In This Training Session

- The Importance of Estimating Pollutant Loads
- Estimating Pollutant Loads
    - Using Literature Values to Estimate Loads
    - Using Monitoring Data to Estimate Loads
    - Watershed Modeling
- Factors to Consider When Selecting a Model
    - Relevance
    - Reliability
    - Ease of Use
    - Utility
- How much data will you need to run a model?
- What watershed components does a model simulate?
- Model Selection and Application Process
- Types of Models Available
- Models Used for Watershed Projects in Minnesota
- A Brief Description of Selected Models
- Capabilities of Models
    - Water Quality Targets or Endpoints of Models
    - Simulation of Land and Water Features
    - Application Considerations
- Model Application Process for the Selected Models
- Watershed Delineation
- Land Use Assignment
- Model Testing
- Presenting Pollutant Loads
- Consider Spatial Scales
- Consider Time Scales
- Common Pitfalls When Selecting /Using Models
- Communicating Modeling Results
- Create an Official File for Modeling Activities
- Summary
- Resources
- References

Acronyms

AGNPS - Agricultural Nonpoint Source
CN – Curve number
DEM - Digital Elevation Model
EMC - Event Mean Concentration
EPA - Environmental Protection Agency
GIS - Geographic Information Systems
GWLF - Generalized Watershed Loading Function
HSPF - Hydrologic Simulation Program–Fortran
MVUE - Minimum Value Unbiased Estimator
SWAT - Soil and Water Assessment Tool

wq-iw3-50-11

TMDL - Total Maximum Daily Load
USGS - United States Geologic Survey
USLE - Universal Soil Loss Equation

The Importance of Estimating Pollutant Loads

An understanding of pollutant loads is necessary in almost all TMDL studies. By the very nature of the name a TMDL deals with load values of one or more pollutants, with load being defined as a unit of mass per some unit of time and/or area. The means for measuring or estimating loads should be evaluated and determined as a part of the study's planning process. If this is not done, you will likely run the risk of not having the necessary data or other information to appropriately estimate the loads needed to complete the study.

There are a number of purposes for estimating loads in a TMDL study. It is important to understand what the needs are and to then select and use the appropriate methods for estimating load, whether it is to characterize current loading, identify sources and allocations, or predict changes in loading over time. Determine what you need to build your model early in your project and through the calculations and modeling phases as well.

Having an understanding of the watershed and waterbody you are studying is an important basis for determining the appropriate method for quantifying the pollutant loads. In some cases, the use of literature values (export coefficients) may be enough to provide a gross estimate of pollutant loading in a watershed. In other cases, more watershed specific load estimates are needed and can be obtained from calculations using monitoring data. And, often, load estimates from modeling are needed to assist in comparing and/or evaluating potential sources and their loads, the movement and transport of the loads in the watershed and water, and the effect of the loads on specific impairment.

The loading analysis provided by modeling provides a more specific numeric estimate of loads from the various pollutant sources in the watershed. By estimating source loads, you can evaluate and compare the relative magnitude of major sources, and the timing and frequency of those loads

The loading analysis can help you allocate pollutant loads, plan restoration strategies, and project future loads under new conditions. This chapter discusses the analysis and modeling techniques commonly used to estimate or to quantify pollutant loads. This part of a project is often missing from current and past watershed projects, even though it can provide some of the most important information. Without knowing where the pollutants are coming from, you can't effectively control them and restore and protect your watershed.

Stakeholders have an interest in the analysis and modeling techniques used to support decision making. Engaging stakeholders in the evaluation and selection of load calculation and modeling techniques can encourage more informed decision making and increase buy-in for the approaches selected. However, the more complex techniques and modeling tools can be more difficult for Project Managers to describe, review, and interpret for the public.

Estimating Pollutant Loads

Three techniques for estimating pollutant loads are described in the following sections. The use of literature values to estimate pollutant loads are presented first. Techniques that directly estimate loads from monitoring data are discussed next. These techniques are best suited to conditions where fairly detailed monitoring and flow gauging are available and the major goal is to determine total pollutant loads from a watershed. And, third, watershed modeling techniques are described, including a discussion of things to consider when selecting a model, the models currently available, and typical steps involved in application of the models.

A. Using Literature Values to Estimate Loads

One of the simplest approaches for estimating pollutant loads involves estimating loads based on export coefficients published in the literature for various types of watersheds and/or land uses. Export coefficients typically represent loading rates for a particular period of time (i.e., load per area of land per year). Loading estimates for a specific project based on literature values should be used with great care if the basis of the coefficients used and the project conditions are not very similar. Care is also needed in their use given the temporal and spatial "averaging" that they represent versus the dynamic temporal and/or spatial variations in environmental conditions.

Loading rates for land uses vary widely throughout the nation depending on many factors such as precipitation, source activity, and soils. Given this variability, it is important to identify values that are realistic for your watershed and to limit their use to the temporal and spatial scales for which they were established.

Published export coefficients and event mean concentration (EMC) data for use in estimating pollutant loading into watersheds were summarized and reviewed by Linn (2004). Some of the references included in that review that are commonly used for export coefficients include:

Beaulac, M.N., and K.H. Reckhow. 1982. *An examination of land use-nutrient export relationships*. Water Resources Bulletin18(6): 1013-1024.

Reckhow, K.H., M.N. Beaulac., and J.T. Simpson. 1980. *Modeling phosphorus loading and lake response under uncertainty: A manual and compilation of export coefficients*. EPA-440/5-80-011. U.S. Environmental Protection Agency, Office of Water Regulations, Criteria and Standards Division, Washington, D.C.

B. Use of Stream Monitoring Data to Estimate Loads

Water quality and flow data collected prior to or during a TMDL study can be used to calculate pollutant loads. Loads calculated with monitored data represent the amounts of materials passing through a given monitoring site. Because the monitoring data represent in-stream conditions, the resulting load estimates represents the combined loadings from all watershed sources upstream of the monitoring point. Loads estimated in this way can be used to:

- evaluate downstream impacts,
- calculate a per acre loading for the area draining to the monitoring point,
- compare the loads with those of other watershed.

These loads cannot be used to:

- attribute loads to particular sources or areas within the drainage area,
- predict how loadings might change in the future.

To calculate loads using data, monitoring typically involves periodic water sampling, analysis of the samples for pollutant concentrations, and continuous flow gaging. Actual loading could be calculated with data if there were near continuous measurements of pollutant concentration and flow. However, most water quality sampling is not continuous given the labor, time, and money costs that would be involved. In lieu of calculating loads for individual units of time, load calculation methods have been developed using various statistical techniques that estimate/predict the relationship between flow and concentration and then estimate loads based on those relationships. It is important to have quality flow data that was collected at an appropriate frequency (daily values are most commonly used) for load calculations to be adequate.

Various approaches, methods, and computer programs have been developed to calculate pollutant loads using monitoring data. Some of the programs use similar approaches and methods with each having its strengths and weaknesses in use and results. Two computer programs (FLUX and LOADEST) are discussed below. A statistical approach using regression analysis is also described.

1. **FLUX**

FLUX was developed by William Walker (1996) for the U.S. Army Corps of Engineers (Walker 1996). It is an interactive computer program that estimates the loads of nutrients, sediment, and other water quality constituents using flow and concentration data. This technique was developed as a companion to the BATHTUB model which was developed to model eutrophication in reservoirs and is now also commonly used for lake modeling (Walker 1985, 1986, 1990). FLUX is designed to estimate monitoring period and annual loads using daily flow values and periodic water quality concentration values based on the relationship between concentration and flow either as a whole or by breaking the data into groups by some means of stratification. FLUX provides six algorithms (equations) for calculating the loads and various error statistics for use in evaluating the best method to use for a given data set. The six estimation algorithms available in FLUX range from simple arithmetic of seasonal values to regression analyses applied to individual daily flows.

Data requirements for FLUX include:

- Water quality parameter concentrations, collected on a weekly to monthly frequency for at least a year
- Flow values (instantaneous or daily mean values) corresponding to the concentration data
- Complete flow record (daily mean flow) for the period of interest.

MPCA watershed monitoring staff uses FLUX as their primary load calculation tool.

2. **LOADEST**

Load Estimator (LOADEST) is a FORTRAN program for estimating constituent loads in streams and rivers. LOADEST was developed by the United States Geological Survey to calculate monitoring period and annual loads based on monitoring data (Runkel, 2004). LOADEST represents the merger of two software packages previously used by USGS into a single, publicly-available, and fully documented package (http://water.usgs.gov/software/loadest/faq/; accessed January 25, 2008).

Given a time series of streamflow, additional data variables, and constituent concentrations, LOADEST assists the user in developing a regression model for the estimation of constituent load. Explanatory variables within the regression model include various functions of stream flow, decimal time, and additional user-specified data variables. The formulated regression model then is used to estimate loads over a user-specified time interval (estimation). Mean load estimates, standard errors, and 95 percent confidence intervals are developed on a monthly and/or seasonal basis.

3. Regression Relationships of Pollutant Concentration and Flow

In lieu of software programs specifically developed to estimate loads, "basic" regression analysis can be used to develop a relationship between sampled pollutant concentrations and corresponding flow values with loads then calculated using predicted pollutant concentrations for days in which a measured concentration is not present multiplied by the flow values for those days. Regression analysis requires that data be normally distributed while pollutant concentration and stream flow data usually are not normally distributed. The log-transformation of water data often results in the data reasonably meeting the normal distribution and other required assumptions for the use of regression analysis of the transformed data. Computation of loads from this analysis requires that the results be transformed back to the regular units of the data. This transformation will result in biased estimates unless a bias correction is used. Cohn and Gilroy (1991) recommended the use the Minimum Variance Unbiased Estimator (MVUE) for use when the distribution of errors is assumed normal and the Smearing Estimator (SM) for situations in which non-normal error distribution is identified (http://co.water.usgs.gov/sediment/bias.frame.html#HDR6, Accessed January 25, 2008).

C. Use of Watershed Modeling to Estimate Loads

Models provide another approach for estimating pollutant loads, providing source load estimates, and evaluating various management alternatives. A model is a set of mathematical equations designed to describe the natural or man-made processes in a watershed system, such as precipitation, runoff, erosion, transport, loading, land use and management practices. The equations are "built" by examining and mathematically describing cause-and-effect relationships between various watershed and hydrologic processes. Through this, models can be used to evaluate the effects of one thing on the others, including estimating future conditions that might occur under various conditions.

Models range from being simple and/or generalized to highly sophisticated and complex. A simple model may be comprised of a simple empirical relationship estimating runoff based solely on amount of precipitation; whereas, a sophisticated model would estimate runoff based on detailed descriptions of many specific processes such as infiltration and evapotranspiration in addition to precipitation. Simple models in the form of equations may be applied with a calculator or spreadsheet, while many models are available as computer software programs.

This section discusses the role of modeling in watershed planning, the types of models available, how to select appropriate models for your watershed study, and setting up and applying models for a watershed.

Factors to Consider When Selecting a Model

Before selecting a model or models, you should define the approach for the specific study. An approach may require one or more models, multiple analysis procedures, and a variety of input data to address the project needs. Selecting the appropriate model application or approach requires an understanding of the range of complexity of the analytic techniques and a clear understanding of the questions to be answered by the analysis. Note that the model application should consider the following:

- Various levels of detail for each component
- More than one model to address different waterbodies, pollutants, or stressors
- An available modeling system; a modification of an existing model; or a local, custom model
- A model documentation plan

Modules #6 and 7 of this series provides guidance in defining the approach and information needs for a TMDL study.

Selection of the model or models that are appropriate for your project depends on several factors, including the water quality parameters of interest, time and spatial scales of concern, types of pollutant sources, data input needs, model output needs, and experience of the modeling staff or team. Computer models can be categorized in many different ways, so it becomes important to understand both the goals and expectations for modeling in a project and the various types and intended uses of available models.

A list of some of ways models are characterized and/or compared follows below:

- simple – complex
- watershed loading – pollutant transport
- water quality – hydrology (quantity)
- empirical – mechanistic
- yearly or seasonal – daily or smaller time steps

Any given model may be characterized by one or more of these groupings.

To select a model and model application, the data and information needs of the project and the needed accuracy and precision of the modeling results has to be defined. To do this, begin by asking and examining the questions that need to be answered. The following are questions that models are typically used to help answer:

- Will the management actions result in meeting water quality standards?
- Which sources are the main contributors to the pollutant load targeted for reduction?
- What are the loads associated with the individual sources?
- Which combination of management actions will most effectively meet identified targets?
- When does the impairment occur?
- Will the loading or impairment get worse under future land use conditions?
- How can future growth be managed to minimize adverse impacts?

To evaluate and answer the questions with the aid of models, model results such as loads, concentrations, flow, or other variables need to be examined and compared against each other to ensure that a particular model can provide information at the expected spatial and temporal scales for the watershed and water quality concerns.

For example:

- A lake eutrophication problem might focus on predicting the total nitrogen and phosphorus load at a seasonal or annual scale.
- A river with an attached algae problem might need models that can predict concentrations of dissolved nitrogen and phosphorus during low-flow conditions.
- An area with beach closures due to pathogens might focus on predicting pathogen counts and the frequency of standards exceedances.
- A concern over sediment in streams might focus on changes in hydrology, stream morphology, and/or sediment loading from erosion-prone areas at a daily time-step.

In each case, the predictions of the model should be evaluated on the basis of the indicators identified for meeting and tracking the goals of the TMDL study. The indicators used will often dictate the level of detail of the study. The model should support the development of source loads and estimates of their magnitude, and it should support the development of the appropriate pollutant load reduction estimates.

In choosing a model application for your watershed, keep in mind four general considerations:

1. **Relevance**
2. **Reliability**
3. **Ease of Use**
4. **Utility**

Each of these considerations is discussed below.

1. Relevance

Even if the model has been reviewed in the literature and has been applied in other watersheds, you need to make certain that it is relevant to the needs of your watershed. For example, a model developed and tested only in urban areas, or even in rural areas that are mostly forested, is apt not to be a good choice for a watershed that consists almost entirely of agricultural row crops or mixed uses. If flow-through tile drains are one of the main pathways through which water reaches the stream in your watershed, a model that does not include artificial drainage is probably not a good choice. For specialized cases, such as tile drainage, a custom modeling application might be needed. Many models have been developed for specific pollutants.

Some specialize in sediment only because reducing erosion was historically the mission of modeling conducted by the U.S. Department of Agriculture (USDA). Many models give results for sediment, nutrients, and perhaps pesticides, but not for microbial contaminants.

Work closely with MPCA staff to determine which model is most appropriate for your TMDL project.

2. Reliability

Models contain mathematical equations which are a simplification of real-world physical processes. In nature, there are thousands to millions of natural processes. No model can perfectly represent the complexity and inter-connectedness of all watershed processes.

Because it's not possible to know in advance how accurate the results of a specific model will be, you need to rely on what others have found. Scientists rely on peer review of journal articles written about the use of a model. A quick rule of thumb is to use only models which have been validated and documented in respected peer-reviewed journals. In so doing, you will benefit from the time other modelers and scientists have spent reviewing the model.

All the models reviewed in this chapter have been validated, at least to some extent. Since all models involve some uncertainty, you should communicate that clearly to stakeholders and the public.

Most models distributed in the public domain have been developed by government agencies (e.g., EPA or USDA) or universities and are freely available.

However, some consultants use proprietary models, which are privately owned software. Such models cannot be checked because the code is not available to others. *MPCA strongly encourages the use of those models that are widely used in the public arena (free).* If your technical team proposes use of proprietary models, prior approval will be required by MPCA staff. The proprietary model must have full and complete documentation of all assumptions and algorithms used in the model. Proprietary models normally require a purchase fee and have limited distribution rights. Limiting distribution and review might affect acceptance by the stakeholders.

3. Ease of Use

Accuracy of prediction is important, but if the model will not answer the questions you need to develop your watershed plan, it will not be useful. Documentation that explains the parameters, how to get them, and reasonable values for them is essential to ensure that the model is usable. New users might need some sort of training to learn how to use the model. Finally, model users sometimes run into questions that are not addressed in the documentation. A model that will be widely used needs to have user support available. The support can be in the form of a person who provides technical assistance or a list server where other users can answer questions.

Obtaining input data is often the most time-consuming and difficult part of running a model. This often comes as a surprise to those who have not used models. Models generally require data on land cover, land management (such as agricultural practices), factors that affect the rate at which water can flow into the soil and recharge ground water (usually geology or soil type), and other information about the land in the watershed. In addition, daily or even hourly weather data, including precipitation and temperature, are usually required. Other weather data that is more difficult to obtain, such as relative humidity and wind speed, might be required. For models to be calibrated, accurate input data are needed. Modeling systems, such as EPA BASINS, have compiled much of the basic data needed to run the models; however, this coarse, national-scale data will not always be accurate enough to give useful results, particularly in small watersheds. While national, publicly available databases are available from USGS and other sources, many parameters are not available nationally and need to be obtained locally.

4. Utility

Using a model to predict the impact of changes in a watershed requires that the model be able to represent those changes. Models represent changes in watershed management in very different ways. You will need to consider what management practices are likely to be applied in your watershed and whether the model can be used to evaluate their benefits. In many cases other analyses are used to supplement a model; sometimes additional spreadsheet calculations can be used to check on the potential load reductions from various methods. In addition, you might want to consider how the model will be used in the future. Will it be used to check future changes in management or as a tool to track progress?

If the model will be used as an ongoing planning tool, remember to consider the complexity of the model and the availability of trained staff to apply the model.

Questions to keep in mind while selecting your model include:

- Is the model appropriate given the scale (both spatial and temporal) of the impairment?
- What predictive capability must the model have to address this particular impairment?

Another consideration to make in selecting a model for a project may be the ease in which the results can be communicated to audiences without modeling expertise. While even the most complex models can be effectively presented for review and comment through public meetings, workshops, and other technical transfer opportunities, simpler modeling approaches, if sufficient for addressing a project's needs, may be easier to communicate to and be understood by the community.

How Much Data Will You Need to Run a Model?

There is no hard and fast rule regarding the kind and amount of data you will need to run a model. Each model has different data input needs. While the need for good quality water data is usually recognized in projects, it is equally important to have high quality land use, management and drainage data for modeling. Be certain to ascertain up-front whether and in what detail the following types of data will be needed for the model you select:

- Soils types
- Topography
- Fertilizers
- Crop types
- Tillage practices
- Crop residue
- Tillage
- Drainage type, extent and location

The more quality data you obtain the better model results will be.

In addition to model input data, some models require the presence of water quality data for use in calibrating and validating them. When this is the case, high quality data is central to obtaining good model outputs. For some of the more complex models, it is desirable to have between 6-15 years of water quality monitoring and meteorological data. Some models may even require up to 30 years of meteorological data for an adequate model run. Some complex, lumped-parameter models such as HSPF require extensive calibration. Ideally, you should have several years of data in order to calibrate a model and several years of data to validate it. Several years of contiguous data are usually preferred.

What Watershed Components Does a Model Simulate?

Watershed models use various equations or techniques to simulate one or more processes in a watershed system and/or water body. Components of the watershed system that are key to many watersheds and, therefore, to many models include:

- **Water Cycle Components:** The description of precipitation, infiltration, evaporation, transpiration, runoff, and other flow pathways is critical for many models. A model will often estimate the amount and timing of runoff from a land area using these components.

These then may be related to erosion and to sediment and pollutant transport. In cold-climate watersheds, it might be important to use a model that can represent snowmelt/runoff conditions. For certain variables of interest (e.g. heat) and some types of stratigraphy (e.g. karst, unconfined sand aquifers), the ability to link infiltration to a vadose zone or groundwater model is also important.

- **Erosion and Sediment Transport Components:** These components involve the description of soil detachment, erosion, and sediment movement from a land area. Models may simulate these simply or in great detail.

- **Pollutant Loading:** This represents the wash-off or infiltration of pollutants from a land area. In generalized modeling approaches, this is estimated as some sort of loading factor, while more detailed modeling techniques link pollutant wash-off and infiltration to hydrology and sediment movement on the land.

- **Stream Transport:** Stream transport components represent the hydraulic functioning of a stream. As a component of watershed models, it is needed, at a minimum, to collect or compile the runoff and pollutants estimates modeled from the various land areas. More detailed models include evaluation of in-stream behavior of sediment and pollutants. Processes may include deposition, resuspension, decay, and transformation.

- **Management Practices:** This component represents the human-affected land use and cover factors usually of critical interest to watershed managers. Management practices can encompass a wide range of activities which may be land-based (e.g., tillage or fertilizer application) or constructed (e.g., stormwater ponds, wastewater treatment facilities). In modeling, land-based management can be generalized (e.g., number of acres treated) or specific and detailed (e.g., individual field-specific practices). Some models include more detailed simulation techniques. For example, a pond analysis might include sediment settling and first-order decay of pollutants.

  The order and degree to which the various components are simulated in models varies considerably. However, land-related components such as land use, soils, and slope are typically described first. These are usually key features that affect runoff, erosion, and pollutant loadings. Management practices present in the watershed are often considered next followed by the stream transport components and water body functioning. Depending on the model, each component is addressed at various levels of detail. As the level of detail increases to address specific factors and processes, the equations used to build the modeling system become more detailed and complex. This then increases the number of parameters that need to be estimated and the detail in which model performance needs to be evaluated. The model selection process, therefore, needs to account for this range in levels of detail so that a model will be consistent with the objectives of the study.

  Model applications for specific watersheds often include a mixture of levels of detail depending on the problems being considered. For example, a modeling analysis supporting an agricultural nutrient management initiative might include very detailed descriptions of land behavior, such as nitrogen use by plants, and a very simplified analysis of stream transport. A study considering the upgrade of a wastewater treatment plant would include a detailed examination of the stream conditions in summer and a very simplified representation of land use activities.

The Model Selection and Application Process

With so many models available, how do you know which one to choose? The development of a modeling analysis involves more than just selecting a modeling tool. The application of a model for decision-making also involves designing and implementing an analysis that addresses the management questions. Typically this involves a combination of data analysis techniques, as described in Chapter 9, and compilation and organization of disparate data sources.

Described below are the key steps for selecting and designing a modeling application for a TMDL project. Throughout the TMDL process you've built an understanding of the watershed—through scoping, stakeholder input, and data collection and analysis. The design of the modeling approach should build on this understanding and help you to better understand the watershed.

1. **Consider the objectives of the analysis.** During the scoping process, the key objectives of the study are identified, as well as the general modeling needs and watershed characteristics. The specific objectives and associated indicators will help to define the pollutants to be modeled and the scale of the model to be used.

2. **Define the specific questions that the modeling will be used to answer.** As discussed earlier in the chapter, before selecting a model, the analyst should first carefully define the questions that the model will be used to answer. The questions should directly relate to the overarching objectives of the study.

   The following are examples of modeling questions:

   - What are the sources of the pollutant load?
   - Where can management practices be targeted to best meet load reduction requirements?
   - What combination of management practices will result in reducing the load to the desired level and meeting water quality goals?

3. **Select the modeling approach that will address the questions.** The modeling approach includes the model(s) to be used, the input data processing requirements and data sources, the model testing locations and data sources, and the output analysis. The modeling approach defines how the model will be applied, not just what the model is. The approach provides the entire plan or road map for analysis and is broader than just the selection of a model. Each of these items should be clearly documented to help ensure that the study needs are met.

4. **Build the model.** After the modeling approach is selected, the input data needed for the model or models needs to be collected and processed. Typical data inputs for models include the following:

   - Land use
   - Tillage practices (for agricultural watersheds)
   - Management practices
   - Soils
   - Cross-sections and topography (to address riverine models as well as watershed models)
   - Activities, management locations, and types
   - Drainage types, extent and locations
   - Meteorologic data—precipitation and temperature

   Each dataset might require some preprocessing before input. For example, land use information might be selectively updated where new development has occurred. Sometimes multiple land use datasets are combined. For example, one data source might provide a more detailed breakdown of forest types and could be used to add detail to more general land use coverage. Some models require developing categories of land use, soil, and slope characteristics.

   Resulting units could include corn fields with B soils (a hydrologic soil group defined by the USDA) and moderate slopes, pasture with C soils and steep slopes, and so on. User's guides and the selected modeling references provide some additional guidance on data preprocessing needs for individual models. Much of the data required for watershed models will be collected as part of your data inventory and monitoring efforts (see Chapters 6 and 7).

5. **Test the model's performance.** The performance of a model is tested in different ways depending on the type of model being used. Some models require extensive calibration procedures to check and adapt the model to adequately match monitored data. HSPF, for example, contains numerous equations that contain coefficients that need to be determined by calibration. Other models require little to no calibration.

   Formal calibration and validation procedures involve separating available monitoring data into two separate time periods for testing. Using one data set, calibration parameters are adjusted, within reasonable ranges, until a "best fit" to the observed data is generated. Using the second dataset, validation is performed by keeping the parameter set constant and testing the performance of the model.

Time periods for calibration and validation are carefully selected to include a range of hydrologic conditions.

When data are limited, you should also compare model results to literature values and data from surrounding watersheds to review the integrity of the results.

Models, such as SWMM, begin to use equations based on watershed or hydraulic process theories which are not intended to be calibrated; however, some calibration and validation should be done for specific uses of such models (i.e., predicting in-stream concentrations with SWMM). Other models include equations that are not intended to be calibrated. In most cases, these models were "pre-calibrated in their development by fitting empirical equations to sets of monitoring data.

All models, especially those being used to predict in-stream pollutant concentrations, should still be "tested" or validated for "reasonable" model outputs. This testing can be completed by comparing model results to monitored data – either project specific or local, regional or national averages and ranges – to determine if the results are "reasonable." When testing models, ask yourself:

- Do the loads seem realistic given observed concentrations and flows or documented loads in nearby watersheds?
- Do the simulation results make sense given the watershed processes? For example, if a watershed model produces monthly loads, do the higher loads occur during the times of higher observed flows and concentrations?
- Or, if a model provides output from both ground water and surface water, do the relative contributions make sense given the topography and geology of the area?

Models are meant to represent the processes affecting runoff and pollutant transport and loading. Use your knowledge of the area to reality-check the model representations and output.

Be aware that distributed parameter models are not designed to provide estimates of an absolute pollutant concentration even though they "often" are used in this way. They are designed to predict relative changes in a watershed to evaluate changes in land use or management rather than predict absolute concentrations or load values.

6. **Apply the model and interpret the results.** The model is applied to evaluate the range of conditions required for addressing the modeling questions. For example, a model might be used to evaluate the nutrient loading over a 10-year period. Output post-processing might include developing weekly and seasonal loading summaries to evaluate weekly or seasonal variations. Multiple model applications might be used to consider changes in land use, installation of management practices, and alterations in cultivation techniques. Output can be processed to support development of essential elements of the TMDL (magnitude of sources, and pollutant load allocations).

7. **Update the model to include new information or refine assumptions.** Once the initial study is complete, the model can be updated periodically to further refine and test performance and update management recommendations, if additional data are collected or new information is obtained.

Selection and execution of an appropriate modeling approach can support the development of a TMDL. Use caution in selecting an approach consistent with the available data, the specific questions to be addressed, and the type of management. Data analysis is an ongoing process in which modeling is only one potential tool. In some cases, simplified techniques or statistical analysis is adequate to evaluate watershed conditions and no formal modeling is required. Throughout the process, focus on using the simplest methods appropriate to answering the questions at hand.

Types of Available Models

Various modeling systems have been developed and used to answer a wide range of environmental questions. This chapter focuses on selected models that are publicly available and have a track record of application and use. The models have been commonly used in TMDL studies and other watershed studies. They represent a range of complexity and are applicable to a variety of pollutants and pollutant sources. An inventory of available models that evaluates the models across a set of key characteristics is provided in Table **11-1**. These characteristics were selected to help differentiate among available tools and to describe areas of emphasis, complexity, and types of pollutants considered.

The key characteristics include the following:

- **Type.** "Landscape only" indicates that the model simulates only land-based processes; "comprehensive" models include land and stream and conveyance routing.
- **Level of complexity.** Complexity in watershed models is characterized by three types of functions. Export functions are simplified rates that estimate loading based on a very limited set of factors (e.g., land use). Loading functions are empirically-based estimates of load based on generalized meteorological factors (e.g., precipitation, temperature). Physically-based models include physically based representations of runoff, pollutant accumulation and wash-off, and sediment detachment and transport. Most detailed models use a mixture of empirical and physically based algorithms.
- **Time step.** Time step is the unit of time (e.g., hourly, monthly) for which a model simulates processes and provides results. The table identifies the smallest time step supported by a model. If larger output time steps are needed, model output can be summarized from smaller time steps.
- **Hydrology.** This criterion identifies whether a model includes surface runoff only or surface and ground water inputs are considered.
- **Water quality.** Water quality capabilities are evaluated based on the pollutants or parameters simulated by the model.
- **Types of management practices.** The types of management practices simulated by the models are indicated in the table. Even if you're not planning to run the model yourself, it's helpful to know the capabilities and requirements of the major types of watershed models so you can "talk the talk" and make informed decisions about how to proceed with your data analysis.

Remember that typically it is not the model itself that causes problems but the matching of the model to local conditions, key assumptions, and interpretation of model outputs.

Additional detailed information on available models is provided in EPA's *Compendium of Tools for Watershed Assessment and TMDL Development* (USEPA 1997c). Although updated versions of some models have been released since the compendium was published, it provides a good starting point for researching available models and understanding their capabilities.

A more recent online database, provided by EPA's Council on Regulatory Environmental Modeling, provides links to model reviews and resources (http://cfpub.epa.gov/crem/).

Models Used for Watershed Projects in Minnesota

A brief description of some of the more commonly used models in Minnesota follow. Before selecting a model for your TMDL project, consult with one of MPCA's modeling experts.

Watershed Models:

- **AGNPS**
- **STEPL**
- **HSPF**
- **SWAT**
- **GSSHA**
- **WARMF**

Urban Models

- **PONDNET**
- **SIMPLE MODEL**
- **SWMM**
- **P8**

Field Scale Models

- **WEPP**
- **GLEAMS**

Receiving Water Models

- **WASP**
- **CE-QUAL-W2**
- **QUAL2K**

These models represent a cross section of simple to more detailed approaches, provide simulation of rural and more urbanized areas, and include a diversity of approaches. These models are used to describe key differentiators and considerations in selecting and applying models. Other models that have specialized capabilities to support TMDL development are available.

A Brief Description of Selected Models

Watershed Models:

- **AGNPS - Agricultural Non-Point Source Pollution Model –** The AGNPS model was developed by USDA's Agricultural Research Service for use in evaluating the effect of agricultural management decisions on a watershed system. AnnAGNPS is currently a continuous simulation watershed model, as opposed to the earlier versions of the model, which were solely storm–event models. The storm–event version was not migrated to the Windows NT® environment.

AGNPS is a distributed parameter model, and has the advantage of providing spatially explicit modeling results, which is difficult or impossible to achieve with lumped parameter watershed models. Agricultural BMP analysis can be performed at the field–scale level, and the model will route the combined effect of BMP implementation to the watershed outlet.

As a distributed parameter model, it was not designed to provide absolute load estimates, but is intended to be used to demonstrate relative change due to changes in management practices.

- **STEPL - Spreadsheet Tool for Estimating Pollutant Load -** STEPL is a collection of spreadsheets compiled within a Microsoft® Excel workbook which estimates annual nitrogen, phosphorus, 5-day biochemical oxygen demand (BOD_5), and sediment loads from watersheds. STEPL was developed to approximate the benefits of implementation of best management practices in the form of reductions in watershed average annual pollutant loads for input into the EPA Grants Reporting and Tracking System (GRTS).

Estimates of the removal efficiency of many commonly used best management practices are included in the STEPL database.

For each watershed, the annual pollutant loading is calculated based on the runoff volume and the pollutant concentrations in the runoff water as influenced by factors such as the land use distribution and management practices.

- **HSPF - Hydrological Simulation Program - FORTRAN** - HSPF is a component of the EPA Better Assessment Science Integrating Point and Nonpoint Sources environmental analysis system. It is a comprehensive package for simulation of watershed hydrology and water quality for both conventional and toxic organic pollutants. HSPF incorporates watershed-scale ARM and NPS models into a basin-scale analysis framework that includes fate and transport in one dimensional stream channels. It is the only comprehensive model of watershed hydrology and water quality that allows the integrated simulation of land

and soil contaminant runoff processes with in-stream hydraulic and sediment-chemical interactions. The result of this simulation is a time history of the runoff flow rate, sediment load, and nutrient and pesticide concentrations, along with a time history of water quantity and quality at any point in a watershed. HSPF simulates three sediment types (sand, silt, and clay) in addition to a single organic chemical and transformation products of that chemical.

- **SWAT – The Soil and Water Assessment Tool** - SWAT is a lumped–parameter agricultural watershed model developed by the USDA’s Agricultural Research Service (ARS) office in Temple, Texas. SWAT grew out of the USDA–ARS model Simulator for Water Resources in Rural Basins (SWRRB), with algorithms from the field–scale model Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS), and other numerous modifications and enhancements. SWAT runs on a minimum daily time-step, and simulates the export of sediment, nutrients, pesticides, and bacteria from watersheds. Agricultural tillage practices effecting crop residue and crop rotation can be explicitly modeled with SWAT. SWAT includes snowmelt capabilities for cold climates, and also includes a climate generator for filling in data gaps or generating alternative climate scenarios.

- **GSSHA - The Gridded Surface Subsurface Hydrologic Analysis** model was developed by the US Army Corps of Engineers in Vicksburg, Mississippi. GSSHA is a distributed–parameter watershed model which simulates hydrology and sediment transport, with nutrient components currently being tested. Both 2 dimensional surface water flow and groundwater flow are simulated using rigorous finite–difference techniques. The watershed is discretized into a network of cells, with surface and subsurface flow occurring between cells. Given the distributed nature of GSSHA, small scale BMPs can be evaluated.

- **WARMF - The Watershed Analysis Risk Management Framework** - WARMF is a mechanistic watershed model that is based on the Integrated Lake–Watershed Acidification Study (ILWAS) model. ILWAS was created for the Electric Power Research Institute to study the acidification of lakes in the eastern United States as part of the National Acid Precipitation Assessment Program. The ILWAS model divides a watershed into land catchments, stream segments, and lake layers.

Land catchments are further divided into canopy and soil layers. These watershed compartments are connected to form a network for hydrologic and water quality simulations. Algorithms from the Water Analysis

Simulation Program (WASP), the Areal Nonpoint Source Watershed Environment Response Simulation (ANSWERS) model, and the Stormwater Management Model (SWMM) were added to the ILWAS code to create a comprehensive watershed model. Such additions allow WARMF to simulate biochemical oxygen demand, coliform bacteria, dissolved oxygen, and nutrients in addition to alkalinity and pH.

Urban Models

- **PONDNET** – PONDNET is an empirical spreadsheet model which estimates the average annual phosphorus removal of stormwater detention ponds. Watershed runoff volumes discharging to the pond are determined using the Simple Method. Runoff phosphorus concentrations are based upon results obtained from the EPA Nationwide Urban Runoff Program (NURP). Removal algorithms for ponds are based upon NURP study results and a second–order removal formulation developed from data obtained from monitoring 60 US Army Corps of Engineers reservoirs.

- **SIMPLE Method** – The Simple Method is a planning level tool that estimates annual urban runoff pollutant loads. It was developed by the Metropolitan Washington Council of Governments in 1987. The Method utilizes an equation which relates watershed pollutant load to rainfall depth, event mean runoff pollutant concentration, percent impervious cover, and area of a particular land use. The Simple Method is based on urban runoff monitoring data from four metropolitan Washington DC area sites and from of 40 monitoring sites in 16 NURP locations across the United States.

Initially, no attempt was made to extend the method to construction, industrial, rural development, or agricultural areas. The method does not include pollutant loads generated by base flow. The Method is applicable to drainage areas of less than one square mile. "Scaling up" the method to watersheds larger than one square mile may produce uncertain results due to appreciable base flow which is not incorporated in the Simple Method.

A major revelation of the method is that analysis of NURP nutrient monitoring showed no significant difference in average nutrient pollutant concentrations between sites and no consistent correlation between pollutant concentrations and storm volume or intensity. Therefore, a single value can be used for the runoff nutrient concentration for an urban area. Suspended sediment monitoring results do not illustrate these statistical properties. Individual monitoring site means and variances were significantly different from each other. Because of the high degree of storm event and site variability, a single value cannot be used for suspended sediment concentration.

- **SWMM - The Storm Water Management Model** - SWMM was originally developed for the Environmental Protection Agency (EPA) in 1971 by Metcalf and Eddy, Inc., Water Resources Engineers, Inc. and the University of Florida. SWMM is a dynamic rainfall-runoff and water quality simulation model, primarily but not exclusively for urban areas, for single-event or long-term (continuous) simulation. The Storm Water Management Model (SWMM) is a comprehensive computer model for analysis of quantity and quality problems associated with urban runoff. Simulations can be performed on Catchments having storm sewers, or combined sewers and natural drainage, for prediction of flows, stages and pollutant concentrations.

In SWMM, flow routing is performed for surface and sub-surface conveyance and ground water systems, including the options of non-linear reservoir channel routing and fully dynamic hydraulic flow routing. In the fully dynamic hydraulic flow routing option, SWMM simulates backwater, surcharging, pressure flow, and looped connections. SWMM has a variety of options for water quality simulation, including traditional buildup and wash-off formulation as well as rating curves and regression techniques.

SWMM incorporates first order decay and particle settling mechanisms in pollutant transport simulations and includes an option of simple scour-deposition routine. The latest version of SWMM simulates overland flow routing between pervious and impervious areas within a subwatershed. Storage, treatment, and other management practices can also be simulated. The model assumes all pollutants entering the waterbodies are sediment adsorbed.

- **P8** - Program for Predicting Polluting Particle Passage Thru Pits, Puddles, and Ponds. P8 is a model for predicting the generation and transport of stormwater pollutants in urban watersheds.

Continuous water balance and mass balance calculations are performed on a user-defined system consisting of watersheds, devices (runoff storage/treatment areas, BMPs), particle classes, and water quality components.

Simulations are driven by continuous hourly rainfall and daily air temperature time series data. The model simulates pollutant transport and removal in a variety of treatment devices (BMPs), including swales, buffer strips, detention ponds (dry, wet, and extended), flow splitters, and infiltration basins (offline and online), pipes, and aquifers. Water quality components include total suspended solids (TSS) (five size fractions), total phosphorus (TP), total Kjeldahl nitrogen (TKN), copper, lead, zinc, and hydrocarbons.

Field Scale Models

- **WEPP – The Watershed Erosion Prediction Project** is a state–of–the–art erosion and sediment transport technology developed by the USDA–Agricultural Research Service National Soil Erosion Research Laboratory at Purdue University. The intent of the development of WEPP was to replace the Universal Soil Loss Equation (USLE) with a mechanistic model incorporating modern hydrologic and erosion science.

The WEPP erosion model is a continuous simulation computer program which predicts soil loss and sediment deposition from overland flow on hillslopes, soil loss and sediment deposition from concentrated flow in small channels, and sediment deposition in impoundments. In addition to the erosion components, it also includes a climate component which uses a stochastic generator to provide daily weather information, a hydrology component which is based on a modified Green-Ampt infiltration equation and solutions of the kinematic wave equations, a daily water balance component, a plant growth and residue decomposition component, and an irrigation component. The WEPP model computes spatial and temporal distributions of soil loss and deposition, and provides explicit estimates of when and where in a watershed or on a hillslope that erosion is occurring so that conservation measures can be selected to most effectively control soil loss and sediment yield.

WEPP was modified in 2006 to reflect forest hydrologic conditions. Work at Washington State University regarding the relatively greater lateral vadose zone flow due to steep slopes, shallow bedrock, and thin soils in mountainous forested regions was incorporated into the model. Different versions or adaptations of WEPP have been created that are somewhat simplified in terms of its use and application. Care should be taken to determine the "version" to be used.

- **GLEAMS - Groundwater Loading Effects of Agricultural Management Systems** is a continuous simulation, field scale model, which was developed as an extension of the Chemicals, Runoff and Erosion from Agricultural Management Systems (CREAMS) model. GLEAMS assumes that a field has homogeneous land use, soils, and precipitation. It consists of four major components: hydrology, erosion/sediment yield, pesticide transport, and nutrients. GLEAMS was developed to evaluate the impact of management practices on potential pesticide and nutrient leaching within, through, and below the root zone. It also estimates surface runoff and sediment losses from the field. GLEAMS was not developed as an absolute predictor of pollutant loading. It is a tool for comparative analysis of complex pesticide chemistry, soil properties, and climate.

GLEAMS can be used to assess the effect of farm level management decisions on water quality. GLEAMS does not simulate frozen soil conditions or snowmelt – a limitation for use of GLEAMS in northern climates.

Receiving Water Models

- **WASP - The Water Quality Analysis Simulation Program** is an enhancement of the original WASP model (Di Toro et al., 1983; Connolly and Winfield, 1984; Ambrose, R.B. et al., 1988). This model helps users interpret and predict water quality responses to natural phenomena and manmade pollution for various pollution management decisions. WASP is a dynamic compartment-modeling program for aquatic systems, including both the water column and the underlying benthos.

  WASP allows the user to investigate 1, 2, and 3 dimensional systems, and a variety of pollutant types. The time varying processes of dispersion, point and diffuse mass loading and boundary exchange are represented in the model. WASP also can be linked with hydrodynamic and sediment transport models that can provide flows, depths velocities, temperature, salinity and sediment fluxes.

- **CE-QUAL-W2** - is a two-dimensional, laterally averaged, finite difference hydrodynamic and water quality model. Because the model assumes lateral homogeneity, it is best suited for relatively long and narrow waterbodies exhibiting longitudinal and vertical water quality gradients. The model can be applied to rivers, lakes, reservoirs, and estuaries. Branched networks can be modeled.

  The model accommodates variable grid spacing (segment lengths and layer thicknesses) so that greater resolution in the grid can be specified where needed. The model equations are based on the hydrostatic approximation (negligible vertical accelerations). Eddy coefficients are used to model turbulence.

  The hydrodynamic time step is calculated internally as the maximum allowable time step that ensures numerical stability. A third-order accurate (QUICKEST) advection scheme reduces numerical diffusion. The water quality portion of the model includes the major processes of eutrophication kinetics and a single algal compartment. The bottom sediment compartment stores settled particles, releases nutrients to the water column, and exerts sediment oxygen demand based on user-supplied fluxes; a full sediment digenesis model is under development.

The following parameters can be modeled by CE-QUAL-W2:

- temperature
- total dissolved solids
- coliform bacteria
- inorganic suspended sediments
- dissolved organic matter
- biochemical oxygen demand
- algae
- detritus
- phosphorus
- nitrogen
- dissolved oxygen
- iron
- alkalinity
- pH
- bottom sediments

- **QUAL2K -** QUAL2K is a hybrid spreadsheet/FORTRAN water quality model. Input datasets are created utilizing pre–developed spreadsheets. The actual water quality model utilizes data from the spreadsheets which runs as a FORTRAN program in the background. Model output is presented in pre–formatted spreadsheets and graphs in the Excel® workbook.

Capabilities of Models

Major factors to consider when selecting a watershed model include:

- Water quality indicators simulated
- Simulation of land and water features (e.g., land use and waterbody types)
- Application considerations (e.g., training required)

The following sections discuss the capabilities and characteristics of selected models for each of these considerations.

Water Quality Targets or Endpoints for Models

The selection of the appropriate model for your watershed and your goals depends on the types of processes you need to simulate. The initial criteria for determining which model is right for your watershed analysis include the water quality targets or goals. Water quality targets are based on specific parameters (e.g., phosphorus, sediment) and typically have an associated magnitude, duration, and frequency.

For example, a target might be established for a monthly sediment load of 20 tons, or bacteria targets might be set as a daily maximum of 400 counts/100 mL.

To better summarize the selected watershed models' applicability to typical water quality targets and to aid in identifying appropriate models for your watershed, Table **11-2** summarizes the models' ability to simulate typical target pollutants and expressions (e.g., load vs. concentration). The table scores the models depending on the time step of the simulation for the target—annual, daily, or hourly.

Simulation of Land and Water Features

After you've initially identified models based on the necessary parameters, it's important to identify the major land and water features or processes that you want to simulate. For example, what types of land uses are in your watershed? Is ground water an important influence on instream water quality? Are there certain types of management measures you want to evaluate in your watershed? The available models simulate different land and water features, and they do so at different levels of detail. Table **11-3** provides a summary of the selected key models' capabilities for simulating a variety of land and water features. The table identifies the following categories:

- **General Land and Water Features Supported:** Rates models according to their ability to simulate general land uses and waterbody types.
- **Special Land Features Supported:** Rates models on the basis of their ability to simulate special land processes such as wetlands, hydrologic modification, urban management practices, and rural management practices.
- **Special Water Features Supported:** Rates models on the basis of their ability to simulate special processes occurring in receiving waterbodies such as air deposition, streambank erosion, algae, and fish. Because the selected models are primarily watershed models, many of the detailed water features are not supported. If these processes are important in your watershed, it might be necessary to investigate receiving water models or other outside analyses to use in combination with your watershed model.

Application Considerations

Another issue to consider when selecting your model is what it takes to apply the model-considerations such as how long it will take to setup and apply the model, how much training you'll need, and how much the model will cost. Table **11-4** rates the selected models based on the practical considerations affecting their application. Models with filled circles are generally easier to use and require less data and time for application.

Model Application Process for the Selected Models

Previous sections discussed the basic features of models, how to select appropriate models for your project, and general steps in applying models. This section discusses the decisions made during model application. Although the models have different features and capabilities, some basic decisions regarding data and data processing are required for every model application.
These are the decisions that result in tailoring the model to your specific site. Each major decision point is discussed, along with some suggestions for how to decide the appropriate level of detail.

For loading analysis you need to think carefully about the area being modeled. A watershed is usually composed of areas with diverse land uses and activities. Some watersheds have regional differences, such as a densely populated areas surrounded by countryside. When applying a model to a watershed, the diversity within the watershed is simplified into major categories so that the loads can be estimated. If the analysis is too detailed, the modeling becomes very difficult to apply and test. If the analysis is too simplified, some important information might be lost. Modeling should build on the detailed understanding of the watershed developed during planning and data analysis.

Watershed Delineation

Although you've already delineated your watershed, you'll likely further divide the watershed into small subwatersheds for modeling and evaluation. Dividing the watershed into subwatersheds is usually the very first step in watershed modeling. A watershed of 10 square miles might be subdivided into 20 subwatersheds about 0.5 square mile each. How do you decide how small to go? That will depend on the watershed characteristics, the type of model you're using, and the management actions that might be considered. Some watershed characteristics to consider when subdividing the watershed include:

- Land use distribution and diversity
- Location of critical areas
- Stream gauging stations and water quality monitoring locations (subwatersheds should match key monitoring locations for testing)
- Location of physical features such as lakes, dams, and point sources discharges
- Changes in topography
- Soil distribution
- Areas where management might change

Table **11-5** provides examples of the number of sub watersheds and average size of subwatersheds for some very large watershed modeling applications using HSPF or LSPC. Why do they vary significantly?

The watershed with the most uniform land uses and a large area was evaluated using large subwatersheds. The watershed with the smallest subwatersheds is in an area that ranges from highly urbanized to rural and has a dense network of monitoring data available for testing.
In this application the local conditions are represented by using smaller watersheds. Each application is unique, and watersheds are defined accordingly.

The number and size of subwatersheds can affect the model selection process. Some watershed models have limitations on the number of subwatersheds or the size of the area the model can simulate.

HSPF, SWMM, and SWAT are typically used for multiple subwatersheds, allowing for the evaluation of geographic distributions of loads. Models such as STEPL do not inherently handle multiple watersheds and therefore is applied to one watershed at a time.

How are subwatersheds delineated? Most applications today use a geographic information system (GIS) to delineate watersheds based on Digital Elevation Models (DEMs) and topographic maps. Some software packages provide auto delineation tools or other aids to help define hydrologic boundaries. Predefined watershed boundaries such as 14-digit hydrologic units can be used.

Land Use Assignment

Land use information is typically provided as a GIS coverage or map with many individual codes that describe detailed land use types. For modeling purposes, these individual codes should be grouped into a more manageable set of dominant land use types. How much combining is done depends on the watershed characteristics.

Factors to consider in deciding on land use grouping include the following:

- Dominant land use types
- Land uses subject to change or conversion
- Land use types where management changes are expected
- Spatial diversity within the watershed
- Availability of information on individual land use types

When grouping land uses, recognize that the summary of pollutant loading will be presented by land use category. Too many categories of land uses can be difficult to model, test, and report. Too few categories can result in oversimplification and generalization of the watershed conditions. Like so many aspects of watershed analysis, this decision depends on the local conditions and the management concerns being evaluated.

When selecting your land use grouping, think about the dominant features of your watershed and how they might change in the future. For example, in a watershed that is dominantly forested, the key land use categories might include various ages of trees (newly established, mature), logging roads, and small residential areas.

Changes under consideration might be forest practices/harvesting techniques, road removal, and road management.
For this watershed most of the detailed land use categories would relate to forest type and practice. In an urban watershed, forest might be grouped into a single category while numerous densities of urban land uses (e.g., commercial, industrial, high-density urban) are represented in more detail.

Model Testing

How do you know if the model is working appropriately? What kinds of tests can be performed to prove that the model is working? Before embarking on a detailed evaluation and the statistical testing of a model, you must first check the fundamental performance of the model. Check whether the model is working, evaluate the basic performance, and adjust or verify inputs if necessary. Then test for accuracy. In the early testing process, most modelers look at graphs of observed and simulated data and generalized summaries of flow and loading prediction. Initially, you're looking for ways to improve the model and identify features that might have been missed during setup. In the later part of model testing, you're looking for proof that the model is working well and providing reasonable results.

Testing involves comparing modeling results with observed data and the evaluation of the results for reasonableness. Practitioners always initially evaluate a model for simple conditions before moving on to utilizing actual monitoring data. For a process–based model, this is usually done at the completion of building the hydrologic/hydraulics components of the model and before water quality components are added. If the hydrology of a process–based model is incorrect, the water quality component will also be incorrect. It is more efficient to determine a hydrologic problem at this stage rather than to have incorrect water quality results and to find out later that the erroneous results were caused by an obvious error in hydrology.

A common method to use for a watershed model is to simulate a single moderate–sized storm event and determine if the outlet hydrograph is reasonable. If the outlet hydrograph rises and fails to recede for an unreasonable length of time, then there is an obvious problem with the watershed storage parameters that must be corrected. Another problem that can commonly occur is that the outlet hydrograph fails to materialize from the storm. It is at this stage of the process where knowledge of the system that the model user is attempting to simulate is critical.

If the model user is very familiar with how to run the model, but knows little or nothing about hydrology and hydraulics, the user will be unable to determine if the model is behaving correctly.

Empirical models can also be tested for reasonableness. An empirical model may have used a large dataset during development with a large range of values from various locations in the country.
Local data may only form a small subset of the national dataset, or may have not been part of the model dataset. Therefore, it is unlikely that the empirical model will predict a given concentration at a given moment in time that matches the locally derived data.

However, the model should produce reasonable responses to forcing functions such a climate and changes in land use. If the model is being used for comparative purposes (i.e., before and after BMP implementation), then the relative change in pollution loading is far more important than the prediction of pollutant concentration. Conversely, if a model is being used to predict a pollutant concentration for comparison to a water quality standard, then calibration to local water quality data is very important.

Presenting Pollutant Loads

As the modeling is completed, the model results need to be analyzed and presented to project stakeholders for use in completing the TMDL study. It is important to consider the information needs and model capabilities when presenting the pollutant load results from the modeling. Space and time are two factors that need to be considered in developing your presentation and communicating the results.

Care should be taken to consider your audience when developing the presentations. Presentations will often use a combination of tables, graphs, maps, charts, text, and spoken words. It is important for all formats to be clear and relatively easy to understand which will generally require significant work, but can mean the difference between success and failure.

Consider Spatial Scales

The scale in which pollutant loads need to be and/or can be evaluated is very important in a project. The discussion of factors affecting model selection in previous sections partially explains this. An example of spatial scale affecting the presentation of loads is in the requirement for TMDLs to provide wasteload allocations for all NPDES sources. With this in mind, the spatial scale of the modeling and modeled load estimates must take this requirement into account. A question to ask and answer is how will the WLAs for municipal separate storm sewer systems (MS4s) be presented.

They could be lumped together, but doing so will make implementation and development of each individual SWPPP more difficult and cumbersome to administer. Or, they could be kept separate, resulting in other constraints.

There are various options for assigning the spatial extent of load estimates; however, care should be exerted in evaluating the capabilities of a model in addressing different scales and determining the spatial scale needed for the TMDL study. You may want to quantify a gross load for the overall watershed or for each land use or even for each land use in each subwatershed, but individual models may not provide appropriate results at all scales. The detail to which you calculate the loads in the watershed will depend on the types and locations of the watershed sources identified during the data analysis and the model specifications.

Common Pitfalls When Selecting/Using Models

Avoid these common pitfalls when selecting and using watershed models:

- Selecting the wrong model for a given situation (location, watershed size, pollutants, time and spatial scales, information needs)
- Beginning the process with insufficient data (climate, land use, drainage, tillage practices, fertilization rates, etc.)
- Your technical team lacks experience using the model
- Underestimating the time needed to build the model
- Underestimating model complexity (need for large input of data, calibration and validation)
- The temptation to interpret output as absolute truth
- Using the model outputs to replace good planning
- Poor translation of the results to the public and stakeholders

Communicating Modeling Results

In addition to the discussion on the presentation of loading estimates above, keep the following suggestions in mind:

- Focus your message on 1-2 model outputs
- Discuss uncertainty in the model outputs
- Present data documenting the model's performance (i.e., does it simulate reality?) when applicable. Refer to discussion of model types and uses for determining applicability.
- Work to keep message as simple and clear as possible (develop 3 key points, use no more than 30 words per PowerPoint slide)
- When addressing skeptics, point out that while models are imperfect, they can be an adequate approximation of current and future water quality conditions. In addition, models can handle many more data sets than the human mind, which can make them more accurate than operating on a hunch.
- Identify strengths and weaknesses of the modeling effort.

Create an Official File for Modeling Activities

Documentation of all work is important as emphasized in each training module. As part of your TMDL project file, be certain to gather all information and data pertinent to your modeling efforts. Include the following materials in your file:

1. Model name and version
2. Documentation of applicability and appropriateness of the model(s) used
3. Model assumptions
4. Input data files
5. If calibration was needed, all calibration data and a detailed explanation of why it was used
6. A validation report explaining how the model replicates real-world conditions when appropriate (i.e., HSPF)
7. A scenario report providing likely water quality outcomes if various BMP solutions are implemented in the watershed

Next Steps in the Development of the Watershed Plan

Once you have calculated source loads for your watershed, you can move on to the next step of the TMDL process – allocating pollutant loads. The loads you have calculated will provide the basis for identifying the necessary load reductions for all major sources of the impairment within the watershed. This important step will set the stage for development of your Implementation Plan, which will include recommended management practices intended to meet load reductions.

Summary

- Models are used to integrate many sets of data in order to understand pollutant impacts to water quality.
- Models allow us to determine water quality impacts under different land management scenarios.
- There are several types of models available for use. It is important to select and use models following their intended uses.
- Models can be reasonably accurate, when they are applied properly and data sets used as inputs are of good quality.
- Any uncertainties about model outputs should be conveyed to stakeholders.
- Where necessary models should be calibrated to increase confidence in their ability to simulate real-world physical processes in a particular watershed (or subwatershed).
- Models should be validated to increase confidence in their outputs.
- Building a model takes a good deal of time – especially data collection and preparation.
- Select a model which best simulated you target pollutant(s) and expressions.
- Select the simplest model possible that meets your needs.
- Be certain to ask detailed questions and consult with MPCA modeling experts before selecting a model.

Resources

Chuck Regan, MPCA, Modeler 651-296-7363
Hafiz Munir, MPCA, Modeler 651-296-9286
Nick Gervino, MPCA, Modeler 651-296-8847
John Erdmann, MPCA, Modeler 651-282-5559

Written by

Greg Johnson, MPCA

References

Lin, J.P. 2004. *Review of Published Export Coefficient and Event Mean Concentration (EMC) Data*, ERDC TN-WRAP-04-03. U.S. Army Corps of Engineers, Engineer Research and Development Center, Vicksburg, MS.

Runkel, Robert L., Charles G. Crawford, and Timothy A. Cohn. 2004. "Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers." *Techniques and Methods Book 4, Chapter A5.* U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia.

USEPA. 2005. *Draft "Handbook for Developing Watershed Plans to Restore and Protect Our Waters*", Chapter 8 (portions in their entirety), 8-1 to 8-41.

Walker, W.W. 1996. *Simplified Procedures for Eutrophication Assessment and Prediction: User Manual*. Chapter 2, FLUX. Instruction Report W. U.S. Army Corps of Engineers, Water Operations Technical Support Program**.**