United States
Environmental Protection
Agency

Office of Water (4305)

EPA- 820-R-12-014
August 2012

AQUATOX (RELEASE 3.1)

MODELING ENVIRONMENTAL FATE AND ECOLOGICAL EFFECTS IN AQUATIC ECOSYSTEMS

TECHNICAL NOTE 2: REQUIREMENTS, SOURCES, AND CONDITIONING OF DATA FOR AQUATOX

AQUATOX Technical Note 2

Requirements, Sources, and Conditioning of Data for AQUATOX

Prepared by

Richard A. Park[1]
and
Jonathan S. Clough[2]

Performed under
EPA Contract EP-C-06-029
Work Assignments No. 4-11, 5-11
with
AQUA TERRA Consultants

and
EPA Contract EP-C-12-006
Work Assignment No. B-01
with
Horsley Witten Group

Submitted to
Marjorie Coombs Wellman
Office of Water/Office of Science and Technology
Standards and Health Protection Division
U.S. Environmental Protection Agency
Washington, DC 20460

August 2012

[1]Eco Modeling, Diamondhead MS 39525
[2]Warren Pinnacle Consulting, Inc., Warren VT 05674

Contents

Table of Figures

Acknowledgments

We wish to thank Marjorie Coombs Wellman and Brenda Rashleigh of the US EPA and Mehmet Umit Taner, presently at the University of Massachusetts Amherst, for their constructive criticism and useful suggestions for additional material, which definitely improved this Note. Work was performed under EPA Contract EP-C-06-029, Work Assignments No. 4-11 and 5-11 with AQUA TERRA Consultants; Tony Donigian's help is gratefully acknowledged. Some examples were also taken from work done with AQUA TERRA under EPA Work Assignment 2-11 (Tenkiller Lake) and SERDP Contract No. W912HQ-07-0026 (Fort Benning). This Note was completed under Contract No. EP-C-12-006 to The Horsley Witten Group, Inc., Nigel Pickering, Work Assignment Leader.

Disclaimer

This document describes ways of converting, conditioning, and using available data in the AQUATOX model. Anticipated users of this document include persons who are interested in using the model for various purposes, including but not limited to researchers and regulators. The model described in this document is not required, and the document does not change any legal requirements or impose legally binding requirements on EPA, states, tribes or the regulated community. This document has been approved for publication by the Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency. Mention of trade names, commercial products or organizations does not imply endorsement or recommendation for use.

Introduction

The aquatic ecosystem model AQUATOX has been designed to be extremely flexible, allowing many potential types of applications. Applications can span a wide range of time scale, geographic scale, stressor type, ecosystem or water body type, as well as a myriad of research or management questions that can be examined. AQUATOX is mechanistic, simulating the many physical, chemical and biological processes, with the result that it is also data intensive.

Because of this flexibility, data requirements for setting up an AQUATOX simulation depend considerably on the site type that is being modeled and the goal of the modeling study. For example, precise accounting of time-series loadings is especially important in rivers and streams; due to low retention time, inflow loadings can be the dominant factor affecting water column conditions. For standing water the importance of daily loadings follows a continuum from low-retention, run-of-the-river reservoirs, where daily loadings are important, to long-retention lakes and reservoirs that can be driven with seasonal data.

AQUATOX can accommodate a variety of modeling goals. A modeling study to assess the effects of changing nutrient levels in a particular water body (such as a site specific nutrient criteria study or a total maximum daily load study) would benefit from an excellent accounting of nutrients in the water column and nutrient boundary conditions, including loadings and initial concentrations in bottom sediments. On the other hand, a study that is primarily to evaluate fate and effects of new chemicals on a national or regional level may not require such detailed nutrient information. In that case, a user may wish to adjust the nutrient loadings in a representative or "canonical" environment to obtain a stable food web and then apply detailed information about chemical loadings in the simulation to assess fate and effects. Most major projects involving changes in land use require an environmental assessment, which could benefit from linking a hydrologic or watershed model and a receiving water model, such as AQUATOX, with the boundary conditions being defined by the hydrologic model.

Identification of environmental stressors that confound an analysis may involve a combination of well-defined drivers and calibration data that span several trophic levels. An emerging use of AQUATOX is in forecasting the effects of climate change, which is facilitated by setting up the model for probabilistic loadings, given the uncertainty of future conditions. Perhaps most demanding is modeling the time-to and extent-of recovery of ecosystems from pollution, which can entail accurate initial conditions, complex sediment-water relations, and repeating time series of environmental drivers for simulations that may represent decades.

The specificity and robustness of model formulations and associated parameter values should also be considered. It has been stated that models cannot be realistic, general, and precise at the same time (Levins 1966); one must choose two out of the three characteristics. While precision is inherently desirable, for many applications it is more important to have generality and realism (Park and Collins 1982). Precision is usually achieved by careful calibration to an extensive set of observed data; however, without generality there is no assurance that the calibration domain won't be exceeded under changing conditions—and most models are expected to forecast behavior under changing conditions. Depending on the goal, the tradeoff between precision and generality should be carefully considered.

Seldom do analysts have the luxury of being able to dictate what data should be collected. Often the model application comes after the original study and must rely on incomplete sets of data collected for other purposes. A prime objective of this Note is to describe ways of converting, conditioning, and using available data and of filling in the gaps with other sources. There is a logical progression in the use of data for any implementation of the model, and this progression is described in detail in the sections that follow. Briefly, the site is characterized by morphometric data, and the loadings at the site boundaries are specified, usually as time series. Initial conditions are given for each of the state variables, and the differential equations are populated with process-level parameter values. The parameter values are then calibrated—adjusted within reasonable limits to obtain simulation output that fits available site data, which may or may not need to be conditioned for purposes of comparison. Finally, sensitivity analyses may be performed to identify key parameters that might be adjusted further to improve the fit; and the model may be validated with data from another site or time period. To illustrate this process, in this Note we cite numerous examples in which data were obtained and conditioned for a variety of studies, in lieu of data collected for the specific applications.

Site Characteristics

Site characteristics and loadings are numerous, and setting them up for an application can be time-consuming. It is recommended that, whenever possible, existing studies be used as templates for new studies; one advantage is that the inputs can be viewed as a checklist and if site data are missing the data in the existing study can be used as defaults. Obviously, one should be careful to choose a comparable study and to note how the new study might differ from that "template." A brief description of each of the example studies is distributed with the AQUATOX software. As new studies are developed, they will be added to the set of example studies available to users.

Length

The model uses site length, which is measured in km, in several ways. Depending on the site type, it can be well defined or arbitrary. If modeling linked segments, it is the length of an individual segment. If modeling a lake or reservoir, it is the fetch or distance across which the wind blows and is used to calculate the depth of the thermocline, unless that estimated depth is defined by the user. If modeling a stream reach that is not linked, it is taken as an arbitrary length. In this case, the reach can be considered to be represented by a point model where the longer the length the greater the retention time. Very short reaches can result in stiff equations, where the step size has to be decreased to simulate the fast through-flow. This can be avoided by arbitrarily increasing the reach length, resulting in a longer retention time and a considerably faster simulation. We quite often use an arbitrary length of 2 km.

Width

A site characteristic with low sensitivity is channel width, measured in m. It is used in computing dynamic depth (see below). In the absence of data it may be estimated by zooming in on an aerial photograph, for example with Google Earth, and using the measuring tool on a representative traverse (Figure 1).

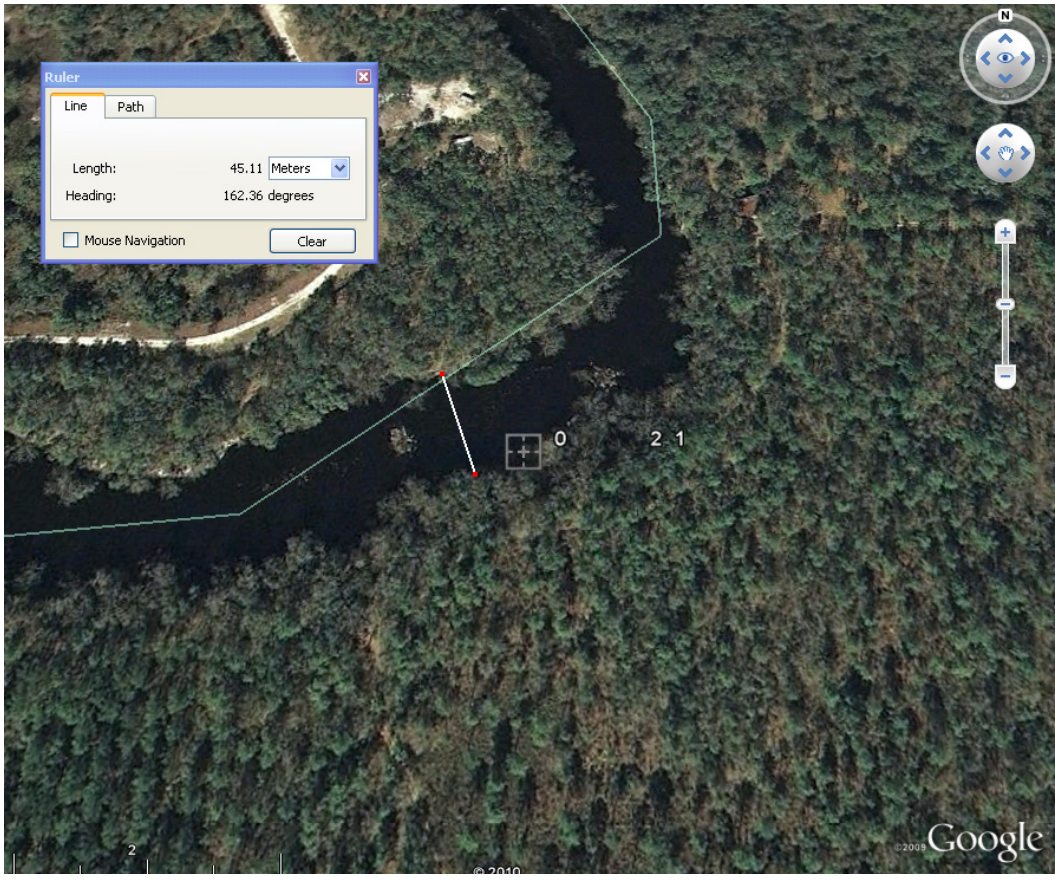

Figure 1. Channel width of Withlacoochee River FL measured using Google Earth.

Depth

Depth, measured in m, is input as "mean depth" and "maximum depth." The distinction between the two is important for lakes, reservoirs, and estuaries because they are used to define general bathymetric relationships and calculate entities such as the area of the euphotic zone. In general, these two parameters are important for computing the light climate at depth. Shallow streams and ponds are not very sensitive to the difference in these two variables, and we often arbitrarily take the mean depth to be one-half the maximum depth.

Time-varying water depth in streams is a function of the flow rate, channel roughness, slope, and channel width using Manning's equation, which is rearranged to yield (U.S. Environmental Protection Agency 2012):

$$Y = \left(\frac{Q \cdot Manning}{\sqrt{Slope} \cdot Width} \right)^{3/5}$$

where:

Y = dynamic mean depth (m),
Q = flow rate (m^3/s);
$Manning$ = Manning's roughness coefficient ($s/m^{1/3}$);
$Slope$ = slope of channel (m/m); and

Width = channel width (m).

The Manning's roughness coefficient represents frictional loss; it is not subject to direct measurement but can be used as a calibration parameter to estimate dynamic depth using observed depth and flow rate (Figure 2). In an example taken from the Parma site on the Lower Boise River, Idaho, time-varying depths were estimated using both default and calibrated roughness coefficients and were compared with observed depths. Although visual inspection indicates that the calibrated depth trend is a better fit, the default trend appears to be acceptable. The default value for the roughness coefficient for natural streams is 0.040; however, you may wish to choose a different value from Table 1 based on site characteristics. The calibrated value in this example is 0.055. A spreadsheet entitled **Depth_estimated_from_flow.xls** is supplied with the example studies as a part of the Release 3.1 distribution.

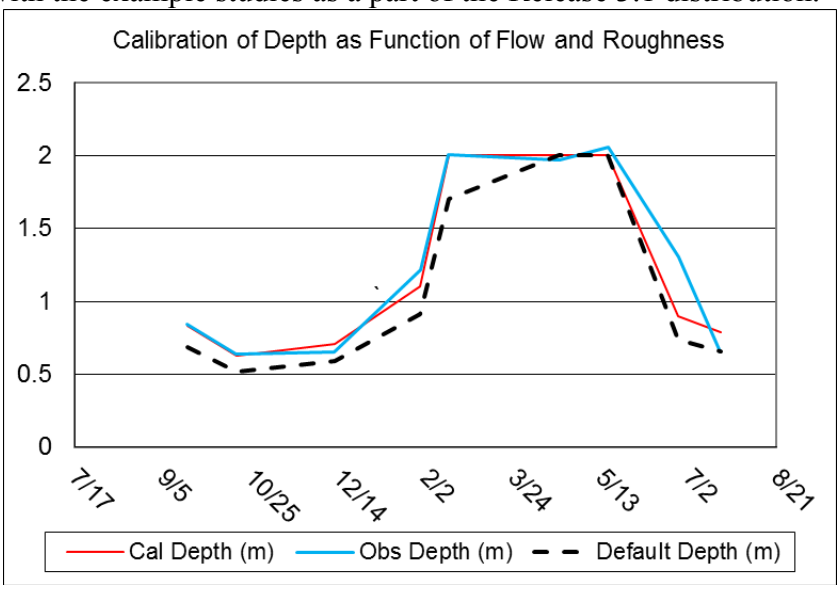

Figure 2. Comparison of calibrated and observed depths of water, Lower Boise River near Parma, Idaho.

Table 1. Values of Manning's roughness coefficient ($s/m^{1/3}$). Based on (Cowan 1956).

Type of Channel	Minimum	Normal	Maximum
Minor alluvial stream			
Clean, straight	0.025	0.030	0.033
Same, but more stones or weeds	0.030	0.035	0.040
Clean, winding, some pools & shoals	0.033	0.040	0.045
Same, but some weeds or stones	0.035	0.045	0.050
Same, but more stones	0.045	0.050	0.060
Sluggish reaches, weedy, deep pools	0.050	0.070	0.080
Very weedy reaches, deep pools	0.075	0.100	0.150
Minor mountain stream			
Bottom: gravel, cobbles, and few boulders	0.030	0.040	0.050
Bottom: cobbles with large boulders	0.040	0.050	0.070

Major river			
Regular section with no boulders or brush	0.025	n/a	0.060
Irregular and rough section	0.035	n/a	0.100

Thermocline Depth

The depth to the thermocline in lakes and reservoirs is normally estimated by AQUATOX using by a regression relationship based on length (U.S. Environmental Protection Agency 2012):

$$\log(MaxZMix) = 0.336 \cdot \log(Length) - 0.245$$

where:

MaxZMix = maximum mixing depth under stratified conditions (thermocline depth) for lake (m); and

Length = maximum effective length for wave setup (m, converted from user-supplied km).

AQUATOX Release 3 also accepts a constant or time series of user-specified thermocline depths; this allows one to specify varying depths based on interpretations of available temperature and dissolved-oxygen isopleths. Or one can force a particular thermocline depth by varying the length of the lake or reservoir. This was done in an earlier implementation of the model for Lake Onondaga NY, in which higher salinity water limits the mixing depth; in that case changing the length from 7.6 km (the true length) to 2.376 km decreased the depth appreciably to that observed for part of the stratified period. The results of all three options are shown in Figure 3.

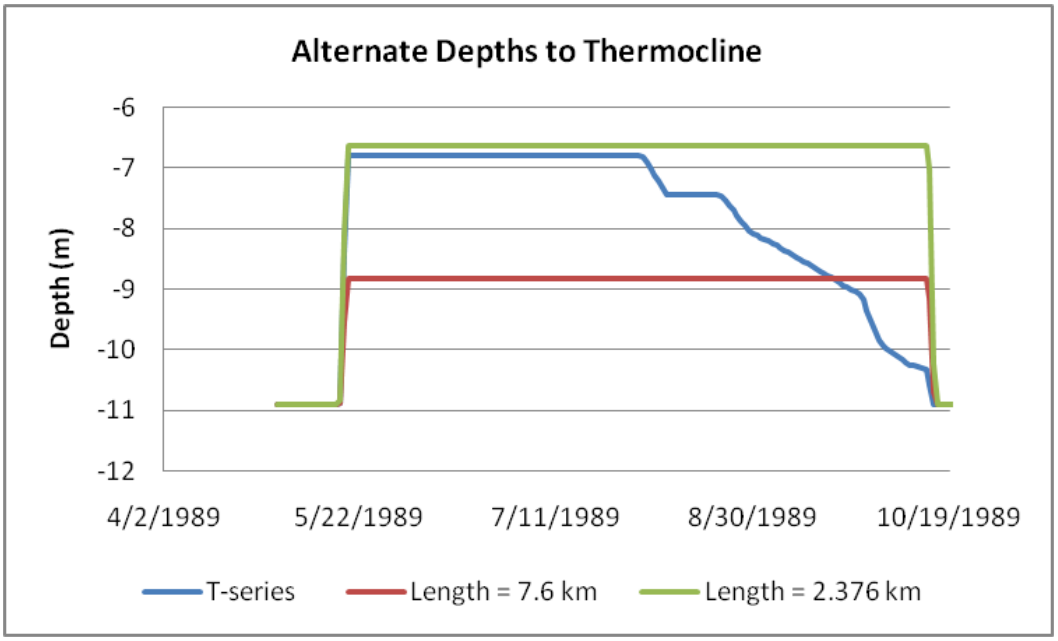

Figure 3. Depths to Lake Onondaga thermocline estimated from the observed site length of 7.6 km, and using a time-series of observed depths and a constant depth imposed by changing the length.

Boundary Conditions

Inflow and Discharge

Probably the most important environmental loading is water flow because it is the carrier for nutrients, suspended sediments, plankton, and organic pollutants. Depending on how it is used with respect to volume, it may be input as inflow, outflow, or both. For any site other than a long-retention lake, it should be input as daily values if possible; that is especially important with respect to streams. In the U.S. observed flow data are often taken from a USGS gage, with the values in cubic feet per second converted to cubic meters per day. If the gage is not located at the site then the flow should be adjusted for the difference in the drainage area. For example, in a Florida study the USGS gage on Holmes Creek is 10 km downstream from the site being modeled. Based on the drainage, about one-third of the gaged watershed is below the site (Figure 4); therefore, the flow loading was reduced to two-thirds of that observed at the gage. Of course, where possible it is better to estimate the flow using a hydrologic model such as HSPF calibrated to the watershed (including both flow and precipitation data), rather than using an approximation of the type just described. As an example, in a Minnesota study HSPF was used to obtain estimated flow at a site at river mile 54 on the Blue Earth River, 36 miles above the USGS gage (Figure 5).

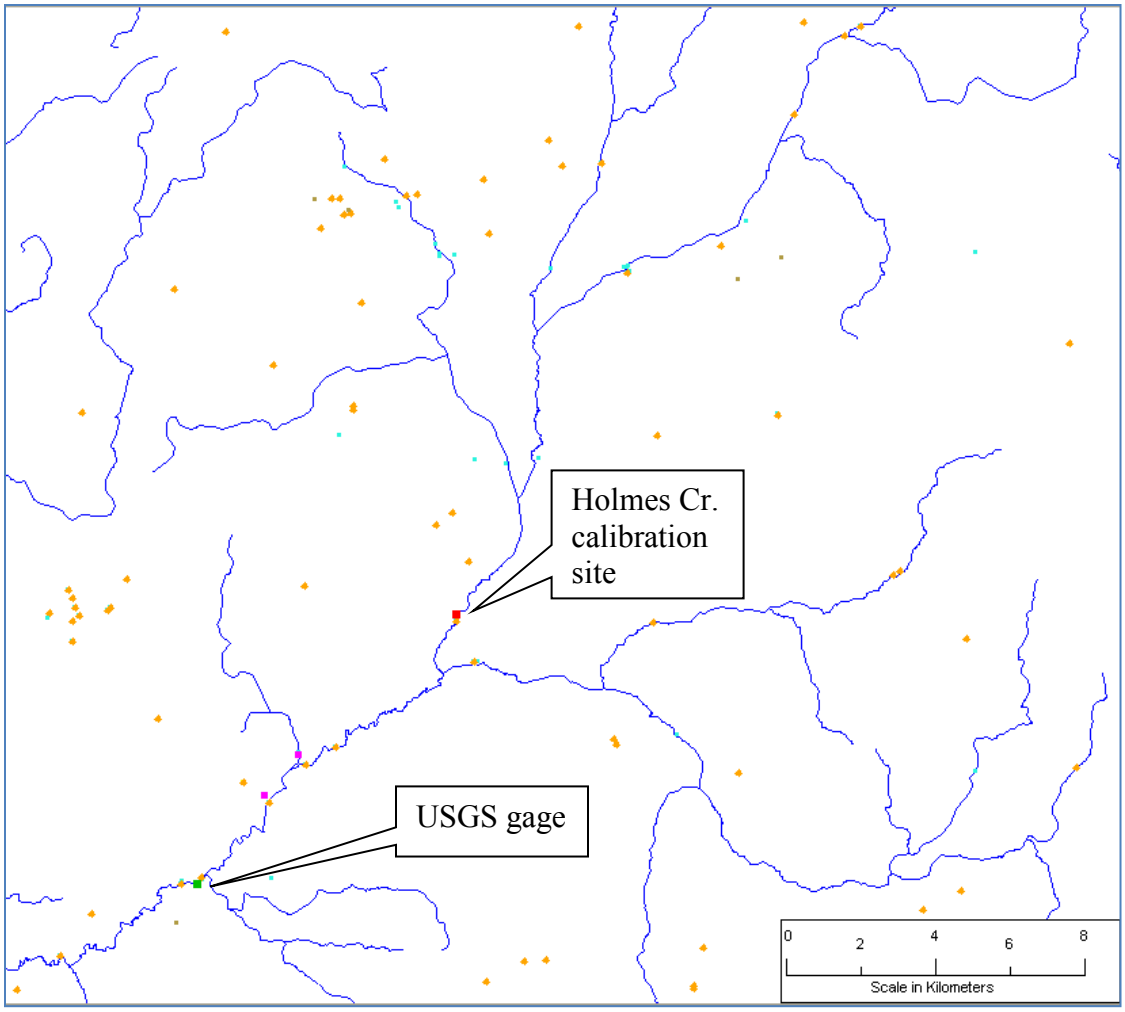

Figure 4. Drainage pattern for Holmes Creek, FL; note tributary from the east below the site being modeled and above the gage.

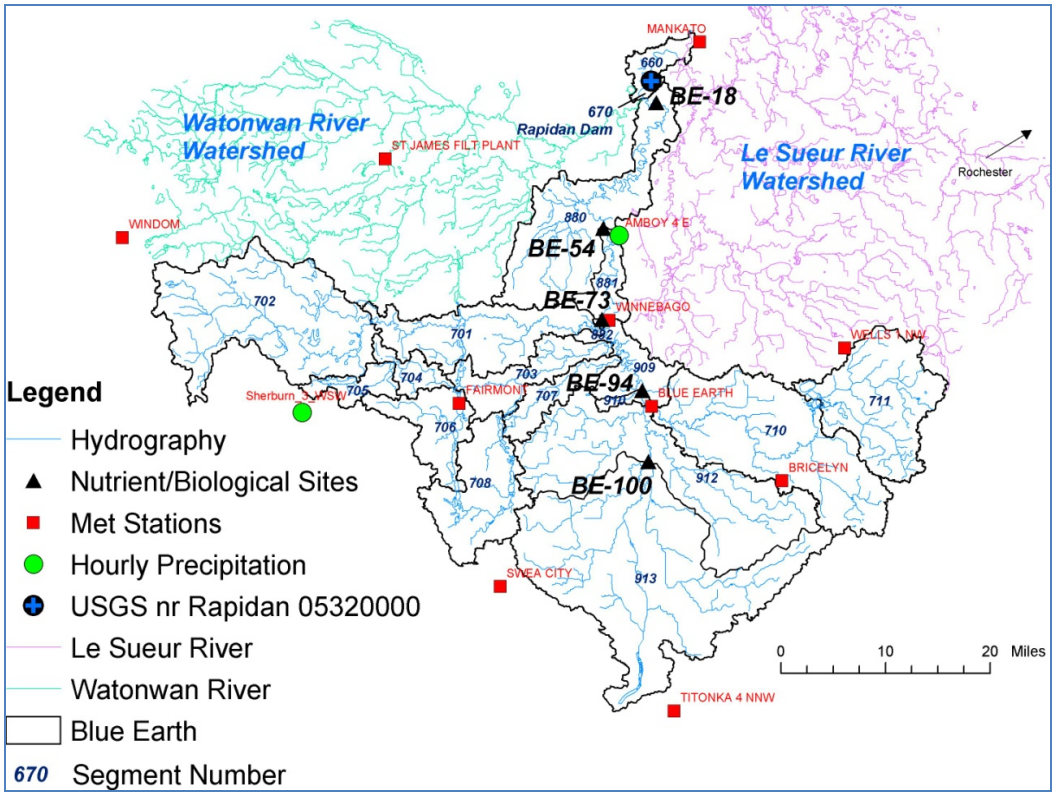

Figure 5. Blue Earth MN watershed showing segments, USGS gage (BE-18), and meteorological stations (Donigian et al. 2005). BE-54 is the site modeled with AQUATOX.

There is a spectrum of reliability for flow data: best is data from an adjacent gage, next best is calibrated flow estimate from a hydrologic model, followed by data adjusted for distance to gage, and finally data on a "comparable" watershed. The latter approach can be useful, for example, if the site of interest has a very small watershed with response times that cannot be represented well by taking a downstream gage and scaling back the flow. A small watershed will be "flashy" in comparison to a larger watershed, so picking a gage on a nearby small watershed may be the best approach.

Flow in a linked-segment reservoir is a special case. The model can accommodate main stem and tributary flow into the segments, flow from one segment to another, overland flow into segments (best computed by a hydrologic model), and withdrawals from the reservoir. Flow data for the linked-segment version will be discussed in further detail below.

Also of concern is the time period being simulated. Preferably a simulation can include a wet year, a dry year, and a normal hydrologic year. The Little Withlacoochee River in Florida is an extreme case, where the river dried up during a couple of droughts (Figure 6). AQUATOX is able to step through dry periods with the implicit assumptions that the time period is short and that there are refuges so that the state variables do not change during that period. A word of caution: if the time series shown in Figure 6 were extended to a longer period of time, AQUATOX would "wrap around" the last year of loadings so that the drought would be repeated. The best way to avoid this, given the example, would be to obtain a longer hydrologic

record; otherwise, one could generate a synthetic time series offline based on a typical year of data.

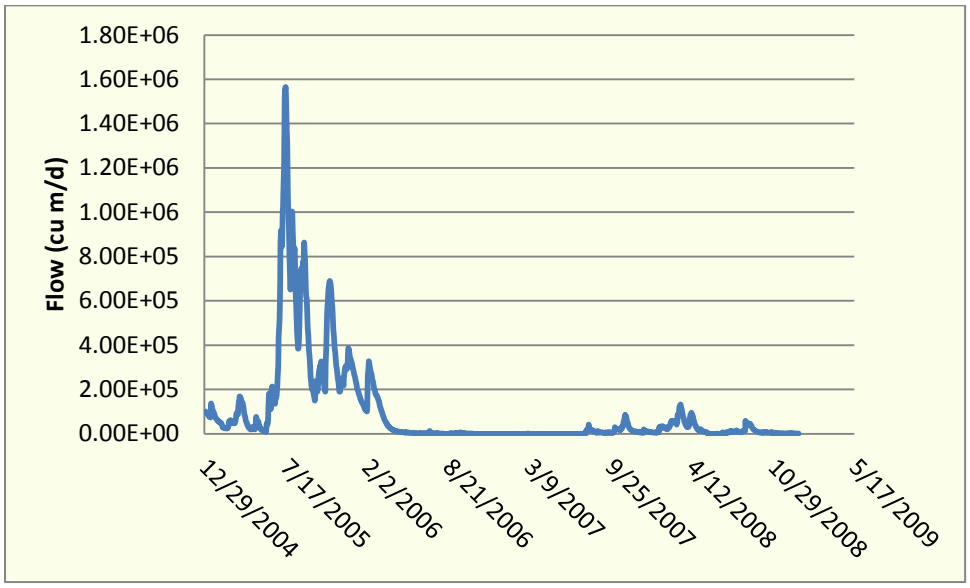

Figure 6. Discharge on the Little Withlacoochee River, FL; note effect of prolonged drought.

Velocity is usually computed by the model based on flow and cross-sectional area, but it can be entered by the user in the site loading screen. The model can be sensitive to velocity, which affects:

- washout and deposition of phytoplankton and scour of periphyton;
- breakage of macrophytes;
- entrainment of zooplankton;
- deposition or scour of organic matter;
- scour and deposition in the sand-silt-clay model; and
- oxygen reaeration.

Linked-Segment Flow

Flow among segments in a lake, reservoir, or estuary is best obtained from a hydrodynamic model such as EFDC (Environmental Fluid Dynamic Code) (Tetra Tech Inc. 2002). In an application to Tenkiller Lake OK, EFDC was used to obtain flow rates between horizontal segments of the reservoir, and exchanges between vertical segments (i.e., epilimnion and hypolimnion) were calculated from observed temperature differences. These flows are reflected in the time-varying volumes of the reservoir segments (Figure 7). Note that the riverine segment is run-of-the-river with low volume and retention time.

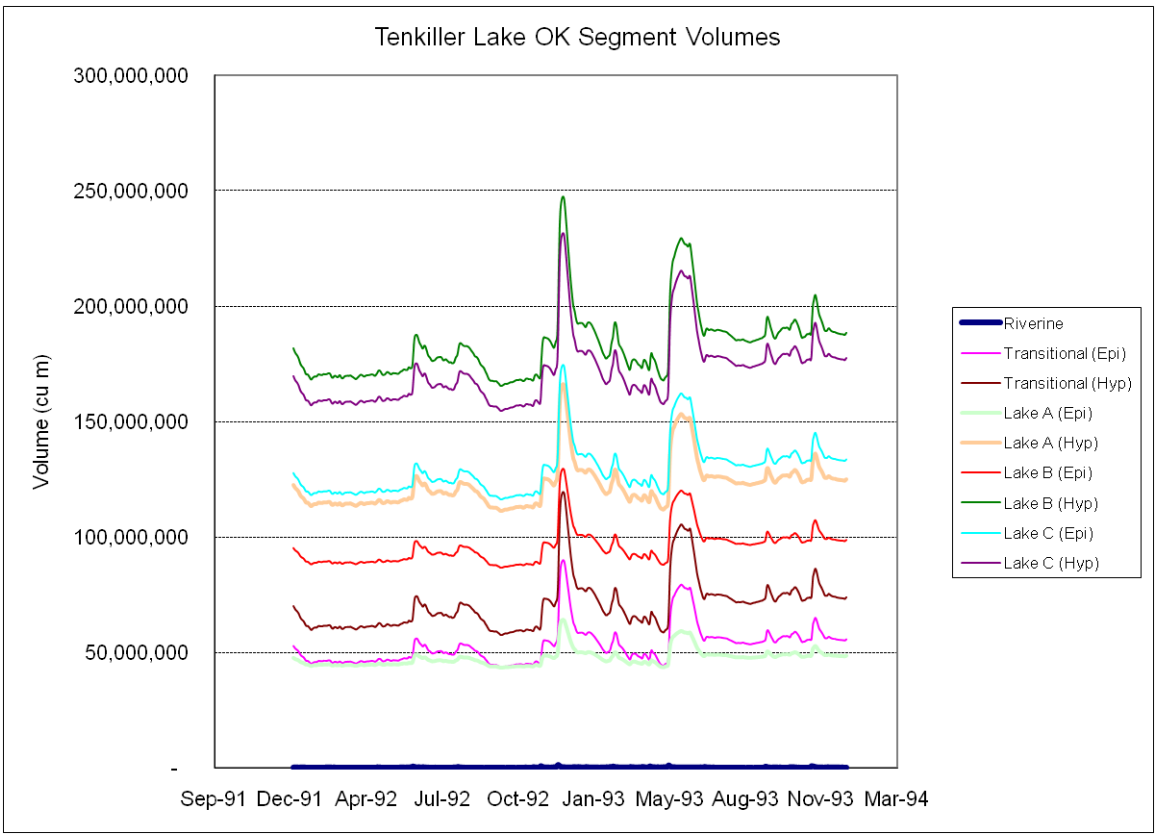

Figure 7. Time-varying segment volumes in Tenkiller Lake, including Lacustrine A, B, and C epilimnion and hypolimnion segments.

For a river, flow among segments can be obtained from a model such as EFDC or, better yet, from a hydrologic model, such as HSPF (Hydrologic Simulation Program-Fortran) (Bicknell et al. 1991) that accounts for time-varying differences in runoff from adjacent watersheds. In an application to the Lower Boise River ID, which is an intensely managed river with diversion canals and return drains for agricultural irrigation, a flow-routing spreadsheet model was developed to keep track of the complexities.

Nutrient Loadings

Nutrients are important drivers of primary productivity and therefore the basis of the aquatic food web. In fact, AQUATOX requires nutrient loadings in order to run. Nutrient loadings can be input in several ways:

- Concentrations (mg/L) in inflow waters (the most common loading)
- Non-point source loadings (g/d)
- Point-source loadings (g/d)
- Direct precipitation, which includes dry fall (g/m^2 d)
- Concentrations (mg/L) in linked tributaries (linked segment version)

Nutrient loadings can be derived from observed data collected at the site over time. Ideally these would be daily values; however, in practice nutrient data are often weekly, monthly, or sporadic

observations; in temperate zones nutrient data in the winter are rare. In a project on multiple streams in Florida quarterly averages were used for many sites, which added a degree of uncertainty and decreased the temporal resolution of the ecosystem responses. AQUATOX calculates nutrient loadings from streams by interpolating the input concentrations and multiplying by the daily discharge. A more flexible and rigorous method is to use the FLUX program (Walker 1999), available from the U.S. Army Corps of Engineers Environmental Laboratory, to calculate the loading based on any one of six methods. Developed for calculating loadings for reservoirs, the procedure can also be used for streams and lakes. Daily estimates of nutrient loadings also can be obtained from a linked hydrologic model such as HSPF in much the same way that water flows can be obtained. In fact, HSPF was used to drive both the Blue Earth River MN (Donigian et al. 2005) and the Tenkiller Lake OK (Donigian et al. 2009) implementations. These linkages have the advantage of facilitating "what if" scenarios where the potential effects of land-use changes can be examined (Carleton et al. 2009, Donigian et al. 2009). The effects of climate changes could also be modeled using these linked models.

Linked segments present almost endless possibilities for disaggregating and loading nutrients from diverse sources. In the Lower Boise River ID study, the river was divided into 13 segments, including channels on opposite sides of Eagle Island (CH2M HILL et al. 2008). Nutrient loadings came from wastewater treatment plants, groundwater, fish hatcheries, tributaries, and drains (return flow from irrigation) (Figure 8). Each of these sources can be subject to perturbations and analysis of impacts in subsequent simulations.

Detrital Loadings

One of the potentially important loads for nutrient and eutrophication analysis, but one which is often missing or in a different form than that used in AQUATOX, is organic matter, or detritus. As stated in Section 5.1 of the *Technical Documentation* (U.S. Environmental Protection Agency 2012): "For the purposes of AQUATOX, the term 'detritus' is used to include all non-living organic material and associated decomposers (bacteria and fungi)…. Detritus is modeled as eight compartments: refractory (resistant) dissolved, suspended, sedimented, and buried detritus; and labile (readily decomposed) dissolved, suspended, sedimented, and buried detritus. This degree of disaggregation is considered necessary to provide more realistic simulations of the detrital food web; the bioavailability of toxicants, with orders-of-magnitude differences in partitioning; and biochemical oxygen demand, which depends largely on the decomposition rates…. In general, dissolved organic material is about ten times that of suspended particulate matter in lakes and streams (Saunders 1980), and refractory compounds usually predominate; however, the proportions are modeled dynamically."

The effort necessary to obtain these loadings depends on how important the detrital components are in a particular study. Oftentimes a general value of 10% Particulate is used (see above paragraph); % Refractory is more dependent upon the site. An extreme case where detrital dynamics are important is the Upper Suwannee River FL—a blackwater river that drains the Okefenokee Swamp. Three different measures of organic matter were available for the river, and those were used to populate time-series loadings (Table 2).

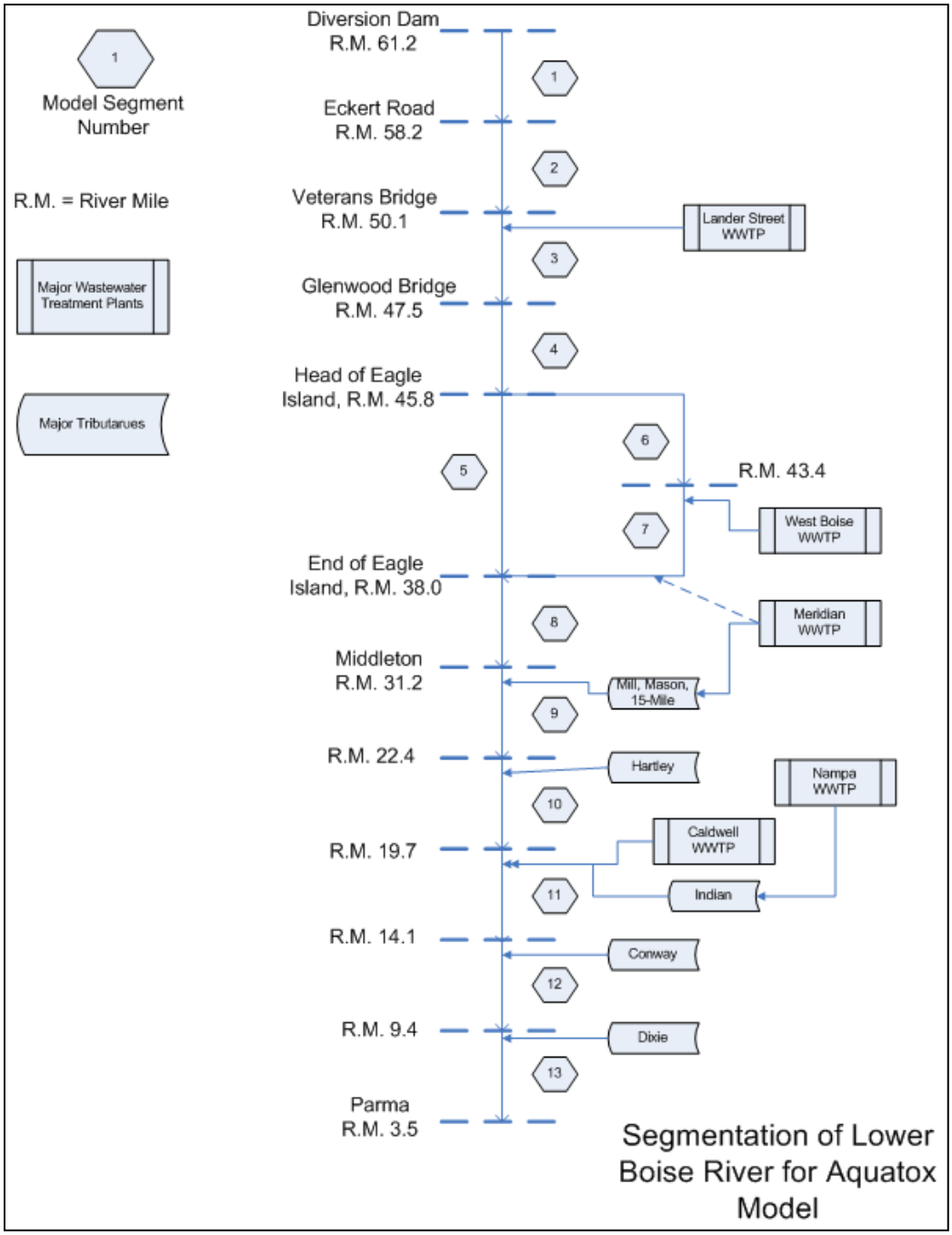

Figure 8. Schematic showing segmentation and major linkages in the Lower Boise River ID model (CH2M HILL et al. 2008).

Table 2. Observed and estimated detrital characteristics, Upper Suwannee River FL.

Date	COLOR* Ptu	TOC* mg/L	BOD_5* mg/L	Est. DOC (mg/L)	Estimated %Particulate
6/4/1990	335	50	1.55	37.70	25

*Observed

The observed Total Organic Carbon (TOC) was used as the primary loading to the Upper Suwannee River, where it was converted to organic matter by the model using a conversion factor of 1.9 grams of organic matter per gram of organic carbon (Winberg 1971). Color was used to estimate Dissolved Organic Carbon (DOC) using a weak regression relationship based on data from numerous Florida sites (Figure 9). In turn, comparing observed TOC to estimated DOC, the % Particulate can be calculated. Of course, each of these conversions should be checked for consistency; for example, the estimated DOC might exceed observed TOC on a particular date, in which case it would need to be constrained.

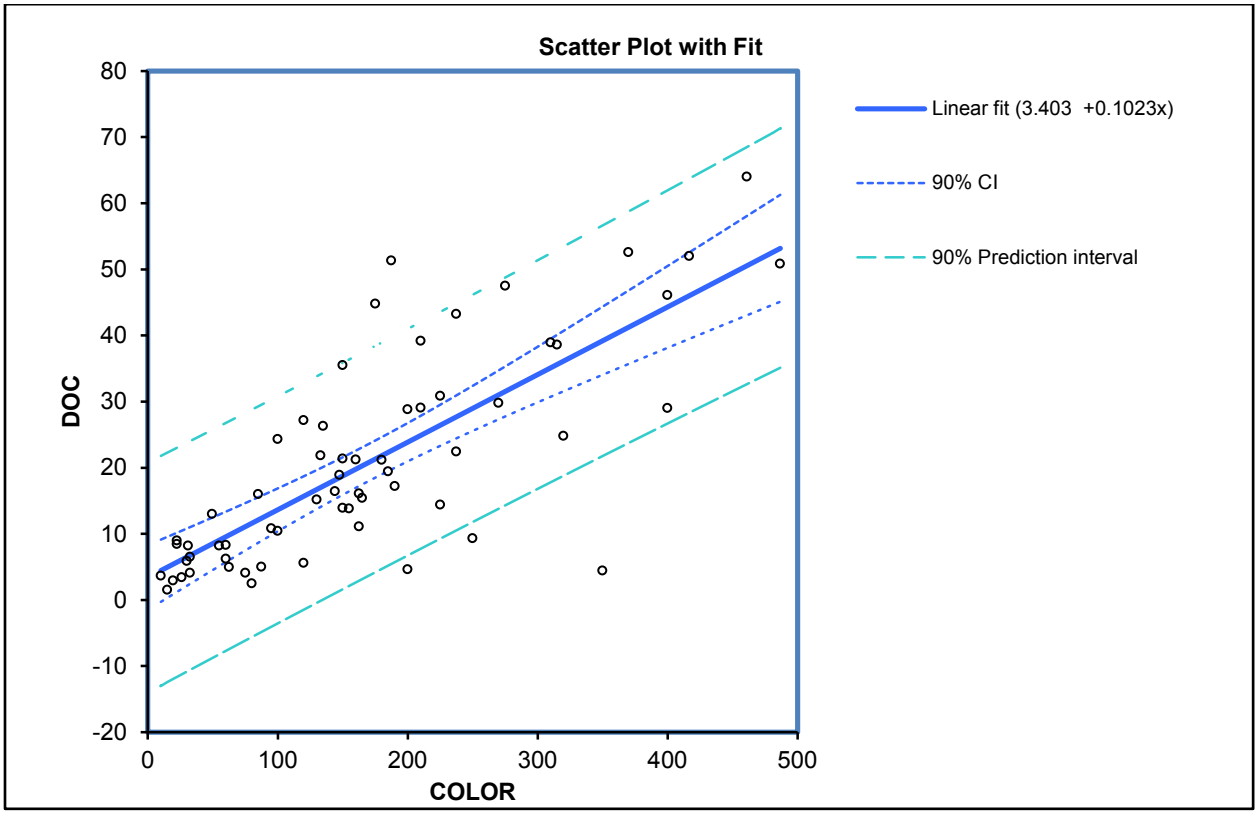

Figure 9. Regression relationship between color and dissolved organic carbon, based on Florida Department of Environmental Protection (DEP) data.

The same guidelines apply to detrital loadings as to nutrients; that is, where possible either observed or simulated daily loadings should be used for low-retention sites. For example, HSPF was used to obtain estimated BOD loadings for Tenkiller Lake based on land use in the watershed (Donigian et al. 2009). There were no calibration data for BOD, so the loadings were used with appreciable uncertainty; however, data collected later showed that the maximum value simulated by HSPF was remarkably close to the observed maximum for later dates.

Total Suspended Solids

TSS can be an important loading, especially as an alternative to modeling sand, silt, and clay. Daily values are highly desirable for streams and reservoirs, where sporadic light limitation and sedimentation can affect algae and zoobenthos. Unfortunately, TSS is often not measured or is measured only occasionally. Two Minnesota streams that differ greatly in TSS exemplify two

approaches to representing this loading, as described in *AQUATOX Technical Note 1* (Park et al. 2009). The Blue Earth River demonstrated a close relationship between TSS and flow, and that relationship was used to obtain daily loadings (Figure 2 in (Park et al. 2009). On the other hand, data from the Crow Wing River, which is located on well-sorted glacial outwash sands, showed no significant variation in TSS values with differing flow rates. In this case, TSS was based on an HSPF calibration that incorporated more complex controls in the watershed such as upstream lakes (see Figure 4 in (Park et al. 2009).

Dissolved Oxygen

Sufficient dissolved oxygen is, of course, vital for the survival of aquatic animals; however, accurate simulation is important only when concentrations drop below about 3 mg/L. Oxygen dynamics are more or less sensitive to loads depending upon the type of water body. Dissolved oxygen loadings in a stream present a challenge. Because of low retention times, loadings may completely dominate concentrations in a particular reach, reflecting upstream conditions rather than processes occurring in the reach; both sources affect observed data and are usually simulated. All loadings can be turned off in the oxygen loading screen in order to isolate in-reach processes, providing a valuable perspective on DO concentrations. However, this is an analytical technique and can distort the results if the oxygen loads are left "turned off".

As an example of how the influence of dissolved oxygen loads can vary, Reach 3 of the Lower Boise River ID is 13 miles downstream from the Diversion Dam (Figure 8), which is the source of daily loadings for the model. There are very few periphyton under conditions of high flow, and water from the dam passes through Reach 3 quickly, accounting for most of the DO; the daily loadings correspond closely to the mean observed in Reach 3 (Figure 10). However, with low-flow conditions the retention time above Reach 3 increases considerably and periphyton are able to grow; therefore, the predicted DO no longer corresponds entirely to the loadings at the dam, and photosynthesis increases the range of predicted diel DO (Figure 11).

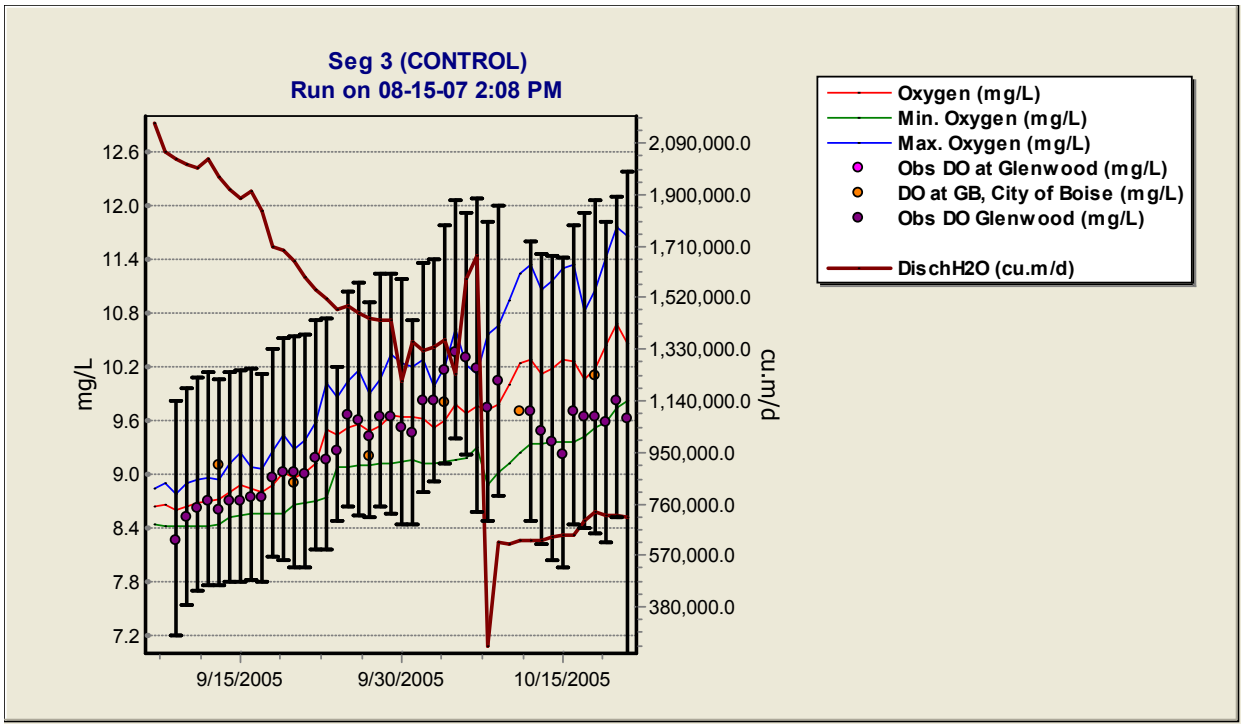

Figure 10. Predicted and observed DO (Y1 axis) and discharge (Y2 axis) in Reach 3, Lower Boise River ID. Black error bars represent daily minimum and maximum oxygen observations at Glenwood.

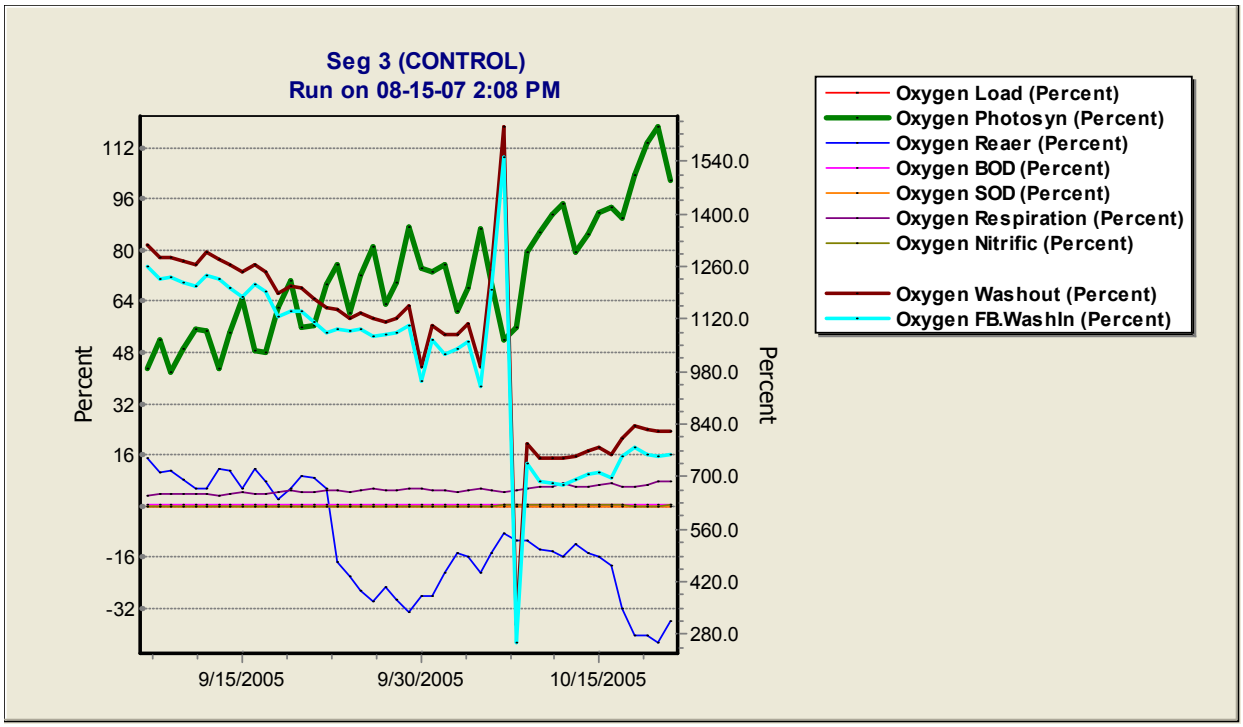

Figure 11. Predicted rates for dissolved oxygen in Reach 3, Lower Boise River ID. Note that Washin and Washout use the Y2 scale.

On the other hand, DO loadings from tributary streams may be an important source of oxygen in lakes and reservoirs during the winter when ice cover may impede photosynthesis and preclude reaeration. Otherwise, DO loadings are seldom critical for a lacustrine epilimnion. The hypolimnion may receive much-needed oxygen from inflowing water, but that may be offset by the BOD from concurrent detrital loadings.

For smaller streams, hourly oxygen loadings may be input in conjunction with an hourly AQUATOX simulation. This can vastly improve the simulation of oxygen fluctuations within each day; however, hourly oxygen data are rarely available.

pH

pH is important in AQUATOX because it affects precipitation of calcium carbonate and sorbed phosphorus and it can control hydrolysis of organic toxicants and the fraction of total ammonia that is in its more toxic form. AQUATOX can predict time-varying pH, but the procedure requires good estimates of alkalinity and dissolved organic matter. Perhaps the best example of computed pH is the simulation of the Upper Suwannee River FL. Alkalinity was held constant at 220 ueq $CaCO_3$ (an input variable on the pH screen), but the principal factor contributing to time-varying pH was the variation in dissolved refractory detritus (mainly humic acids). The algorithm constrains the minimum pH to 3.75, which was achieved with higher concentrations of dissolved organic matter (Figure 12). However, it is often better to use observed values of pH, and that is our practice with most studies; it is commonly available.

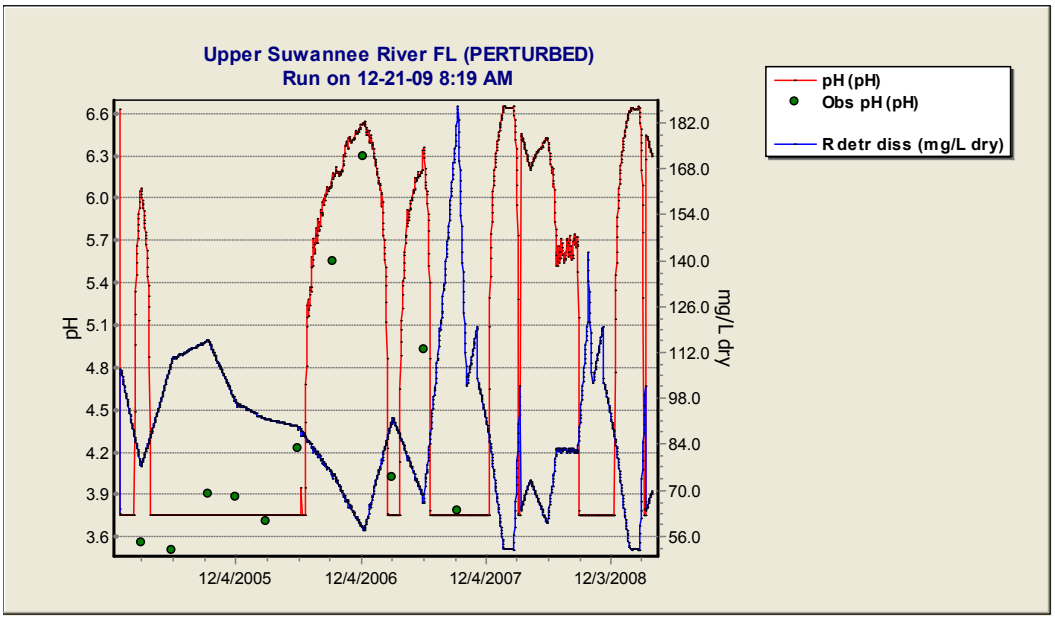

Figure 12. Predicted pH at Upper Suwannee River FL site, and inverse relationship to dissolved organic matter.

Light

Seldom are light observations available for a site. However, broad seasonal patterns of solar radiation will usually suffice; with correction for riparian shading if necessary (see below). AQUATOX computes time-varying light given the annual mean and range.

A very useful source is a NASA Web page, which can be used to estimate solar radiation and other climate characteristics worldwide. The page changes frequently, so search on "**NASA**, **Surface meteorology and Solar Energy."** An example that uses NASA data is given for the Upper Suwannee River site (Table 3). From the table one can obtain the annual mean, and by scanning for the annual maximum and minimum values the range can be calculated. The values are in kWh/m^2 d, which AQUATOX will convert to Ly/d when entered in the site screen. It is not uncommon to find values in other data sets expressed as $\mu Ein/m^2$ s, which the model will also convert.

Table 3 Monthly averaged insolation incident on a horizontal surface ($kWh/m^2/d$) (NASA 2010).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2004	3.21	2.83	5.35	6.29	6.58	5.56	5.97	5.07	4.17	4.15	3.36	2.98	4.64
2005	3.10	3.42	4.44	5.97	6.28	5.16	5.84	5.21	5.54	4.05	3.65	2.85	4.63
2004 - 2005	3.16	3.12	4.89	6.13	6.43	5.36	5.91	5.14	4.86	4.10	3.51	2.92	4.63

In narrow streams with dense riparian vegetation, light can be very limiting during the growing season. The user can also specify shading as the fractional coverage by riparian canopy, either as a constant or as a time series. In the absence of data, sources of aerial photographs such as Google Earth provide a means to obtain an approximate canopy value by zooming in on a site, sometimes with accompanying photographs (Figure 13).

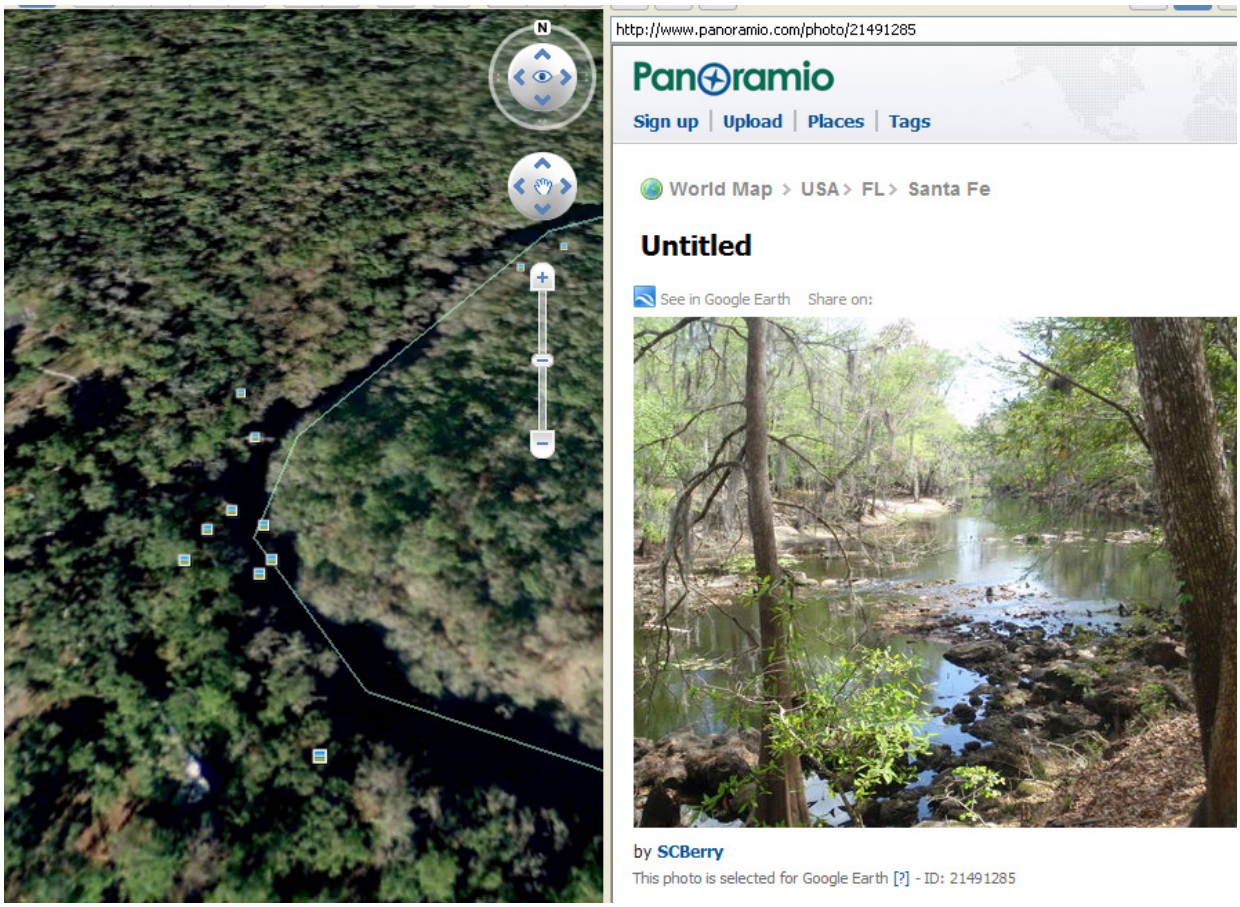

Figure 13. Google Earth aerial photo of Santa Fe River FL site with accompanying photograph, useful in estimating fraction of canopy (0.5).

Temperature

Temperature is an important driving variable, affecting many physical and all chemical and biologic processes; and this is confirmed by sensitivity analysis. However, it seldom changes very rapidly in aquatic systems, so that a sinusoidal seasonal approximation based on user-specified annual mean and range is usually sufficient. Alternatively, time series of observed values can be specified. Hydrodynamic-thermal stratification models can be used to compute vertical temperature profiles based on climate and hydrological drivers. If the site is a stream, loadings can be obtained from a hydrologic model such as HSPF. Figure 14 illustrates the difference between simulated temperature loadings obtained from HSPF and those calculated using the annual mean and range. The simulated temperature would appear to be preferable, but the sinusoidal curve based on minimal data seems to be adequate.

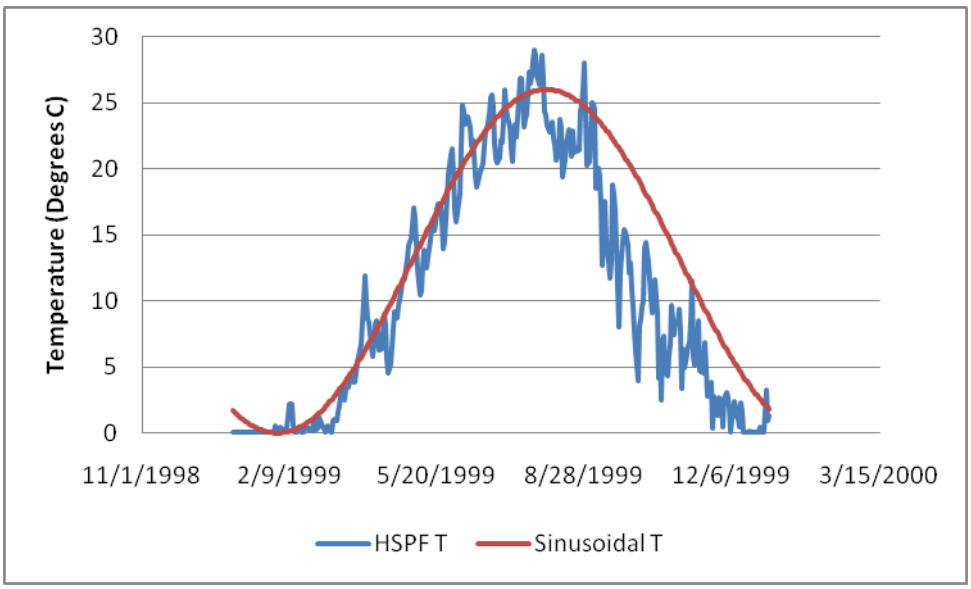

Figure 14. Temperature in Blue Earth River MN based on HSPF predictions and annual mean and range.

Organic Chemicals

There are numerous applications of AQUATOX involving potentially toxic organic chemicals, with varying levels of precision required for the chemical loadings. Simulating time to recovery from historic pollution often involves painstaking specification of initial conditions including concentrations (µg/kg) in buried sediment layers and associated pore waters (µg/L). Analysis of whether a chemical has a high potential of being of concern may just require an observed concentration used as a constant. This was done with Blue Earth River MN, using a single reported atrazine concentration, to make sure that the herbicide was not affecting the algal calibrations. A common application is environmental risk analysis of pesticides and other toxicants, often using a representative site (sometimes called a "canonical environment"). Three types of analyses can be performed by AQUATOX on the same study with minimal effort:

- Pulsed loadings of toxicant corresponding to runoff during storm events, provided by the analyst or predicted by a model such as PRZM (Suárez 2005) or SWAT (Arnold et al. 1998)
- Initial toxicant concentration with no loadings (uncheck "Use Dynamic Loadings" and ensure that constant loadings are set to zero)
- Initial concentration held constant

All three can be performed with the Ohio Stream Chlorpyrifos studies that are provided as examples in the AQUATOX installation. In the "Pulsed" study the initial condition is set to the maximum observed (Figure 15), thus when the dynamic loadings are turned off there will be the single initial concentration of chlorpyrifos. Setting the simulation to "keep toxicant constant" will result in a constant concentration of 0.4 µg/L, unless the initial condition is changed.

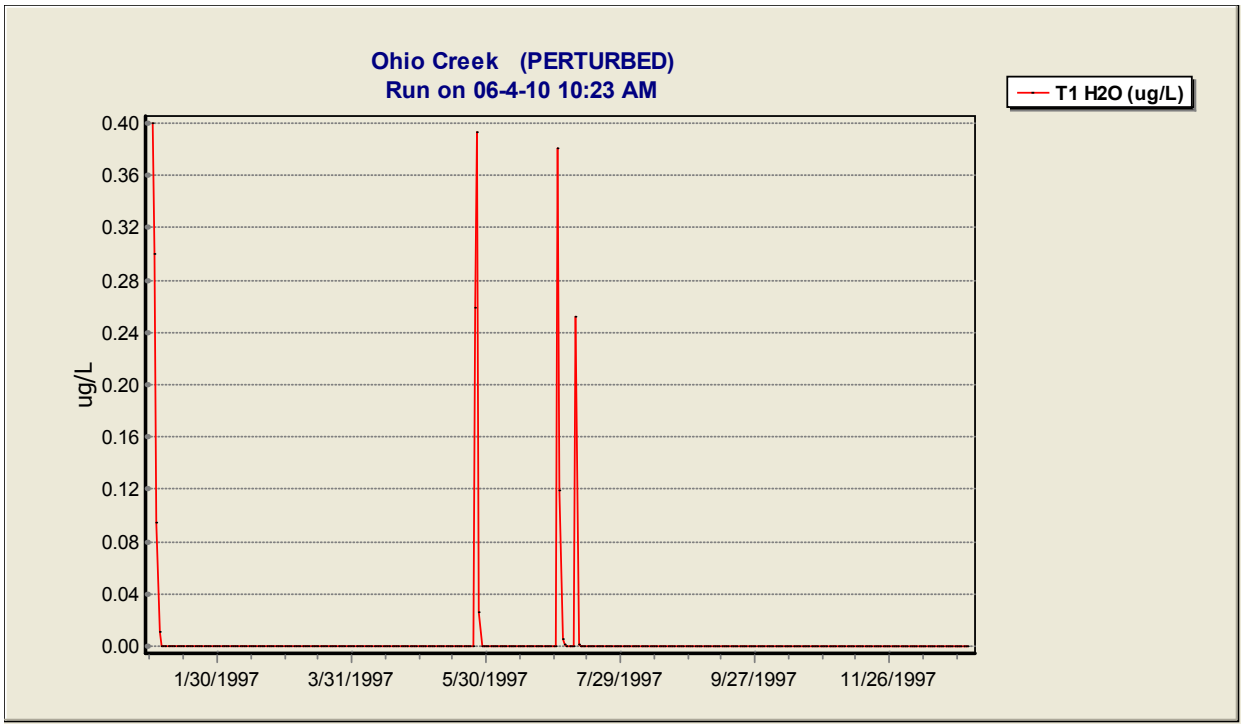

Figure 15. Concentration of chlorpyrifos in the Ohio stream with pulsed loadings.

Biotic Loadings

Usually biotic loadings are set to an arbitrarily small number to serve as a "seed" to simulate recolonization following extinction, and therefore allow the species to establish itself under improved environmental conditions. Such a "seed" is appropriate for algae and aquatic insects. The analyst should consider whether recolonization can occur for other invertebrates and fish; for example, if fish are driven to extinction in a pond they may not repopulate the site. On other hand, biotic loadings may be quite important for a reservoir, where upstream productivity in an unmodeled tributary can provide a seasonal or continuous supply of organisms to the reservoir. For example, the Illinois River transports significant biomass of sestonic algae (measured as chlorophyll *a*) into the riverine segment of Tenkiller Lake OK. AQUATOX models individual algal groups, and combines them for an estimate of chlorophyll; but frequently algal species data are not available, as was the case in the Tenkiller. The chlorophyll value was converted to biomass (mg/L) and split evenly among four phytoplankton groups (Donigian et al. 2009), as seen in Figure 16. Stocking can also be an important factor in lakes and reservoirs managed for fishing. In a study of DeGray Lake AR stocking records (in numbers of fish per acre) were available for hybrid striped bass (Fourt et al. 2002), and those were converted and used in modeling that reservoir (Figure 17).

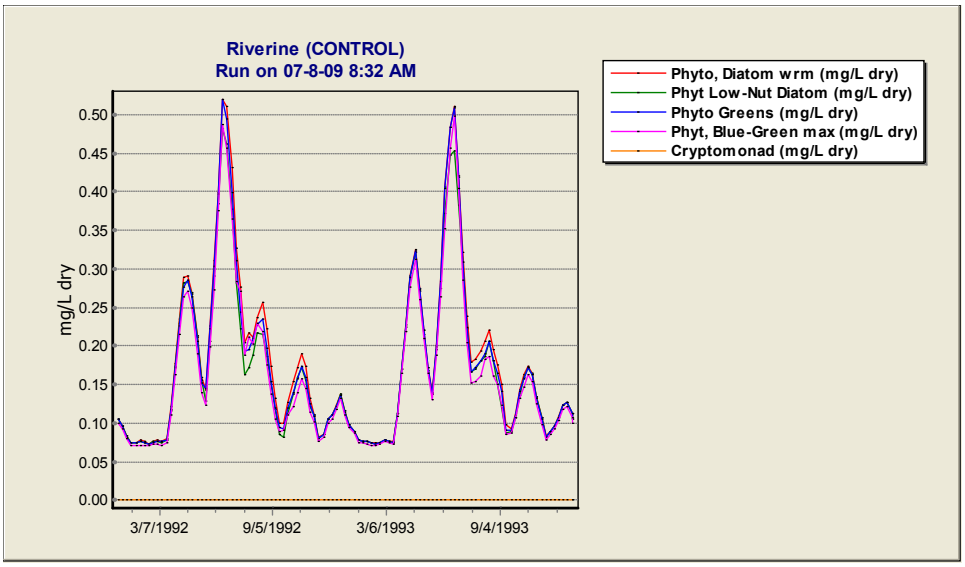

Figure 16. Simulated algal biomass in Riverine segment of Tenkiller Lake OK. It is driven almost entirely by loadings based on splitting of chlorophyll *a* data from the Illinois River among four groups.

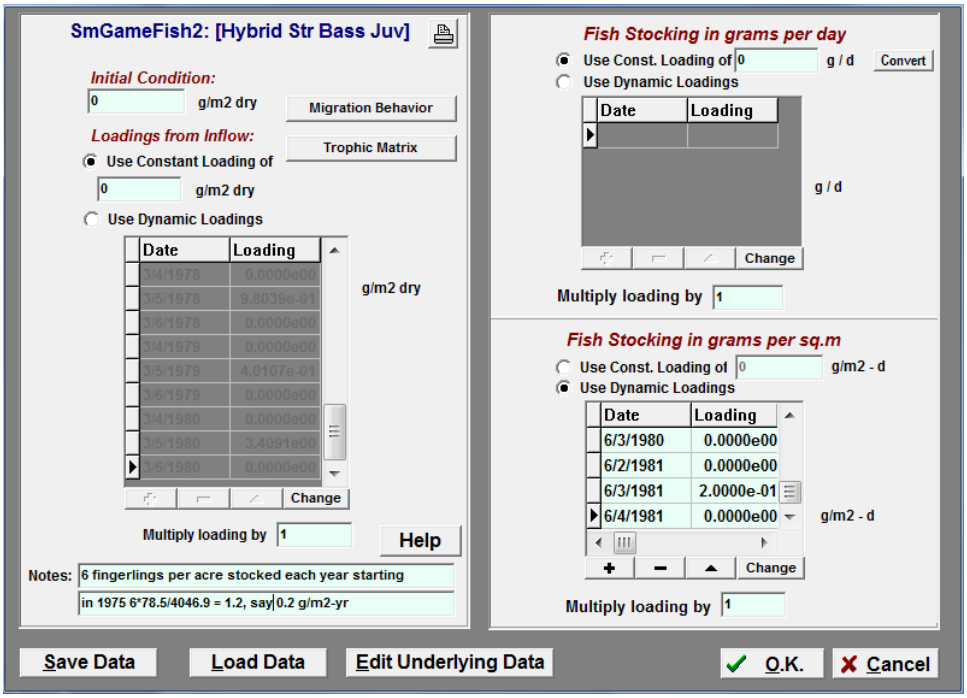

Figure 17. Stocking hybrid striped bass in DeGray Lake.

Initial Conditions

AQUATOX requires initial conditions for all state variables, but depending upon the site type, they may not be important. In particular, initial conditions for nutrients, suspended detritus, and phytoplankton are quickly replaced by loadings from upstream in a river. Initial conditions are

often uncertain; or, if good values are available, they are from the growing season. Of course, you can begin a simulation on the date with the best data, but seldom are data available for all state variables. We prefer to begin a simulation in winter (for simulations of temperate ecosystems) and allow the system to "spin up" as it goes into the growing season. Furthermore, unless it represents a short-lived mesocosm, a simulation should run for at least a year, and preferably several years. That way the initial conditions may be uncertain, but stable fluctuations can be attained and compared to available data. To that end, one can specify "Run model in Spin-up Mode" and the initial conditions for the biotic state variables will be set using the end values; the same may be done with nutrient and detrital state variables. Once the model is calibrated, the spin-up mode should be turned off; in fact, a warning is issued at the beginning of each run if spin-up mode is being used.

Parameters

Unlike driving variables, which are characteristics of a site, model parameters are intrinsic characteristics of the species or chemical of interest. Obtaining biotic, chemical, and toxicity parameters can be difficult. The AQUATOX libraries of parameters cover a wide range of species and chemicals, and should be consulted first. However, one may wish to try different values as part of the calibration process, or one may wish to add state variables that are not in the library. There are three compendiums of parameters that were assembled for the US Corps of Engineers about 30 years ago. We have found them so useful that we have made them available in pdf format on the AQUATOX Web site: http://water.epa.gov/scitech/datait/models/aquatox/data.cfm.

The report by (Collins and Wlosinski 1983) covers phytoplankton, zooplankton, zoobenthos, and fish. Data are in tabular form (for example, Table 4), and the user can obtain representative values and statistics to define distributions for uncertainty analysis.

Table 4. Part of a table giving observed values for maximum photosynthesis and temperature at time of the observation (Collins and Wlosinski 1983).

Table 5

Gross production rates of phytoplankton (1/day)

SPECIES	TPMAX	TEMP °C	REFERENCE
DIATOMS			
Asterionella formosa	0.81	20	Holm and Armstrong 1981
Asterionella formosa	0.69	10	Hutchinson 1957
Asterionella formosa	1.38	20	Hutchinson 1957
Asterionella formosa	1.66	25	Hutchinson 1957
Asterionella formosa	1.71	20	Fogg 1969
Asterionella formosa	0.28	4	Talling 1955
Asterionella formosa	0.69	10	Talling 1955
Asterionella formosa	1.38	20	Talling 1955
Asterionella formosa	2.2	20	Hoogenhout and Amesz 1965
Asterionella formosa	1.9	18.5	Hoogenhout and Amesz 1965
Asterionella japonica	1.19	22	Fogg 1969
Asterionella japonica	1.3	18	Hoogenhout and Amesz 1965
Asterionella japonica	1.7	25	Hoogenhout and Amesz 1965
Biddulphia sp.	1.5	11	Castenholz 1964
Coscinodiscus sp.	0.55	18	Fogg 1969
Cyclotella meneghiniana	0.34	16	Hoogenhout and Amesz 1965
Cyclotella nana	3.4	20	Hoogenhout and Amesz 1965
Detonula confervacea	0.62	2	Smayda 1969

Useful parameter tables for zooplankton and fish were developed under contract to the Corps of Engineers (Leidy and Jenkins 1977, Leidy and Ploskey 1980). Much of the data are summarized by Collins and Wlosinski (1983), but there are lots of tables on fish in particular that are very useful. For example, maximum biomass by fish species is listed for many reservoirs (Table 5), providing both a parameter (carrying capacity) and a reality check on AQUATOX simulations. This particular set of data (which can be converted by g/m^2 = pounds/acre * 0.028) is based on recovery of fish from coves or enclosed open-water areas following application of rotenone.

Table 5. Part of a table giving observed values for standing crop of fish; data are pounds/acre (Leidy and Jenkins 1977).

Appendix G

Fish Carrying Capacity Arranged by Species and Major Reservoir Groups

Species or Species Group	Carrying Capacity	Carrying Capacity Biomass in Pounds per Acre Supported by Each Food Compartment				
		Detritus	Benthos	Zooplankton	Fish	Terrestrial
	Gulf and South Atlantic Drainage Area					
Gars	0.6				0.6	
Bowfin	0.5				0.5	
Gizzard shad	25.5	24.2	1.3			
Threadfin shad	6.4	4.5	1.9			
Pickerels	0.8				0.8	
Carp	18.1	10.9	5.4	1.8		
Minnows	0.7		0.1	0.6		
Carpsuckers						
Suckers	5.2	4.2	0.3	0.8		
Hog suckers						
Buffalofishes						
Redhorses	5.2		5.2			
Bullheads	2.5	0.7	1.4		0.5	
Catfishes	6.8		1.2		5.5	
Madtoms						
Silversides						
Temperate basses	0.8				0.8	
Sunfish	18.8	0.9	12.8		1.7	3.4
Black basses	10.0		0.8		8.6	0.6
Crappies	8.2	0.3	1.7	1.5	4.8	0.6
Perches	1.4		0.3	0.3	0.9	
Freshwater drum						
All other species	1.2		1.2			
Total	112.8	45.7	33.6	4.9	24.7	4.0

Another useful fish database is FishBase (note: several URLs are given in the succeeding examples; if any link gets broken, then use Google to search on the key words): www.fishbase.org/search.php. FishBase has 32,200 species at this time with more being added. This database gets 33 million hits a month, so the response time is sometimes slow. However, as of this writing, there are seven mirrors listed across the top of the Web page, so look for one that is faster. FishBase is most useful as a source of information on feeding preferences, but it also gives general information on distribution, size, and habitats (Figure 18).

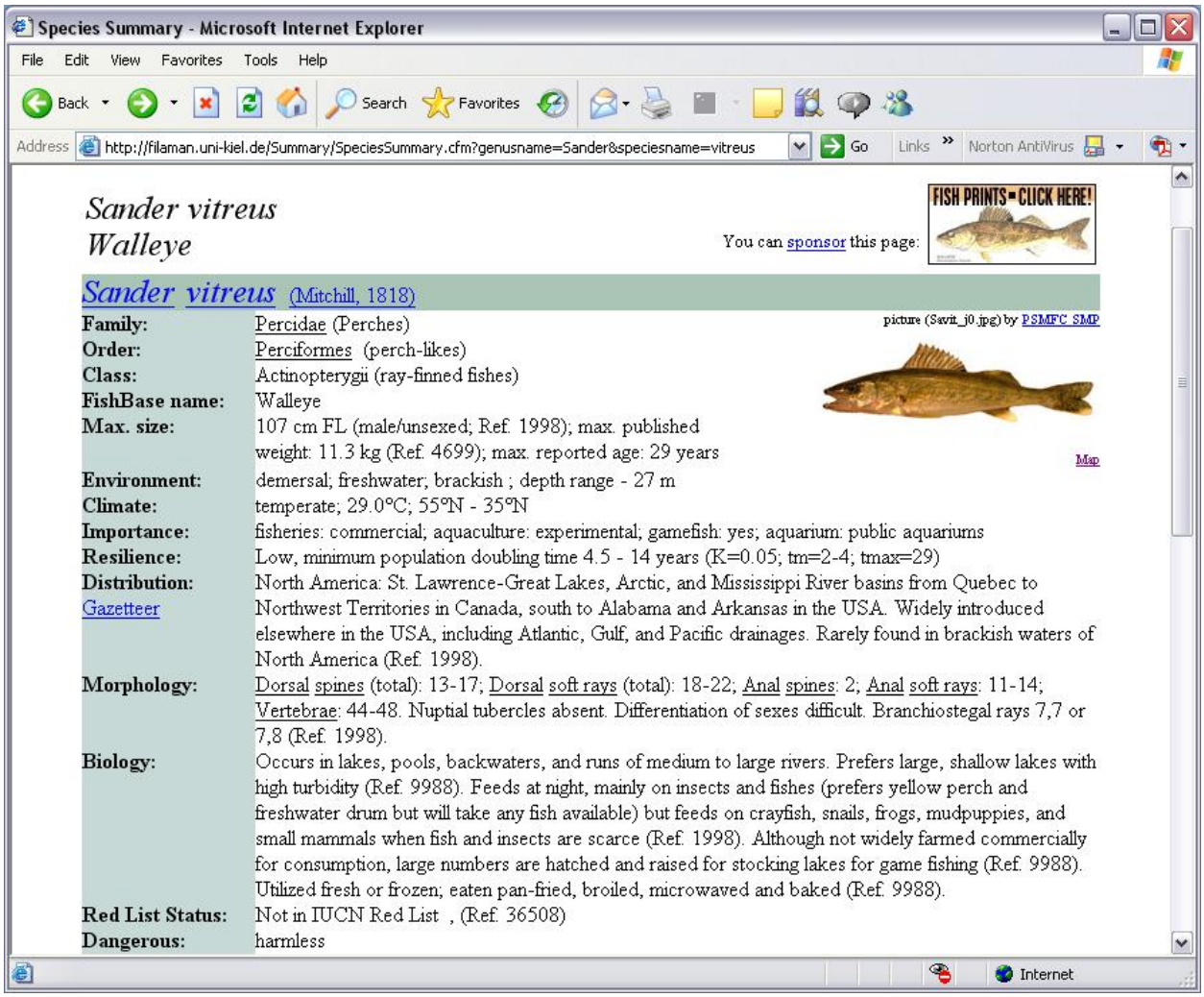

Figure 18. Typical species page in FishBase.

The Wisconsin Bioenergetics Model (Hewett and Johnson 1992, Hanson et al. 1997) is the source of the allometric parameters (i.e., those that vary with the size and weight of the fish, such as respiration). These parameters are already incorporated into the fish records in the AQUATOX library. However, the user may wish to refer to one of the original references.

ECOTOX is an extensive database published on CD-ROM by Elsevier (Jørgensen et al. 2000). However, the database is out of print and may no longer be available. The database was first published as a book twenty years earlier (Jørgensen 1979), and it may be available in libraries. It should not be confused with the ecotoxicological database of the same name developed by the Duluth Laboratory of the US EPA (see below).

ToxRefDB is a new US EPA portal for fate and toxicity data. The URL and the description on the Web page follows: http://actor.epa.gov/toxrefdb/faces/Home.jsp. "ToxRefDB (Toxicity Reference Database) captures thousands of in vivo animal toxicity studies on hundreds of chemicals." Unfortunately, the information presented is so voluminous and organized for purposes other than model parameterization so that it may be counterproductive for our purposes. For example, there are 722 pages of information on chlorpyrifos, yet LC50s for aquatic

organisms are mingled with rat studies and are scattered over many pages. Furthermore, only one octanol-water partition coefficient (KOW) value is given.

Far more useful for purposes of model setup is the US EPA ECOTOX database (Figure 19), where data can be filtered and output in Excel format:
http://cfpub.epa.gov/ecotox/

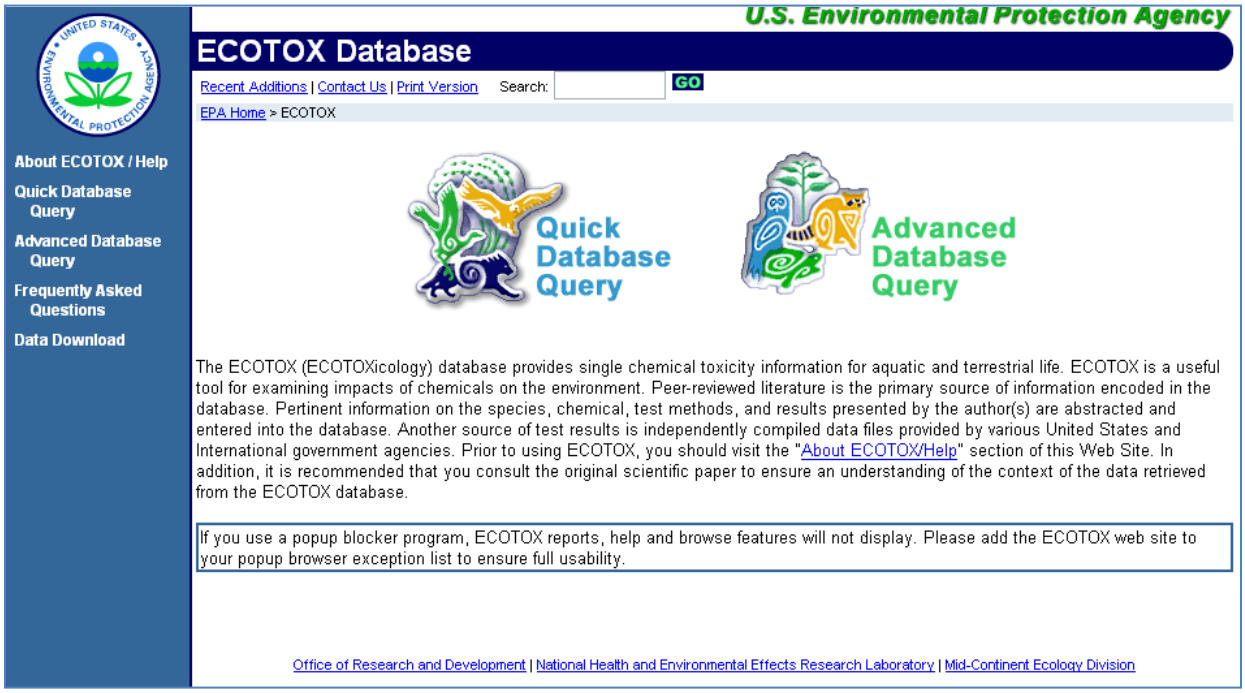

Figure 19. ECOTOX Web site.

The Web-ICE database, developed by US Office of Research and Development (Raimondo et al. 2007), is incorporated in AQUATOX (U.S. Environmental Protection Agency 2009, 2012) and can be used to extend the toxicity data to species that are not in ECOTOX. Details on using Web-ICE in AQUATOX are given in the User's Manual and the context-sensitive Help files.

The Agricultural Research Service (ARS) Pesticide Properties Database: www.ars.usda.gov/Services/docs.htm?docid=14199 has been a useful source of chemical fate parameters. Many of the physical-chemical parameters needed to represent older pesticides in AQUATOX can be found in this ARS database (Figure 20). The biggest drawbacks are that it is restricted to pesticides, and it was last updated in 2001.

name:CHLORPYRIFOS CASRN: 2921-88-2
molecular formula: C9H11CL3NO3PS
molecular weight : 350.62
physical state : S
(L=liquid; G=gas; S=solid)
reference: 9ACHB2

Key to sources: (M)Manufacturer, (R)eview, (H)andbook, (E)xperiment,
(C)alculated, (U)nknown, (P)EPA data, (W)auchope
* denotes a selected value where multiple values of a property are listed.

-value-	-medium-	-temp-	-pH-	-source-	-reference-
Boiling point(deg C):					
-					
Melting point(deg C):					
42 - 43.5				H	9ACHB2
Decomposition point(deg C):					
-					
Heat of vaporization(deg C):					
--RATE CONSTANTS--					
Hydrolysis (per day):					
0.009*		25	5	M	6DOWCH
0.0236*		25	7	M	6DOWCH
0.0440*		25	9	M	6DOWCH
Photolysis (per day):					
Vapor pressure (mPa):					
2.5*		25		H	9ACHB2 1983 ED.
12		35		H	9ACHB2 1983 ED.
2.3		20		R	9HRLCP
2.5(av.2.5,2.4,2.7)		25		M	6DOWCH

Figure 20. Part of a record on chlorpyrifos in the ARS Pesticide Properties Database.

In addition to the sources given above, there are a number of books that can be used to obtain parameter values for ecosystems (Goldman and Carpenter 1974, Horne and Goldman 1994, Wetzel 2001) and for organic chemicals (Verscheuren 1983, Schwarzenbach et al. 1993, Schnoor 1996). Of course, a search engine, such as Google, can query the open and "gray" scientific literature for both historic and recent research.

Calibration and Validation Data

As stated in the *Technical Documentation*, (Rykiel 1996) defines calibration as "the estimation and adjustment of model parameters and constants to improve the agreement between model output and a data set" while "validation is a demonstration that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model." The purpose of calibration is to obtain the best goodness-of-fit to observed data while using general, defensible parameter values. Validation is a test of the robustness of the calibration by evaluating its application to an independent set of observed data. Data on state variables, and indices based on state variables, while used in calibration and validation, may also be used to specify initial conditions; but if so, they cannot be used for goodness-of-fit tests as

well. Driving-variable data are external data that are introduced at the boundaries of the simulation (see **Boundary Conditions**) and cannot be used directly for calibration and validation.

As noted in the Introduction, models are often applied after the fact, using data that were collected for other purposes. This especially can be a problem with data used for calibration and validation. Data may be inadequate when sampling is infrequent and/or sporadic. For example, samples taken every two weeks may completely miss an algal bloom. Problems also occur when data do not specify the precise area sampled so that they can be normalized to a unit area or volume. Furthermore, data on numbers of individuals (density) are common, but ideally they would be paired with mean weights so that biomass can be calculated.

Nutrients

Nutrient data are usually easy to collect so that coverage is often relatively extensive, and their use should be straightforward. However, two commonly encountered complications may be:

- the form of the nutrient data—total soluble phosphorus (TSP) should be distinguished from total phosphorus (TP), which may include that associated with particulate and dissolved organic matter (Figure 21); and nitrate may or may not include nitrite (Figure 22).
- data that are below the detection limit—the user may indicate such data, and AQUATOX will plot those points with a special symbol (Figure 23). Although goodness-of-fit cannot be determined directly, a suitable fit would be indicated if the model output is also below the detection level.

Sometimes other forms of nutrient data are provided, which may increase uncertainty in the model setup. If only soluble reactive phosphorus (SRP) is provided then the contributions of soluble unreactive phosphorus (SUP) are ignored. SUP includes phosphorus temporarily tied up in organic complexes, which are subject to cleaving by algal enzymes and UV light (Wetzel 2001). Total Kjeldahl nitrogen data are often available; these consist of organic nitrogen plus ammonia. At present, AQUATOX does not accept data in this form; however, if the amount of nitrogen tied up in organic compartments can be determined, the user can calculate the ammonia concentration offline.

In the examples from the Lower Boise River, Idaho, the model was not calibrated to the nutrient data, but rather, the data reflect concentrations in a river reach that is strongly affected by wastewater treatment effluent (specified as loadings). Given the low retention times, the loadings often overwhelm the in-stream processes. However, the upstream release of water decreases considerably during the winter when no longer needed for irrigation, and some degree of calibration may be possible with the longer retention. This is true for ammonia (Figure 23), which was not calibrated and exhibits a poor fit during winter flow conditions. A more striking example is given for dissolved oxygen, discussed below.

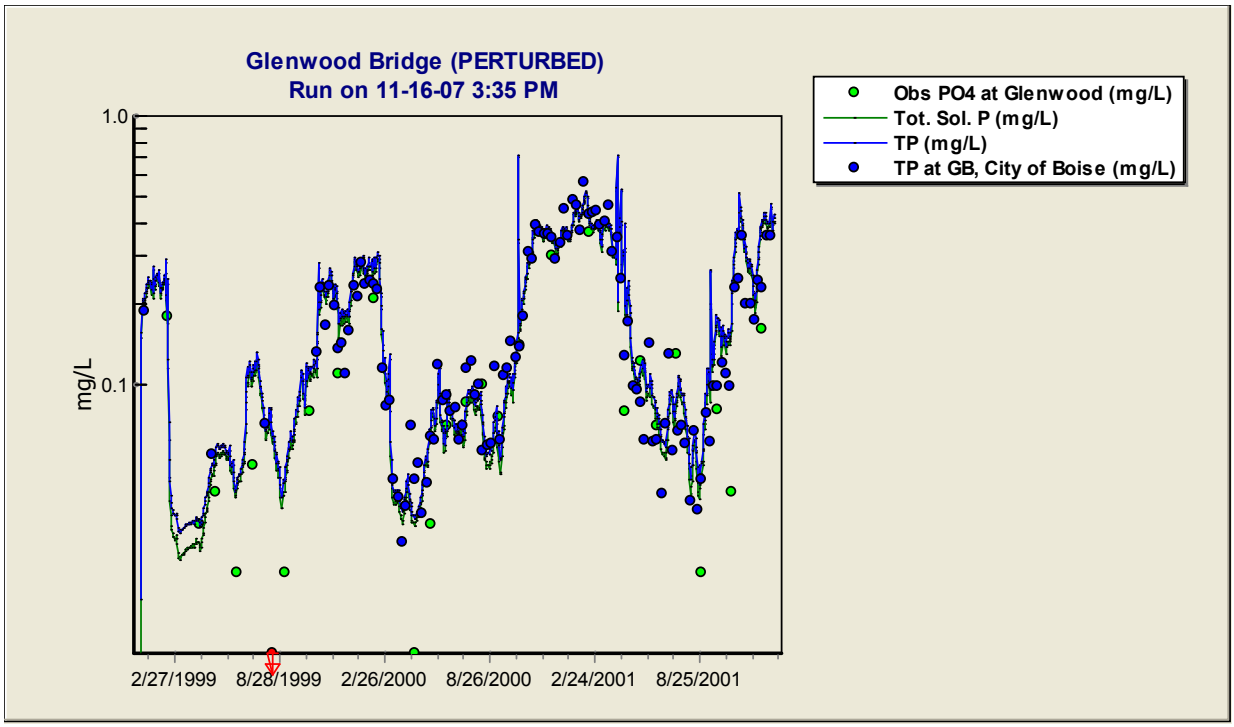

Figure 21. Predicted and observed TSP and TP at Glenwood Bridge, Boise ID (Lower Boise River). Note the predicted TSP and TP are superimposed most of the time, implying that the model predicts that almost all P is inorganic.

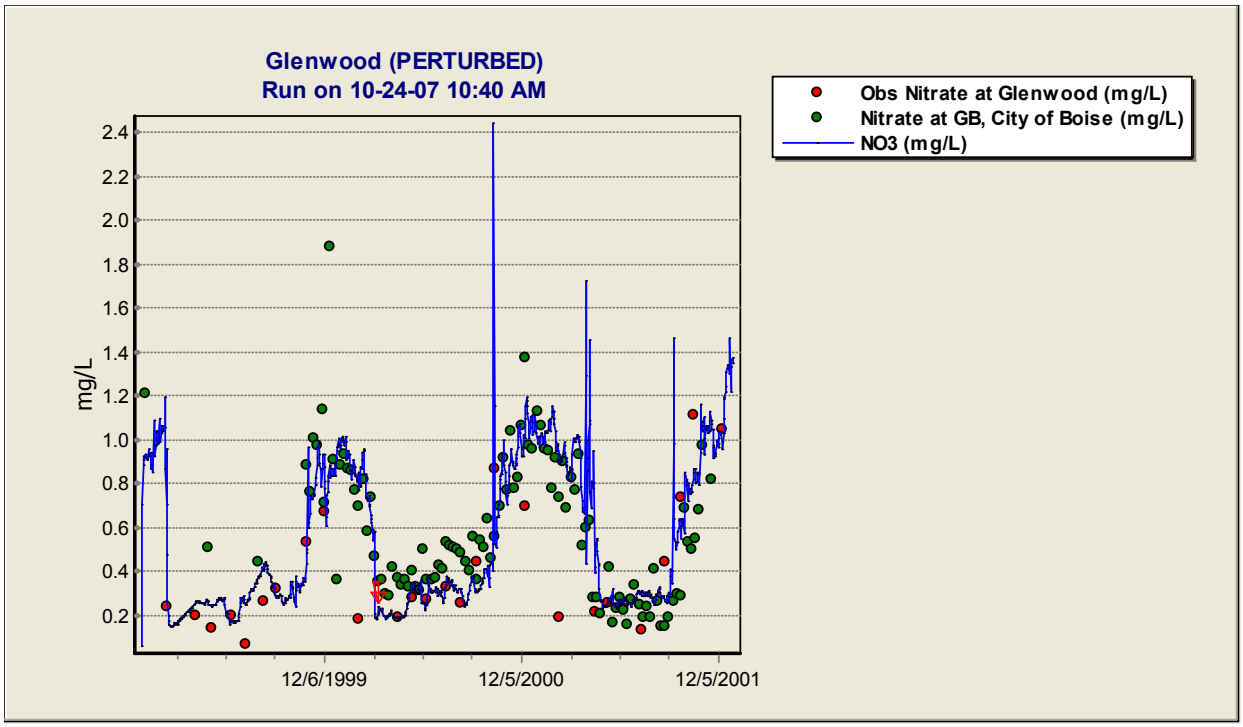

Figure 22. Predicted nitrate-nitrite and observed nitrate at Glenwood Bridge, Boise ID (Lower Boise River); observed data are from two different sources.

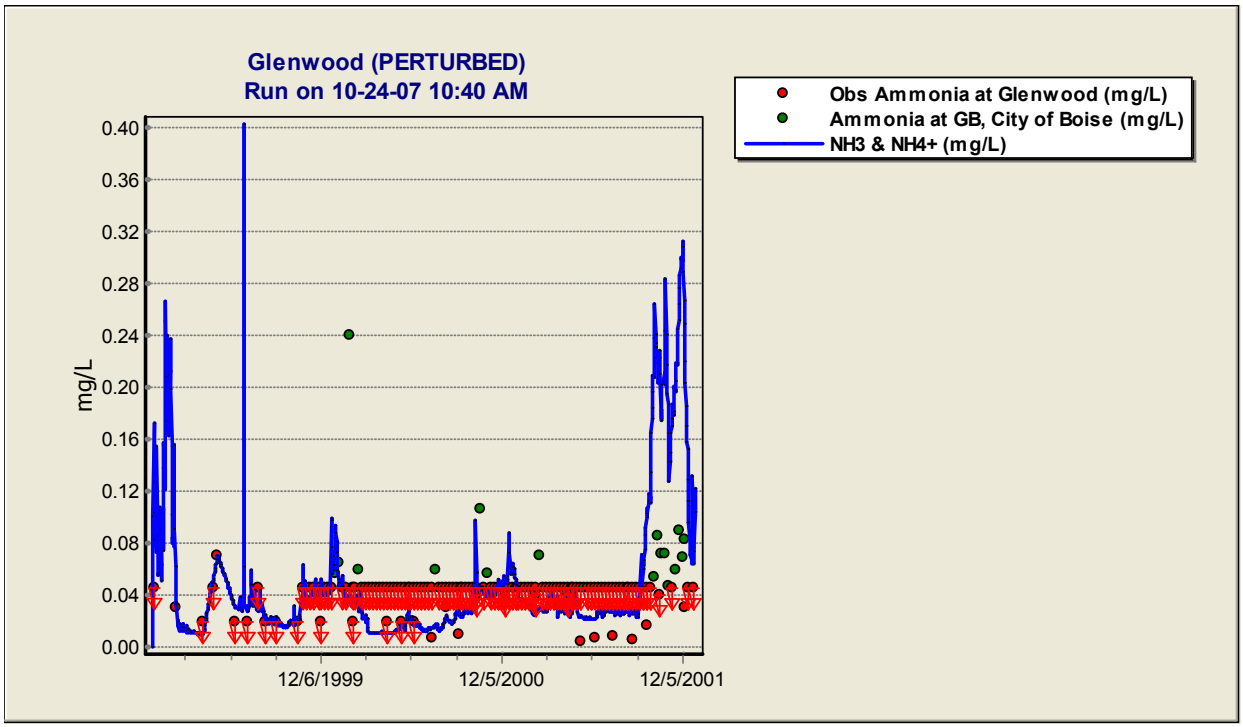

Figure 23. Predicted and observed ammonia at Glenwood Bridge, Boise ID (Lower Boise River). The symbol with a down arrow indicates a data point that is below detection limit.

Dissolved Oxygen

Dissolved oxygen can vary vertically within the water column, going from supersaturated near the surface to anoxic at depth in some lakes and many reservoirs. This is a source of uncertainty in the application of a model that only represents the epilimnion and hypolimnion as two well-mixed layers. Ideally, observed data would be averaged for each of the two layers. In practice, values often are arbitrarily chosen from "representative" depths. It is not surprising that the fit is often better for an anoxic hypolimnion (Figure 24), as actual conditions would be expected to be more uniform than in the epilimnion.

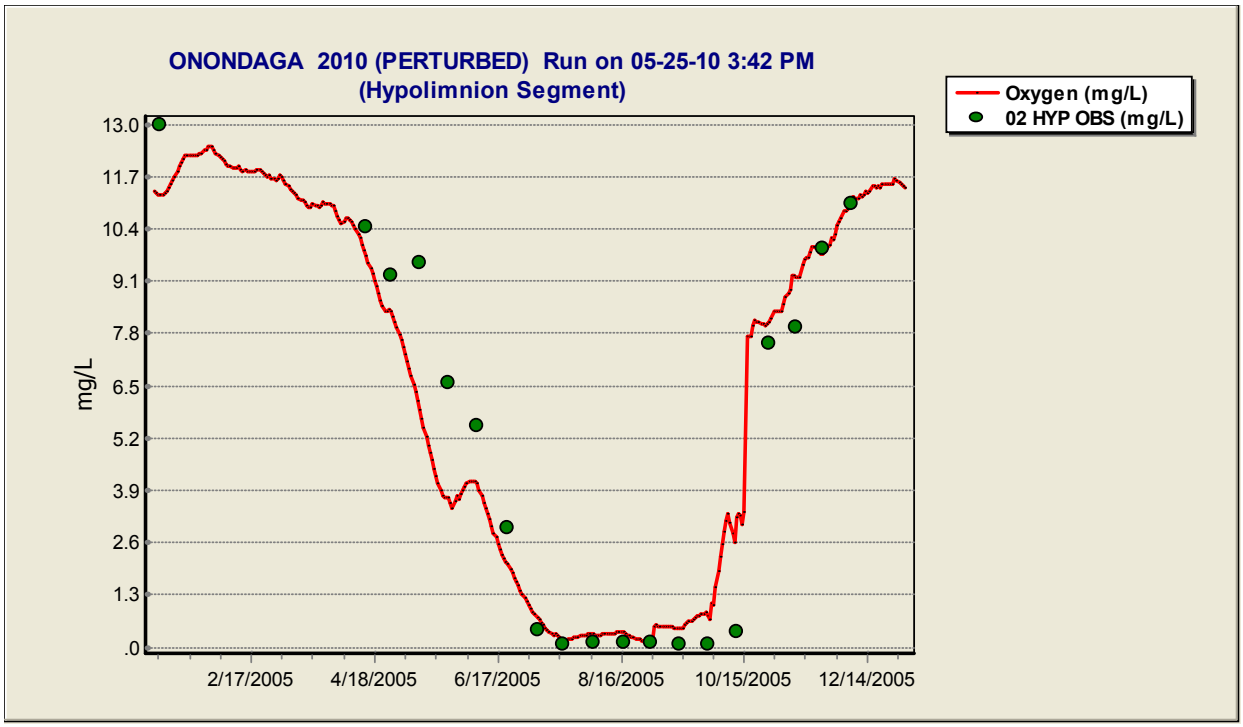

Figure 24. Observed and predicted dissolved oxygen in the hypolimnion of Lake Onondaga NY.

Biomass

Observed data on biomass of organisms are often difficult to obtain, especially for invertebrates and fish. (Stevenson 1996) gives a good discussion of the different measurements of algal biomass. The most accurate, but laborious, method for obtaining biomass of both phytoplankton and periphyton is to compute the biovolume of each algal species. As described by (Hill and Boston 1991), they "measured the dimensions of at least 15 individuals with an ocular micrometer. The averages of these dimensions were used in volumetric formulae of appropriate geometric shapes. Biovolume ($\mu m^3/cm^2$) for each algal species was calculated by multiplying cell density ($cells/cm^2$) by estimated cell volume." Biovolume can then be converted to biomass (g/m^3) for use in AQUATOX. Table 4 in (Reynolds 1984) lists biovolumes and other common measures of biomass for common species.

The most common surrogate for algal biomass is chlorophyll *a*. AQUATOX assumes a constant relationship between chlorophyll *a* and biomass, depending on the algal group and whether or not it is periphyton. Biomass is output and is also converted to chlorophyll *a* so that the predicted values can be compared with observations, whether expressed as biomass or chlorophyll *a*. It is not unusual to have two or three replicate samples; we prefer to plot each point so that the spread of values can be assessed (Figure 25). Of course, if sufficient samples were collected to calculate standard deviations then those should be plotted. Care must be taken to determine the representativeness of the sampling site, especially for periphyton in streams where the substrate may be very heterogeneous. For example, if periphyton samples are taken from cobbles, it is reasonable to normalize the values for the percent of the site with cobbles in order to compare with the AQUATOX results, which are for the entire site. That was done by the user for the

example shown in Figure 25 (the percent riffles site characteristic is not used by AQUATOX to automatically normalize the results).

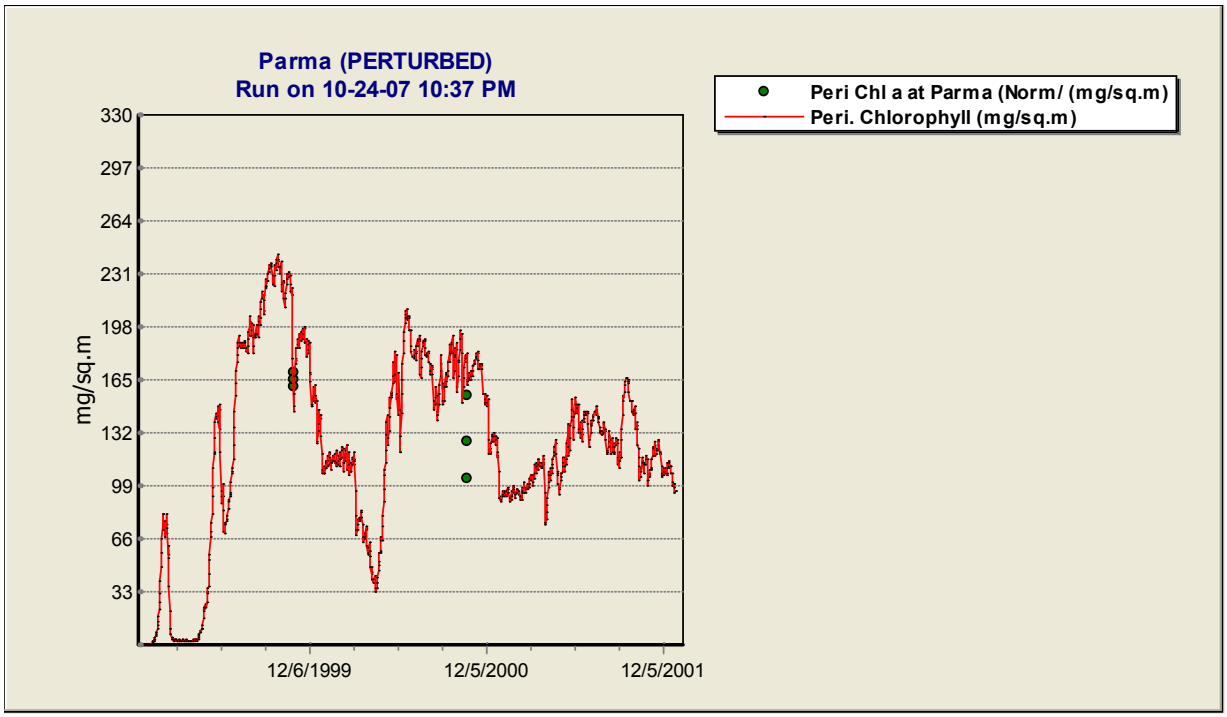

Figure 25. Predicted and observed periphytic chlorophyll *a* at Parma site, Lower Boise River ID.

AQUATOX computes algal biomass as ash-free dry weight (AFDW). In a study of the Cahaba River, Alabama, AFDW data were collected specifically for use in the model. Furthermore, the aquatic moss *Fontinalis* was analyzed and modeled separately (Figure 26).

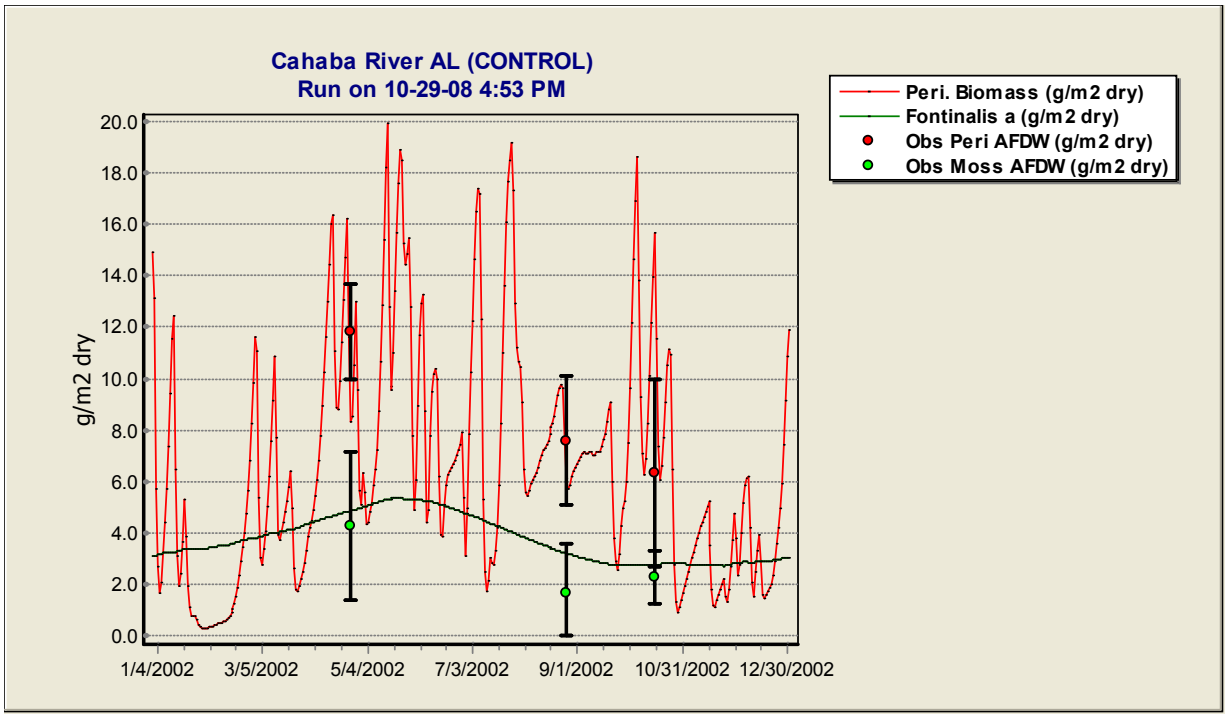

Figure 26. Predicted and observed AFDW biomass of periphyton and moss, Cahaba River AL.

A quick method for characterizing periphyton biomass is to estimate thickness as part of a rapid bioassessment (Barbour et al. 1999). Thickness measurements are ranked: "0" indicates a rough surface with no algae, "1" for no algae visible but surface is slimy, "2" for 0.5 mm to 1 mm, "3" indicates an algal cover of greater than 1 mm but less than 5 mm, "4" for 5-20 mm and "5" for algae over 20 mm (Barbour et al. 1999). These can be converted to AFDW (g/m^2) values based on polynomial regression of paired thickness and AFDW data in the literature (Stevenson et al. 2007) (Figure 27). The conversion yields estimates with a relatively large error component, which should be made explicit in plots by plotting error bars.

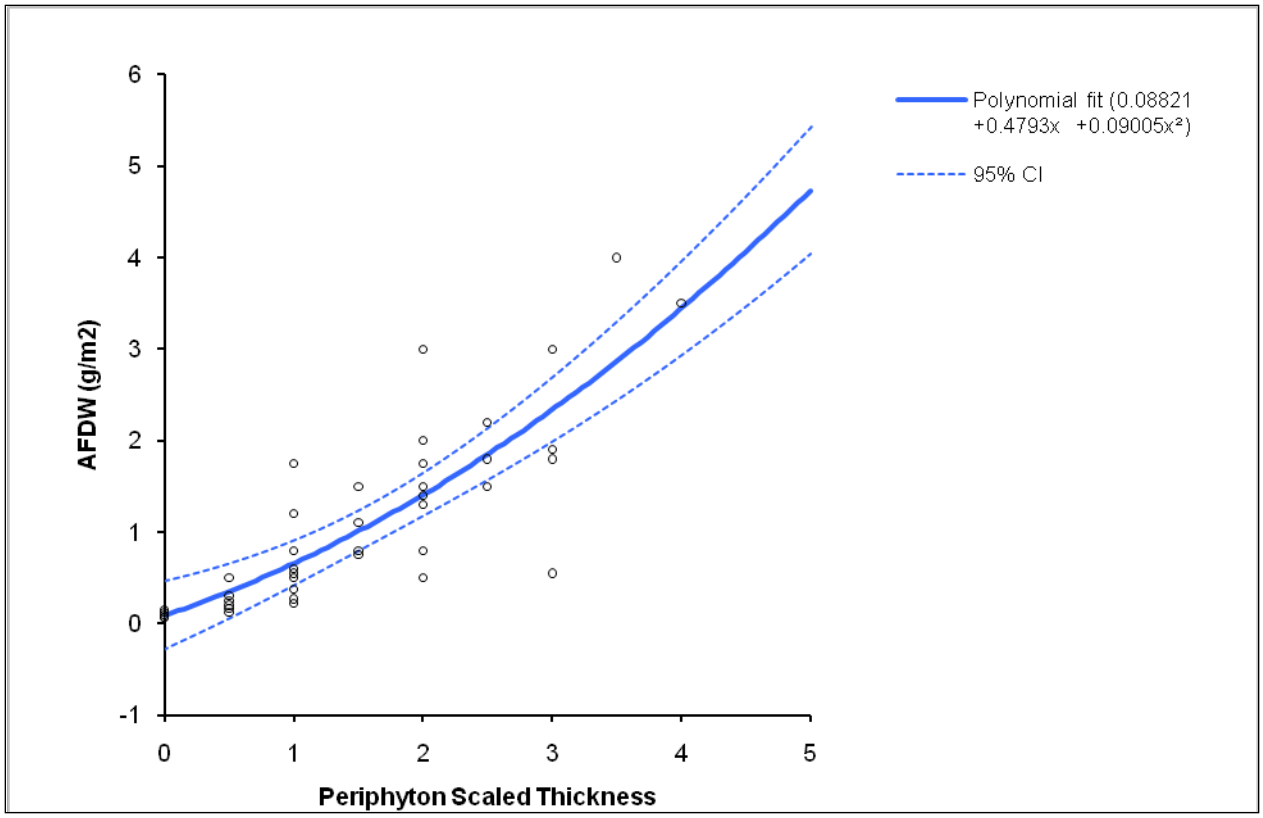

Figure 27. Regression of AFDW on diatom thickness rank, based on data of Stevenson et al. (2007); R2 = 0.77.

Another challenge in using data collected for other purposes occurs then the data were taken from periphytometers—artificial substrates placed in the stream for a given period of time. Depending on the design, the periphytometers may or may not exclude grazers. Although natural substrates are preferred, periphytometer data can be used if grazers were not excluded; if they have been excluded then the data will not be usable for purposes of modeling algal biomass, as grazers can be a significant control on periphyton biomass.

Density (Numbers of Individuals)

Invertebrate and fish data are often reported as numbers of individuals or density. Presumably the data are collected from a unit area so that they are internally consistent, although sometimes the area is not reported—especially for fish collected by electrofishing. If the area is given and if mean weights are reported, then the data can be converted to biomass (g/m^2) and compared directly with model output. However, if only density is reported then one can still compare trends indirectly. For example, in a microcosm study conducted at Duluth MN (U.S. Environmental Protection Agency 1988), a pond enclosure was dosed with 6.3 µg/L chlorpyrifos. Data on chironomids were reported as numbers/sample, which were compared with predicted biomass by plotting the latter on the Y1 axis and the density on the Y2 axis (Figure 28). Of course, such comparisons are only approximate because the weights of the individuals may change considerably over the period being simulated, especially if it involves fast-growing aquatic insect larvae. Nevertheless, in this example the comparison was informative and

provided a limited validation of the ecotoxicological submodel of AQUATOX. Similarly, fish data are often reported as catch per unit effort (CPUE). Here again trends can be compared by plotting biomass on the Y1 axis and CPUE on the Y2 axis.

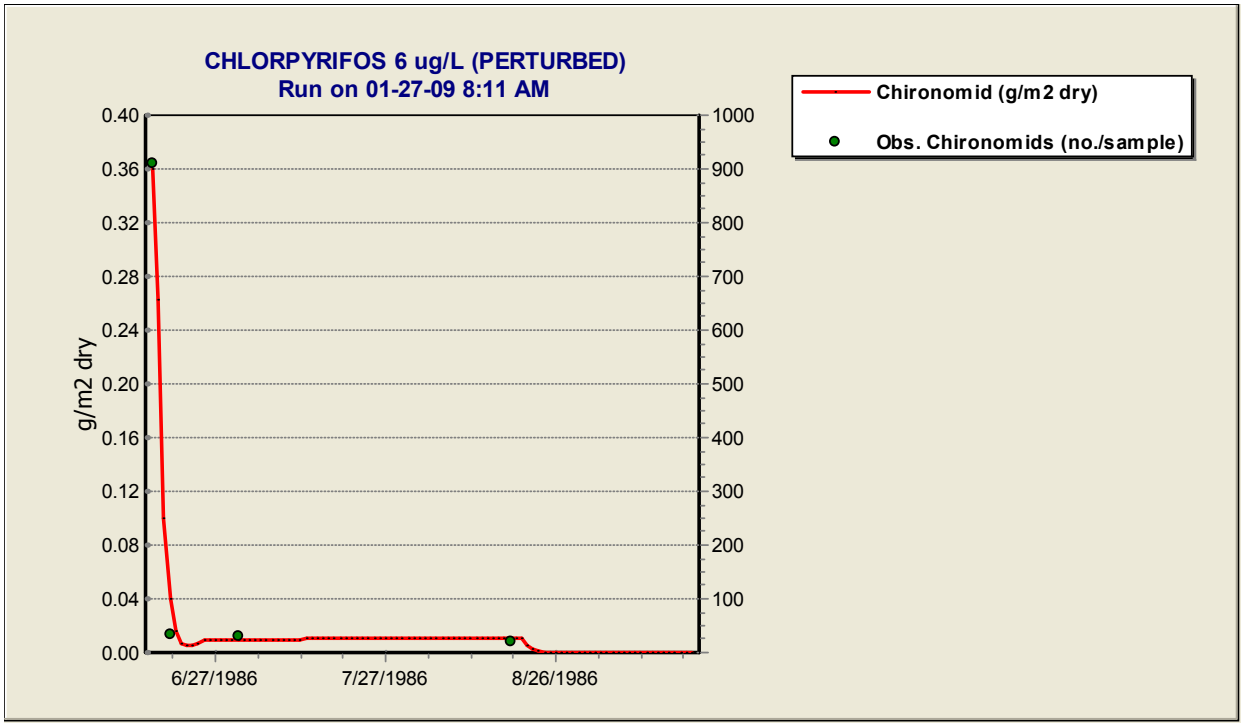

Figure 28. Predicted biomass of chironomids compared with observed numbers/sample (Y2 axis) in pond microcosm at Duluth MN dosed with 6.3 µg/L chlorpyrifos.

Percent Composition and Other Metrics

Species data may be available as percent composition. AQUATOX can convert biomass of many biotic groups to percentages, both for purposes of comparison to reported values and as metrics (such as percent blue-greens) used by decision makers. There are a couple of pitfalls to this approach. First, percentages involve closed-number systems that are sensitive to values of other state variables. In Figure 29 the percent blue-greens is partly an inverse function of the amount of diatom biomass at any given time.

Second, reported percent compositions are usually based on numbers of individuals, but AQUATOX computes them on the basis of biomass, making comparisons between observations and predictions difficult. Biomass-based percentages are actually more accurate because there is a common basis; percentages based on numbers can be biased by numerous small organisms compared to a few large organisms. In computing percentages of invertebrates, AQUATOX excludes mussels because one large individual can exceed all the biomass of the other invertebrates at a site.

Metrics based on predicted percent aquatic insect larvae may not coincide temporally to observed data. AQUATOX simulates emergence of insects based on computed growth rates. That approximation may throw off the timing and cause a dip in biomass that is out of phase with

the actual emergence. Therefore, the magnitude of the predicted metric, such as percent chironomids (Figure 30), is much more important than the timing.

Gross primary productivity (GPP) is a metric that represents overall primary productivity (expressed as oxygen production) at a site and can be summarized over time by calculating the mean and standard deviation (Figure 31). Of course, the predicted and observed statistics should be calculated for the same period.

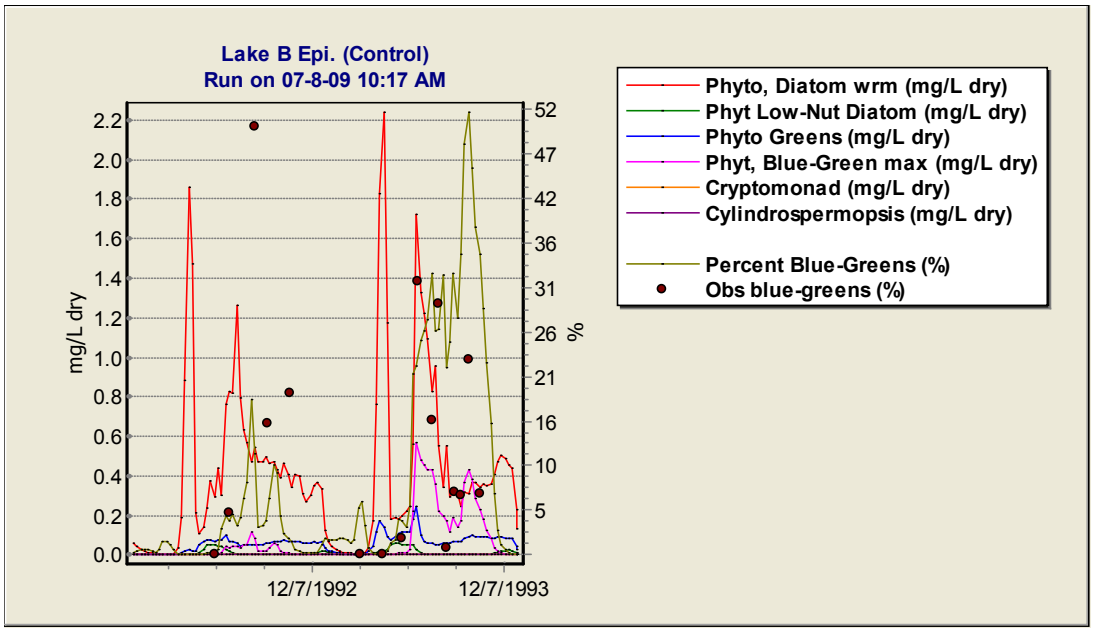

Figure 29. Predicted and observed percent blue-greens plotted with algal biomass, especially diatoms.

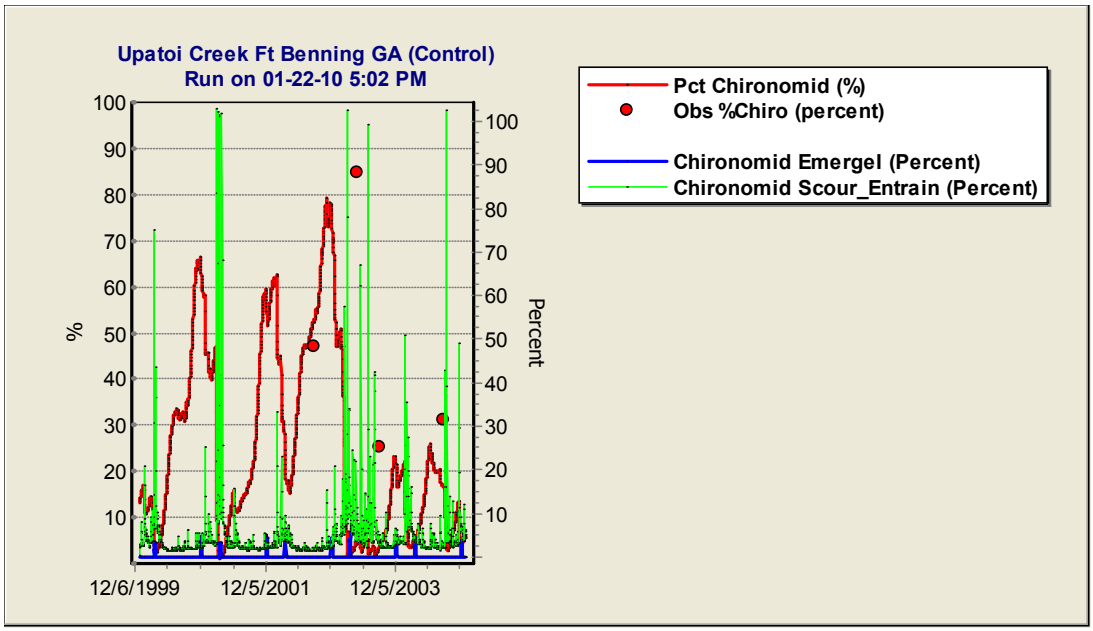

Figure 30. Predicted and observed percent chironomids in Upatoi Creek GA. The steep drops are predicted due to a combination of emergence and, much more important, scour from high flow.

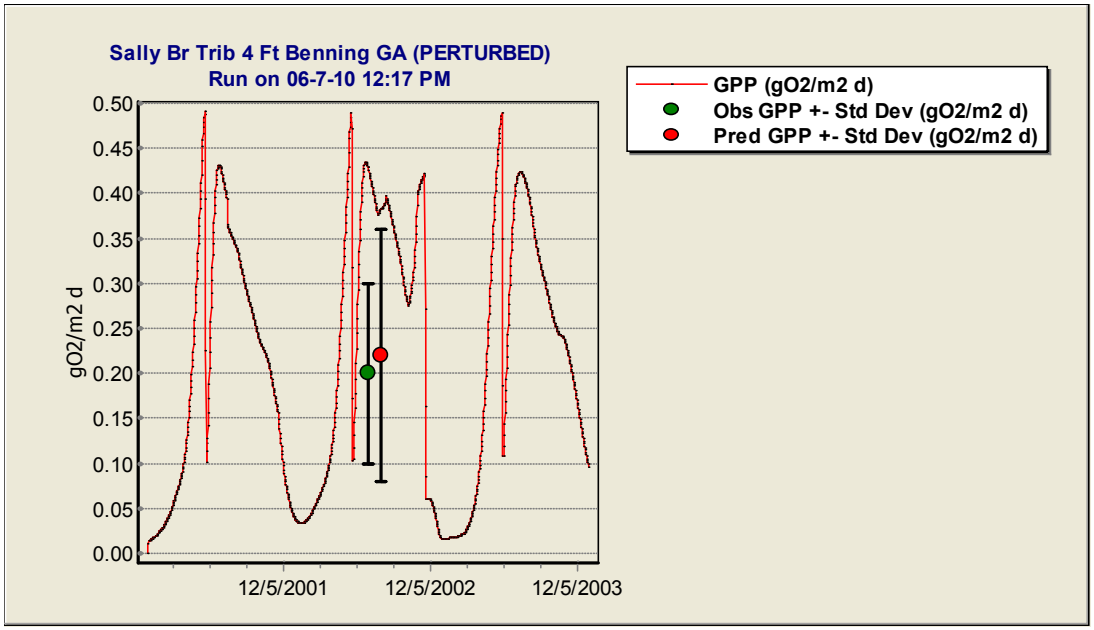

Figure 31. Predicted and observed gross primary productivity (GPP) in Sally Branch Tributary #4, Fort Benning GA. Statistics (mean ± 1 standard deviation) are for the entire three-year period, and therefore their position on the horizontal axis is arbitrary.

Organic Chemicals

Persistent bioaccumulative organic compounds may be simulated for both short and long periods of time. Unless the goal is to simulate acute effects, the model should be run at least to equilibrium. A quick way to determine the length of time required for such a model run is to check the chemical-parameter input screen where an estimate for equilibrium in fish is given

based on the octanol-water partition coefficient. For example, according to the estimate in the chemical-parameter screen, PCB 1254 should reach equilibrium in fish in 366 days. This result is exemplified by a simulation of Galveston Bay TX (Figure 32). The results shown use PCB exposure data for a Massachusetts bay in a calibrated Galveston Bay Texas AQUATOX simulation (exposure data were not available for Galveston Bay). The only serious mismatch between predicted and observed PCB concentrations is for polychaetes, which apparently are not getting adequate exposure to PCBs in the sediments within this simulation. Note the concentrations use the customary units of µg/kg wet weight, in contrast to concentrations in sediments that are expressed as µg/kg dry weight.

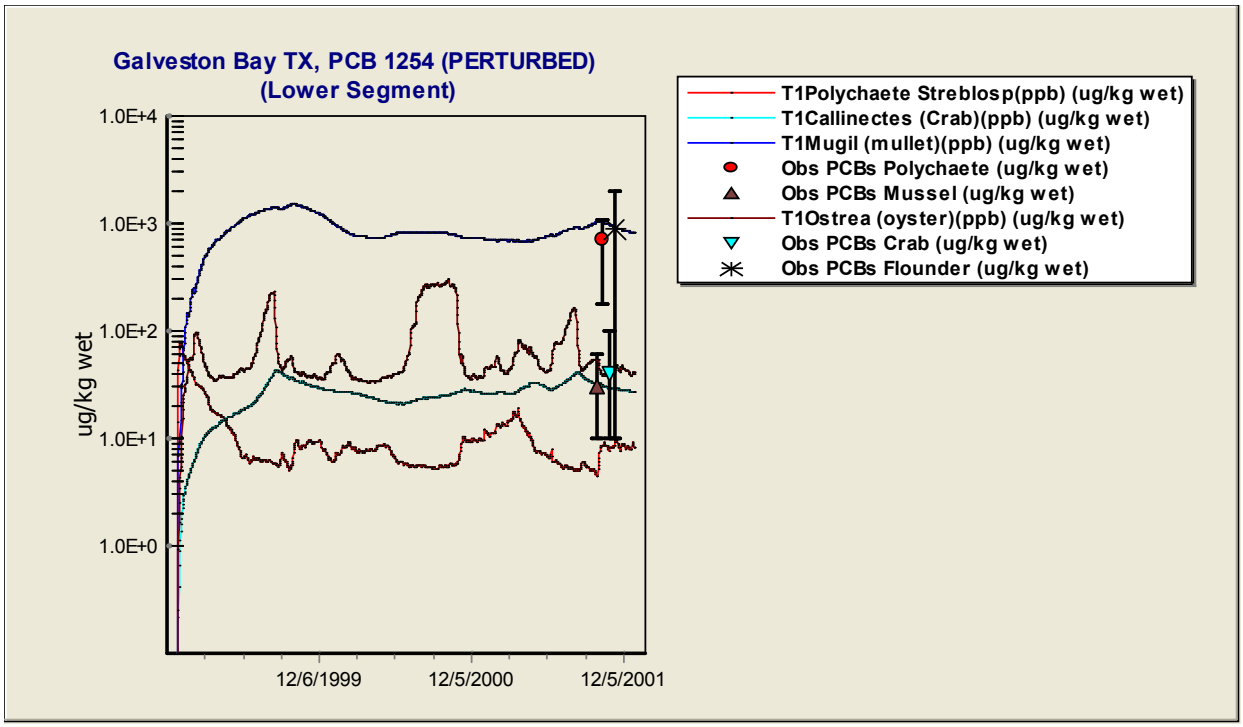

Figure 32. Predicted concentrations of PCBs in Galveston Bay TX animals using exposure data and observed concentrations from a Massachusetts bay.

A common way to express bioaccumulation is as a bioaccumulation factor (BAF), which is the concentration in the organism divided by the concentration in the water. However, one must be aware of what the "concentration in the water" means, because 50% or more could be bound to colloidal detritus (Oliver and Niimi 1988). AQUATOX provides the alternative to "Include complexed toxicant in BAF calculations" when setting up the study. Essentially, this option allows the analyst to use legacy data collected before the relationship to colloidal organics was recognized. The data of (Oliver and Niimi 1988) were used in a study of PCBs in Lake Ontario. The observed data include the complexed PCBs in the calculation, so that option was chosen for the simulated values too (Figure 33).

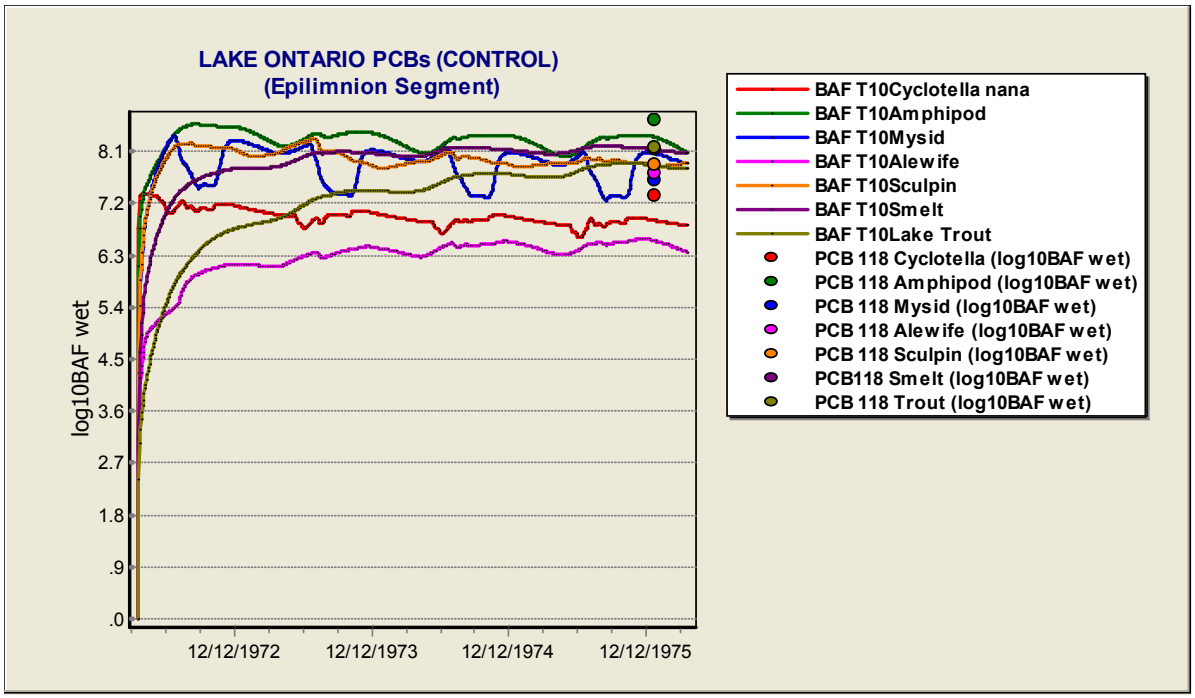

Figure 33. Predicted and observed BAFs for PCB 118 in Lake Ontario; complexed PCBs were used in the computation of BAFs.

Sensitivity Analysis

AQUATOX includes a built-in nominal range sensitivity analysis (Frey and Patil 2001), which varies each user-specified parameter and loading by plus and minus a given percentage and ranks them according to how much each input affects the target output. Furthermore, one can perform statistical sensitivity analysis in which parameters are tested one at a time using an appropriate distribution of parameter values. By performing sensitivity analyses on parameters and loadings once the model is initially calibrated important parameters deserving additional attention can be identified, and alternate parameter values can be identified. These may then be used to improve the agreement between the observed and predicted values of the calibration target. For example, nominal range sensitivity and statistical sensitivity analysis were performed on Onondaga Lake NY. One of the sensitive parameters for chlorophyll *a* was the maximum photosynthetic rate (*PMax*) for diatoms (nominally *Cyclotella*) (Figure 34). This parameter was then used in a statistical sensitivity analysis, where a normal distribution of diatom *PMax* values was used, based on the compilation of values by (Collins and Wlosinski 1983); the results are shown in Figure 35.

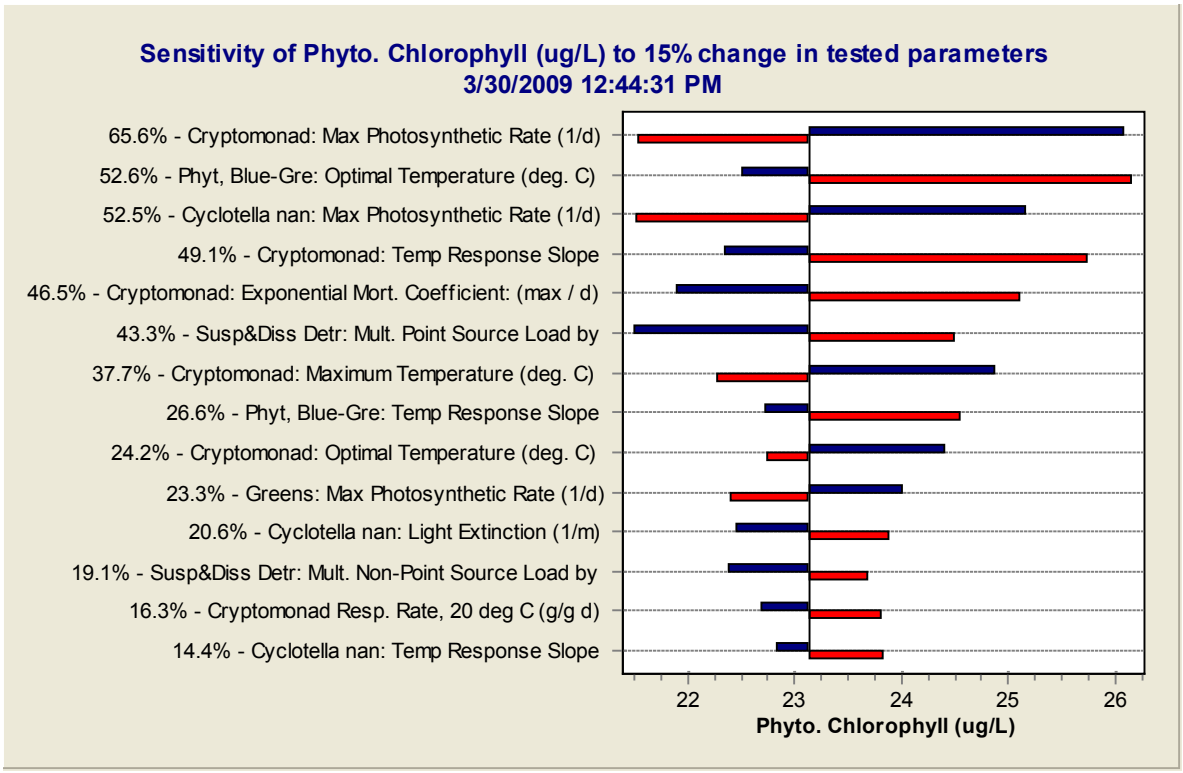

Figure 34. Sensitivity of Phytoplanktonic Chlorophyll a in Lake Onondaga NY.

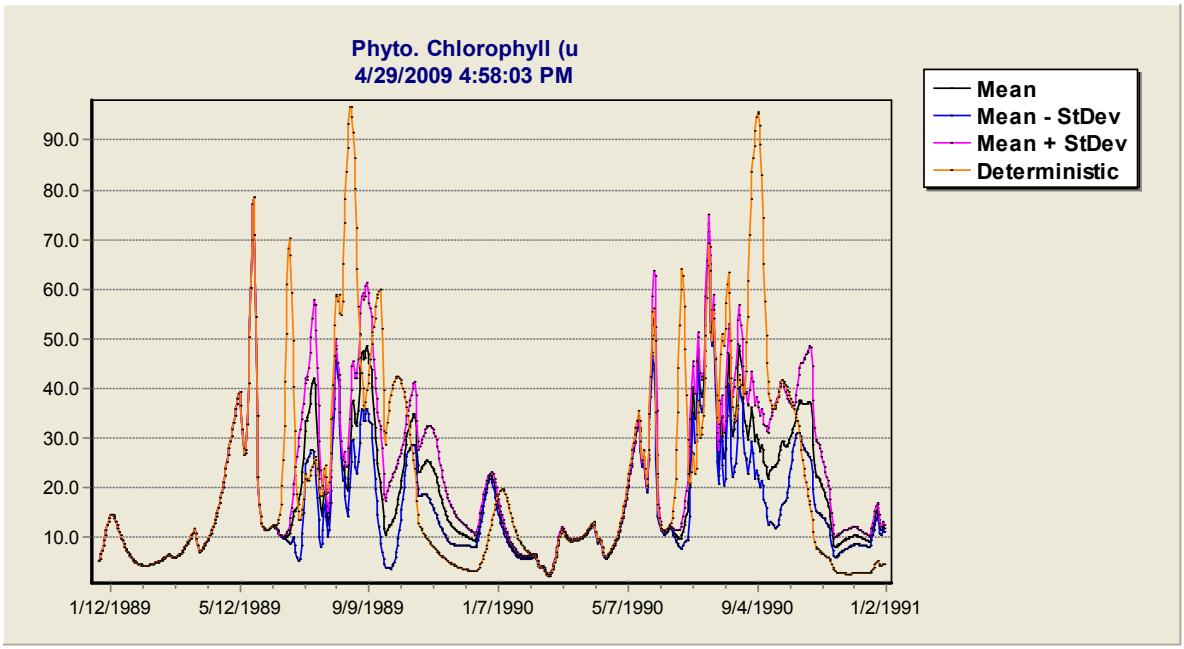

Figure 35. Sensitivity of chlorophyll a to diatom PMax in Lake Onondaga NY.

The mean results seemed to be a better fit than the deterministic, so the model was run with *PMax* = 1.6 (the mean of observed values) instead of the prior value of 3.4. The results (perturbed) did seem better than those with the original value (control), especially in the second year when a clearing event with low observed chlorophyll *a* was simulated (Figure 36, Figure

37). Of course, any new parameter value should be within a reasonable range, not just one that gives a better fit.

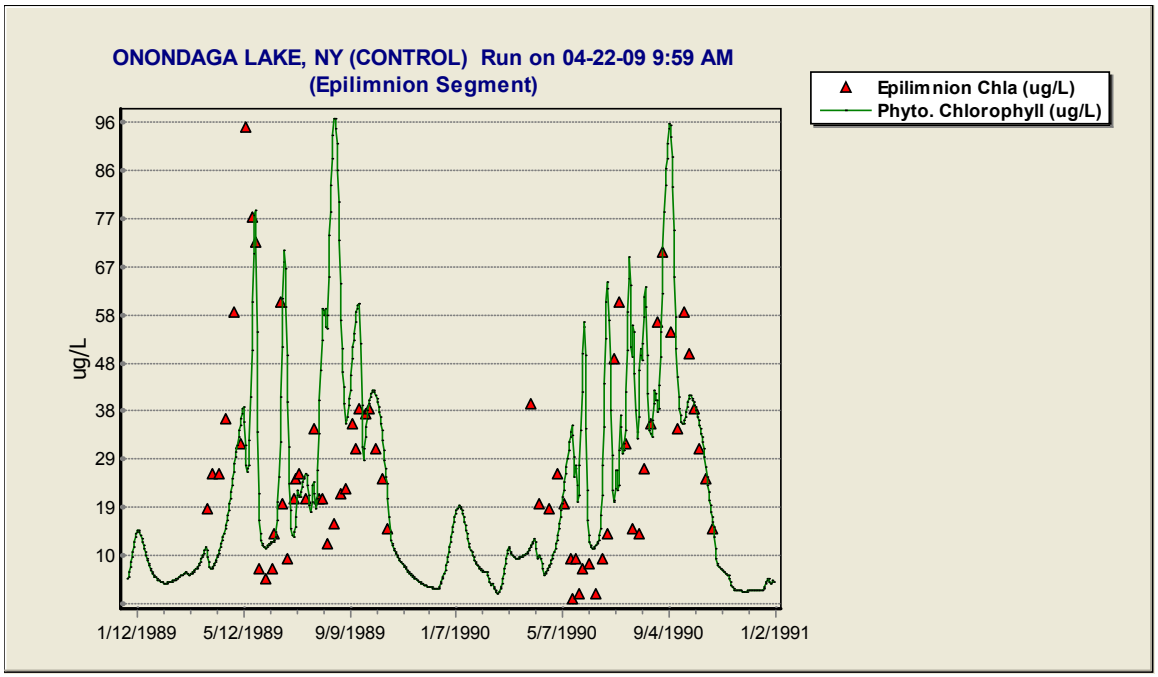

Figure 36. Original calibration of chlorophyll *a* with diatom *PMax* = 3.4 in Onondaga Lake NY.

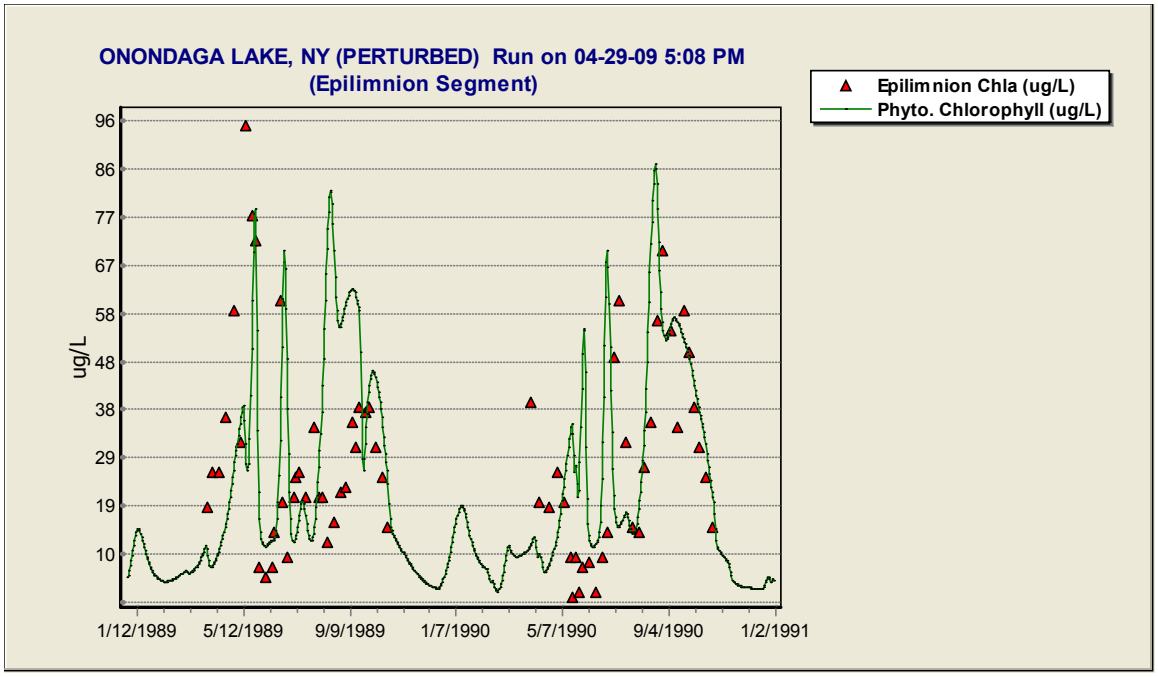

Figure 37. Chlorophyll *a* result with diatom *PMax* = 1.6 in Onondaga Lake NY. Note better fit in 1990.

Summary

Seldom do modelers have the luxury of being able to dictate what data should be collected before setting up the model. Often the model application comes after the original study and must rely on incomplete sets of data collected for other purposes. One of the objectives of this Note is to describe ways of using available data and of filling in the gaps with other sources.

Site characteristics include length and width, to which the model is not usually sensitive; and depth and thermocline depth, which can be important. Site loadings are often very important and include, roughly in order of importance, time-varying inflow and discharge, flow among linked segments, organic chemicals, nutrients, several detrital compartments (based on organic matter, organic carbon, or BOD), and total suspended solids; and seasonally-varying light, temperature, dissolved oxygen, pH, and biotic loadings. The analyst may of necessity use diverse data sources such as USGS for flow and nutrient data from a representative gage, NASA for solar radiation estimates, and state and local laboratories for data that often require conversion to units used in AQUATOX.

Because models are often applied using data collected for other purposes, data for calibration and validation may be inadequate or problematic. Nutrient data are often extensive; however, certain chemicals, such as Kjeldahl nitrogen, may be analyzed to the exclusion of other forms, such as nitrate, making it difficult to evaluate nutrient mass balance in the model. Biomass is most important for calibrating ecosystemic responses, yet temporal coverage is often sporadic or confined to one season. Furthermore, biotic data are often collected as environmental indices and cannot be converted to biomass per unit area or volume without making tenuous assumptions. Often the analyst is left comparing simulated biomass trends with trends in numbers of individuals (especially for invertebrates) and catch per unit effort (for fish).

Obtaining parameter values can be difficult, especially since they are scattered among many published papers and gray-literature reports. Fortunately, there are compendiums of parameter values assembled in support of models. Unfortunately, several of these were published decades ago and have not been kept up to date. The exceptions are FishBase and EPA's ECOTOX, which continue to track the literature and grow in size. Now, Internet search engines are able to find parameter values almost instantaneously in both the open literature and in obscure reports and other sources.

The model is usually not sensitive to initial conditions, so that one may use a spin-up period to achieve equilibrium. Finally, it is a good idea to perform sensitivity analyses on parameters and loadings once the model is initially calibrated; AQUATOX has several tools to facilitate this. Important parameters deserving additional attention can be identified, and alternate parameter values can be tentatively identified.

In conclusion, the model has extensive data requirements, and those requirements are often not satisfied with data collected for purposes of the model. This paper describes ways of coping with commonly encountered data issues. Data can usually be obtained from a variety of sources, and those data can be converted and conditioned to forms acceptable to the model. Furthermore, AQUATOX can accept various data types, facilitating application to numerous water body types and modeling goals.

References

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. Large−area Hydrologic Modeling and Assessment: Part I. Model Development. J. American Water Res. Assoc. **34**:73−89.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002, U.S. Environmental Protection Agency; Office of Water, Washington, D.C.

Bicknell, B. R., J. C. Imhoff, J. L. Kittle, A. S. Donigian, and R. C. Johanson. 1991. Hydrological Simulation Program - FORTRAN — User's Manual for Release 10 (Pre-release Draft Version). U.S. EPA Technology Development and Applications Branch in cooperation with USGS Water Resources Division, Office of Surface Water., Washington, D.C.

Carleton, J. N., R. A. Park, and J. S. Clough. 2009. Ecosystem Modeling Applied to Nutrient Criteria Development in Rivers. Environmental Management **44**:485–492.

CH2M HILL, Eco Modeling, Warren Pinnacle Consulting, and Boise City Public Works. 2008. Numeric Nutrient Criteria Development for the Lower Boise River Using the AQUATOX Model. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Collins, C. D. and J. H. Wlosinski. 1983. Coefficients for Use in the U.S. Army Corps of Engineers Reservoir Model, CE-QUAL-R1. Tech. Rept. E-83-15, Environmental Laboratory, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Miss.

Cowan, W. L. 1956. Estimating hydraulic roughness coefficients. Agricultural Engineering **37**:473-475.

Donigian, A. S., J. T. Love, J. S. Clough, R. A. Park, J. N. Carleton, P. A. Cocca, and J. C. Imhoff. 2005. Nutrient Criteria Development with a Linked Modeling System: Watershed and Ecological Model Application and Linkage.*in* 2005 TMDL Conference. Water Environment Federation, Philadelphia PA.

Donigian, A. S., R. A. Park, and J. S. Clough. 2009. Tenkiller Ferry Lake Modeling Case Study - HSPF and AQUATOX. Office of Water/Office of Science and Technology, U.S. Environmental Protection Agency, Washington DC.

Fourt, R., S. Wooldridge, and D. Brader. 2002. Striped Bass Management Plan. Arkansas Game and Fish Commission.

Goldman, J. C. and E. J. Carpenter. 1974. A kinetic approach to the effect of temperature on algal growth'. Limnology and Oceanography **19**:756-776.

Hanson, P. C., T. B. Johnson, D. E. Schindler, and J. F. Kitchell. 1997. Fish Bioenergetics 3.0. Center for Limnology, University of Wisconsin, Madison.

Hewett, S. W. and B. L. Johnson. 1992. Fish Bioenergetics 2 Model. University of Wisconsin Sea Grant Institute, Madison.

Hill, W. R. and H. L. Boston. 1991. Community development alters photosynthesis-irradiance relations in stream periphyton. Limnol. Oceanogr. **36**:1375-1389.

Horne, A. J. and C. R. Goldman. 1994. Limnology. Second edition. McGraw-Hill, New York.

Jørgensen, L. A., S. E. Jørgensen, and S. N. Nielsen. 2000. ECOTOX: Ecological Modelling and Ecotoxicology. Elsevier Science B.V.

Jørgensen, S. E., editor. 1979. Handbook of Environmental Data and Ecological Parameters. Pergamon Press, Oxford.
Leidy, G. R. and R. M. Jenkins. 1977. The Development of Fishery Compartments and Population Rate Coefficients for Use in Reservoir Ecosystem Modeling. CR-Y-77-1, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Miss.
Leidy, G. R. and G. R. Ploskey. 1980. Simulation Modeling of Zooplankton and Benthos in Reservoirs: Documentation and Development of Model Constructs. E-80-4, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Miss.
Levins, R. 1966. The strategy of model building in population biology. American Scientist **54**:421-431.
NASA. 2010. Surface Meteorology and Solar Energy Interannual Variability. NASA Atmospheric Science Data Center.
Oliver, B. G. and A. J. Niimi. 1988. Trophodynamic Analysis of Polychlorinated Biphenyl Congeners and Other Chlorinated Hydrocarbons in the Lake Ontario Ecosystem. Environ. Sci. Technol. **22**:388-397.
Park, R. A., J. N. Carleton, J. S. Clough, and M. C. Wellman. 2009. AQUATOX Technical Note 1: A Calibrated Parameter Set for Simulation of Algae in Shallow Rivers. EPA-823-R-09-003, U.S. Environmental Protection Agency, Washington D.C.
Park, R. A. and C. D. Collins. 1982. Realism in Ecosystem Models. Perspectives in Computing **2**:18-27.
Raimondo, S., D. N. Vivian, and M. G. Barron. 2007. Web-based Interspecies Correlation Estimation (Web-ICE) for Acute Toxicity: User Manual. Version 1.1. EPA/600/R-07/071, U.S. Environmental Protection Agency, Gulf Breeze, FL.
Reynolds, C. S. 1984. The Ecology of Freshwater Phytoplankton. Cambridge University Press, Cambridge.
Rykiel, E. J., Jr. 1996. Testing ecological models: the meaning of validation. Ecological Modelling **90**:229-244.
Saunders, G. W. 1980. Organic Matter and Decomposers. Pages 341-392 *in* E. P. Le Cren and R. H. Lowe-McConnell, editors. The Functioning of Freshwater Ecosystems. Cambridge University Press, Cambridge.
Schnoor, J. E. 1996. Environmental Modeling: Fate and Transport of Pollutants in Water, Air, and Soil. John Wiley & Sons, Inc., New York.
Schwarzenbach, R., P. M. Gschwend, and D. M. Imboden. 1993. Environmental Organic Chemistry. John Wiley & Sons, New York.
Stevenson, R. J. 1996. An Introduction to Algal Ecology in Freshwater Benthic Habitats. Pages 3-30 *in* R. J. Stevenson, M. L. Bothwell, and R. L. Lowe, editors. Algal Ecology: Freshwater Benthic Ecosystems. Academic Press, San Diego.
Stevenson, R. J., Y. Pan, J. P. Kociolek, and J. C. Kingston, editors. 2007. Advances in Algal Biology: A Commemoration of the Work of Rex Lowe. Springer, Amsterdam.
Suárez, L. A. 2005. PRZM-3, A Model for Predicting Pesticide and Nitrogen Fate in the Crop Root and Unsaturated Soil Zones: Users Manual for Release 3.12.2. EPA/600/R-05/111, U.S. Environmental Protection Agency, Athens GA.
Tetra Tech Inc. 2002. Draft User's Manual for Environmental Fluid Dynamics Code Hydro Version (EFDC-Hydro). U.S. Environmental Protection Agency, Atlanta, GA.

U.S. Environmental Protection Agency. 1988. The Effects of Chlorpyrifos on a Natural Aquatic System: A Research Design for Littoral Enclosure Studies and Final Research Report. Duluth MN.

U.S. Environmental Protection Agency. 2012. AQUATOX (Release 3.1) Modeling Environmental Fate and Ecological Effects in Aquatic Ecosystems, Volume 2: Technical Documentation. EPA-820-R-12-015, U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Washington DC.

Verscheuren, K. 1983. Handbook of Environmental Data on Organic Chemicals. Second edition. Van Nostrand Reinhold, New York.

Walker, W. W. 1999. Simplified Procedures for Eutrophication Assessment and Prediction: User Manual. Instruction Reporl; W-96-2, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Wetzel, R. G. 2001. Limnology: Lake and River Ecosystems. Third edition. Academic Press, San Diego.