USGS
science for a changing world

Prepared in cooperation with the Great Lakes Restoration Initiative

Estimated Total Phosphorus Loads for Selected Sites on Great Lakes Tributaries, Water Years 2014–2018

U.S. Department of the Interior
U.S. Geological Survey

Cover image. The Great Lakes of North America - Superior, Michigan, Huron, Ontario, and Erie, a true-color Terra MODIS image from the NASA Visible Earth collection, acquired on April 8, 2005. Visible Earth is part of the EOS Project Science Office at NASA Goddard Space Flight Center.

Estimated Total Phosphorus Loads for Selected Sites on Great Lakes Tributaries, Water Years 2014–2018

By G.F. Koltun

Prepared in cooperation with the Great Lakes Restoration Initiative

Open-File Report 2020–1145

U.S. Department of the Interior
U.S. Geological Survey

U.S. Geological Survey, Reston, Virginia: 2021

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit https://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit https://store.usgs.gov/.

Suggested citation:
Koltun, G.F., 2021, Estimated total phosphorus loads for selected sites on Great Lakes tributaries, water years 2014–2018: U.S. Geological Survey Open-File Report 2020–1145, 13 p., https://doi.org/10.3133/ofr20201145

Koltun, G.F., 2021, Model Archive—Estimated Total Phosphorus Loads for Selected Sites on Great Lakes Tributaries, Water Years 2014–2018: U.S. Geological Survey data release, https://doi.org/10.5066/P9WEW32M.

ISSN 2331-1258 (online)

Contents

Abstract 1
Introduction 1
Study Methods 2
Regression Equations and Estimated Total Phosphorus Loads 5
Summary 12
Acknowledgments 13
References Cited 13

Figures

1. U.S. Geological Survey streamgage sites on Great Lakes tributaries for which total phosphorus loads were estimated 4
2. Estimated cumulative total phosphorus loads for water years 2014–2018 11
3. Estimated flow-weighted mean concentrations of total phosphorus for water years 2014–2018 12

Tables

1. Selected characteristics of U.S. Geological Survey streamgage sites on Great Lakes tributaries for which total phosphorus loads were estimated 3
2. Model parameter estimates and summary statistics of surrogate-model regression equations for instantaneous total phosphorus flux 6
3. Model parameter estimates and summary statistics of UV-flow-model regression equations for instantaneous total phosphorus flux 8
4. Estimated cumulative and average annual total phosphorus loads, average streamflows, flow-weighted mean concentrations, average annual yields, and cumulative flow volumes at selected streamgage sites 10

Appendix Tables

1.1. Estimated annual total phosphorus loads and flow-weighted mean concentrations at selected U.S. Geological Survey gage sites on Great Lakes tributaries.
1.2. Estimated monthly total phosphorus loads at selected U.S. Geological Survey gage sites on Great Lakes tributaries.

(Available for download at https://doi.org/10.3133/ofr20201145.)

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
	Length	
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	2.590	square kilometer (km^2)
	Volume	
cubic foot (ft^3)	0.02832	cubic meter (m^3)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
cubic foot per second per square mile ($[ft^3/s]/mi^2$)	0.01093	cubic meter per second per square kilometer ($[m^3/s]/km^2$)
	Mass	
pound, avoirdupois (lb)	0.4536	kilogram (kg)
ton, short (2,000 lb)	0.9072	metric ton (t)

International System of Units to U.S. customary units

Multiply	By	To obtain
	Length	
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
	Area	
square kilometer (km^2)	0.3861	square mile (mi^2)
	Volume	
liter (L)	0.2642	gallon (gal)
cubic meter (m^3)	264.2	gallon (gal)
cubic meter (m^3)	35.31	cubic foot (ft^3)
	Flow rate	
cubic meter per second (m^3/s)	35.31	cubic foot per second (ft^3/s)
cubic meter per second per square kilometer ($[m^3/s]/km^2$)	91.49	cubic foot per second per square mile ($[ft^3/s]/mi^2$)
cubic meter per second (m^3/s)	22.83	million gallons per day (Mgal/d)
cubic meter per day per square kilometer ($[m^3/d]/km^2$)	0.0006844	million gallons per day per square mile ($[Mgal/d]/mi^2$)
	Mass	
kilogram (kg)	2.205	pound avoirdupois (lb)
metric ton (t)	1.102	ton, short [2,000 lb]

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F = (1.8 \times °C) + 32.$$

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C = (°F - 32) / 1.8.$$

Abbreviations

FWMC	flow-weighted mean concentration
GLRI	Great Lakes Restoration Initiative
TP	total phosphorus
USGS	U.S. Geological Survey
UV	unit values

Estimated Total Phosphorus Loads for Selected Sites on Great Lakes Tributaries, Water Years 2014–2018

By G.F. Koltun

Abstract

Monthly and annual total phosphorus loads were estimated for water years 2014 through 2018 for 23 streamgaged (gaged) sites on tributaries to the Great Lakes. Processing and regression methods described by Robertson and others (2018) were used with discrete and continuous data collected during water years 2011 and 2018 to update regression models for estimating instantaneous flux with the same form of equations as published by Robertson and others (2018). Monthly and water year average fluxes for all but two of the 23 gage sites were estimated using a weighted combination of results from surrogate models (which have streamflow, turbidity, and seasonal indicators as explanatory variables) and unit-value (UV)-flow models which have only UV streamflow and seasonal indicators as explanatory variables. Two of the gage sites had extensive periods of missing turbidity records, so average flux estimates for those stations were based solely on results from UV-flow models.

For most sites, estimated loads of total phosphorus were computed and summed for water years 2014–2018. The cumulative loads were used to compute yields and flow-weighted mean concentrations for water years 2014–2018. The estimated cumulative total phosphorus loads for water years 2014–2018 ranged from 112 to 11,500 metric tons. The Maumee River site (U.S. Geological Survey gage number 04193500) had the largest estimated cumulative load for water years 2014–2018 and the third largest estimated flow-weighted mean concentration. In fact, the estimated cumulative load at the Maumee River site was more than three times larger than the second largest estimated cumulative load.

Estimated average annual total phosphorus yields and flow-weighted mean concentrations for water years 2014–2018 ranged from 0.016 metric tons per square kilometer to 0.771 metric tons per square kilometer and 0.033 milligram per liter to 0.466 milligram per liter, respectively. The Cattaraugus Creek gage site (U.S. Geological Survey gage number 04213500) had the highest estimated average annual total phosphorus yield and flow-weighted mean concentration. The average annual total phosphorus yield at the Cattaraugus Creek gage site was almost twice as large as the second largest estimated yield.

Introduction

Concern for environmental issues affecting the Great Lakes—such as eutrophication and harmful algal blooms—and a desire to engage in proactive restoration led to the authorization (Title 33 U.S. Code § 1268) and 2010 publication of the first Great Lakes Restoration Initiative (GLRI) plan (GLRI, 2010). In response to the GLRI, the U.S. Geological Survey (USGS) began a program to monitor suspended sediment and nutrients at 30 streamgage (gage) locations, representing approximately 46 percent of the U.S. drainage to the Great Lakes (Robertson and others, 2018). The monitoring program included periodic collection of discrete stream-water samples that were analyzed for suspended sediment and nutrients, and continuous measurement of streamflow and selected water-quality parameters (turbidity, water temperature, specific conductance, pH, and dissolved oxygen). Robertson and others (2018) described the data collection methods in detail.

Robertson and others (2018) used data from the USGS GLRI monitoring program to estimate loads of suspended sediment and selected nutrients for water years 2011–2013 (a water year is defined as the 12-month period beginning October 1 and continuing through September 30 of the following year and is designated by the calendar year in which it ends). Linear regression techniques were used to develop two forms of regression models, which in turn were used to estimate loads. Regression models, referred to by Robertson and others (2018) as "UV-flow models," incorporated as explanatory variables unit-value (UV) streamflow data and time-based indicators of season. A second form of regression model, referred to as a "surrogate model," incorporated UV streamflow data and indicators of season but also included measures of one or more continuous water-quality parameters that covary in a systematic way with suspended sediment and (or) nutrient concentrations. Along with streamflow, the continuous water-quality measures function as surrogates for suspended sediment and (or) nutrient concentrations in regression models.

Since Robertson and others (2018) estimated loads of suspended sediment and selected nutrients for water years 2011–2013, five additional water years of data (2014–2018) have been collected for many of the original 30 monitoring locations. This report describes the results of a study that uses data collected during water years 2011–2018 to develop updated regression models for total phosphorus (TP) flux at monitoring locations where sufficient data exists and then use those models to estimate monthly and annual loads for water years 2014–2018. Loads of constituents other than TP may be computed in the future; however, TP was chosen to be addressed first because of its close association with ongoing issues of eutrophication and harmful algal blooms in the Great Lakes (Great Lakes Commission, 2020).

Study Methods

Methods described by Robertson and others (2018) were used to develop regression models for estimating instantaneous TP flux at 23 of the 30 sites discussed in Robertson and others (2018); table 1, fig. 1). Regression models were not developed for 7 of the original 30 sites because the data required to update the surrogate models and estimate annual loads were not collected because of changes in the GLRI monitoring network. For consistency, the same explanatory variables used by Robertson and others (2018) for estimating TP flux were used in this analysis; however, an additional 5 years of data (from water years 2014 through 2018) were used for model calibration. All data used for model calibration were retrieved from the USGS National Water Information System: Web Interface (NWISWeb; U.S. Geological Survey, 2016). The data were processed using nearly the same R-language (R Core Team, 2018) code and libraries used by Robertson and others (2018). Minor changes were made to the code to facilitate more automated processing and to improve computational speed through use of parallel-processing techniques. In addition, the version of R used at the time of analysis (version 3.6.1) required the use of more recent versions of some libraries. R source code and selected model outputs are available as a USGS data release (Koltun, 2020).

The data processing steps are discussed in detail by Robertson and others (2018) and are briefly outlined as follows:

1. Retrieve discrete TP data, unit- (subdaily) and daily value time series of streamflow, and UV time series of turbidity from NWISWeb.
2. Merge discrete and time-series data by date/time into a single R data frame.
3. Fill in days without UV streamflow data with daily value streamflow data and use time-based interpolation to fill in any partial-day gaps.
4. Resample or interpolate all UV data to a 5-minute time interval.
5. Fill in missing periods of turbidity data less than 2 hours in duration by time-based linear interpolation.
6. Process streamflow data to convert negative flows to small positive flows while over time conserving the total volume of water transported downstream. This was done by replacing negative streamflow values with a value of 1 cubic foot per second and then using an iterative process to reduce bracketing positive flow values by an amount equal to the sum of the small positive values substituted for the intervening negative streamflow period plus the sum of the negative flow values that were replaced.
7. Develop regression models for the natural logarithm of instantaneous TP flux with rloadest (Lorenz and others, 2015) using (1) the natural log of streamflow and sine and cosine functions of date/time as explanatory variables (UV-flow model); and (2) the natural logarithms of streamflow and turbidity, and sine and cosine functions of date/time as explanatory variables (surrogate model). The sine and cosine functions of time serve as surrogates for seasonality and are computed using $2\pi T$ as the argument to the functions, where T is the date/time in decimal years.
8. Review regression diagnostics to assess statistical significance of the model and model parameters, model fit, and to look for evidence of heteroscedasticity, lack of normality, and (or) serial correlation of residuals.
9. Split the UV data into two mutually exclusive datasets. The first dataset contains UV data for days when both flow and turbidity measures were available for the entire day, and the second contains data for days when only flow measures were available for the entire day.
10. Use the datasets created in step 9 with the UV-flow and surrogate regression models developed with rloadest to estimate bias-corrected monthly and annual average fluxes and their associated 95 percent confidence intervals for water years 2014–2018.
11. Combine the results from the UV-flow and surrogate regression models by computing weighted averages of the monthly and annual average flux estimates where the weights were equal to the number of days in the month or year on which the average flux estimates were based. Weighted standard errors of estimate—which were used to compute confidence limits on the flux estimates—were also determined as described in Roberts and other (2018).

Two of the gages (station numbers 04085427 and 04165500) had extensive periods of missing turbidity records, so flux estimates for those two stations were based solely on the UV-flow models.

Table 1. Selected characteristics of U.S. Geological Survey streamgage sites on Great Lakes tributaries for which total phosphorus loads were estimated.

[USGS, U.S. Geological Survey; mi^2, square miles; km^2, square kilometers; MN, Minnesota; WI, Wisconsin; MI, Michigan; IN, Indiana; OH, Ohio; NY, New York; ND, not determined]

USGS station number	Station name	Decimal latitude	Decimal longitude	Drainage area (mi^2)	Drainage area (km^2)
04024000	St. Louis River at Scanlon, MN	46.7033	–92.4188	3,430	8,884
04027000	Bad River near Odanah, WI	46.4866	–90.6963	597	1,546
04040000	Ontonagon River near Rockland, MI	46.7208	–89.2071	1,340	3,471
04067500	Menominee River near Mc Allister, WI	45.3258	–87.6633	3,930	10,179
040851385	Fox River at Oil Tank Depot at Green Bay, WI	44.5286	–88.0100	6,330	16,395
04085427	Manitowoc River at Manitowoc, WI	44.1062	–87.7160	526	1,362
04087170	Milwaukee River at Mouth at Milwaukee, WI	43.0244	–87.8983	872	2,258
04092750	Indiana Harbor Canal at East Chicago, IN	41.6492	–87.4687	ND	ND
04101500	St. Joseph River at Niles, MI	41.8292	–86.2597	3,666	9,495
04108660	Kalamazoo River at New Richmond, MI	42.6509	–86.1067	1,950	5,050
04119400	Grand River near Eastmanville, MI	43.0242	–86.0264	5,290	13,701
04142000	Rifle River near Sterling, MI	44.0725	–84.0200	320	829
04157005	Saginaw River at Holland Avenue at Saginaw, MI	43.4220	–83.9519	6,060	15,695
04165500	Clinton River at Moravian Drive at Mt. Clemens, MI	42.5959	–82.9088	734	1,901
04166500	River Rouge at Detroit, MI	42.3731	–83.2547	187	484
04176500	River Raisin near Monroe, MI	41.9606	–83.5310	1,042	2,699
04193500	Maumee River at Waterville, OH	41.5001	–83.7127	6,330	16,395
04199500	Vermilion River near Vermilion, OH	41.3820	–82.3168	262	679
04200500	Black River at Elyria, OH	41.3803	–82.1046	396	1,026
04208000	Cuyahoga River at Independence, OH	41.3953	–81.6298	707	1,831
04213500	Cattaraugus Creek at Gowanda, NY	42.4633	–78.9342	436	1,129
04231600	Genesee River at Ford Street Bridge, Rochester, NY	43.1417	–77.6163	2,474	6,408
04249000	Oswego River at Lock 7, Oswego, NY	43.4517	–76.5053	5,100	13,209

Figure 1. U.S. Geological Survey streamgage sites on Great Lakes tributaries for which total phosphorus loads were estimated.

Regression Equations and Estimated Total Phosphorus Loads

The surrogate model equation for estimating the instantaneous flux of TP (F_{TP}) in units of kilograms per day has the following form:

$$F_{TP}=e^{(b_0+b_1(\ln(Q))+b_2(\sin(2\pi T))+b_3(\cos(2\pi T))+b_4(\ln(Turb)))} \quad (1)$$

and the UV-flow model equation has the following form:

$$F_{TP}=e^{(b_0+b_1(\ln(Q))+b_2(\sin(2\pi T))+b_3(\cos(2\pi T)))} \quad (2)$$

where:

Q is instantaneous streamflow in cubic feet per second,

T is time in decimal calendar years, and

Turb is the instantaneous turbidity in formazin nephelometric units.

The model forms were chosen to be consistent with those used by Robertson and others (2018). The model parameter estimates (b_0–b_4 in the surrogate model or b_0–b_3 in the UV-flow model) along with summary statistics associated with the models are listed in tables 2 and 3 for the surrogate and UV-flow models, respectively.

The regression equations with the parameter estimates listed in tables 2 and 3 were used within rloadest to estimate monthly and water-year average fluxes (in kilograms per day; appendixes 1 and 2, respectively). Loads were computed by multiplying average fluxes by the number of days in the applicable accounting period (month or water year). When the calibration dataset was entirely uncensored, a bias correction factor referred to by Likeš (1980) as phi was used to correct for retransformation bias. When using adjusted maximum likelihood estimation (Cohn, 1988; Cohn and others, 1992) with censored calibration datasets, first-order bias in the model coefficients was corrected for using methods described by Shenton and Bowman (1977). Even after bias correction methods are applied, regression models can produce biased estimates (Stenback and others, 2011). Consequently, the total of the loads predicted by the regression model for each site was compared to the total of the observed loads (both computed only for times when concentrations were measured). The ratio of the total of predicted loads to the total of observed loads expressed as a percentage (B_p) was computed and reported as an indicator of the residual potential bias in the regression-based loading estimates.

Because of the complexity of computing and applying the bias correction factors described by Likeš (1980) and Shenton and Bowman (1977) outside of rloadest, a third bias correction factor (frequently referred to as a “smearing estimator”) was computed using methods described by Duan (1983) and reported for potential future use. The smearing estimator has been determined to perform nearly as well as some unbiased estimators (Helsel and Hirsch, 2002) and it only requires the assumption that the regression residuals are independent and homoscedastic.

Estimated monthly and water-year loads of TP and their 95 percent confidence limits were computed for water years 2014–2018 (appendices 1 and 2). With the exceptions of the Clinton River (04165500) and Manitowoc River (04085427) gage sites, all load estimates were based on a weighted combination of surrogate model (table 2) and UV-flow model (table 3) estimates. The load estimates for the Clinton and Manitowoc Rivers were based solely on results from the UV-flow models because there were lengthy periods of missing turbidity data. In addition, load estimates were not computed for water year 2017 for the Indiana Harbor Canal site (04092750) because the streamflow record was incomplete for that year. Estimated loads of TP were summed for water years 2014–2018 and used to compute average annual yields (mass per unit area) and flow-weighted mean concentrations (FWMCs; mass per total flow volume).

Table 2. Model parameter estimates and summary statistics of surrogate-model regression equations for instantaneous total phosphorus flux.

[USGS, U.S. Geological Survey; b_0—b_4 are coefficients in the equation 1 for regressor variables listed in parentheses; PRESS, prediction residual error sum of squares; R^2, coefficient of multiple determination in percent; SEP_{ap}, approximate standard error of prediction; N, number of observations; B_p, potential load bias in percent; BCF, bias correction factor (smearing estimator); MN, Minnesota; WI, Wisconsin; MI, Michigan; IN, Indiana; OH, Ohio; NY, New York; ND, not determined; bolded coefficient values were not different from zero at alpha=0.05]

USGS station number	Station name	b_0 (intercept)	b_1 (ln(Q))	b_2 (sin(2πT))	b_3 (cos(2πT))	b_4 (Turb)	Residual variance	PRESS	R^2	N	N uncensored	SEP_{ap}	B_p	BCF	Period of record
04024000	St. Louis River at Scanlon, MN	−3.4280	1.0382	0.0082	−0.1628	0.4903	0.0565	8.88	97.74	152	152	0.246	1.03	1.033	2011-05-01 to 2018-09-10
04027000	Bad River near Odanah, WI	−4.5231	1.0347	**−0.0859**	**0.0106**	0.7630	0.2196	ND	96.38	157	153	0.489	−2.95	1.183	2011-03-20 to 2018-09-04
04040000	Ontonagon River near Rockland, MI	−5.3384	1.0680	**−0.0597**	**−0.0034**	0.7837	0.0568	3.22	96.70	39	39	0.308	−9.17	1.026	2011-07-14 to 2018-09-10
04067500	Menominee River near Mc Allister, WI	−1.7452	0.8460	**0.0718**	-0.1787	0.4835	0.1388	18.63	85.80	129	129	0.388	−4.09	1.083	2011-04-28 to 2018-09-11
040851385	Fox River at Oil Tank Depot at Green Bay, WI	−2.5446	1.0134	**−0.1514**	−0.1747	0.4155	0.0692	12.54	94.27	176	176	0.270	−1.82	1.035	2011-03-28 to 2018-09-12
04085427	Manitowoc River at Manitowoc, WI	−2.3275	1.0430	**−0.0751**	−0.1664	0.4459	0.1445	ND	94.89	153	153	0.397	8.73	1.068	2011-03-19 to 2018-07-31
04087170	Milwaukee River at Mouth at Milwaukee, WI	−2.1438	0.9861	**−0.0138**	0.1700	0.3606	0.1726	20.25	87.08	112	112	0.433	0.16	1.090	2011-03-24 to 2018-09-11
04092750	Indiana Harbor Canal at East Chicago, IN	−2.2783	0.9832	−0.2463	−0.1965	0.4557	0.0883	6.13	89.77	62	62	0.328	−0.28	1.044	2011-05-12 to 2018-09-04
04101500	St. Joseph River at Niles, MI	−2.2262	0.9559	−0.1880	−0.1158	0.4988	0.1085	13.06	91.59	112	112	0.349	−12.86	1.058	2011-04-07 to 2018-09-11
04108660	Kalamazoo River at New Richmond, MI	−4.9861	1.3080	-0.1209	−0.3192	0.3091	0.0521	6.74	89.89	123	123	0.239	−1.75	1.027	2011-04-18 to 2018-09-13
04119400	Grand River near Eastmanville, MI	−3.7211	1.1800	−0.1378	−0.1977	0.2923	0.0538	6.75	94.47	119	119	0.243	−0.43	1.027	2011-04-18 to 2018-09-12
04142000	Rifle River near Sterling, MI	−4.2126	1.0073	0.1247	**0.0461**	0.7473	0.0891	10.99	96.60	118	118	0.312	3.61	1.049	2011-07-12 to 2018-09-05
04157005	Saginaw River at Holland Avenue at Saginaw, MI	−2.9090	0.9675	**−0.0214**	**0.0096**	0.5729	0.0918	10.39	95.71	108	108	0.318	-0.86	1.048	2011-11-29 to 2018-09-04
04165500	Clinton River at Moravian Drive at Mt. Clemens, MI	−2.0691	0.9331	−0.1305	**−0.0528**	0.5225	0.1838	30.80	91.26	159	159	0.447	−7.90	1.117	2011-11-16 to 2018-06-05

Table 2. Model parameter estimates and summary statistics of surrogate-model regression equations for instantaneous total phosphorus flux.—Continued

[USGS, U.S. Geological Survey; b_0—b_4 are coefficients in the equation 1 for regressor variables listed in parentheses; PRESS, prediction residual error sum of squares; R^2, coefficient of multiple determination in percent; SEP_{ap}, approximate standard error of prediction; N, number of observations; B_p, potential load bias in percent; BCF, bias correction factor (smearing estimator); MN, Minnesota; WI, Wisconsin; MI, Michigan; IN, Indiana; OH, Ohio; NY, New York; ND, not determined; bolded coefficient values were not different from zero at alpha=0.05]

USGS station number	Station name	b_0 (intercept)	b_1 (ln(Q))	b_2 (sin(2πT))	b_3 (cos(2πT))	b_4 (Turb)	Residual variance	PRESS	R^2	N	N uncen-sored	SEP_{ap}	B_p	BCF	Period of record
04166500	River Rouge at Detroit, MI	−3.3125	0.8853	**−0.0077**	**−0.0866**	0.7772	0.1295	28.09	95.36	211	211	0.369	1.46	1.077	2011-05-16 to 2018-09-04
04176500	River Raisin near Monroe, MI	−3.5277	1.1595	**−0.1296**	**−0.0306**	0.4050	0.2501	ND	92.98	141	140	0.522	10.31	1.120	2011-05-25 to 2018-07-18
04193500	Maumee River at Waterville, OH	−3.3384	1.1113	−0.1304	0.1009	0.4442	0.0401	6.78	98.98	163	163	0.207	1.48	1.020	2011-04-19 to 2018-09-26
04199500	Vermilion River near Vermilion, OH	−3.4551	1.1150	**−0.0745**	−0.1547	0.5062	0.0777	9.75	99.04	120	120	0.291	1.04	1.038	2011-06-29 to 2018-09-04
04200500	Black River at Elyria, OH	−3.0509	1.0564	−0.1696	**−0.0714**	0.4833	0.0653	9.51	99.05	135	135	0.270	−4.20	1.028	2011-06-29 to 2018-09-04
04208000	Cuyahoga River at Independence, OH	−1.6458	0.8863	**−0.0181**	**0.0052**	0.4560	0.1498	17.52	93.27	109	109	0.410	−14.46	1.077	2011-08-17 to 2018-09-05
04213500	Cattaraugus Creek at Gowanda, NY	−7.5490	1.4095	−0.2417	−0.2495	0.7838	0.3185	ND	97.07	346	345	0.589	6.92	1.392	2011-07-06 to 2018-09-19
04231600	Genesee River at Ford Street Bridge, Rochester, NY	−4.6376	1.1263	**−0.0201**	−0.1706	0.5947	0.1890	35.44	93.81	179	179	0.451	−9.57	1.109	2011-04-05 to 2018-09-11
04249000	Oswego River at Lock 7,Oswego, NY	−3.7904	1.1530	−0.2636	−0.2191	0.1753	0.1151	ND	88.69	121	120	0.354	0.10	1.045	2011-03-09 to 2018-08-07

Table 3. Model parameter estimates and summary statistics of UV-flow-model regression equations for instantaneous total phosphorus flux.

[USGS, U.S. Geological Survey; b_0—b_3 are coefficients in the equation 2 for regressor variables listed in parentheses; PRESS, prediction residual error sum of squares; R^2, coefficient of multiple determination in percent; SEP_{ap}, approximate standard error of prediction; N, number of observations; B_p, potential load bias in percent; BCF, bias correction factor (smearing estimator); MN, Minnesota; WI, Wisconsin; MI, Michigan; IN, Indiana; OH, Ohio; NY, New York; ND, not determined; bolded coefficient values were not significantly different from zero at alpha=0.05]

USGS station number	Station name	b_0 (intercept)	b_1 (ln(Q))	b_2 (sin(2πT))	b_3 (cos(2πT))	Residual variance	PRESS	R^2	N	N uncen-sored	SEP_{ap}	B_p	BCF	Period of record
04024000	St. Louis River at Scanlon, MN	−5.5017	1.4123	−0.0956	−0.1756	0.1028	17.48	95.71	166	166	0.328	−19.42	1.055	2011-05-01 to 2018-09-10
04027000	Bad River near Odanah, WI	−6.5875	1.7255	**0.0289**	−0.4058	0.4777	ND	92.44	187	183	0.708	3.21	1.032	2011-03-20 to 2018-09-04
04040000	Ontonagon River near Rockland, MI	−7.0963	1.7485	**0.0156**	−0.1561	0.2606	22.55	89.84	83	83	0.534	20.35	1.219	2011-07-14 to 2018-09-10
04067500	Menominee River near Mc Allister, WI	−4.2246	1.1899	**0.0363**	−0.3076	0.1883	29.63	80.35	154	154	0.444	−3.82	1.123	2011-03-15 to 2018-09-11
040851385	Fox River at Oil Tank Depot at Green Bay, WI	−1.3092	1.0118	−0.3853	−0.4536	0.1641	38.34	98.84	230	230	0.412	−9.52	1.091	2011-03-17 to 2018-09-12
04085427	Manitowoc River at Manitowoc, WI	−2.4664	1.3015	−0.2898	−0.4556	0.2651	53.77	90.17	199	199	0.525	7.08	1.145	2011-03-19 to 2018-09-11
04087170	Milwaukee River at Mouth at Milwaukee, WI	−1.5928	1.0128	0.1003	−0.1175	0.2104	72.93	98.61	342	342	0.465	−2.55	1.110	2011-03-24 to 2018-09-11
04092750	Indiana Harbor Canal at East Chicago, IN	−1.4025	1.0087	−0.1679	−0.1371	0.1638	23.92	98.65	142	142	0.416	−6.19	1.083	2011-03-31 to 2018-09-04
04101500	St. Joseph River at Niles, MI	−8.0945	1.7619	**−0.1068**	−0.2234	0.2257	31.78	82.05	137	137	0.489	−12.84	1.124	2011-04-07 to 2018-09-11
04108660	Kalamazoo River at New Richmond, MI	−4.5250	1.3393	−0.0968	−0.4581	0.0795	14.16	86.11	174	174	0.289	0.23	1.041	2011-04-18 to 2018-09-13
04119400	Grand River near Eastmanville, MI	−4.3563	1.3311	−0.1487	−0.3466	0.0849	12.74	91.93	146	146	0.300	−0.70	1.045	2011-04-18 to 2018-09-12
04142000	Rifle River near Sterling, MI	−8.5946	2.0437	**0.0584**	**−0.0261**	0.2650	41.38	91.89	153	153	0.527	7.86	1.154	2011-07-12 to 2018-09-05
04157005	Saginaw River at Holland Avenue at Saginaw, MI	−2.0446	1.0565	0.2200	−0.1447	0.2695	47.32	96.05	171	171	0.532	−11.50	1.149	2011-11-29 to 2018-09-04
04165500	Clinton River at Moravian Drive at Mt. Clemens, MI	−3.6492	1.4270	−0.2537	−0.2676	0.2922	61.35	85.72	210	210	0.546	−0.85	1.185	2011-07-19 to 2018-09-04

Table 3. Model parameter estimates and summary statistics of UV-flow-model regression equations for instantaneous total phosphorus flux.—Continued

[USGS, U.S. Geological Survey; b_0—b_3 are coefficients in the equation 2 for regressor variables listed in parentheses; PRESS, prediction residual error sum of squares; R^2, coefficient of multiple determination in percent; SEP_{ap}, approximate standard error of prediction; N, number of observations; B_p, potential load bias in percent; BCF, bias correction factor (smearing estimator); MN, Minnesota; WI, Wisconsin; MI, Michigan; IN, Indiana; OH, Ohio; NY, New York; ND, not determined; bolded coefficient values were not significantly different from zero at alpha=0.05]

USGS station number	Station name	b_0 (intercept)	b_1 (ln(Q))	b_2 (sin(2πT))	b_3 (cos(2πT))	Residual variance	PRESS	R^2	N	N uncen-sored	SEP_{ap}	B_p	BCF	Period of record
04166500	River Rouge at Detroit, MI	−3.4075	1.4072	−0.3438	−0.7036	0.3395	78.52	87.76	227	227	0.593	40.38	1.227	2011-05-16 to 2018-09-04
04176500	River Raisin near Monroe, MI	−4.6362	1.5267	−0.1777	−0.1886	0.3262	ND	90.98	172	171	0.585	12.68	1.162	2011-05-25 to 2018-09-05
04193500	Maumee River at Waterville, OH	−3.8737	1.3917	−0.1522	0.0736	0.1242	24.35	96.81	191	191	0.361	−3.30	1.062	2011-03-08 to 2018-09-26
04199500	Vermilion River near Vermilion, OH	−3.9169	1.5443	−0.2050	−0.3477	0.2761	41.00	96.71	144	144	0.541	36.44	1.142	2011-06-29 to 2018-09-04
04200500	Black River at Elyria, OH	−2.9360	1.3607	−0.1635	−0.2113	0.1381	21.63	98.00	152	152	0.382	5.05	1.070	2011-06-29 to 2018-09-04
04208000	Cuyahoga River at Independence, OH	−4.3691	1.5104	−0.1690	−0.2377	0.2989	47.49	87.77	154	154	0.563	−7.19	1.162	2011-04-0 to 2018-09-05
04213500	Cattaraugus Creek at Gowanda, NY	−11.9827	2.5249	−0.3289	−0.3080	0.5606	ND	94.92	353	352	0.767	64.94	1.814	2011-07-06 to 2018-09-19
04231600	Genesee River at Ford Street Bridge, Rochester, NY	−7.9542	1.8332	−0.3075	−0.3537	0.3776	74.10	87.96	191	191	0.629	−16.57	1.231	2011-04-05 to 2018-09-11
04249000	Oswego River at Lock 7, Oswego, NY	−3.8974	1.1950	−0.2161	−0.2578	0.1202	ND	87.71	154	153	0.355	−1.74	1.051	2011-03-09 to 2018-09-05

The estimated cumulative TP loads for water years 2014–2018 ranged from 112 metric tons (t) at the Rifle River gage near Sterling, Michigan (04142000), to about 11,500 t at the Maumee River gage at Waterville, Ohio (04193500) (table 4, fig. 2). Even though the Maumee River gage has a drainage area that is equal to the Fox River gage (040851385), the estimated water year 2014–2018 cumulative load at the Maumee River gage was more than three times larger and also was more than five times larger than the third largest estimated cumulative load (at the Genesee River gage at Rochester, New York [04231600]). The water year 2014–2018 volume of flow at the Maumee River gage was only about 12 percent larger than at the Fox River gage, so neither drainage-area size nor flow volume account for the large difference in cumulative load.

Loads vary as a function of both flow and concentration. Large drainage basins tend to contribute larger loads than small drainage basins because large drainage basins also tend to contribute larger volumes of water. Because of the drainage-area-size effects on loads, yields and FWMCs frequently are used when comparing nutrient contributions from drainage basins of varying size. Similarly, because of flow-related effects on loads, FWMCs frequently are used to help account for year-to-year variations in flow when assessing trends in nutrient delivery over time.

Estimated average annual TP yields for water years 2014–2018 ranged from 0.016 metric tons per square kilometer (t/km^2) at the St. Louis River gage at Scanlon, Minnesota (04024000), to 0.771 t/km^2 at the Cattaraugus Creek gage at Gowanda, New York (04213500; table 4). TP yields were not

Table 4. Estimated cumulative and average annual total phosphorus loads, average streamflows, flow-weighted mean concentrations, average annual yields, and cumulative flow volumes at selected streamgage sites on Great Lakes tributaries, water years 2014–2018.

[USGS, U.S. Geological Survey; cfs, cubic feet per second; mg/L, milligrams per liter; km^2, square kilometers; m^3, cubic meters; ND, not determined]

USGS station number	Estimated cumulative load (metric tons)	Estimated average annual load (metric tons)	Average streamflow (cfs)	Estimated flow-weighted mean concentration (mg/L)	Average annual yield (metric tons/km^2)	Total volume of flow ($m^3 \times 10^{-6}$)
04024000	707	141	2,810	0.056	0.016	12,600
04027000	1,030	206	852	0.270	0.344	3,810
04040000	1,060	212	1,480	0.160	0.158	6,630
04067500	585	117	3,930	0.033	0.030	17,600
040851385	3,400	680	5,840	0.130	0.107	26,088
04085427	631	126	446	0.317	0.240	1,990
04087170	427	85.4	899	0.106	0.098	4,015
04092750[a]	183	45.8	494	0.104	ND	1,767
04101500	1,510	302	3,920	0.086	0.082	17,500
04108660	726	145	2,360	0.069	0.074	10,500
04119400	2,140	428	5,240	0.092	0.081	23,400
04142000	112	22.4	364	0.069	0.070	1,630
04157005	2,260	452	4,830	0.105	0.075	21,600
04165500	705	141	667	0.237	0.192	2,980
04166500	113	22.6	163	0.155	0.121	729
04176500	780	156	802	0.218	0.150	3,580
04193500	11,500	2,300	6,530	0.393	0.362	29,200
04199500	509	102	271	0.420	0.389	1,210
04200500	609	122	415	0.328	0.307	1,860
04208000	1,230	246	1,110	0.249	0.349	4,960
04213500	1,680	336	807	0.466	0.771	3,600
04231600	2,400	480	2,880	0.187	0.194	12,800
04249000	1,730	346	7,820	0.050	0.068	34,900

[a]Values reported for this station are based only on data for water years 2014–2016 and 2018

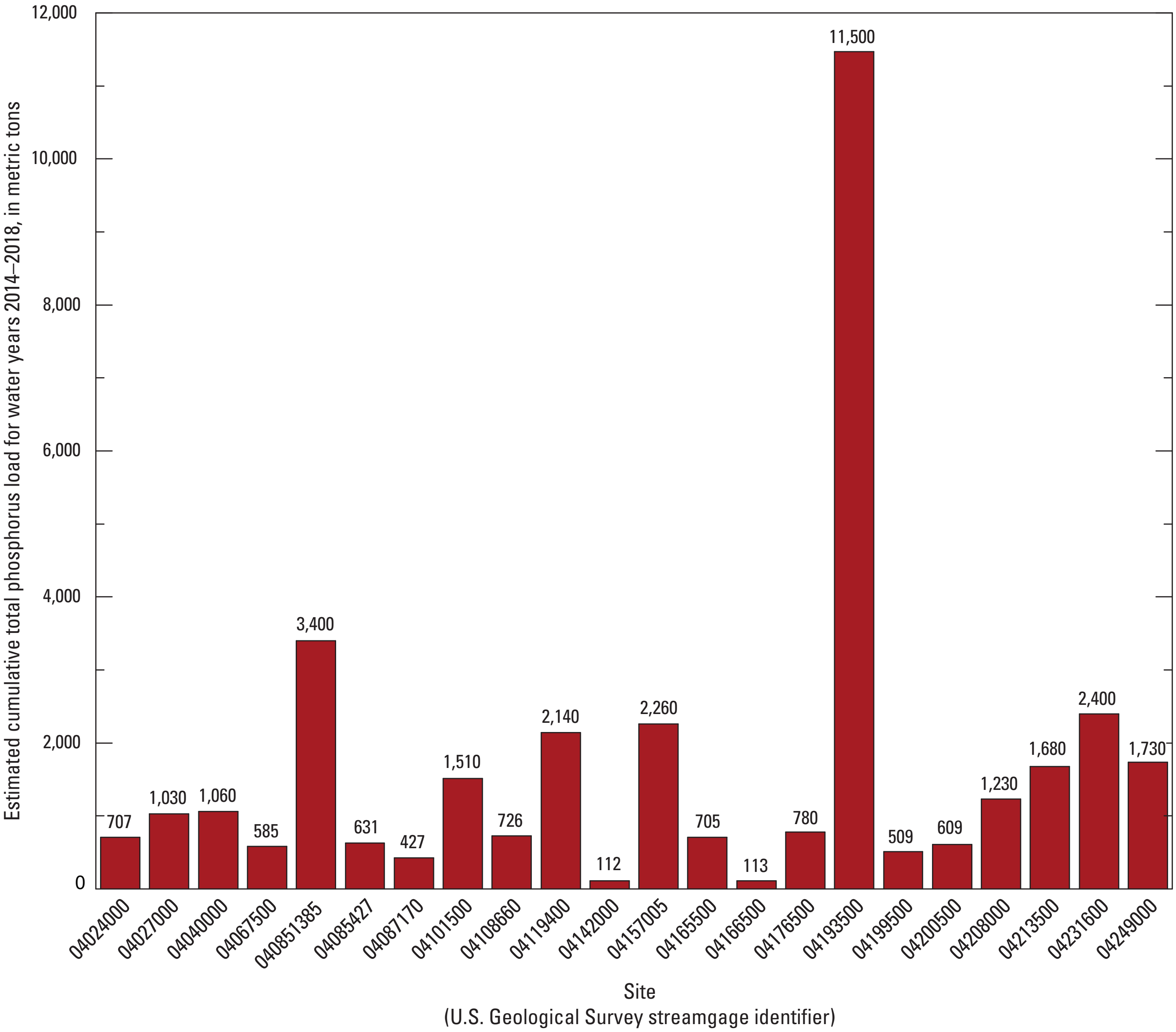

Figure 2. Estimated cumulative total phosphorus loads for water years 2014–2018 at selected U.S. Geological Survey streamgage sites on Great Lakes tributaries.

determined for the Indiana Harbor Canal at East Chicago, Indiana (04092750), because the drainage area is indeterminate. The estimated yield at the Cattaraugus Creek gage was almost twice as large as the second largest yield (at the Vermilion River gage near Vermilion, Ohio [04199500]). In spite of its large estimated yield, six other gage sites (04119400, 040851385, 04157005, 04193500, 04231600, and 04249000) had larger estimated cumulative loads for water years 2014–2018 than Cattaraugus Creek, due in part to the relatively small volume of water passing the Cattaraugus Creek gage site.

Estimated FWMCs of TP for water years 2014–2018 ranged from 0.033 milligrams per liter (mg/L) at the Menominee River gage near McAllister, Wisconsin (04067500), to 0.466 mg/L at the Cattaraugus Creek gage (04213500; table 4, fig. 3). Estimated FWMCs of TP exceeded 0.4 mg/L at only one other gage site (Vermilion River gage [04199500]). In spite of the high FWMC of TP at the Vermilion River gage, its cumulative load for water years 2014–2018 is less than all but three of the other 21 gages (04092750, 04142000, and 04166500) due in part to its relatively small drainage area and volume of flow.

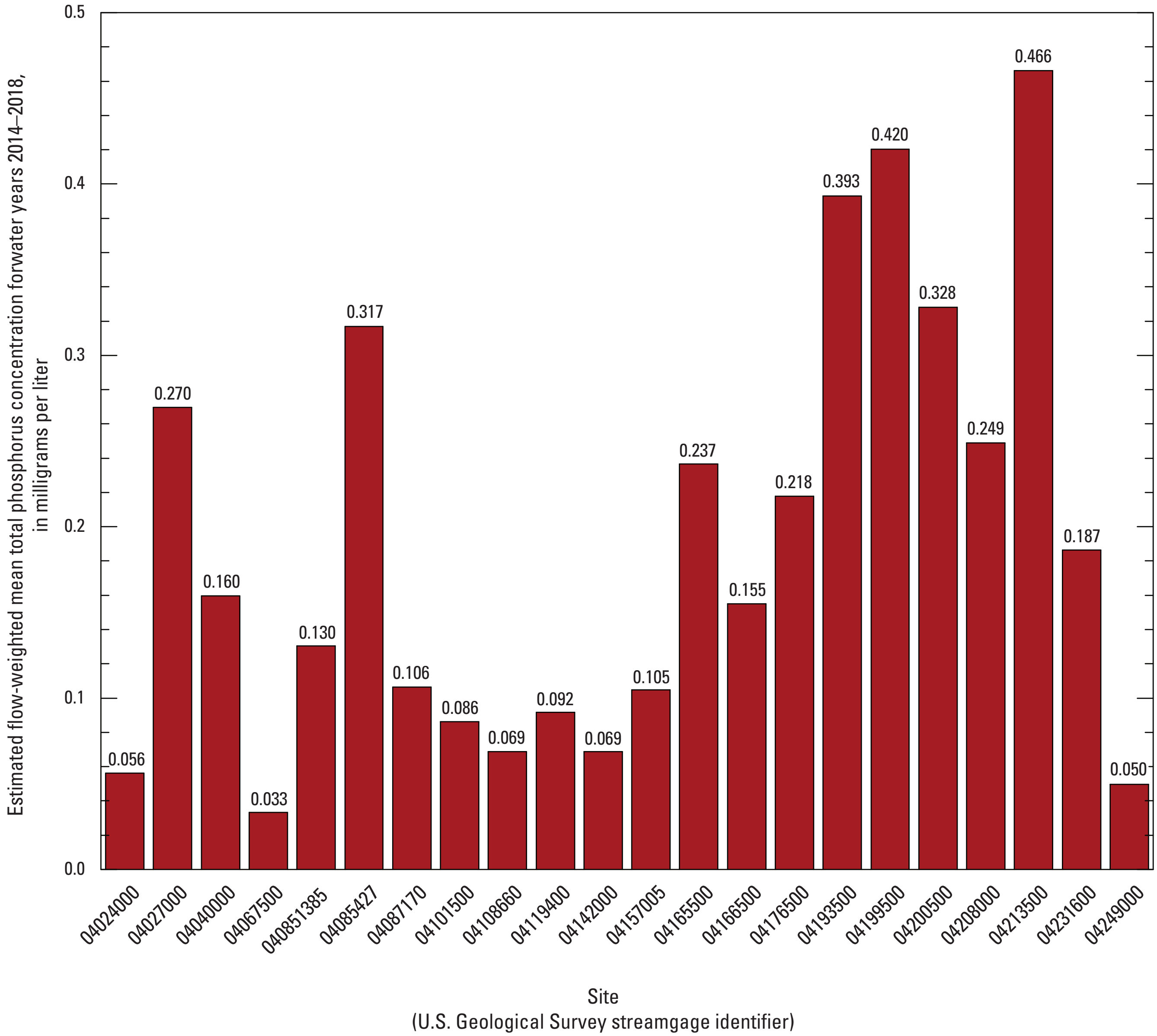

Figure 3. Estimated flow-weighted mean concentrations of total phosphorus for water years 2014–2018 at selected U.S. Geological Survey gage sites on Great Lakes tributaries.

Summary

Regression and processing methods described by Robertson and others (2018) were used with discrete total phosphorus concentration data and continuous water-quality and streamflow data collected during water years 2011 through 2018 to update regression models for estimating instantaneous total phosphorus flux at 23 streamgaged (gaged) sites on tributaries to the Great Lakes. Monthly and water-year average fluxes for all but 2 of the 23 gage sites were estimated using a weighted combination of results from surrogate models, which have streamflow, turbidity, and seasonal indicators as explanatory variables, and unit-value (UV)-flow models which have only UV streamflow and seasonal indicators as explanatory variables. Average fluxes for two sites were estimated with only UV-flow models because of extensive periods of missing turbidity data. Loads were computed by multiplying average fluxes by the number of days in the applicable accounting period (month or water year).

For all the gage sites (except the Indiana Harbor Canal that had missing streamflow data in 2017), estimated loads of total phosphorus were computed from flux estimates and summed for water years 2014–2018. The cumulative loads were used to compute average annual yields

and flow-weighted mean concentrations for water years 2014–2018. The estimated cumulative total phosphorus loads for water years 2014–2018 ranged from 112 to 11,500 metric tons. Estimated average annual total phosphorus yields and estimated flow-weighted mean concentrations for water years 2014–2018 ranged from 0.016 metric tons per square kilometer to 0.771 metric tons per square kilometer and 0.033 milligrams per liter to 0.466 milligrams per liter, respectively. The Maumee River gage at Waterville, Ohio (U.S. Geological Survey gage number 04193500), had the largest estimated cumulative load for water years 2014–2018, more than 3 times larger than the site with the second largest cumulative load. The Cattaraugus Creek gage at Gowanda, New York (U.S. Geological Survey gage number 04213500), had the highest estimated average annual total phosphorus yield and flow-weighted mean concentration.

Acknowledgments

The author wishes to thank Dale Robertson and Laura Hubbard (USGS, Upper Midwest Water Science Center), and David Lorenz (USGS, retired) for sharing their R code and providing information on the computational workflow that they used.

References Cited

Cohn, T.A., 1988, Adjusted maximum likelihood estimation of the moments of lognormal populations from type I censored samples: U.S. Geological Survey Open-File Report 88-350, 34 p., accessed January 2019, at https://pubs.usgs.gov/of/1988/0350/report.pdf.

Cohn, T.A., Gilroy, E.J., and Baier, W.G., 1992, Estimating fluvial transport of trace constituents using a regression model with data subject to censoring: Proceedings of the Joint Statistical Meeting, Boston, August 9–13, 1992, p. 142–151, accessed January 2019, at https://www.researchgate.net/publication/265100089_Estimating_Fluvial_Transport_of_Trace_Constituents_Using_a_Regression_Model_with_Data_Subject_to_Censoring/link/555bf18208ae91e75e76a855/download.

Duan, N., 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610, https://www.jstor.org/stable/2288126, accessed January 2019.

Great Lakes Commission, 2020, About Nutrients: Great Lakes Regional Sediment Management Team web page, accessed November 2020, at https://greatlakesrsm.net/about-nutrients/.

Great Lake Restoration Initiative, 2010, Great Lakes Restoration Initiative Action Plan I (FY2010-2014): Great Lakes Restoration Initiative Report, 41 p., accessed January 2019, at https://www.glri.us/sites/default/files/glri-action-plan-fy2010-fy2014-20100221-41pp.pdf.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 510 p., accessed January 2019 at http://pubs.er.usgs.gov/publication/twri04A3.

Koltun, G.F., 2021, Model archive—estimated total phosphorus loads for selected sites on Great Lakes tributaries, water years 2014–2018, U.S. Geological Survey data release, https://doi.org/10.5066/P9WEW32M.

Likeš, J., 1980, Variance of the MVUE for lognormal variance: Technometrics, v. 22, no. 2, p. 253–258, https://www.jstor.org/stable/1268465, accessed January 2019.

Lorenz, D., Runkel, R., and De Cicco, L., 2015, River load estimation, rloadest package: U.S. Geological Survey Software, accessed January 2019, at https://github.com/USGS-R/rloadest.

R Core Team, 2018, R: A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed January 2019 at http://www.R-project.org/.

Robertson, D.M., Hubbard, L.E., Lorenz, D.L., and Sullivan, D.J., 2018, A surrogate regression approach for computing continuous loads for the tributary nutrient and sediment monitoring program on the Great Lakes: Journal of Great Lakes Research, v. 44, no. 1, p. 26–42, accessed January 2019 https://doi.org/10.1016/j.jglr.2017.10.003.

Shenton, L.R., and Bowman, K.O., 1977, Maximum likelihood estimation in small samples: London, Charles Griffin and Co., 186 p.

Stenback, G.A., Crumpton, W.A., Schilling, K.E., and Helmers, M.J., 2011, Rating curve estimation of nutrient loads in Iowa Rivers: Journal of Hydrology (Amsterdam), v. 396, no. 1-2, p. 158–169, accessed January 2019 https://doi.org/10.1016/j.jhydrol.2010.11.006.

U.S. Geological Survey, 2016, National Water Information System data: USGS Water Data for the Nation web page, U.S. Geological Survey, accessed December 2019 at https://doi.org/10.5066/F7P55KJN.

For additional information contact:
Director, Ohio-Kentucky-Indiana Water Science Center
U.S. Geological Survey
6460 Busch Boulevard Ste 100
Columbus, OH 43229-1737

Publishing support provided by:
Indianapolis Publishing Service Center
U.S. Geological Survey Science Publishing Network

Koltun—**Estimated Total Phosphorus Loads for Selected Sites on Great Lakes Tributaries, Water Years 2014–2018**—Open-File Report 2020–1145

ISSN 2331-1258 (online)
https://doi.org/10.3133/ofr20201145