AWRA 2009 SUMMER SPECIALTY CONFERENCE
Snowbird, Utah
June 29-July 1, 2009

WATERSHED MODELING FOR WATER QUALITY TRADING

B. T. Neilson, C. Bandaragoda, M.E. Baker, J. S. Horsburgh, D. K. Stevens[1]

ABSTRACT: Water quality trading has been proposed as a potential solution for addressing water quality impairments in a cost effective manner. Determining if a trading program is feasible requires key information about both point and nonpoint loads and the amount of each load reaching a location of interest, or delivery ratio. While point source loading estimates are relatively simple to determine, nonpoint source loads are much more difficult. Additionally, delivery ratios, which are primarily dependent on instream processes and withdrawals, can be difficult to estimate. In this paper, a modeling framework is proposed for developing seasonal farm/field loads and the associated delivery ratios required to support a water quality trading program. Our results show this approach to be an objective and scientifically based foundation for the development of water quality trading programs that include nonpoint sources of pollution.

INTRODUCTION

Water quality trading can potentially be a cost effective solution to mitigating key sources of pollution and provide a means to adaptively manage water resources. Setting up a water quality trading program, however, requires the quantification of both point and nonpoint sources of pollution and an understanding of how pollutants behave in the system. As stated in guidance provided by the EPA (EPA 2004), the steps required to complete a suitability analysis for trading a particular pollutant include determining watershed loading profiles, understanding the effects of load timing, and considering water quality equivalence.

A watershed loading profile is a description of the current pollutant loads within a watershed by source and location. Point source loads are generally regulated and monitored, which makes them relatively easy to quantify. Nonpoint sources of pollution, however, are not regulated through a permitting process, require voluntarily implementation of management actions, and are difficult to identify, measure, and estimate. For integration and consistency with the Total Maximum Daily Load (TMDL) program, a loading profile must include reductions in point and nonpoint sources set by a TMDL developed for the focus water body. The current loads less the load reductions required by the TMDL result in target loads. The timing of these loads relative to a critical time period for a water body is also an important consideration for both trading and TMDL development. If a water quality management program only considers annual loads from sources, the receiving water body may not meet the instream water quality standard during limiting times throughout the year.

Development of a watershed loading profile and testing the feasibility of a water quality trading program in a highly agricultural area with demonstrated water quality impairments require an objective and scientifically sound method for estimating time variable loads and their associated delivery ratios to key receptor points. Towards this end, we have developed a modeling framework for determining seasonal current total phosphorus (TP) loads at the spatial scale of specific farm/fields (i.e., seasonal field current loads) and seasonal fractions of the field TP loads that reach a receptor point (i.e., seasonal delivery ratios). With this information, seasonal field delivered loads can be calculated, and the modeling results can be used to inform each farm or field owner of the amount of TP they own and can potentially sell in a trading program (i.e., seasonal tradable loads).

MODELING FRAMEWORK

[1] Respectively: Assistant Professor, Civil and Environmental Engineering, Utah Water Research Laboratory, Utah State University, Logan, UT, 8200 Old Main Hill, Logan, UT, 84322-8200, 435-797-7369, bethany.neilson@usu.edu; Research Consultant, Silver Tip Solutions, Mukilteo, WA; Assistant Professor, Department of Geography & Environmental Systems, University of Maryland, Baltimore County, Baltimore, MD; Research Professor, Civil and Environmental Engineering; Utah Water Research Laboratory, Utah State University, Logan, UT; Professor, Civil and Environmental Engineering, Utah Water Research Laboratory, Utah State University, Logan, UT.

For a modeling application to more fully support a water quality trading program, it must be able to simultaneously capture physical hydrology at the watershed scale, while representing the spatial variability of loads at the farm/field scale. Depending on the goal of a typical watershed management application, watershed models are generally developed for either broad watershed or field specific applications. A number of existing watershed models have been used in watershed management and TMDL development. Examples include Hydrologic Simulation Program FORTRAN (HSPF) (Bicknell et al. 2000), Soil and Water Assessment Tool (SWAT) (Neitsch et al. 2005), and Generalized Watershed Loading Function model (GWLF) (Haith and Shoemaker 1987). The utility of these and other modeling approaches in water quality trading has been recognized to assist in estimating information that supports trading ratios (e.g., delivery ratios, spatial and temporal loads, uncertainty ratios, etc.). Model selection, modeling approaches, and application of results in a trading program, however, have been highly variable (Chesapeake Bay Program Nutrient Trading Negotiation Team 2000; Fang et al. 2005; Horan et al. 2002; Idaho Department of Environmental Quality 2003).

To more accurately understand and assess water quality trading potential, estimates of field scale loads linked to whole basin behavior is required. Additionally, time variable delivery ratios to receptor points of interest need to be determined. These information requirements led us to the development of a Water Quality Trading (WQT) modeling framework that coupled a number of models, modeling approaches, and processing tools to provide the necessary information to facilitate water quality trading. The framework includes: TOPNET (Bandaragoda et al. 2004) for simulating watershed hydrology; variable source area (VSA) calculations (Lyon et al. 2004) for resolving spatial areas contributing surface runoff and pollutant loading; a subbasin Loading Model component based on the VSA calculations, event mean concentrations (EMCs) for determining pollutant concentrations in surface runoff, and spatially distributed land use information; and a Water Body Response (WBR) component that incorporates QUAL2E (Brown and Barnwell 1987) to determine delivery ratios.

TOPNET Component

TOPNET (Bandaragoda *et al.*, 2004) is a semi-distributed, rainfall-runoff model that combines TOPMODEL (Beven et al. 1995; Beven and Kirkby 1979) with a kinematic wave channel routing algorithm (Goring 1994) to create a hydrologic model that can be applied over large watersheds using small subbasins as model elements. Semi-distributed models can have an advantage over lumped models in variable terrain when the landscape is subdivided to reduce within-element heterogeneity. This may result in a more accurate representation of the spatial distribution of physical characteristics of the basin. A key contribution of TOPMODEL is the parameterization of the soil moisture deficit, or depth to water table, using a topographic index to identify hydrologically similar areas with respect to saturation excess. TOPNET uses TOPMODEL concepts for the representation of subsurface storage controlling the dynamics of saturated contributing areas and baseflow recession. To form a complete model, potential evapotranspiration, interception, and soil zone components are represented, as well as a terrain-based system for delimiting streams, model catchments, and estimation of model parameters (Bandaragoda et al. 2004).

Variable Source Area (VSA) Component

The purpose of the VSA component was to locate the origin of predicted saturation excess flow, and to link surface runoff generation and landuse in the estimation of surface water quality. Using the calibrated results of TOPNET, a combination of parameterized and state variables for each subbasin was used to determine probable nonpoint source pollution locations. Calculations use estimates of the following state variables specific to each subbasin: watershed area, drainable porosity, and the distribution of topographic index (slope/contributing area) values. In addition, this component uses time series of effective precipitation, predicted average water table depths, and saturation excess runoff estimated by TOPNET. The saturated fraction of subbasin area contributing to runoff was estimated using effective precipitation and average watershed storage following the approach of Steenhuis *et al.* (1995) and Lyon *et al.* (2004). Effective precipitation was derived from interpolated measures of total precipitation and potential evapotranspiration. Following Wang *et al.* (2006), average watershed storage was calculated as the product of drainable porosity and average water table depth in each subbasin. A time series of saturated subbasin fractions was then used to map the origin of predicted saturation excess runoff according to the spatial distribution of topographic index values.

Mapped values of the topographic index were combined with a landuse raster and subbasin boundaries within a geographic information system. The geographic intersection of these inputs was used to create a database of individual topographic index values and their landuse by subbasin. In order to relate saturated watershed fractions to specific mapped locations within a watershed, Lyon *et al.* (2004) proposed a function relating saturated fractions to a cumulative probability distribution of the topographic index, valid as long as the difference between the overall storage deficit and each local storage deficit is a function of the average topographic index and each local index value. However, as originally conceived, TOPMODEL assumes that the realized zone of saturation will be small and the overall watershed storage deficit (S) is the integral of *all* local storage deficits ($\int s_i$), when in fact, S should be the integral of all *unsaturated* local storage deficits ($\int s_i^*$). Thus, we employed the solution proposed by Saulnier and Daltin (2004) to compensate for this analytical bias in relating saturated watershed fractions and critical values of the topographic index.

A combination of saturated watershed fraction estimates and the topographic index distribution were used to generate a time series of critical topographic index values required for pixel saturation. Assuming that saturated pixels contributed equally to runoff during each TOPNET timestep, predicted saturation excess flow (m^3/d) was divided by the number of contributing saturated pixels to represent flow per-unit-area. Cumulative yields attributed to each 30-m pixel were tracked by season for the entire time series and converted to seasonal averages for the simulation period.

Loading Model Component

The Loading Model component provides estimates of TP loads from diffuse sources within each subbasin, or distributed model element. Two flow components, surface runoff (combination of saturation and infiltration excess) and baseflow, are provided by TOPNET for individual subbasins for each model timestep (1 day). The time series of critical topographic index values provided by the VSA calculations were linked to landuse classes through a database of mapped attributes. From this table it was possible to determine the area and proportional fraction of land uses contributing to surface runoff at each time step. In the current application, all surface runoff is treated as saturation excess (i.e., infiltration excess is treated as saturation excess). Surface runoff TP concentrations are based on the land cover from which the surface flow originates and are assigned using event mean concentrations (EMCs) for TP. Since an EMC is the average TP concentration in surface runoff generated from a particular land cover, and since this model component simulates TP concentrations at the subbasin outlets, a subbasin average EMC (EMC_{avg}) was defined to represent the flow weighted average EMC (across all land cover classes) for TP in each modeled subbasin. To quantify the effects of base flow loads, average base flow concentrations (C_b) are input to the loading model component. Once the surface runoff and base flow TP concentrations are known, the average subbasin outlet TP concentrations are given by the flow weighted average of the surface runoff concentrations and the base flow concentrations. The average subbasin outlet TP concentration and the total flow (surface flow plus base flow) were calculated for each modeled subbasin at each time step.

Water Body Response (WBR) Component

The purpose of the WBR component is to accumulate the TP loads from the subbasins, and route them downstream to the receptor point while accounting for intervening point loads, diversions, and chemical and biological processes that might affect the fate of TP. The WBR accepts inputs from the subbasins based on the TOPNET, VSA, and Loading Model Components. The WBR component inputs also include external data concerning weather, point loads, diversions, and reservoir releases. The instream routing model component uses the public domain QUAL2E model (Brown and Barnwell 1987) for steady-state calculation of 15 water quality measures. Outputs from QUAL2E enable computation of the delivery ratio for the reach within a subbasin for each day. The degree to which the load from a particular field influences the constituent concentration at the receptor point depends on what happens to the constituent as it traverses the intervening river reaches. Total P is taken up into algae, settled in particulate form, and modified by benthic sources/sinks. TP is also removed when water is diverted for agricultural or municipal/industrial use. The significance of these losses are situation dependent in the sense that many of the attenuating processes are nonlinear and the fraction of the TP that is taken up is dependent on the concentrations of TP and algae.

The attenuation of TP occurring within a stream reach is calculated as the fraction (from 0 to 1) of the mass flow that remains at the downstream terminus of a reach relative to the mass flow that enters a reach via the upstream reach or any point loads. The delivery ratio (*DR*) is defined as the load leaving the reach divided by the load entering the reach. This means that the load leaving a reach is equal to the delivery ratio multiplied by the load entering the reach ($Load_{leaving\ reach} = DR\ x\ Load_{entering\ reach}$). The individual delivery ratio values are found using the QUAL2E model for the representative reach within a subbasin. In this way, we can estimate the overall delivery ratio as the product of the subbasin-based *DR* values along the reach path leading from an upstream location to any downstream location as shown in Equation (1).

$$DR_{overall} = \prod_i DR_{subbasin\ i} \qquad (1)$$

APPLICATION OF MODEL RESULTS

To establish a water quality trading program, each individual stakeholder in a watershed needs to know the amount of pollutant they have available to trade. This quantity, designated here as the tradable load, is a function of: 1) the field current load 2) the target load (based on the TMDL or other watershed management requirements), 3) the associated delivery ratio to a receptor point and the delivered load, and 4) the ability to mitigate a load. With the resulting seasonal estimates of loads and delivery ratios that are based on a many year average of daily modeling results, variability over different hydrologic conditions can be accounted for. It is recognized that there are many other factors that may be important in determining tradable loads (e.g., more explicit uncertainty estimates or ratios), however, this section of the paper presents a simplified method to determine tradable loads and illustrates how the information provided by the modeling framework can be used in the context of a water quality trading program in a nonpoint source dominated watershed.

Farm/Field Current Loads - Once the results of the cumulative seasonal saturation excess runoff per pixel are generated from the hydrologic and VSA components of the modeling framework, appropriate EMCs assigned by landuse in the loading component of the framework were multiplied by the water yields per pixel to determine the seasonal load per pixel. The field loads were then determined by summing the values for all 30-m pixels contained within the boundary of each field. The fields can then be aggregated to determine the farm loads, which are made up of multiple fields. An example of aggregated seasonal current field loads taken from a case study in the Bear River Basin in Utah is shown in Figure 1. These loads are the 15 year average of the seasonal load (calculated by averaging the daily loads over a season).

Target and Delivered Loads-To calculate the amount of phosphorus a stakeholder has to trade, a seasonal delivered load must first be calculated. In many cases, this calculation will be dependant on the target load specified by a TMDL or other watershed management programs. For example, if a TMDL for a watershed specified a need for an overall 20% decrease in agricultural nonpoint source pollution, before a stakeholder knows the amount of a pollutant available for trade, they must account for this prior required reduction according to Equation (2).

$$\begin{aligned}\text{Field Target Load}_{\text{Season}} &= \text{Field Current Load}_{\text{Season}} \\ &\quad - (\%\ \text{TMDL Reduction}) * \text{Field Current Load}_{\text{Season}}\end{aligned} \tag{2}$$

The seasonal load delivered to a receptor point from each field is calculated by multiplying seasonal target loads by the seasonal delivery ratio for the subbasin the field is located within (Equation (3)). Where a TMDL reduction is not required, the current load or field load can be multiplied by the delivery ratio to determine the field delivered load (Equation (4)).

$$\text{Field Delivered Load}_{\text{Season}} = \text{Field Target Load}_{\text{Season}} * \text{Delivery Ratio}_{\text{Season}} \tag{3}$$

$$\text{Field Delivered Load}_{\text{Season}} = \text{Field Load}_{\text{Season}} * \text{Delivery Ratio}_{\text{Season}} \tag{4}$$

Tradable Loads - With the calculation of the field delivered loads, the final step is to determine the amount of the load available to trade. The amount a field owner has to trade is called the field tradable load, which is defined here as the fraction of the field delivered load that can be offset through the implementation of management practices. As shown in Equation (5), this can be determined by multiplying the field delivered load by a % BMP reduction.

$$\text{Field Tradable Load}_{\text{Season}} = \text{Field Delivered Load}_{\text{Season}} * \%\ \text{BMP Reduction} \tag{5}$$

The % BMP reduction must be determined for a specific BMP and a combination of the resulting anticipated reductions based on literature values and existing conditions of the field (e.g., soils and crop types, prior BMPs implemented, etc.). A local trading facilitator would likely be the optimal selection to facilitate the estimation of this % BMP reduction. Once the field tradable loads are determined, each field or farm owner knows the amount of total phosphorus they own and can sell to another point or nonpoint source trying to offset a portion of their load.

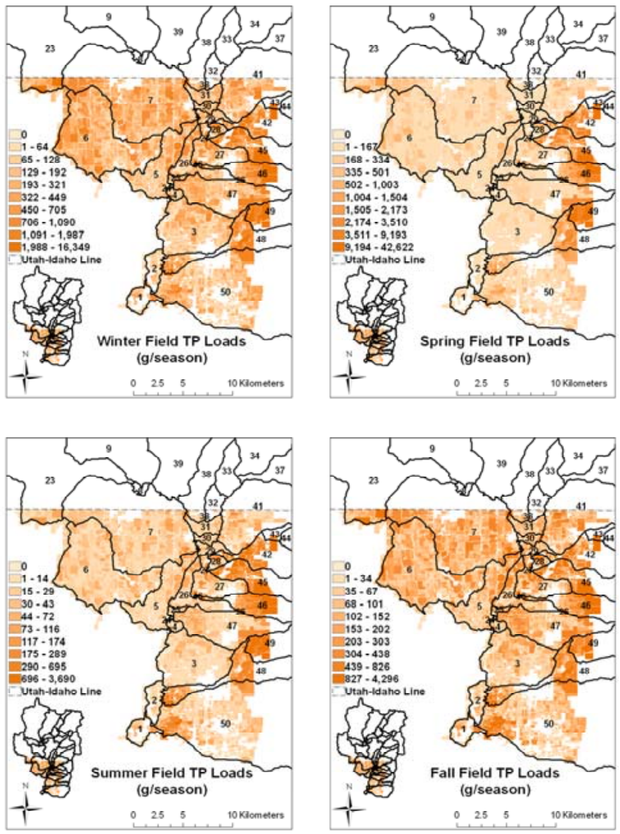

Figure 1. Field total phosphorus loads for each season for a portion of the Bear River basin in Utah.

LIMITATIONS OF THE FRAMEWORK

Throughout the development of this modeling framework and approach, a number of limitations were identified. Under the current framework, land surface loads are not routed within the subbasin of origin and therefore, no assimilation occurs at the local scale. All assimilation occurs once the load enters the WBR component. Given the current implementation, it is also necessary that this system be applied to watersheds that have relatively short travel times from top to bottom due to the hydrology and loading models calculating daily average flows and loads at each drainage outlet and QUAL2E calculating the steady-state response to those daily average loadings. This requires the effects of one day’s load to clear the stream reach before the next day’s load is applied. Some other limitations that require future research and development include: the need to incorporate infiltration excess flow into the VSA component instead of assuming that saturation excess represents the dominant source of surface runoff and the need to enforce hydrologic connectivity of saturated cells to a waterbody in order to transport sediment and phosphorous and contribute to surface loads.

CONCLUSIONS

The growing number of impaired waterbodies in the United States necessitates new approaches and incentives to control pollution. The potential solution of setting up water quality trading programs, however, requires the quantification of both point and nonpoint sources of pollution and an understanding of how pollutants behave in individual hydrologic systems. The modeling framework presented in this paper to assist in estimating field current loads and the corresponding overall delivery ratios is an objective and scientifically based mechanism to support a water quality trading program that can evolve as the demands within the watershed change. Example calculations have additionally shown how the information estimated using this model framework can be utilized to calculate the delivered loads and eventually, the amount of pollution a field owner could potentially trade, or tradable loads. The results from this effort are currently being used to determine whether a water quality trading program is feasible in a study watershed in Utah.

REFERENCES

Bandaragoda, C., Tarboton, D. G., and Woods, R. (2004). "Application of Topnet in the Distributed Model Intercomparison Project." *Journal of Hydrology*, 298, 178-201.

Beven, K., Lamb, R., Quinn, P., Romanowicz, R., and Freer, J. (1995). "Topmodel." Computer Models of Watershed Hydrology, V. P. Singh, ed., Water Resources Publications, Highlands Ranch, CO, 627–668.

Beven, K. J., and Kirkby, M. J. (1979). "A physically based variable contributing area model of basin hydrology." *Hydrological Sciences Bulletin* 24(1), 43–69.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L., Jobes, T. H., and Donigian, A. S. (2000). "Hydrological Simulations Program: FORTRAN (HSPF): User's Manual for Release 12.0 ", E. R. Laboratory, ed., United States Environmental Protection Agency, Athens, GA.

Brown, L. C., and Barnwell, T. O. (1987). "The Enhanced Stream Water Quality Model QUAL2E and QUAL2E-UNCAS: Documentation and User manual." United States Environmental Protection Agency, Athens, GA.

Chesapeake Bay Program Nutrient Trading Negotiation Team. (2000). "Chesapeake Bay Program Nutrient Trading Guidance Document - Draft." Chesapeake Bay Program

EPA. (2004). "Water quality trading assessment handbook: Can water quality trading advance your watershed's goals?" O. o. Water, ed., National Service Center for Environmental Publications, 110.

Fang, F., Easter, K. W., and Brezonik, P. L. (2005). "Point-Nonpoint Source Water Quality Trading: A Case Study in the Minnesota River Basin." *Journal of the American Water Resources Association*, 41(3), 645-658.

Goring, D. G. "Kinematic shocks and monoclinal waves in the Waimakariri, a steep, braided, gravel-bed river " *International Symposium on Waves: Physical and Numerical Modelling*, University of British Columbia, Vancouver, Canada, 336–345.

Haith, D. A., and Shoemaker, L. L. (1987). "Generalized Watershed Loading Functions for Stream Flow Nutrients." *Water Resources Bulletin*, 23(3), 471-478.

Horan, R. D., Shortle, J. S., and Abler, D. G. (2002). "Point-Nonpoint Nutrient Trading in the Susquehanna River Basin." *Water Resources Research*, 38(5), 13.

Idaho Department of Environmental Quality (IDEQ). (2003). "Pollutant Trading Guidance - Draft." Idaho Department of Environmental Quality.

Lyon, S. W., Walter, M. T., Gerard-Marchant, P., and Steenhuis, T. S. (2004). "Using a topographic index to distribute variable source area runoff predicted with the SCS curve number equation." *Hydrological Processes*, 18, 2757-2771.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Williams, J. R. (2005). "Soil and Water Assessment Tool: Version 2005." Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Blackland Research Center, Texas Agricultural Experiment Station, Temple, TX.

Saulnier, G., and Datin, R. (2004). "An anlayitical solution to a bias in the TOPMODEL framework balance. ." *Hydrological Processes* 18, 1195-1218.

Soil Survey Staff. (2007). "Soil Survey Geographic (SSURGO) Database for Survey Area Natural Resources Conservation Service." United States Department of Agriculture.

Steenhuis, T., Winchell, M., Rossing, J., Zollweg, J., and Walter, M. (1995). "SCS runoff equation revisited for variable-source runoff areas." *Journal of Irrigation and Drainage Engineering*, 121, 234-238.

Wang, J., Endreny, T. A., and Hassett, J. M. (2006). "Power-function decay of hydraulic conductivity for a TOPMODEL-based infiltration routine." *Hydrological Processes* 20, 3825-3834.