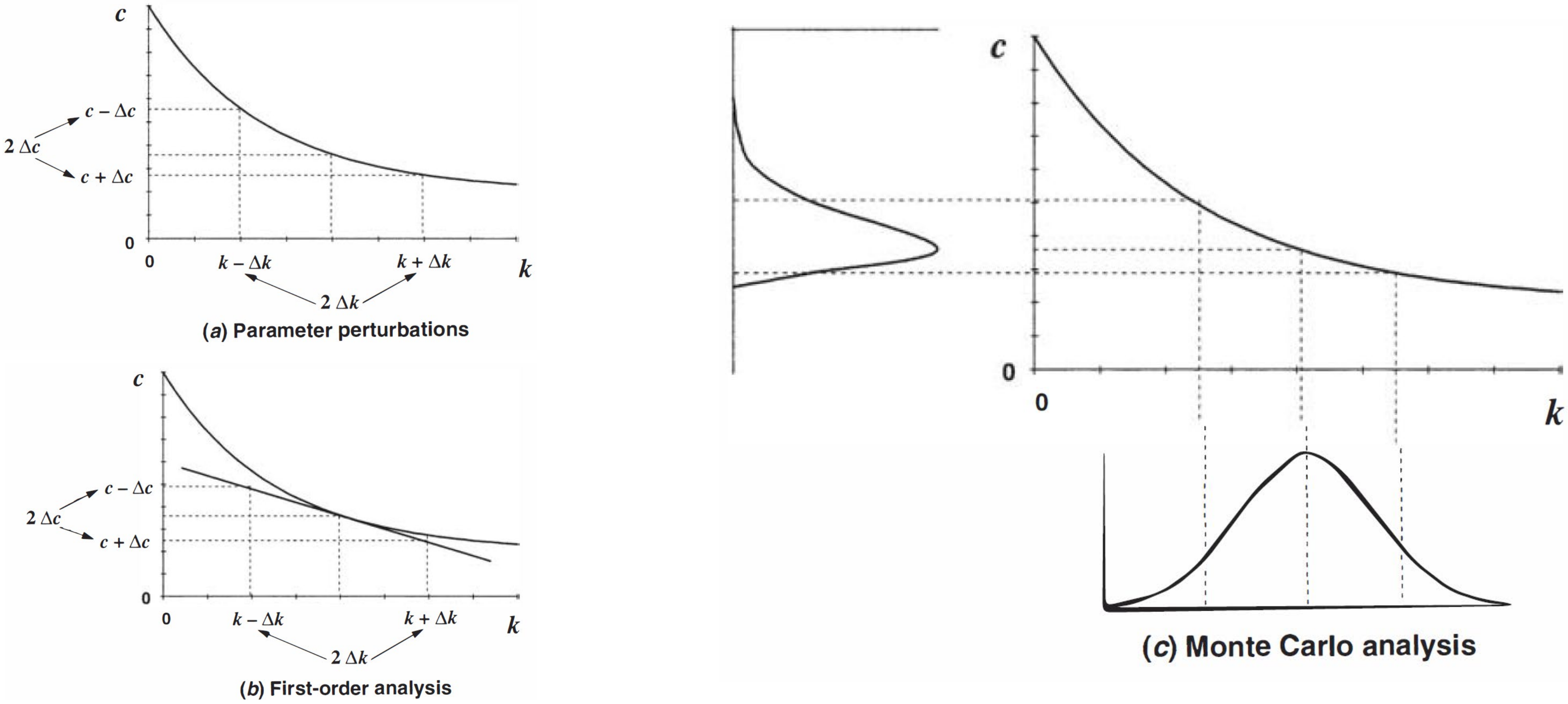


FIGURE 18.4
Graphical depictions of three methods for assessing model sensitivity.

Monte Carlo Analysis

Here are a few ideal probability distributions that are commonly used to describe parameter variability in water-quality modeling.

Uniform distribution assumes equal probability of occurrence between each bound. **Normal** and **triangular** demonstrate the most likely value at the center.

These centered distributions can also be off-centered or “skewed”.

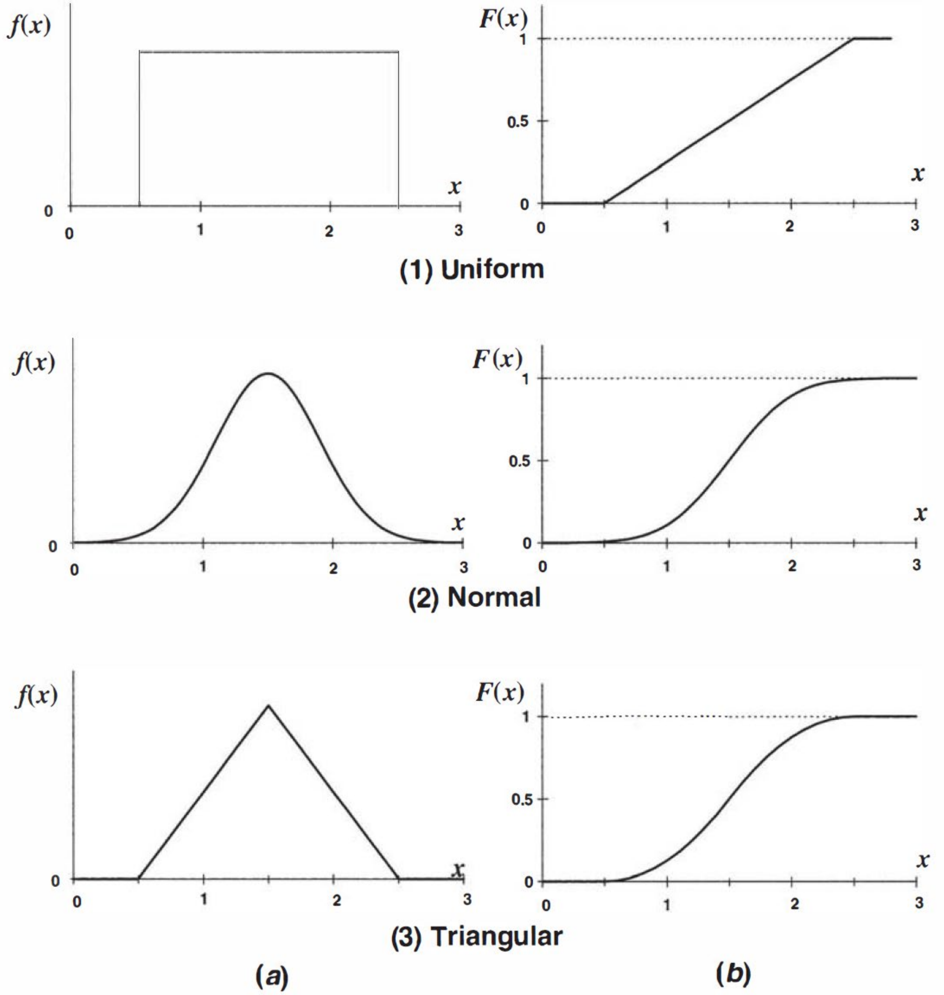


FIGURE 18.5
(*a*) Probability density and (*b*) cumulative distribution functions for three distributions commonly used to characterize uncertainty of water-quality modeling parameters: (1) uniform, (2) normal, and (3) triangular distributions.

Monte Carlo Analysis

The cumulative distribution function (c.d.f.) represents the integral of the p.d.f.,

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

This integral specifies the probability that the parameter will be less than x. The two distributions are related inversely by differentiation.

$$f(x) = \frac{dF(x)}{dx}$$

i.e. the frequency of occurrence is equal to the rate of change of the c.d.f. with respect to the rate of change of the parameter.

These distributions for parameters are used to simulate distributions of concentration.

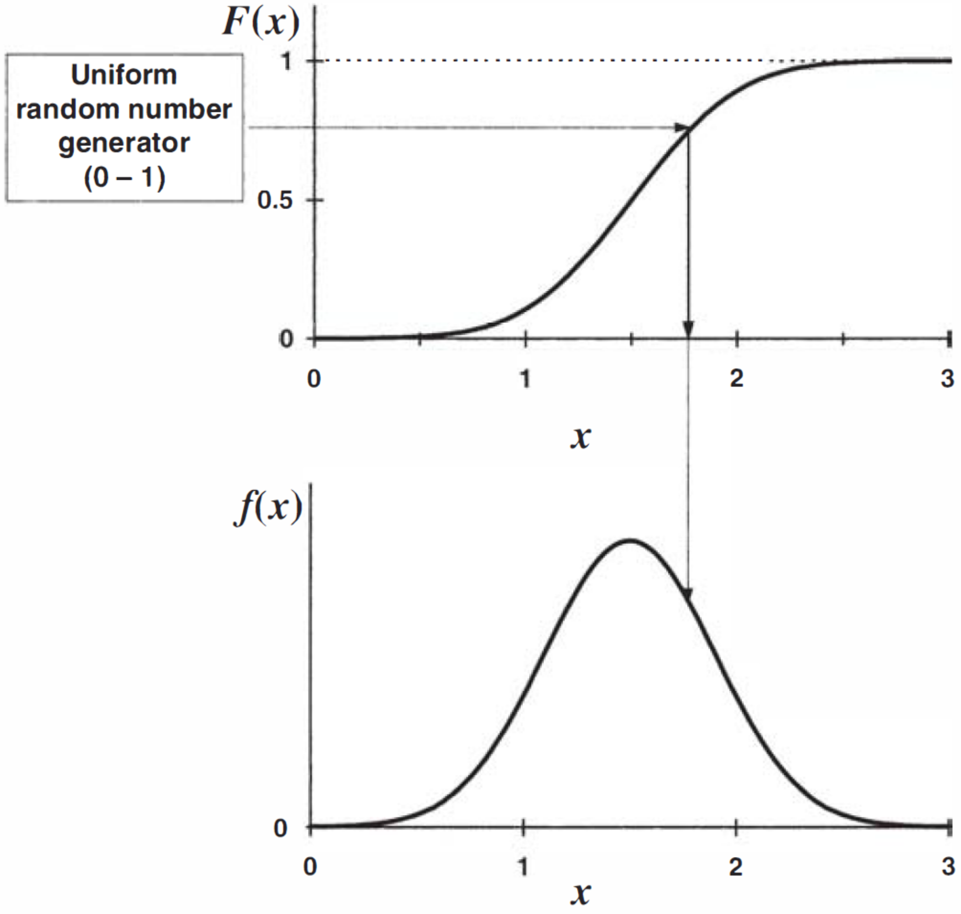


FIGURE 18.6
Graphical depiction of how the c.d.f. is used in conjunction with random numbers to generate values of model coefficients in Monte Carlo simulation.

EXAMPLE 18.3. MONTE CARLO ANALYSIS. For the case from Example 18.1, determine the expected uncertainties of the concentrations due to the uncertainty of the parameter k_{12}. Assume that the parameter follows the triangular distribution

$$f(x) = 0.4444(x - 1) \qquad 1 \leq x < 2.5$$

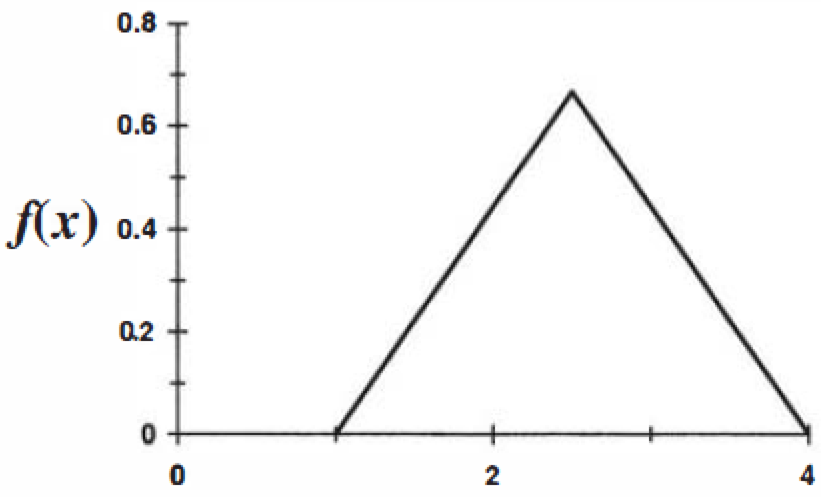
$$f(x) = -0.4444(x - 4) \qquad 2.5 \leq x < 4$$

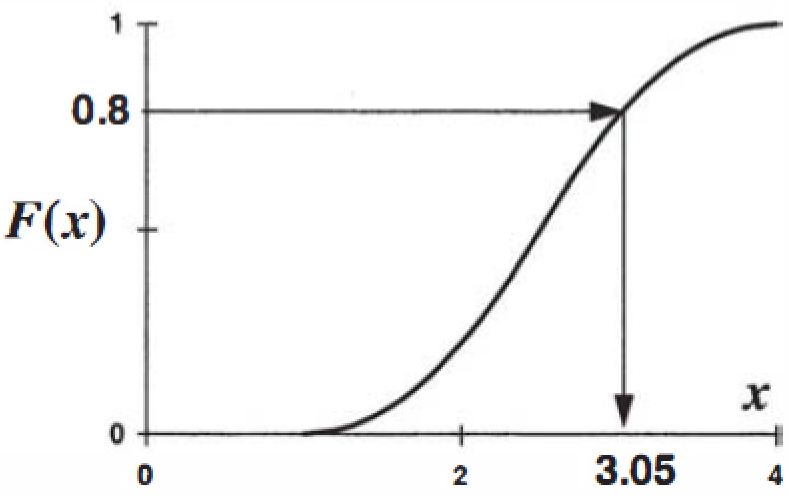
These can be integrated to develop equations for the c.d.f.,

$$F(x) = 0.2222(x - 1)^2 \qquad 1 \leq x \leq 2.5$$

$$F(x) = -0.4444(0.5x^2 - 4x + 5.75) \qquad 2.5 \leq x \leq 4$$

Both functions are displayed below:



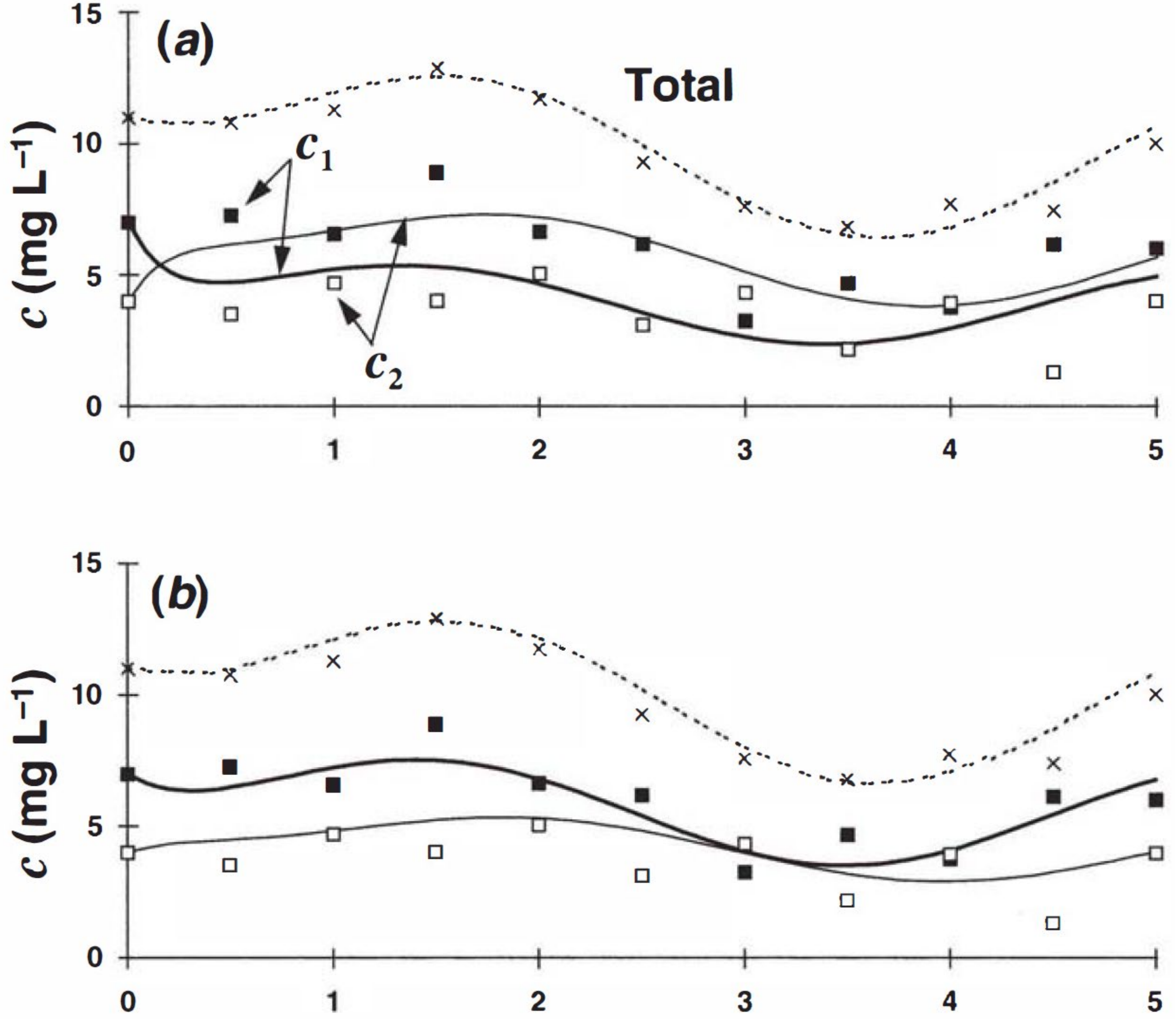


Assessing Model Performance

Previously we describe the aim of calibration as obtaining the best fit. Assessing this best fit can be subjective and objective.

Subjective assessment is based on visual comparison of simulation with data.

Objective assessment involves some quantitative measure of the quality of fit (usually a measure of error).



Assessing Model Performance

There are several measures that can be developed to assess fit. The follow is a focus on minimizing the sum of squares of the residuals:

$$S_r = \sum_{i=1}^{n} \left(c_{p,i} - c_{m,i}\right)^2$$

where $c_{p,i}$ = the *i*th model prediction of concentration and $c_{m,i}$ = *i*th measured concentration.

One can use trial-and-error or numerical optimization methods. Each has value.

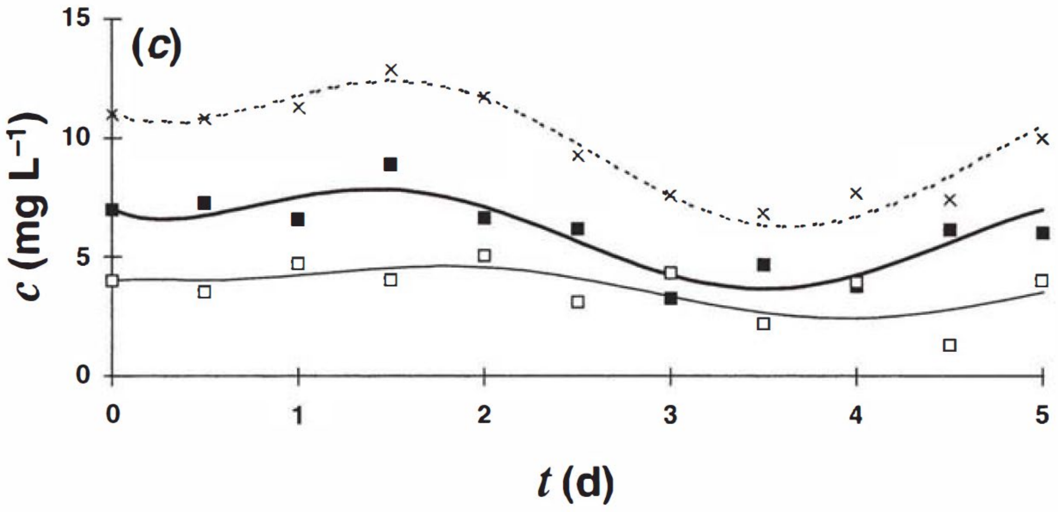


FIGURE 18.7
Calibration. (*a*) The original fit using average values for the parameters; (*b*) improved fit after using the measured value for k_{12}; (*c*) final calibration.

EXAMPLE 18.4. MODEL CALIBRATION. For the same system studied in Example 18.1, the following data were collected during a 5-d sampling survey:

t (d)	c_1 (mg L^{-1})	c_2 (mg L^{-1})	c_T (mg L^{-1})
0.0	7.0	4.0	11.0
0.5	7.3	3.5	10.8
1.0	6.6	4.7	11.3
1.5	8.9	4.0	12.9
2.0	6.6	5.0	11.7
2.5	6.2	3.1	9.3
3.0	3.3	4.3	7.6
3.5	4.7	2.2	6.8
4.0	3.75	3.9	7.7
4.5	6.1	1.3	7.4
5.0	6.0	4.0	10.0
Mean	6.03	3.64	9.67

The concentration of the inflowing stream c_{in} during the period of study can be represented by the sine function,

$$c_{in} = 10 + 5\sin\left(\frac{\pi}{2}t\right)$$

In addition in situ experiments were performed at two times during the study to estimate k_{12} directly. These yielded estimates of 1.05 d^{-1} and 1.55 d^{-1} for that parameter. Use this information (and any other techniques at your disposal) to estimate the three parameters k_{12}, k_{21}, and v_s.

Segmentation and Model Resolution

Segmentation is the process of dividing space and matter into increments. e.g. into volumes with mass-balance equations, each with different chemical and biological forms.

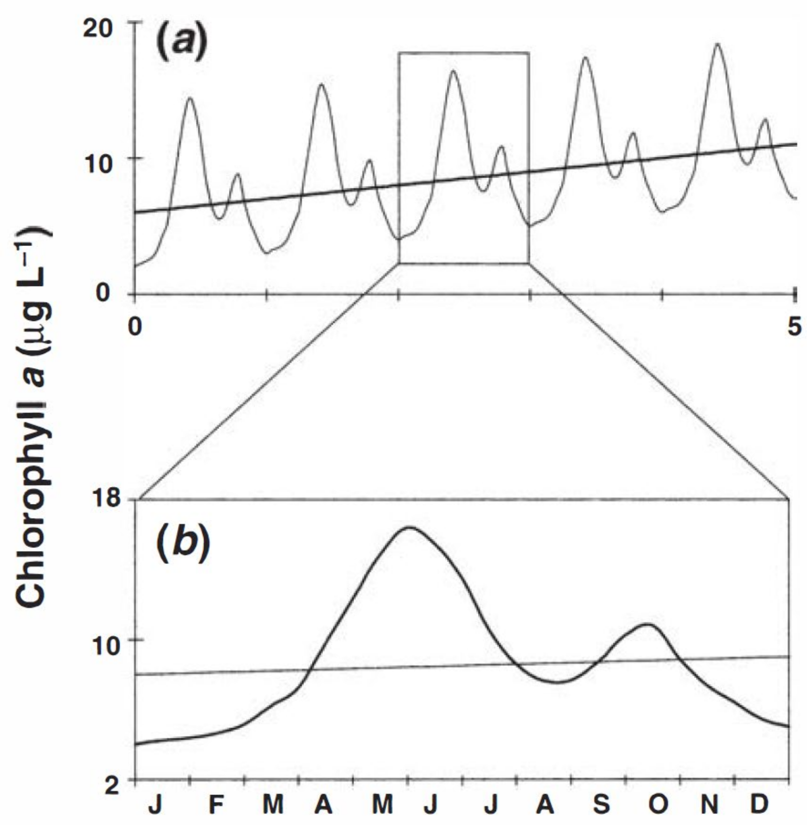


FIGURE 18.8
Plot of chlorophyll *a* concentration (μg L^{-1}) versus time. (*a*) Depiction of the long-term trend (straight line) along with the underlying seasonal cycle (wavy line); (*b*) depiction of the seasonal cycle (wavy line) along with the underlying long-term trend (straight line).

There is also a temporal aspect to segmentation which can increase in temporal focus by using shorter “finite periods” or time steps for the mass-balance computation. See graph showing variability and scale.

Segmentation and Model Resolution

The degree to which space, time, and matter are segmented is called ***model resolution***. Analogous to photography, where the lens is adjusted to bring different parts of the field of view into focus.

At times the foreground is important, other times distant details are of interest.

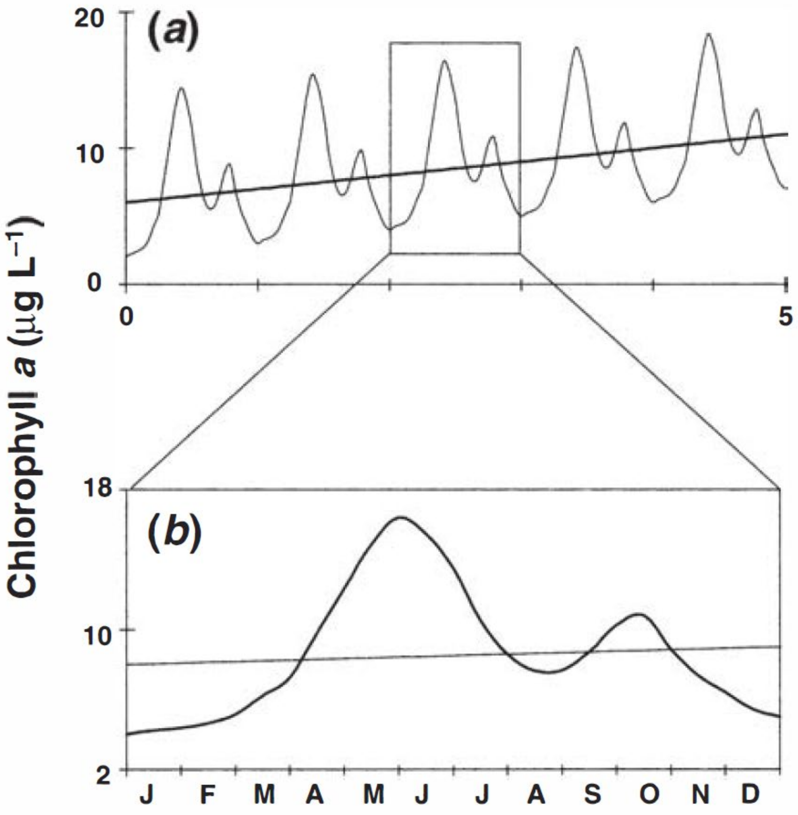


FIGURE 18.8
Plot of chlorophyll *a* concentration ($\mu g\ L^{-1}$) versus time. (*a*) Depiction of the long-term trend (straight line) along with the underlying seasonal cycle (wavy line); (*b*) depiction of the seasonal cycle (wavy line) along with the underlying long-term trend (straight line).

Segmentation and Model Resolution

Physical characteristics of the system can dictate the required level of segmentation.

The issue of concern for a planner may be such that spatial and temporal aggregation results in negligible loss of relevant planning information.

Other times (e.g. bacterial contamination of beaches, might require finer scale approach).

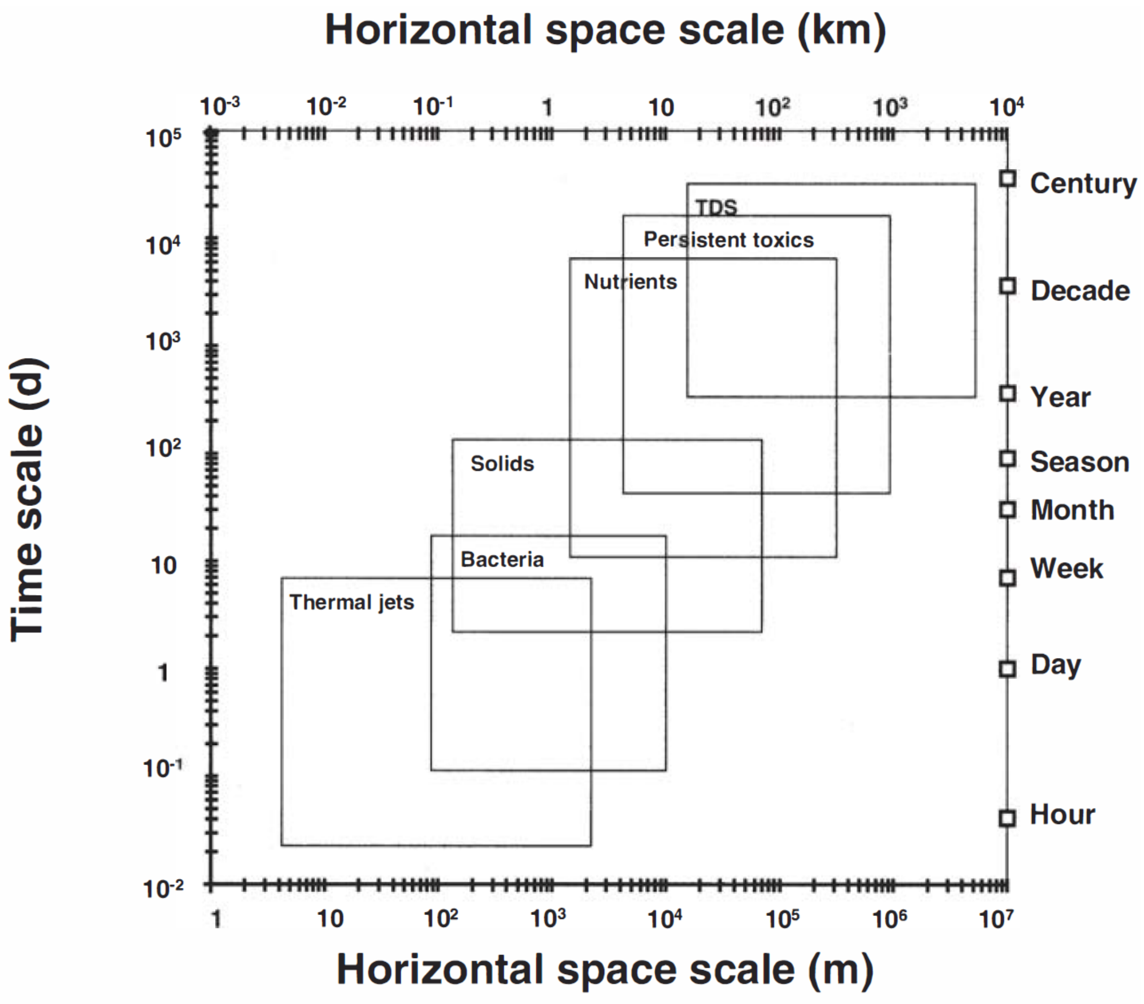


FIGURE 18.9
Approximate time and space scales of water-quality problems.