Estimating Phosphorous Load from a Large Watershed in the Great Lakes Basin

Weichun Tao, Carlo DeMarchi,
Department of Geological Sciences
Case Western Reserve University
Cleveland, Ohio 44106-7216, USA
carlo.demarchi@case.edu, wxt84@case.edu

Chansheng He
Department of Geography
Western Michigan University
Kalamazoo, Michigan 49008-5424, USA

Thomas H. Johengen
School of Natural Resources & Environment
University of Michigan
Ann Arbor, Michigan 48109-1041, USA

Craig Stow
Great Lakes Environmental Research Laboratory
National Oceanic and Atmospheric Administration
Ann Arbor, Michigan 48105-2945, USA

Abstract— **Common ways to quantify watershed nutrient loads include estimating the annual or seasonal loads using simple relations between discharge and load, such as the ratio estimator, and fitting complex nutrient transport models to the observed concentrations. The former approach produces quite uncertain estimates at low temporal resolution when based on typically infrequent routine monitoring data. The second approach may produce more reliable estimates, even at high temporal resolution, but requires a lot of time and auxiliary data. The approach explored in this paper uses linear combination of river discharge at the time of estimate and for antecedent periods to quantify Total Phosphorous (TP) concentration, yielding high resolution load estimates sufficiently reliable for a variety of applications**

Keywords- Phosphorus load estimation, Saginaw Bay, Great Lakes, Regression, Nutrient loads.

I. INTRODUCTION

Phosphorous load from human activities in the draining basin is the major cause of river, lake, and swamp eutrophication. Consequently, estimating phosphorous (TP) loads from watersheds at temporal scales varying from annual and seasonal, for policy analysis, to daily, for driving detailed water quality models of the recipient water bodies, is often essential in water resources management. Yet, water quality monitoring frequency is often insufficient for reliably assessing annual pollutants loads, let alone daily watershed outputs. Building full nutrient generation and transport models is the best response to this challenge, yet it is a complex and time consuming endeavor. On the other hand, simple regression models can provide accurate TP quantification, even at fine temporal scales and in absence of frequent measurements, without the costs of developing a full transport model.

Theoretically, estimating nutrient outputs from watersheds should be relatively easy, since load is the product of concentration and discharge, two easily measurable quantities. However, while discharge data are frequently provided by river gages at daily or sub-daily frequency, water quality data are normally less frequent, especially if they are not the product of ad-hoc campaigns. Thus, concentration measurements are often insufficient to represent the concentration variability, causing uncertainty in load estimates. Methods for estimating watershed loads can be divided into two categories. Methods employing simple relations between discharge and load, such as Averaging methods [1-4], Ratio estimators [2-6], and Regression methods [2, 3, 7, 8], and methods employing complex nutrient transport models such as the Soil and Water Assessment Tool [9] or Agricultural Nonpoint Source Pollution Model [10].

According to Quilbé et al. [3], (i) averaging methods are accurate only when concentration measurements are frequent enough to sample the entire flow range; (ii) the ratio estimator is less sensitive to river and pollutant characteristics than regression methods, but requires more data to achieve the same level of precision; (iii) regression methods can give the best results for sediments and total P if stream flow and concentration data are strongly correlated for a wide range of stream flow values. Complex nutrient transport models may supply more reliable high frequency estimates, have the capability of tracking the sources of pollutants, and of simulating alternative pollution prevention policies, but require a lot of effort an data to be deployed. In this paper we explore a regression model that takes advantage of infrequent, but long-term, water quality data to produce reliable TP estimates, even at high frequency, and its application to the Saginaw River Basin, Michigan (Figure 1).

A two-step effort is currently underway to better understanding the nutrients loads entering Saginaw Bay. In the first step, we use the regression models to evaluate current nutrient loads using the relatively infrequent water quality measurements and daily discharge data available at a few points in the basin. In the second step, we will adapt a distributed water quality model simulating pollution generation and transport to the watershed, using experimental data and the results of the models produced in the first steps for calibration and validation [11-13]. In Section II, we describe the watershed and available data, and the model development. In Section III we compare the estimated load with the target load to the bay, and determine the contribution of the sub-watersheds to the total load.

II. METHODS

A. *Study area*

The Saginaw River and its sub-basins (green shades in Figure 1) make the largest (16,680 km^2) tributary of Saginaw Bay, a large (2,770 km^2) and shallow (average

Proceedings of the 2010 International Conference on Challenges in Environmental Science and Computer Engineering (CESCE 2010), March 6-7, 2010, Wuhan, China. IEEE Computer Society's Conference Publishing Services, Los Alamitos, California USA, pp. 427-430.

depth 5.1 m) gulf in western Lake Huron, which is used for drinking water supply, fishing, tourism, and navigation. The basin is an important base for industrial supply and food production, with agriculture and forests being the two major land uses. Over the years, the extensive agricultural land use and associated runoff, improper manure management, industrial pollution, and inadequate sewage systems have led to high nutrient runoff, eutrophication in the bay, loss of wildlife habitat, and beach closures. To improve this situation, the Great Lakes Water Quality Agreement between the United States and Canada established a target Total Phosphorus (TP) load of 440 metric tons/yr for Saginaw Bay. Consequently, efforts to control excessive phosphorus inputs were implemented and early assessment indicated diminishing eutrophication symptoms in response to phosphorus load reductions. However, eutrophication symptoms, including algal blooms and nuisance algal beach deposits, have returned to Saginaw Bay with underscoring the need for continued long-term phosphorus load evaluation.

B. Data Sources

To estimate TP loads, Saginaw river and sub-watersheds data for TP concentration 1997-2007 were obtained from Michigan Department of Environmental Quality except Saginaw up for which data was from USGS. Total Phosphorus concentration in 2008 was provided by the Cooperative Institute for Limnology and Ecosystem Research. Table I shows the total number of available concentration samples.

Flow data for Saginaw river and sub-watersheds at sample site from 1997 to 2008 were obtained from U.S. Geological Survey (USGS) database.

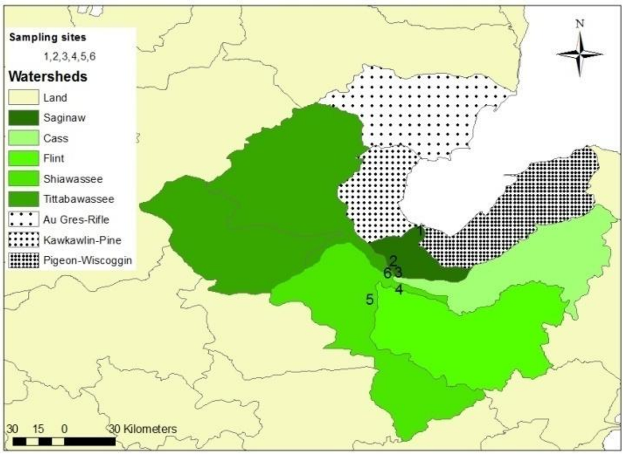

Figure 1. The Saginaw Bay basin and sampling sites used for model calibration

C. Concentration Eestimation

*1)*Coefficient of determination (r^2)

Regression method does not require extensive data but the quality of prediction depends on the quality of the correlation between flows and concentrations. Quilbé et al. [10] proposed the coefficient of determination (r^2) to select the regression method, which means that the higher the r^2 is, the more the variability of concentration should be explained by stream flow. The coefficient in this case shown in Table II indicates that by splitting data into two seasons the quality of prediction may be improved especially for Saginaw River Basin Outlet.

Table I. AVAILABLE IN-STREAM WATER QUALITY DATA

	Flint (4)	**Tittabawasee(6)**	**Shiawassee(5)**	**Cass(3)**	**SaginawUp(2)**	**SaginawRiver BasinOutlet(1)**
97					4	
98		13	13		4	8
99					3	
00	3	3	6	3	4	
01	4	4	4	12	4	12
02	4	12	4	4	4	12
03	11	4	4	4	4	12
04	4	4	4	4	4	12
05	4	4	12	4	2	12
06	4	4	4	12	2	12
07	4	12	4	4		12
08	8	8	8	8	32	31

Table II. R^2 CONCENTRATIONS AND FLOWS, TIME

	Daily flow(Q*t*)		***t***
	Annual	***Two seasons***	
Saginaw River Basin Outlet	0.446	0.510	0.046
Flint	0.005	0.111	0.171
Cass, Shiawassee, Tittabawassee	0.234	0.270	0.035
Saginaw Up	0.140	0.269	0.000

*2)*Temporal trend

Table 2 shows there is a weak temporal trend probably due to long-term changes in watershed conditions (e.g., expansion of no-till agricultures; adoption of better wastewater treatment techniques; build-up of fertilizers in agricultural soil, etc.). Particularly for Flint watershed, the coefficient of determination (r^2) reaches 0.171. Therefore, we explored the use of time as explanatory variable in the TP concentration estimation model.

*3)*Concentration estimate

The regression models we considered here (Table III) express TP concentration as: 1) average concentration; 2) a linear function of the same-day average discharge (Q); 3) a power function of Q; 4) a linear combination of Q and the average discharge in a previous period, the length of which depends on the watershed size (Qave), to take into consideration the difference between the rising and receding phases of floods; 5) a linear combination of Q, the average discharge in the previous five days (Q5), to take into consideration the difference between the rising and receding phases of floods, and the average discharge in the previous ten days (Q10) to account for the flushing effect of previous storms. In order to capture the variable relation between discharge and TP concentration with the limited number of samples available each year, we combined data from a long time period. As table III shows, model 4) and 5) outperformed the other three models with model 5) being slightly better. As mentioned above, in model 6) we superimposed the linear temporal trend mentioned above for the TP to Model 5).

III. Results and Discussion

A. Model performance

The result shows model 6) was better able to track TP dynamics, providing excellent results especially for the Saginaw River, where data were more abundant: correlation (Table IV) between modeled and observed daily concentration in 1998-2008 reaches 0.84 for the entire Saginaw River watershed, 0.63 for the combined Cass, Shiawassee, and Tittabawassee Rivers, and 0.62 for the Flint River, while correlation in daily load is above 0.96 for all watersheds.

B. Estimated load

Fig. 2 indicates that the TP loads from the Saginaw River have been higher than the target TP load to the bay most of the years. Considering that Saginaw River carries around 80-90% of the TP load to the Bay with the rest contributed by the AuGres-Rifle, Kawkawlin-Pine, and Pigeon-Wiscoggin Rivers as well as the atmospheric deposition, it is clear that the target TP load has been met only during the driest years and that the average TP load is well above 500 metric tons P per year.

C. Contribution of sub-watersheds

By considering the difference between the TP loads carried by the tributaries and the load at the Saginaw River outlet is possible to determine the contribution of the different portions of the basin to the TP load entering the Bay (Fig. 3). While the largest fraction of Saginaw River TP load originates in the largest sub-basin (Tittabawassee), Fig. 3 highlights the importance of the urban discharges at the coastal cities of Saginaw and Bay as TP source (in average 23% of the total), that feature a TP load per square kilometer four times higher than the average level of other sub-watersheds (Fig. 4) showing the preponderance of pollution from urban sources.

Table III. Regression models explored for TP concentration for Saginaw River and tributaries and correlation between observed and modeled concentrations

Model #	Regression equation	1998-08 TP concentration correlation at sampling sites: Saginaw River Basin Outlet (1)	Flint River (4)	Cass, Shiawassee, Tittabawassee (3, 5, 6)	Saginaw River Upstream (2)
1)	$C(t) = a$	0.03	0.05	0.01	0.35
2)	$C(t) = a + b \cdot Q$	0.71	0.33	0.52	0.52
3)	$C(t) = aQ^b$	0.62	0.32	0.49	0.48
4)	$C(t) = a + b \cdot Q + c \cdot Q_{ave}$	0.82	0.51	0.62	0.74
5)	$C(t) = a + b \cdot Q + c \cdot Q_5 + d \cdot Q_{10}$	0.82	0.57	0.61	0.74
6)	$C(t) = a + b \cdot Q + c \cdot Q_5 + d \cdot Q_{10} + e \cdot t$	0.84	0.62	0.63	0.76

IV. Conclusion

Estimation of total phosphorus loads from the drainage basin is fundamental for preventing the eutrophication of lakes and wetlands. The challenge in load estimation consists in making the best use of available information, which often includes daily or subdailystreamflow data, but only infrequent water quality data.. Although the relation between discharge and TP concentration is not very accurate, the approach shown here take advantage of infrequent, but long-term, water quality samples to produce reliable estimates at high temporal resolution. In particular we show the advantage of combining temporal trend and splitting data to two seasons as well as taking into consideration the difference between the rising and receding phases of floods, and account for the flushing effect of previous storms. We also showed that such models can provide the answer to several policy questions concerning nutrients load generation in the watershed as well as supply reliable inputs to models simulating the water quality in the recipient water bodies.

Table IV. Performance of model 6) at Sampling Sites.

		(1) C	(1) L	(4) C	(4) L	(3,5,6) C	(3,5,6) L	(2) C	(2) L
R	T1	0.84	0.96	0.62	0.98	0.63	0.97	0.76	0.96
R	T2	0.84	0.95	0.61	0.99	0.56	0.99	0.63	0.99
R	T3	0.87	0.99	0.30	0.91	0.72	0.95		
R	T4	0.72	0.98	0.47	0.69	0.31	0.94	0.87	0.99
Bias	T1	0.04	0.07	0.00	0.00	0.00	0.05	0.00	0.00
Bias	T2	0.05	0.07	0.01	0.03	0.02	0.03	0.01	0.13
Bias	T3	-0.05	0.04	-0.12	-0.15	-0.08	0.10		
Bias	T4	0.07	0.10	0.14	0.09	0.08	-0.11	-0.01	-0.08
RMSE/Average	T1	0.29	0.53	0.29	0.28	0.41	0.61	0.34	0.55
RMSE/Average	T2	0.28	0.61	0.27	0.18	0.39	0.37	0.39	0.75
RMSE/Average	T3	0.31	0.35	0.35	0.53	0.46	0.63		
RMSE/Average	T4	0.29	0.31	0.30	0.34	0.48	1.01	0.29	0.38

C: concentration; L: load; T1:98-08; T2:98-05;T3:06-07;T4:08

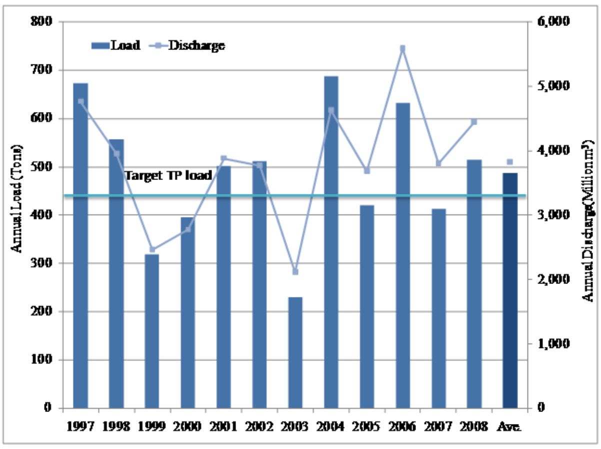

Figure 2. Saginaw River's Annual TP Load Estimates (Model 6).

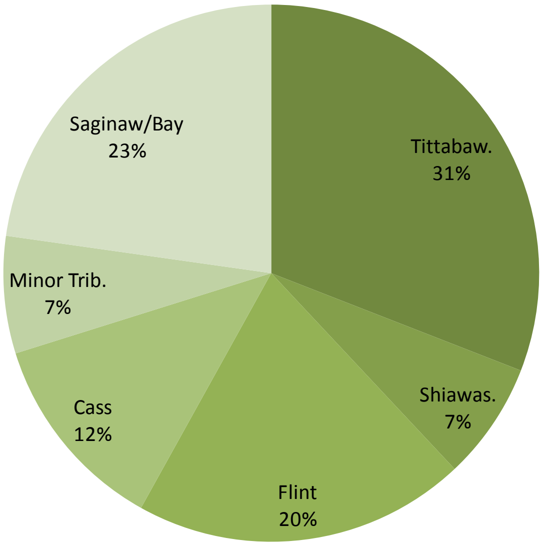

Figure 3. Average contribution of sub-watersheds to the total load of Saginaw River in 1997-2008 (model 6).

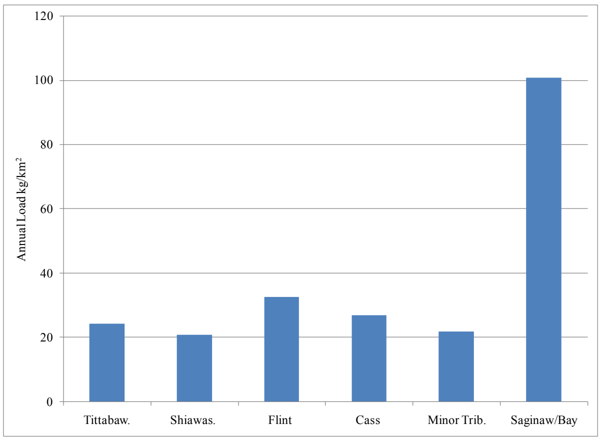

Figure 4. Average annual load per area of sub-watersheds (model 6).

REFERENCES

[1] Littlewood, I.G., Watts, C.D. and Custance, J.M. Systematic application of United Kingdom river flow and quality databases for estimating annual river mass loads (1975-1994). Science of The Total Environment (1998), pp. 210-211, 21-40.

[2] Phillips, J.M., Webb, B.W., Walling, D.E. and Leeks, G.J.L. Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. Hydrological Processes 13(7), (1999), pp.1035-1050.

[3] Quilbé, R., Rousseau, A.N., Duchemin, M., Poulin, A., Gangbazo, G. and Villeneuve, J.P. Selecting a calculation method to estimate sediment and nutrient loads in streams: Application to the Beaurivage River (Québec, Canada). Journal of Hydrology 326(1-4), (2006), pp. 295-310.

[4] Johnes, P.J. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. Journal of Hydrology 332(1-2), (2007), pp.241-258.

[5] Salles C., Tournoud, M.G. and Chu, Y. Estimating nutrient and sediment flood loads in a small Mediterranean river. Hydrological Processes 22(2), (2008), pp. 242-253.

[6] Dolan, D.M., Yui, A.K. and Geist, R.D. Evaluation of river load estimation methods for total phosphorus. Journal of Great Lakes Research 7(3), (1981), pp.207-214.

[7] Ferguson, R.I. River loads underestimated by rating curves. Water Resources Research, 22(1), (1986), pp.74–76.

[8] Cohn, T.A., Caulder, D.L., Gilroy, E.J., Zynjuk, L.D. and Summers, R.M. The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. Water Resources Research 28(9), (1992), pp. 2353–2363.

[9] Neitsch, S.L., J. G. Arnold, J. R. Kiniry, and J. R. Williams. 2005. *Soil and Water Assessment Tool. Theoretical documentation. Version 2005*. Temple (TX): Agricultural Research Service.

[10] Young, R. A., C. A. Onstad, D. D. Bosch, and W. P. Anderson, 1989. AGNPS: a non-point-source pollution model for evaluating agricultural watersheds. *Journal of Soil and Water Conservation*, 44(2):168-173.

[11] Croley T.E. II and He C., "Watershed surface and subsurface spatial intraflows", *Journal of Hydrologic Engineering*, Vol. 11, No. 1, (2006), pp 12-20.

[12] DeMarchi, C., T.S. Croley II, C. He, and T.S. Hunter, 2009. Application of a distributed watershed hydrology and water quality model in the Great Lakes basin. In *Proceedings of the 7th International Symposium on Ecohydraulics 2009*, Concepcion, Chile, January 12-16, 2009.

[13] He, C. and C. DeMarchi, 2010. Modeling Spatial Distributions of Point and Nonpoint Source Pollution Loadings in the Great Lakes Watersheds. *International Journal of Environmental Science and Engineering,* 2(1): 24-30.