Load Estimation Techniques

7

A pollutant *load* is the mass or weight of pollutant transported in a specified unit of time from pollutant sources to a waterbody. The loading rate, or *flux*, is the instantaneous rate at which the load is passing a point of reference on a river, such as a sampling station, and has units of mass/time such as grams/second or tons/day (Richards, 1997). Mathematically, the load is the integral over time of the flux.

Pollutant load estimation is a fundamental element in the development of many watershed management plans. Reliable estimates of the quantity of pollutants delivered from various sources within a watershed are needed to develop a watershed plan that will address the identified water quality problems or issues. Establishing the link between an identified water quality problem and the sources causing the problem often entails a mass balance analysis, a quantitative accounting of the sources and sinks of the pollutants of interest.

There are many reasons for developing management plans, including the development and implementation of a total maximum daily load (TMDL) pursuant to the requirements of section 303(d) of the Clean Water Act (see Highlight). For those waters either not supporting or not projected to support designated uses even after the implementation of point source or other required pollution controls, a TMDL is needed. The components of TMDL development are:

1. Problem Identification
2. Identification of Water Quality Indicators and Target Values
3. Source Assessment
4. Linkage Between Water Quality Targets and Sources
5. Allocation
6. Follow-up Monitoring and Evaluation Plan
7. Assembling the TMDL

It is important to note that TMDL development is a very site-specific process. Therefore, these components are not necessarily sequential steps but can be conducted concurrently or iteratively depending upon the situation (EPA, 1999b).

In source analysis for a TMDL, the relative contributions of different sources are assessed. An estimate of pollutant loads from both point sources and nonpoint sources is essential to this analysis, as is the ability to determine if the load reduction needed to meet water quality standards can be achieved under different management scenarios (e.g., implementation of the management measures). The load allocation for nonpoint sources (and the wasteload allocation for point sources) is determined from an analysis that links the desired endpoints (e.g., achievement of a water quality standard) to various management alternatives that could be applied to the identified sources.

Clean Water Act
Total Maximum Daily Load (TMDL) Program

Section 303(d) of the Clean Water Act and EPA's implementing regulations at 40 CFR Section 130.7 require States to develop TMDLs for their waterbodies that do not or are not expected to meet applicable water quality standards after the application of technology-based point source or other required pollution controls. EPA's regulations at 40 CFR Section 130.2 define some of the elements of the TMDL programs. These include:

- **Loading capacity** – The greatest amount of loading that a water can receive without violating water quality standards.
- **Load allocation** – The portion of a receiving water's loading capacity that is attributed either to one of its existing or future nonpoint sources of pollution or to natural background sources.
- **Wasteload allocation** – The portion of a receiving water's loading capacity that is allocated to one of its existing or future point sources of pollution.
- **Total maximum daily load (TMDL)** – The sum of the individual wasteload allocations for point sources, load allocations for nonpoint sources, natural background, and a margin of safety. TMDLs can be expressed in terms of either mass per time, toxicity, or other appropriate measure that relate to a State's water quality standard. A margin of safety is required as part of each TMDL to account for the uncertainty about the relationship between the pollutant loads and the quality of the receiving waterbody.
- **Water quality-limited segments** – Those water segments that do not or are not expected to meet applicable water quality standards by the next listing even after the application of technology-based effluent limitations for point sources as required by sections 301(b) and 306 of the Clean Water Act. Technology-based controls include, but are not limited to, best practicable control technology currently available and secondary treatment.
- **Margin of Safety** – Element of a TMDL that accounts for uncertainty and lack of knowledge. A margin of safety may be expressed as unallocated assimilative capacity or conservative analytical assumptions used in establishing the TMDL and its maximum allowable pollutant load.

EPA Protocols for TMDL Development

Protocol for Developing Pathogen TMDLs: First Edition, January 2001, EPA 841-R-00-0002.
www.epa.gov/owow/tmdl/pathogen_all.pdf

Protocol for Developing Nutrient TMDLs: First Edition, November 1999, EPA 841-B-99-007.
www.epa.gov/owow/tmdl/nutrient/pdf/nutrient.pdf

Protocol for Developing Sediment TMDLs: First Edition, October 1999, EPA 841-B-99-004.
www.epa.gov/owow/tmdl/sediment/pdf/sediment.pdf

The following sections present some basic information regarding monitoring and modeling to estimate pollutant loads. References to more detailed treatments of the topics are included as well. Additional information on TMDL is available at *www.epa.gov/owow/tmdl/*.

Estimating Pollutant Loads Through Monitoring

Every monitoring effort should have clearly stated objectives. The estimation of pollutant loads is a general objective that should be refined to clarify the monitoring needs. The specific reasons why the pollutant loads are to be estimated could affect decisions regarding the required precision and the conditions under which monitoring should be conducted. For example, if the pollutant is bacteria and the watershed management concerns are associated with the instantaneous value and the 30-day geometric mean (of 5 or more samples), then the sampling protocol should consider multiple samples at a sufficient frequency to calculate the geometric mean as well as evaluate the various conditions under which loading occurs (wet and dry weather). On the other hand, if nutrients are causing accelerated eutrophication in a reservoir then it may only be important to estimate seasonal loads. The time scales and frequency of monitoring needed will be a function of the critical conditions and the receiving water response to the loading of the pollutant of concern.

The averaging period for loading estimates may be hourly, daily, monthly, or longer depending upon site-specific conditions and needs. The variability of loads within the average period of interest and the certainty with which water quality standards violations need to be documented will drive decisions regarding sampling design and frequency. The importance of clearly stated objectives is described more fully in existing monitoring guides (EPA, 1997a; EPA, 1991c; USDA-NRCS, 1996b). Due to the importance of statistical considerations, those designing monitoring plans are strongly encouraged to seek assistance from a trained statistician with experience in water monitoring.

Components of a Load

To estimate pollutant loading, it is necessary to sum the flux, which is commonly expressed as mass per unit time, over the period of interest. Since the flux varies with time, this summing process can be expressed in integral form as shown in the first equation of the following text box. Since flux cannot be measures directly, flux is often expressed as the product of concentration and flow (see second equation of the text box). Thus the three basic steps for estimating pollutant load are:

- ❒ measuring water discharge (e.g., cubic meters per second),
- ❒ measuring pollutant concentration (e.g., milligrams per liter), and
- ❒ calculating pollutant loads (multiplying discharge times concentration over the time frame of interest).

Since concentration and flow vary with time, the key challenge in measuring loads is to determine when to sample to obtain the best estimate at least cost. Richards (1997) points out that it is not uncommon for 80 to 90% or more of the annual load to be delivered during the 10% of the time which corresponds with high fluxes. Depending on the constituent being evaluated, fluxes during snowmelt and storm events are often many times greater than those during periods of low flow (i.e., dry weather conditions). Thus, monitoring programs must be

Load and Flux

The pollutant load is the integral over time of the flux:

$$Load = k \int_t flux(t)\, dt$$

where k is a constant for converting units, and t is time.

Since we cannot measure flux directly, we measure it as the product of concentration and discharge.

$$Load = k \int_t c(t) q(t) dt$$

where c(t) is the concentration at time=t, and q(t) is the water discharge at time=t.

designed with full consideration given to both periods of pollutant flux. The following equations present the mathematical relationship between load, flux, and time.

Measuring Water Discharge

The major options for monitoring stream discharge are flumes, weirs, natural channels, and existing structures (USDA-NRCS, 1996b; Brakensiek et al., 1979). Device selection for *stream discharge* is a function of site-specific conditions such as slope, sediment load, and stream size. Selection of a device for *runoff* measurement depends on peak runoff rate, runoff variability, the extent to which trash and debris are carried in the runoff, icing conditions, and other factors (Brakensiek et al., 1979). Discharge monitoring approaches, and the selection, implementation, and use of various devices are described by Brakensiek et al. (1979) and USDA-NRCS (1996b).

For established gaging stations, flow measurements are relatively inexpensive to make, and are available almost on a continuous basis (Richards, 1997). It is, however, likely that gaps in the flow record will still occur as a result of equipment failure, operational errors, or extreme flow events. Methods to fill gaps in flow records are described by Brakensiek et al. (1979) and USGS (Rantz et al., 1982).

Measuring Pollutant Concentration

Periodic measurements of pollutant levels in water are used in load estimation. The frequency of the measurements required to adequately characterize pollutant concentrations over time is often difficult to determine. Pollutants such as nitrate-nitrogen often do not vary greatly over weekly or monthly intervals while pollutants such as fecal coliform can vary by several orders of magnitude during a week depending on hydrologic and other conditions. The vast majority of nonpoint source load estimations will require storm event sampling. The choice of sampling frequency for load estimation is a complex function of watershed hydrology, pollutant(s) of interest, land use/management, the duration of monitoring and the water resource type. Periodic measurements in the field (in situ or sample analysis with a field kit) or laboratory measurements performed on collected water samples are typically used to provide the pollutant concentration values that will be used in load estimation.

Water sampling approaches have been categorized in several ways, some based more upon the equipment used, and others based more upon the statistical design employed (USDA-NRCS, 1996b; EPA, 1979; EPA, 1991c). Grab, point, composite, integrated, continuous, random, systematic, and stratified sampling are frequently described in the literature. In practice, sampling involves a decision regarding the population and population units to be sampled (e.g., instantaneous concentration at single point or integrated over depth, average concentration at single point or integrated over depth for a specified time interval or flow interval), a determination of the statistical approach to be used (e.g., simple random sampling, stratified random sampling, systematic sampling), and a choice of sampling equipment and configuration (e.g., grab sample taken manually or automatically with a mechanical sampler, time-weighted or flow-weighted sampling with a programmed mechanical sampler).

For any given watershed, the best approach for estimating loads will be determined based upon the needs and characteristics of the watershed. Still, some general rules-of-thumb should be considered (USDA-NRCS, 1996b; Richards, 1997).

- ❒ **Accuracy and precision increase with increased frequency of sampling.**
- ❒ **Grab, Point, or Instantaneous Samples** — may be insufficient to determine loads unless concentrations are correlated to discharge which is measured continuously.
- ❒ **Depth-Integrated and Width-Integrated Grab Samples** — can account for stratification in concentration with depth or horizontally across a stream, but still depends upon correlation to discharge for suitability in load estimation.
- ❒ **Time-Weighted Composite Samples** — not generally sufficient for load estimation since they may not adequately reflect changes in discharge and concentration during the period over which samples are composited.
- ❒ **Flow-Weighted Composite Samples** — well-suited to load estimation, but difficult to collect since stage-discharge relationship is needed and a “smart sampler” is needed to trigger sampling as a function of flow rate. Projecting sample size and number of bottles needed is difficult.
- ❒ **Systematic Sampling** — as efficient as, or more efficient than, simple random sampling if the sampling interval is not equal to a multiple of any strong period of fluctuation in the sampled population (e.g., sampling weekly on the day when a particular pollutant is always at its peak level due to scheduling by a discharger).

- ❒ **Stratified Random Sampling** — with most samples taken during periods of high flow, can be of great importance in providing increased precision for a given number of samples.

Types of Water Samples

Grab Sample — A single sample taken at one place a single time.

Composite Sample — A series of grab samples, usually collected in the same location but at different times, combined to form one sample for analysis. Composite samples are usually:

Flow-Weighted – Sample is taken after a specified quantity of water has passed the monitoring station (e.g., draw 10 ml sample every 750,000 liters of flow); or

Time-Weighted – A pre-determined sample volume is taken at a predetermined time interval (e.g., draw 10 ml sample every 15 minutes).

Integrated Sample — Subsamples are taken at various depths or distances from the stream bank, and integrated into a single sample.

Continuous Sample — Probes are used to continuously record contaminant concentration in stream. Not widely applicable to nonpoint source programs.

For many TMDLs, the daily pollutant load may be the population unit of greatest importance. In these cases, sampling should emphasize obtaining accurate estimates of daily loads for the pollutant of interest. Since TMDLs establish maximum wasteload and load allocations that can be discharged without violating water quality standards, the monitoring effort should provide the data necessary for determining whether or not quality standards are met. For example, if water quality standards are more likely violated under low-flow (dry weather) conditions, then the monitoring should provide reliable data regarding low-flow loads. Conversely, in cases where water quality standards are violated during high-flows (wet weather or snowmelt) or as a result of loads from high flows, the monitoring should emphasize high-flow monitoring. In other cases, such as those in which annual or seasonal loads are critical, high quality estimates of low-flow and high-flow loads may be equally important.

Sampling location should be determined based upon the monitoring objectives, water resource characteristics, and source characteristics. For example, it may be appropriate to sample at the outlets of tributaries to a lake, or above and below a farm or set of farms, depending upon whether the objective is to estimate lake loading from tributary watersheds or stream loading from an individual farm or farms. Additional information regarding sampling location can be found in existing guides (EPA, 1997a; USDA-NRCS, 1996b; Ponce, 1980).

Detailed discussions of statistical sampling approaches (e.g., random sampling) can be found in several sources (EPA, 1997a; Richards, 1997; USDA-NRCS, 1996b; Gilbert, 1987). Older sampling equipment is described by Brakensiek, et al. (1979), while USDA-NRCS (1996b) provides an overview of more current devices, including a helpful list of references regarding sampling equipment.

Calculating Pollutant Loads

The pollutant load is the integral of flux over time, but flux cannot be measured directly (Richards, 1997). In Figure 7-1 the flux is calculated as the product of concentration and discharge, with appropriate conversion units. Each calculated flux is a discrete value that is assumed to apply across the sampling interval, which is 24 hours in this hypothetical example (daily composites). The cumulative load in Figure 7-1 is determined by adding the calculated fluxes over all sampling intervals.

Because there will be more discharge data than concentration data in almost all chemical monitoring efforts, there will be a need to make estimates of concentration, and therefore pollutant flux, for periods between water quality observations (Richards, 1997). Figure 7-2 illustrates how missing values can greatly affect the calculated load estimates. Load A is the same load as shown in Figure 7-1, whereas Load B was calculated after deleting every other concentration value used to calculate Load A.

Data gaps can be filled by estimating missing concentration values for pairing with the flow data, or by adjusting the load estimate made from the observations where both flow and concentration were measured (Richards, 1997). Flow data typically form the basis for making flux estimates for periods during which water quality (concentration) data are lacking.

Some of the methods for estimating pollutant loads include numeric integration, the worked record procedure, averaging approaches, the flow interval technique, ratio estimators, regression approaches, and flow-proportional sampling (Richards, 1997). A review of evaluative studies of loading approaches has resulted in the following points of consensus (Richards, 1997):

- **Averaging methods** (e.g., for monthly or quarterly loads) are generally biased, and the bias increases as the size of the averaging window increases and/or the number of samples decreases. For example, an annual load determined by adding four quarterly loads will generally be

Figure 7-1. Flux and cumulative load over time.

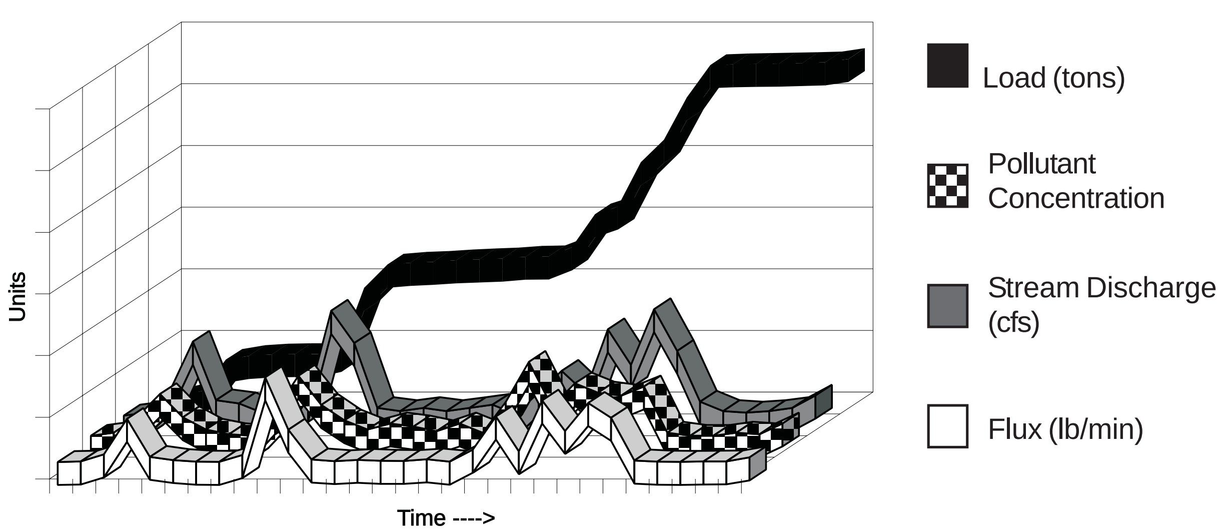
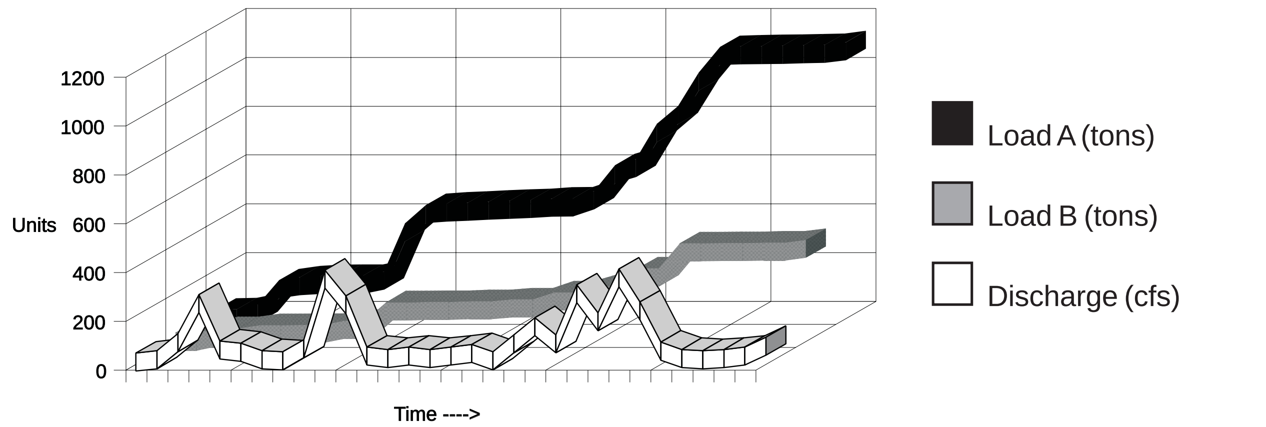


Figure 7-2. Effect of missing concentration data.



more biased than an annual load determined by adding 12 monthly loads.

- In most studies, **ratio approaches** performed better than **regression approaches**, and both performed better than **averaging approaches**.
- **Regression approaches** can perform well if the relationship between flow and concentration is well-defined, linear throughout the range of flows, and constant throughout the year.

Greater detail and illustrative examples regarding averaging approaches, regression approaches, ratio estimators, and sampling approaches can be found in Richards (1997).

Estimating Pollutant Loads Through Modeling

Types of Models Available

Water runoff, sediment delivery, and nutrient loading can be estimated using watershed models. Match modeling objectives, staff expertise, data requirements, and available budget for proper model selection.

Loading models include techniques primarily designed to predict pollutant movement from the land surface to waterbodies (EPA, 1997d). *Watershed loading models* range from simple loading rate assessments in which loads are a function of land use type only, to complex simulation techniques that more explicitly describe the processes of rainfall, runoff, sediment detachment, and transport to receiving waters. Some loading models operate on a watershed scale, integrating all loads within a watershed, and some allow for the subdivision of the watershed into contributing subbasins.

Field-scale models, which have traditionally specialized in agricultural systems, are loading models that are designed to operate on a smaller, more localized scale. Field-scale models have often been employed to aid in the selection of management measures and practices. For example, a dynamic simulation model was used to predict the long-term patterns of phosphorus export from fields under a variety of management scenarios (Cassell and Clausen, 1993). The process model simulated the annual inputs and outputs of phosphorus, and was determined by the authors to be useful for simulating long-term patterns. Process models such as this one, however, are dependent upon local export coefficients and a thorough understanding of pollutant transport processes.

Methods for Estimating Pollutant Loads (Richards, 1997)

Numeric Integration — Total load is calculated as the sum of the individual loads calculated for each sample.

Worked Record Procedure — Chemical observations are plotted onto a detailed hydrograph, and smooth curves are drawn through chemical data points based upon analyst's experience with the relationship of concentration and flow.

Averaging Approaches — Calculation that uses averaging of concentration and/or flow to estimate loads. For example, analyst might multiply average weekly suspended solids concentration by daily flow to estimate daily loads for the week.

Flow Interval Technique — Semi-graphical technique that calculates "interval loads" as the product of average flux for a range of daily flow values times the number of days in which flows were within the particular flow range.

Ratio Estimators — Total loads are estimated using a known relationship between the less-frequently sampled parameter of interest and a more-frequently sampled parameter (e.g., discharge) to fill gaps in the data record for the parameter of interest.

Regression Approaches — Relationship is established between concentration and flow based on samples taken, and then applied to estimate concentration for days not sampled.

Flow-Proportional Sampling — Mechanical approach in which representative samples are taken to determine concentration for a known discharge. Pollutant load is calculated as the sum of the sample concentrations multiplied by the measured discharge.

Other types of models include *receiving-water models*, which emphasize the response of a waterbody to pollutant loadings, flows, and ambient conditions, and *ecological models* that simulate biological communities and their response to stressors such as toxics and habitat modification (EPA, 1997d). *Integrated modeling systems* link models, data, and a user interface within a single system. The advent of geographic information systems (GIS) has facilitated the development of and expanded the capabilities of integrated modeling systems.

The emphasis of this section will be on watershed loading models. The reader is encouraged to seek additional information regarding field-scale, ecological, and integrated models in existing documents (EPA, 1997d; EPA, 1992b). The reader can also consult Chapter 5 of this manual for information on *field-* and *watershed-scale* models.

Watershed Loading Models

Watershed loading models are configured and characterized in several ways (see Modeling Jargon), but they can be grouped into three general categories: *simple methods, mid-range models*, and *detailed models* (EPA, 1997d). The defining characteristics of models are the degree to which processes (and complexities of systems) are simplified and the time scale that is used for analysis and display of output information.

Simple methods are generally used to provide quick and easy identification of critical pollutant sources in the watershed. Detailed watershed models represent the other extreme, featuring costly and time-consuming efforts to provide quantitative estimates of pollutant loads from a range of management alternatives. Richards (1997) cautions that modeling of agricultural settings is often inadequate to evaluate the success of management practices in reducing loads because there are mixed land uses that change annually and these land uses have different loading rates. An additional concern is that most models fail to adequately address stream channel and bank dynamics, including the impact of management practices on these factors. Some detailed models such as GLEAMS, however, attempt to capture the variability associated with cropping practices and rotations in the agricultural setting.

Mid-range watershed models are generally midway between the cost, complexity, and accuracy of simple methods and detailed watershed models. Mid-range models provide qualitative estimates of management alternatives (EPA, 1997d).

Figure 7-3 shows examples of models and integrated modeling systems for load estimation. EPA's *Compendium of Tools for Watershed Assessment and TMDL Development* has additional details regarding the capabilities, limitations, and data requirements for these and other models (EPA, 1997d).

Simple Watershed Methods

Uses

- ❐ Support assessment of relative significance of sources
- ❐ Guide decisions for management plans
- ❐ Focus continuing monitoring efforts

Features

- ❐ Typically derived from empirical relationships between physiographic characteristics of the watershed and pollutant export
- ❐ Often applied using a spreadsheet or hand-held calculator

Pros

- ❐ Rapid
- ❐ Minimal data requirements (large-scale aggregation; low resolution)
- ❐ Minimal effort

Cons

- ❐ Output is typically mean annual values or storm loads
- ❐ Rough estimates of loadings
- ❐ Very limited predictive capability
- ❐ Low transferability to other regions due to empirical basis
- ❐ Do not consider degradation and transformation processes
- ❐ Few incorporate detailed representation of pollutant transport within and from watershed
- ❐ Cannot adequately account for most management practices

Figure 7-3. Load estimation models.

Simple Methods

- EPA Screening
- Simple Method
- Regression Method
- SLOSS-PHOSPH
- Federal Highway Administration Model
- Watershed Mangement Model

Mid-Range Models

- SITEMAP
- GWLF
- Urban Catchment Model
- Automated Q-ILLUDAS
- AnnAGNPS
- SLAMM

Detailed Models

- STORM
- DR3M-QUAL
- SWRRBWQ
- SWMM
- HSPF

Field-Scale Loading Models

- CREAM/GLEAMS
- Opus
- WEPP

Integrated Modeling Systems

- PC-VIRGIS
- WSTT
- LWMM
- GISPLM
- BASINS

Mid-Range Watershed Models

Uses

- ❐ Assist in defining target areas for pollution mitigation programs on watershed basis
- ❐ Support relative comparisons of management alternatives

Features

- ❐ Compromise between empiricism of simple methods and complexity of detailed mechanistic models
  - Use simplified relationships for the generation and transport of pollutants
  - Greater reliance on site-specific data than for simple methods
  - Can address land use patterns and landscape configurations in watersheds
- ❐ Typically require some calibration with additional data sets
- ❐ Often tailored to site-specific applications (e.g., agriculture only)

Pros

- ❐ Can assess seasonal or inter-annual variability of loadings, and long-term water quality trends
- ❐ Those with continuous simulation can compare storm-driven loads over a range of storm events or conditions
- ❐ Those with GIS interface facilitate parameter estimation
- ❐ Relatively broad range of regional applicability
- ❐ Usually include detailed input-output features to simplify processing
- ❐ Often have built-in graphical and statistical capabilities

Cons

- ❐ Use of simplifying assumptions can limit accuracy of predictions
- ❐ Most do not consider degradation and transformation processes
- ❐ Few incorporate detailed representation of pollutant transport within and from watershed
- ❐ Can not account for most management practices

Detailed Watershed Models

Uses

- ❒ If properly applied, can provide accurate estimates of pollutant loads and impacts on water
- ❒ Identify causes of problems rather than simply describing overall conditions

Features

- ❒ Use storm event or continuous simulation to predict flow and pollutant concentrations for a range of flow conditions (small calculation time steps)
- ❒ Algorithms more closely simulate the physical processes of infiltration, runoff, pollutant accumulation, instream effects, and ground/surface water interaction

Pros

- ❒ Input/output have greater spatial and temporal resolution than simple and mid-range models
- ❒ Detailed hydrologic simulations can be used to design potential control actions
- ❒ Linkage to biological modeling is possible
- ❒ Those with new interfaces and GIS linkages facilitate use of models
- ❒ Provide relatively accurate predictions of variable flows and water quality at any point in a watershed if properly applied and calibrated

Cons

- ❒ Considerable time and expenditure required for data collection and model application
- ❒ Complex — not easily utilized by untrained staff
- ❒ Require rate parameters for flow velocities, settling, decay, and other processes
- ❒ Input data file preparation and calibration require professional training and adequate resources

Planning and Selection of Models

Setting modeling objectives should be the first step in developing a modeling approach. In some cases, the objectives may be achievable using a simple model, but in other cases it may be necessary to perform complex modeling involving more than one model. Criteria that apply in selecting a model may include the value of the resource under consideration, data needs, hardware needs, cost, accuracy required, type of pollutants/stressors, management considerations such as long-term commitment to the modeling effort, availability of trained personnel, user experience with the model, and acceptance of the model (EPA, 1997d). It is also important in many cases to involve stakeholders from the outset of modeling exercises to increase the potential for broad acceptance of modeling results.

The following steps can be used to define the modeling approach (EPA, 1997e):

1. Use available information to develop a good understanding of watershed characteristics, watershed problems, and watershed hydrology.
2. Consult with program and project managers to develop a clear understanding of project needs and modeling objectives.
3. Select a model or models that best meet the project needs and modeling objectives.

4. Choose the processes to be simulated and the level of complexity, and focus on the processes that govern the problems of concern.
5. Segment the watershed to the desired degree of complexity including the number of subwatersheds, reaches, and land use categories.
6. Choose a simulation process such as single-event or continuous simulation based upon the specified modeling objectives and the system being modeled.
7. Select the time step and imulation time frame necessary to meet the modeling objectives.

Modeling Jargon

Terms You Should Know When Communicating With Modelers

Deterministic models — Mathematical relationships based on physical or mechanistic processes are represented in the model. For example, runoff output is produced in response to precipitation input.

Empirical models — Mathematical relationships in the model (i.e., coefficients for parameters) are based upon measured data rather than theoretical relationships. Must be calibrated.

Steady-state models — Mathematical model of fate and transport that uses constant values of input variables to predict constant values (e.g., receiving water quality concentrations).

Dynamic models — Mathematical model describing the physical behavior of a system or process and its temporal variability.

Hydrodynamic models — Mathematical model that describes circulation, transport, and deposition processes in receiving waters.

Physical models — The building of a scale model of the system and testing it.

Distributed parameter models — Incorporate the influences of the spatially variable, controlling parameters (e.g., topography, soils, land use) in a manner internal to its computational algorithms (EPA, 1982b). Allows simultaneous simulation of conditions at all points within the watershed. Also facilitates incorporation of equations that represent unique processes that occur at only specific points in the watershed.

Lumped parameter models — Use average values for characterizing the influence of specific, non-uniform distributions of each parameter (e.g., soil type, cover, slope steepness).

Calibrated models — Require calibration with measured data for each site-specific application.

"Uncalibrated" or measured-parameter models — Can be used without calibration. Use measured or estimated parameters.

Event-based simulation — Modeling of individual storms. Does not simulate, or account for, periods between storms.

Annualized — Modeling of a longer time series than individual storms. Event-based model outputs can be annualized.

8. Design a model calibration and validation process, including data requirements.
9. Evaluate the assumptions and limitations of the modeling approach.
10. Develop a post-processing data analysis and data interpretation plan.

For applications to nonpoint source problems, the key features of nonpoint sources of pollution need to be fully considered, including but not limited to the following:

1. Hydrology (i.e., rainfall, snowmelt, and sometimes irrigation) drives the process.
2. Pollutant sources are land-based and distributed, with pollutant loads often highly variable in both space and time.
3. Land use types range from highly urbanized to undisturbed forest.
4. Management measures and practices vary from non-structural (e.g., nutrient management) to structural (e.g., waste storage ponds).
5. Land management and land cover change over time, including seasonal fertilization, tillage, crop growth, road maintenance, and off-season inactivity.

Additional considerations and details regarding modeling approach, model selection, and data requirements can be found in existing guidance documents (EPA, 1997d; EPA, 1985).

Model Calibration and Validation

The analyst must evaluate how the model will be used to address management or future conditions. The adequacy of the calibration and validation can be evaluated based on consideration of the type of changes expected to occur, the types of management expected, and the loading and assimilation processes that dominate the system. In some cases, changes in land use distribution can be modeled well by a calibrated system. In other cases, a new land use, such as a new crop, may require that supplemental calibration be performed to account for its unique features. Detailed discussions of model calibration and validation steps and procedures can be found in existing documents (EPA, 1997d; EPA, 1993b; EPA, 1989b; EPA, 1985; ASCE, 1993; Haan et al., 1995; Donigian, 1983).

A very important consideration in estimating nonpoint source loads is the quality and representativeness of the water quality data used in model *calibration*. A water quality data set that does not include a representative sample of high-flow events is unlikely to yield a calibration that is relevant to the concern addressed in the modeling effort. For example, if the goal is to determine the extent to which phosphorus loads are reduced through the implementation of management measures in a watershed dominated by agricultural nonpoint source impacts, it is important that runoff conditions are represented adequately in the calibration.

It is also important that the water quality data used in model calibration cover the same range of wet and dry conditions that are to be used in model validation and prediction. For example, measured loads to New York's Owasco Lake were greater than estimates generated by a simple unit-area loading method due largely to the fact that the measured loads were based on sampling during wet years (Heidtke and Auer, 1993). The simple model used in this example does not explicitly represent rainfall runoff processes, and is therefore very sensitive to the conditions under which it is developed. An adjustment of loading coefficients based upon data from the wet years would likely result in over-prediction of long-term average annual loads.

Calibration — process of adjusting model input parameters to cause model output values to more closely agree with corresponding observed values.

Validation — comparison of model results with an independent data set (without further adjustment).

Verification — examination of the numerical technique in the computer code to ascertain that it truly represents the conceptual model and that there are not inherent numerical problems.

Successful model *validation* should not be blindly interpreted to prove that a model has predictive capabilities. In some cases, the calibration and validation data sets may come from the same period prior to implementation of control measures and practices. For example, if a data set from a period prior to implementation of measures or practices is arbitrarily split in half, with half of the data used for calibration and the other half used for validation, then validation merely confirms that the model can represent conditions prior to implementation of controls. If the measures and practices are intended to change pollutant loads through source reduction, delivery reduction, and/or runoff attenuation, then post-implementation water quality and flow may (and are expected to) respond very differently to precipitation events as compared to pre-implementation conditions. Thus, the model has not really been proven as a predictive tool because the ability to forecast a change in water quality and flow has not been tested with a data set that reflects the changed response to precipitation. Even if the calibration and validation data sets are determined to be independent through statistical analyses, the predictive capabilities are not proven through successful validation unless the validation data set is derived from or reflects conditions of the modeled "future" condition. This is not to say, however, that validation is not important. Successful validation will increase the credibility of modeling results, but the results must be interpreted with care.

Model Calibration and Validation

A good calibration using bad data is a bad calibration.

❐ Ensure that the water quality data used in the calibration and validation process are representative of the true distribution of water quality conditions in the watershed.

- *Don't* use data sets with only low-flow concentrations to simulate high-flow conditions.
- *Do* use data sets with concentration values covering the range of flow and land management conditions in the watershed.

❐ Land use and land management data should be logically linked both to the water quality parameters simulated and to the sources and management measures and practices that will be implemented.

- *Don't* calibrate nutrient concentrations against general land use variables that cannot be logically linked to nutrient management.
- *Do* incorporate to the extent possible data that reflect long-term crop rotations, erosion control, nutrient control, management at other significant sources, and the control of other pollutants that will be managed and simulated in the modeling.

Unit Loads

Several simple methods (see "Simple Watershed Methods" on p. 234) for watershed loading determination use unit loads, or unit-area loads, to represent pollutant contributions from various land uses. Unit loads are expressed as mass per unit area per unit time. One concern associated with unit-load approaches is the availability of good local data regarding the unit loads for watershed-specific physical, chemical, and climatological conditions (Heidtke and Auer, 1993). In the absence of local data, unit loads are approximated using values that may come from nearby studies or studies conducted in distant regions, thus introducing error to the analysis.

Scale should be considered when selecting unit loads, or export coefficients. A study of 210 paired observations of total phosphorus (TP) export taken from 38 studies showed that TP export in agricultural catchments is not a linear function of catchment area, but instead varies as the 0.77 power of drainage basin area (T.-Prairie and Kalff, 1986). This decline in unit-area export was attributable to the TP export from row crops and pasture catchments. However, the study found that the unit-area export of TP from forested catchments did not change as catchment size increased.

Addressing Uncertainty in Modeling Predictions

Because models simplify the real world, the predictions from a model are uncertain, and quantification of the prediction uncertainty should be included in the modeling approach (EPA, 1980). Prediction uncertainty is caused by natural process variability, and bias and error in sampling, measurement, and modeling. Reliably estimated prediction uncertainty can be useful to the planner as a means for judging the value of the prediction and assessing the risk of not achieving management objectives (e.g., meeting the load allocation of a TMDL). Modeling may also result in "unquantified supplemental uncertainty," which is uncertainty introduced through such things as the use of inappropriate export coefficients. This uncertainty, which is unknown to the analyst, is unquantified, and therefore introduces hidden planning risks.

To address the high variability of pesticide loads, a Monte Carlo simulation approach was developed and applied to estimate atrazine and carbofuran loads from hypothetical corn fields in Georgia and Iowa (Haith, 1985). The approach incorporated mathematical models of weather, hydrology, and soil chemistry. One advantage of this approach is the ability to generate a frequency distribution of pollutant loads rather than just a single value, thus allowing an assessment of the probability that any given single value for the pollutant load will occur.

Because of the complexity of quantifying modeling uncertainty, modelers are encouraged to consult with trained statisticians to devise the best approach for their modeling applications. Detailed examples of uncertainty analyses can be found in existing documents (EPA, 1980; EPA, 1989b; Haan, 1989; Beck, 1987).

Model Applications Using GIS Technology

A unit-load approach for estimating phosphorus loads to Owasco Lake in New York used geographic information system (GIS) technology to distribute land-based attributes within the watershed (Heidtke and Auer, 1993). The GIS enabled the modelers to match unit loads with the appropriate areas within the watershed in a distributed manner. GIS technology was also used to facilitate watershed modeling with models such as AGNPS (Agricultural Non-Point Source Pollution) (Line et al., 1997) and SWAT (Soil and Water Assessment Tool) (Engel et al. 1993).

EPA's BASINS (Better Assessment Science Integrating Point and Nonpoint Sources) is an integrated modeling system for performing watershed- and water-quality-based studies (EPA, 2001d). BASINS is intended to facilitate examination of environmental information, support analysis of environmental systems, and provide a framework for examining management alternatives. BASINS includes assessment tools, spatial data, and watershed and water quality modeling components, with GIS providing the integrating framework. An example illustrating the application of BASINS to estimating the impacts of agricultural management measures and practices is given in the BASINS Highlight.

Using BASINS to Develop a TMDL for Fecal Coliform Bacteria

Problem: The Lost River in the state of West Virginia exhibits water quality impairment due to elevated levels of fecal coliform bacteria. Suspected sources of contamination include cattle grazing and feedlots, poultry houses, failing septic systems, geese, wild turkey, and deer, as well as point source dischargers. Section 303(d) of the Clean Water Act and EPA's Water Quality Planning and Management Regulations (40 CFR Part 130) require states to develop Total Maximum Daily Loads (TMDLs) for waterbodies that are not meeting designated uses under technology-based controls. The TMDL process establishes the allowable loadings of pollutants or other quantifiable parameters for a waterbody based on the relationship between pollution sources and instream water quality conditions.

Approach: The U.S. EPA Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) system Version 2.0 (EPA, 1998) and the Nonpoint Source Model (NPSM) were selected to predict the significance of fecal coliform sources and fecal coliform levels in the Lost River watershed. To obtain a spatial variation of the concentration of bacteria along the Lost River, the watershed was subdivided into 11 subwatersheds. This allowed analysts to address the relative contribution of sources within each subwatershed to the different segments of the river. The watershed subdivision was based on a number of factors, including the locations of flow monitoring stations, the locations of stream sampling stations, the locations of feedlots and poultry houses, and land use coverage. To develop a representative linkage between the sources and the instream water quality response in the 11 reaches of the Lost River, model parameters were adjusted to the extent possible for both hydrology and bacteria loading.

Results: Output from NPSM indicates violations of the 200 cfu/100 mL geometric mean standard throughout the Lost River watershed for the existing conditions using the representative time period (October 1990 through September 1991). After applying the load allocations, the NPSM model indicated that all 11 subwatersheds were in compliance with the fecal coliform bacteria standard. The model analysis indicates that water quality standards will be achieved if fecal coliform loads from pastureland are reduced by 38 percent, loads from forestland are reduced 12.8 percent, and loads from cropland are reduced by 37 percent. No change in the point source load was required. The load reductions at the source are expected to be sufficient to meet the 30-day geometric mean, on a daily basis, throughout the year. The margin of safety, an evaluation of the uncertainty in the TMDL, was included implicitly in the model setup and formulation. Conservative assumptions included loads associated with wildlife, septic systems, and existing BMP implementation. Further refinement and corresponding higher accuracy in the analysis could be achieved by more detailed source characterization (actual daily or monthly manure application rates), further evaluation of the viability and dieoff of fecal coliform in the various types of manure, and continued data collection and calibration.

Attainment of the load reductions is expected through implementation of manure storage and application guidelines, crop and pasture management, and wildlife management. No explicit modeling of the BMP effectiveness was performed. Follow-up monitoring is expected to track water quality improvements.