An R package for estimating river compound load using different methods

Veronica NAVA[a*], Martina PATELLI[a], Marco ROTIROTI[a], Barbara LEONI[a]

Affiliation:

[a]Department of Earth and Environmental Sciences, University of Milano - Bicocca, Piazza della Scienza 1, 20126 Milano, Italy; M.P. m.patelli3@campus.unimib.it; M.R. marco.rotiroti@unimib.it; B.L. barbara.leoni@unimib.it

*Corresponding author: Veronica NAVA, Department of Earth and Environmental Sciences, University of Milano - Bicocca, Piazza della Scienza 1, 20126 Milano, Italy, phone +390264482705, email: v.nava15@campus.unimib.it

Abstract

The load quantification of solutes and suspended materials in rivers provides meaningful ecological information about watershed functionality. High-frequency measurements of flow are often available, whereas concentration data are commonly recorded at low frequencies. Different calculation methods have been developed by various authors to provide unbiased load estimation. We provide a new R package (`RiverLoad`) that implements several of the most widely used load estimation algorithms. The package provides an easy-to-use tool to rapidly calculate the load for various compounds and to compare different methods. The package also supplies additional functions to easily organize and analyze the data. A bootstrapping was performed on two example datasets to illustrate the reliability of the methods at different sampling frequencies. The `RiverLoad` package should make it easier to obtain load data and to compare different estimation algorithms. However, attention must be paid when selecting the method to avoid consistent error in the load estimation.

Keywords: averaging method; load estimation; R package; ratio estimator; regression method; river.

Software availability

Name of software: RiverLoad

Developers: Nava V. [aut, cre], Patelli M. [ctb], Rotiroti M. [ctb], Leoni B. [ctb]

Software required: R

Availability: Freely available through CRAN (cran.r-project.org), development version at https://github.com/VeronicaNava/RiverLoad

1 Introduction

The river load data have a key role in many research and monitoring programs concerning water-quality evaluation, as the load provides an integrated measure of inputs and biogeochemical processes within a watershed (Aulenbach et al., 2016; Cooper and Watts, 2002; Craven et al., 2017; Letcher et al., 2002; Shrestha and Kazama, 2007; Wolfs et al., 2015). Precise and accurate estimates of river nutrients and suspended solid loads are relevant not only to assess the water quality but also to calculate the proportion between point and non-point sources, to highlight critical areas that require effective management strategies, to calibrate catchment-scale models, and to evaluate long-term trends in river load (Chu et al., 2008; Elwan et al., 2018; Littlewood, 1995; Newham et al., 2004; Polyakov et al., 2007; Quilbé et al., 2006).

The total load is the product of solute concentration and discharge integrated over time. While discharge is often measured in a continuous manner, the concentration of most compounds is usually measured at discrete points in time, usually at low frequencies (Aulenbach et al., 2016; Moatar and Meybeck, 2005; Preston et al., 1989; Webb et al., 2000). Consequently, the load estimation may be difficult and various techniques have been developed for this aim (Aulenbach et al., 2016; Moatar and Meybeck, 2005; Phillips et al., 1999; Preston et al., 1989). The different calculation methods can be divided into three groups. The first is represented by the *averaging methods*, simple interpolation methods which use averages as representative measure of concentration, flow, or load for a given time interval. The implicit assumptions are that the data must be independent and identically distributed. The samples often do not cover the entire range of flow and concentration values and, consequently, load estimates can be biased (Preston et al., 1989; Quilbé et al., 2006). However, averaging approaches have shown relatively high precision in some surveys and might be suitable in special situations, for example when the purpose is to detect a temporal change in the load (Richards, 1998). The *ratio estimators* form the second group of the load calculation methods. These estimators attempt to correct for the conditions at the time of sampling; the mean load is adjusted by the ratio of the long-term mean discharge to the average daily discharge of days on which samples have been collected (Aulenbach and Hooper, 2006; Cooper and Watts, 2002; Lee et al., 2016). These methods are derived from the ratio estimator developed by Beale (1962) and are suitable when large amount of flow data are and few concentration data are available (Quilbé et al., 2006). Ratio estimators assume that there is a positive linear relationship between instantaneous fluxes and instantaneous flows (with origin at zero) and the variance in instantaneous fluxes increases with the variance in instantaneous flows. Both of these conditions are often satisfied, at least approximately, by relationships between load and discharge (Preston et al., 1989; Richards, 1998). The last group is represented by the *regression methods (*or *rating curves)*, in which concentration over time is determined using a regression model defining an empirical relationship between streamflow and concentration (Aulenbach and Hooper, 2006; Lee et al., 2016; Preston et al., 1989; Quilbé et al., 2006). The load estimation depends on the accuracy of the proposed model and the model predicts the average concentration response for the conditions present, and therefore does not attempt to match the observed concentrations at any given time (Aulenbach et al., 2016; Aulenbach and Hooper, 2006). Generally, log-log regressions are applied because flow and concentration are assumed to be described by a bivariate lognormal distribution (Preston et al., 1989; Worrall et al., 2013).

The selection of the appropriate method depends on the frequency and distribution of sampling, watershed size, the variability in flow, and the strength and form of the relationship between concentration and discharge (Aulenbach et al., 2016).

In this work, we present a new R package, called `RiverLoad`, to provide a useful tool to perform different calculation methods to estimate load, from the concentration values of various chemical constituents and flow records. We provide twelve functions to perform different algorithms previously elaborated and reported by various authors (Dolan et al., 1981; Moatar and Meybeck, 2005; Phillips et al., 1999; Preston et al., 1989; Quilbé et al., 2006; Smith et al., 2016). The `RiverLoad` package allows the load to be estimated, but also allows to organize the database in an easy way and to obtain statistical parameters. The aim of the package is to provide an accessible, user-friendly tool to quickly get an estimation of load, also with limited databases, and to easily compare different methods of calculation. The release package is available on CRAN and the full open-source code is freely accessible for examination and extension online.

2 Conceptual background: algorithms for load estimation

2.1 Averaging methods

2.1.1 Method 1: Time-Weighted Q and C

Method 1 considers the mean of concentration and the mean of flow of the different samples to obtain the load value, with the following equation:

$$L = K\left(\sum_{i=1}^{n}\frac{C_i}{n}\right)\left(\sum_{i=1}^{n}\frac{Q_i}{n}\right)$$

where C_i (g m^{-3}) is the instantaneous concentration associated with individual samples, Q_i (m^3 s^{-1}) is the instantaneous discharge at time of sampling, n is the number of samples collected, K is a conversion factor to account for the measurement units, and thus its value depends on the measurement units in which flow and concentration data are reported (e.g., Littlewood, 1995; Moatar and Meybeck, 2005; Phillips et al., 1999; Walling and Webb, 1985; Worrall et al., 2013).

Various studies have reported that this estimator is precise, given similar results with different subsamples from the same dataset, but sometimes biased, resulting in an underestimation of the actual load (Ferguson, 1987; Quilbé et al., 2006; Walling and Webb, 1981a).

2.1.2 Method 2: Discharge-weighted C

Method 2 is a simple interpolation method that involves the mean value of loads over a certain time period where both concentration and flow are measured (Dolan et al., 1981; Littlewood, 1995). All concentration and flow pairs are equally weighted (Worrall et al., 2013) and the load is calculated with the following equation:

$$L = K\left(\sum_{i=1}^{n}\frac{C_i Q_i}{n}\right)$$

This method seemed to have a large bias for discrete concentration data (Quilbé et al., 2006).

2.1.3 Method 3: Mean discharge-weighted C

Method 3 is based on the hypothesis of constant concentration around a sample and the load is estimated as follows:

$$L = K' \sum_{i=1}^{n} C_i \overline{Q_{i,i-1}}$$

where $Q_{i,i-1}$ (m^3 s^{-1}) is the mean discharge for the interval between samples *i* and *i-1*, derived from frequent flow records, and *K'* is a conversion factor to account for the period of load estimation (Moatar and Meybeck, 2005; Preston et al., 1989).

2.1.4 Method 4: Time-Weighted C

The fourth method, developed by Dolan et al. (1981), is the product of the means of sampled concentrations and the annual discharge:

$$L = K\bar{Q}\left(\sum_{i=1}^{n} \frac{C_i}{n}\right)$$

where $\bar{Q}$ (m^3 s^{-1}) is the annual mean discharge, derived from frequent flow records. Unlike the previously reported procedure, this algorithm uses all the available flow data (Moatar and Meybeck, 2005; Quilbé et al., 2006). This estimator is reported to be precise, given similar results with different subsamples from the same dataset, but sometimes biased, resulting in an underestimation of the actual load (Ferguson, 1987; Quilbé et al., 2006; Walling and Webb, 1981b). A previous study highlighted that this method gives good load estimation for specific compounds, such as particulate-P (Moatar and Meybeck, 2005).

2.1.5 Method 5: Time and discharge weighted

Method 5 weighs the mean daily load by the mean of all measured flows and estimates the load as follows:

$$L = K \frac{\sum_{i=1}^{n} C_i \, Q_i}{\sum_{i=1}^{n} Q_i} \bar{Q}$$

This estimator was found to be less biased than method 1 and method 4, but resulted in large variability in load estimations (Quilbé et al., 2006).

2.1.6 Method 6: Linear interpolation of C

Method 6 is based on the linear interpolation of the concentration values; then, the values drawn are multiplied by the flow records to obtain the load estimation as follows:

$$L = K'' \sum_{j=1}^{n} C_j^{int} Q_j$$

where C_j^{int} (g m^{-3}) is the daily concentration linearly interpolated between two measured samples, Q_j (m^3 s^{-1}) is the mean daily discharge, and *K''* is a conversion factor to account for the period of load estimation (Moatar and Meybeck, 2005). Interpolation procedures essentially involve the assumption that the values of concentration or discharge obtained from instantaneous samples are representative of a much longer period of time and it is important to take this into account when applying this algorithm (Moatar and Meybeck, 2005). This method seemed to be accurate and precise, and in a previous study, it is highlighted to be suitable for nitrate and soluble reactive phosphorus load estimation (Chu et al., 2008; Moatar and Meybeck, 2005).

2.2 Ratio estimators

2.2.1 Beale Ratio estimation

The Beale Ratio Estimator (Beale, 1962) is a method that has been shown to produce robust and statistically unbiased results (Quilbé et al., 2006). The mean daily load, calculated as the product of concentration and flow of days on which samples are taken and then averaged, is multiplied by a flow ratio, which is derived by dividing the average flow as a whole by the average flow recorded in the chemical sampling days. A bias correction factor is included in the calculation, to compensate for the effects of the correlation between discharge and load (Richards, 1998):

$$L = Q\frac{\bar{l}}{\bar{q}}\left[\frac{1+\frac{1}{n}\left[\frac{Cov(l,q)}{\bar{l}\,\bar{q}}\right]}{1+\frac{1}{n}\left[\frac{Var(\bar{q})}{\bar{q}^2}\right]}\right]$$

where Q ($m^3\ s^{-1}$) is the total flow for the considered period, $\bar{q}$ ($m^3\ s^{-1}$) is the mean flow for times when chemical compounds were measured, and $\bar{l}$ ($g\ s^{-1}$) is the mean flow for times when samples were collected. The term in square brackets is the bias correction term. This method does not assume a normal distribution and it is not recommended to be used with log-transformed data (Richards, 1998; Worrall et al., 2013).

2.3 Regression methods

2.3.1 Log-log rating

The most common regression equation for load estimation is the log-log linear rating curve derived from the relationship between the values of concentration and river flow at the time of sampling:

$$log_{10}(C) = a + b \cdot log_{10}(Q)$$

where *a* and *b* are the intercept and the slope of the least square regression line, respectively.

This relationship is applied to the high frequency discharge record to generate the daily concentration. The values obtained are used to calculate the load by summing, over a specific period, the product of daily concentration and daily streamflow (Phillips et al., 1999; Quilbé et al., 2006), as follows:

$$L_r = \sum_{i=1}^{n} C_i Q_i$$

2.3.2 Ferguson rating curve

Ferguson (1986) recommended a correction to the previous method to obtain an unbiased estimator L_{cr}:

$$L_{cr} = L_r \cdot e^{2.651\, s^2}$$

where s is the standard error of the estimate of the rating curve in $\log_{10}$ units (Phillips et al., 1999; Quilbé et al., 2006; Worrall et al., 2013).

3 `RiverLoad` package features

The load estimation algorithms of the package are listed in Table 1. In Table 2, we reported the additional functions provided to organize the initial matrix (Table 2a-b), to perform descriptive statistics on flow records (Table 2c-f), and to analyze regression models (Table 2g-i).

Table 1. List of the different load estimation algorithms provided by `RiverLoad` package. The reference number of the method ("Method no."), the name of the function in the package ("Function name"), the algorithm of the method ("Algorithm"), the method typology ("Class"), and selected references for the different calculation ("References") are reported.

Method no.	Function name	Algorithm	Class	References
1	`method1`	$L = K\left(\sum_{i=1}^{n}\frac{C_i}{n}\right)\left(\sum_{i=1}^{n}\frac{Q_i}{n}\right)$	Interpolation	(Moatar and Meybeck, 2005; Phillips et al., 1999; Verhoff et al., 1980; Walling and Webb, 1985)
2	`method2`	$L = K\left(\sum_{i=1}^{n}\frac{C_i Q_i}{n}\right)$	Interpolation	(Moatar and Meybeck, 2005; Phillips et al., 1999; Quilbé et al., 2006; Rodda and Jones, 1983; Walling and Webb, 1985; Worrall et al., 2013)
3	`method3`	$L = K'\sum_{i=1}^{n} C_i \overline{Q_{\iota,\iota-1}}$	Interpolation	(Moatar and Meybeck, 2005; Phillips et al., 1999; Walling and Webb, 1985, 1981a)
4	`method4`	$L = K\bar{Q}\left(\sum_{i=1}^{n}\frac{C_i}{n}\right)$	Interpolation	(Moatar and Meybeck, 2005; Ongley, 1973; Quilbé et al., 2006; Walling and Webb, 1985)
5	`method5`	$L = K\frac{\sum_{i=1}^{n} C_i\, Q_i}{\sum_{i=1}^{n} Q_i}\bar{Q}$	Interpolation	(Moatar and Meybeck, 2005; Quilbé et al., 2006; Verhoff et al., 1980; Walling and Webb, 1985)
6	`method6`	$L = K''\sum_{j=1}^{n} C_j^{int} Q_j$	Interpolation	(Moatar and Meybeck, 2005; Williams et al., 2015)
7	`beale.ratio`	$L = Q\frac{\bar{l}}{\bar{q}}\left[\frac{1+\frac{1}{n}\left[\frac{Cov(l,q)}{\bar{l}\,\bar{q}}\right]}{1+\frac{1}{n}\left[\frac{Var(\bar{q})}{\bar{q}^2}\right]}\right]$	Ratio	(Beale, 1962; Elwan et al., 2018; Lee et al., 2016; Phillips et al., 1999; Quilbé et al., 2006)
8	`beale.period`	$L = Q\frac{\bar{l}}{\bar{q}}\left[\frac{1+\frac{1}{n}\left[\frac{Cov(l,q)}{\bar{l}\,\bar{q}}\right]}{1+\frac{1}{n}\left[\frac{Var(\bar{q})}{\bar{q}^2}\right]}\right]$	Ratio (different covariance and variance for month and year)	(Beale, 1962; Elwan et al., 2018; Lee et al., 2016; Phillips et al., 1999; Quilbé et al., 2006)
9	`rating`	$C = aQ^b$; $L_r = \sum_{i=1}^{n} C_i Q_i$	Regression	(Phillips et al., 1999; Quilbé et al., 2006; Walling and Webb, 1981a)

10	`rating.period`	$C = aQ^b$; $L_r = \sum_{i=1}^{n} C_i Q_i$	Regression (different regression for month and year)	(Phillips et al., 1999; Quilbé et al., 2006; Walling and Webb, 1981a)
11	`ferguson`	$C = aQ^b$; $L_{cr} = L_r \cdot e^{2.651\, s^2}$	Regression	(Ferguson, 1986; Phillips et al., 1999; Preston et al., 1989; Quilbé et al., 2006; Worrall et al., 2013)
12	`ferguson.period`	$C = aQ^b$; $L_{cr} = L_r \cdot e^{2.651\, s^2}$	Regression (different estimation for month and year)	(Ferguson, 1986; Phillips et al., 1999; Preston et al., 1989; Quilbé et al., 2006; Worrall et al., 2013)

Table 2. List of the additional useful functions provided by `RiverLoad` package, with the name of the function ("Function name") and a brief description ("Description").

Function name	Description
a) `db.intersect`	Merge concentration and flow data on the basis of the "datetime" column, maintaining only the date and time in which both the data are available
b) `db.union`	Combine concentration and flow data on the basis of the "datetime" column, maintaining all the flow data and returning 'NA' when concentration data are not available
c) `daily.mean`	Return the mean daily flow
d) `monthly.mean`	Return the mean monthly flow, not differentiated by year
e) `monthly.year.mean`	Return the mean monthly flow, differentiated by year
f) `annual.mean`	Return the mean annual flow
g) `CQregression`	Return R^2 of the regression between concentration and flow for the application of `rating` and `ferguson` functions.
h) `rsquared.period`	Return R^2 for monthly or annual regression for the application of the `rating.period` and `ferguson.period` functions
i) `reg.inspection`	Return the regression parameters: coefficients and the associated *p-value*, R^2, degrees-of-freedom
j) `residual.plot`	Return the residual plot of the specified compound of the regression analyses by `rating` and `ferguson` function

3.1 Input data arrangement

The input data must include flow and concentration records. The input matrix must have at least three columns. The first is the column with the date and the hour of the sampling and must be labeled with "datetime". The date and hour information have to be in the standard format yyyy-mm-dd HH:MM:SS (ISO 8601). The second column must contain the flow records, expressed in $m^3 s^{-1}$, and must be labeled with "flow". The following columns must contain the available concentration data of one or more compounds, expressed in $g\ m^{-3}$. The labels can be freely chosen by the user. For the datetime in which the concentration records are not available, 'NA' must be reported.

The package provides a useful function, named `db.union` (Table 2a), to easily create the requested matrix. In many cases, the concentration and the flow data are available in two distinct matrices and the matching operation can be time-consuming, especially for an extended dataset. This function allows the data to be merged, maintaining all the available flow records with the scattered concentration values based on date and time information. Two arguments are required: the first is the matrix with flow records, the second is the matrix with concentration data. The flow matrix must contain a first column with the datetime information, labeled "datetime" and expressed in the standard format yyyy-mm-dd HH:MM:SS (ISO 8601), and a second flow column, in $m^3 s^{-1}$, labeled "flow". The concentration matrix must contain the "datetime" column in the standard format (yyyy-mm-dd HH:MM:SS) and the concentration data with user-defined labels (see Figure S1).

In addition, `RiverLoad` allows the concentration and flow records to be matched in a second manner. The function `db.intersect` (Table 2b) pairs the data maintaining the rows in which both the concentration and the flow records are available. In this case, no 'NA' value is reported. The arguments requested are equivalent to that of the `db.union` function.

All data must be quality assured/quality controlled (QA/QC) before use for load estimation (Winslow et al., 2016). We strongly suggest manual inspection of all the data to identify data gaps, anomalies and potential pitfalls in the dataset.

3.2 Explorative analyses of flow

Supplementary functions permit explorative analyses of the flow data, calculating the mean value over different periods. The function called `daily.mean` (Table 2c) allows a daily mean of flow to be obtained. The function `monthly.mean` (Table 2d) returns a monthly mean not differentiated by year; therefore, if more year records are reported, the function provides a single mean value for the same month in different years. Meanwhile, the `monthly.year.mean` function (Table 2e) returns a monthly mean with a different value for different years. Finally, `annual.mean` gives the annual flow mean. The argument of all of these functions must be the flow matrix, with "datetime" and "flow" columns, or the matrix previously obtained with the `db.union` function. An optional argument is "sd" to obtain the standard deviation.

3.3 How to obtain load estimation

The package provides twelve load estimation algorithms, listed in Table 1. All of these functions require the same arguments. The first argument is the matrix described previously, user-created or obtained by the `db.union` function, with flow records and scattered concentration data (Figure S1). The second mandatory argument is the number of compounds, for which the concentration data have been reported and load estimation must be performed. Indeed, it is possible to simultaneously perform the calculation of different compounds measured at the same time. For example, if the user wants to estimate the load for total phosphorus, total nitrogen and soluble reactive phosphorus, he/she must specify “3” as number of compounds. The last is an optional or mandatory argument, depending on the function considered, indicating that the period throughout the estimation must be performed. If this argument is missing, the default calculation is done on the time period occurring from the first and to the last flow record reported in the matrix included in the function. Otherwise, this argument can assume two different specifications: “year” for an annual estimation of load, and “month” to obtain a load estimation every month. Monthly and annual load estimation with regression methods and the Beale ratio can be performed in two different ways:

a) the functions `rating`, `ferguson`, and `beale.ratio` estimate the load calculating a single log-log linear rating curve, value of variance and bias correction factor throughout the period spanned by the streamflow data;
b) the functions `rating.period`, `ferguson.period`, and `beale.period` estimate the load, calculating the monthly or annual log-log linear rating curve, value of variance, and bias correction factor. For these functions, the third argument is mandatory, as they require the specification of the period of estimation, i.e., “month” or “year”.

The output returns a matrix with the load estimation for the different compounds in grams on the time specified.

3.4 Specifications for regression methods

A statistically significant correlation between concentration and flow is mandatory to perform regression methods (Quilbé et al., 2006). `RiverLoad` provides the function called `CQregression` (Table 2g) to obtain R^2 before the application of the `rating` and `ferguson` function, and the `rsquared.period` function (Table 2h) to obtain the R^2 before `rating.period` and `ferguson.period` application. Aulenbach et al. (2016) suggested using a regression method when the coefficient of determination (R^2) is higher than 0.3.

For `rating` and `ferguson` functions it is possible to obtain, beyond load estimation, statistical data using the function `reg.inspection` (Table 2h). This function requires the same arguments as the algorithm functions: the matrix with flow records and scattered concentration data, and the number of compounds for which the analysis must be performed. The output returns the slope and intercept coefficients and their related *p-value*, the R^2, and the residual degrees-of-freedom.

In addition, the `residual.plots` function permits the diagnostic plots returned by `plot.lm` to be obtained: a plot of residuals against fitted values; a Scale-Location plot of $\sqrt{|residuals|}$ against fitted values; a Normal Q-Q plot; a plot of residuals against leverages. The first argument is the usual matrix with flow records and scattered concentration data and the second is the number of specific compounds for which the user would obtain the plots, as the function returns the graphs for one compound at a time. The user can indicate a file path as an optional argument, specifying the folder in which the output plot will be saved and the name of the file with a .jpeg extension. Otherwise, the plot will be shown in the R window.

A workflow displaying the functioning of the `RiverLoad` package is reported in Figure 1. Further practical examples of the standard application of the package are included in the Supplementary Materials with three datasets, which are also embedded in the package structure.

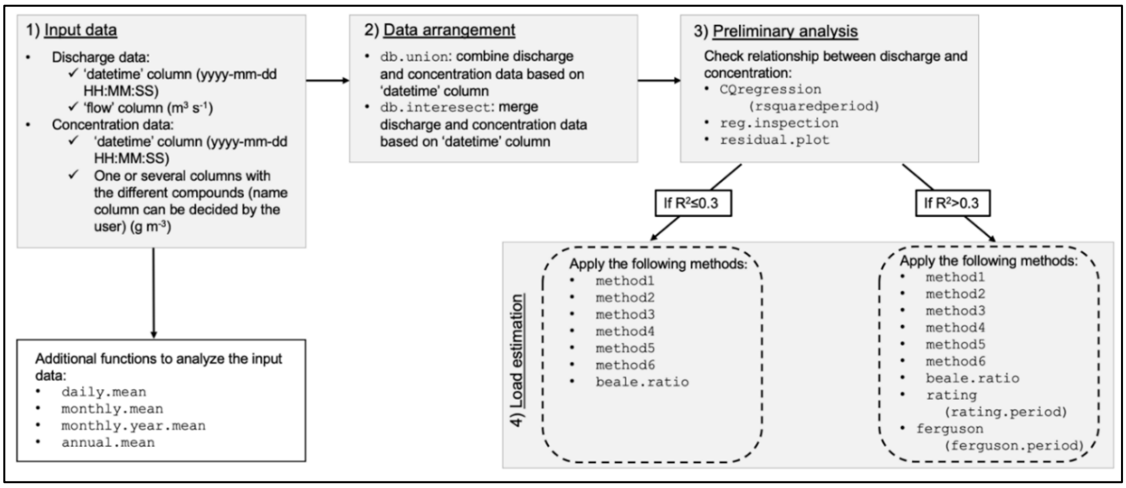

Fig. 1. An example workflow of the `RiverLoad` package.

4 Test cases: a bootstrap approach

To determine the reliability of the various methods provided in the `RiverLoad` package, two further example datasets of rivers, with different watershed extension and flow features, are reported and analyzed below. These datasets have been carefully selected to demonstrate the behavior and the utility of the software. The actual or 'true' load must be known in order to determine the error of load estimates using the different methodologies (Aulenbach and Hooper, 2006); thus, high-frequency datasets were selected. A daily frequency is assumed to be sufficient for capturing most of the significant variability within a given year (Moatar and Meybeck, 2005; Preston et al., 1989). We designed a bootstrap experiment to evaluate the effects of weekly, fortnightly and monthly sampling intervals on the accuracy and precision of the different load estimation procedures, seeing whether the load estimates were close to converging on the true load. We artificially decimated the dataset, by randomly choosing the concentration data collected in a fixed number of sampling days within the defined frequency. Ten thousand replications were performed for each test cases using the 'boot' package in R (Canty and Ripley, 2017). We estimated the load for total phosphorus (TP) for both the datasets to allow an immediate comparison between them. Different compounds could have different behaviour, thus we anticipate that the outcomes reported should be extended to the other compounds with caution.

4.1 Dataset description

The datasets are provided by the National Center for Water Quality Research, Heidelberg University within the Heidelberg Tributary Loading Program (Heidelberg University, 2019), that was initiated in 1974 as part of state and federal programs to restore Lake Erie (United States). All stream flow measurements used by the Heidelberg Tributary Loading Program are provided by the United States Geological Survey (USGS). Within this database, two water bodies, belonging to the Lake Eire watershed but with different basin extension and discharge, were selected: Maumee River and Honey Creek.

With the largest watershed of any Great Lakes river (21538 km^2), the Maumee officially begins at the confluence of the St. Joseph and St. Mary's rivers in Fort Wayne, Indiana, draining all or part of 17 Ohio counties, two counties in Michigan, and five more in Indiana. The International Joint Commission set a phosphorus load limit for Lake Erie of 11 000 t per year; the Maumee River itself discharges about 24% of this value, with fertilizer appearing to be the major source, as the watershed is largely agricultural (Moog and Whiting, 2002). The dataset reported includes the measurement collected from 1st January 2013 to 31st December 2013 (n=365). Discharge and total phosphorus concentration data are reported in Fig. 2.

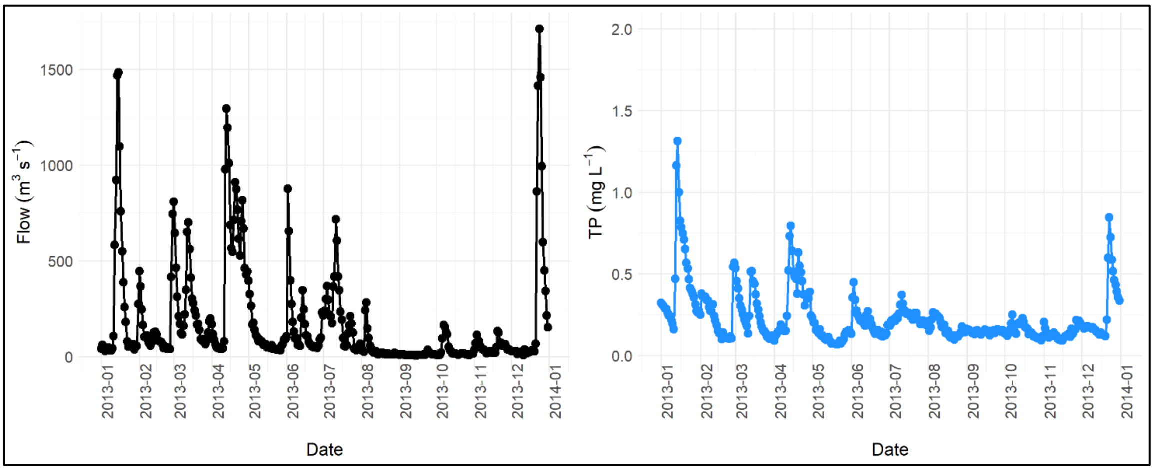

Fig. 2. Time series for Maumee River dataset included in `RiverLoad` package showing variation in (a) discharge (m^3 s^{-1}) and (b) total phosphorus concentration (mg L^{-1}) from 2013-01-01 to 2013-12-31.

Honey Creek is a major tributary of the Sandusky River (USA), which drains 463.4 km^2 of land area in North Central Ohio. The average annual runoff for Honey Creek is 13.7 km^3. This value is about average for subwatersheds in the Sandusky Basin and similar to those for surrounding watersheds. For Honey Creek, 64.1% of the 242.9 miles of streams are first-order streams. Generally, February, March and April are the months with the highest average discharges while August, September and October have the lowest discharges (Loftus et al., 2006). In Fig. 3, flow and concentration data from January 2015 to December 2015 are reported.

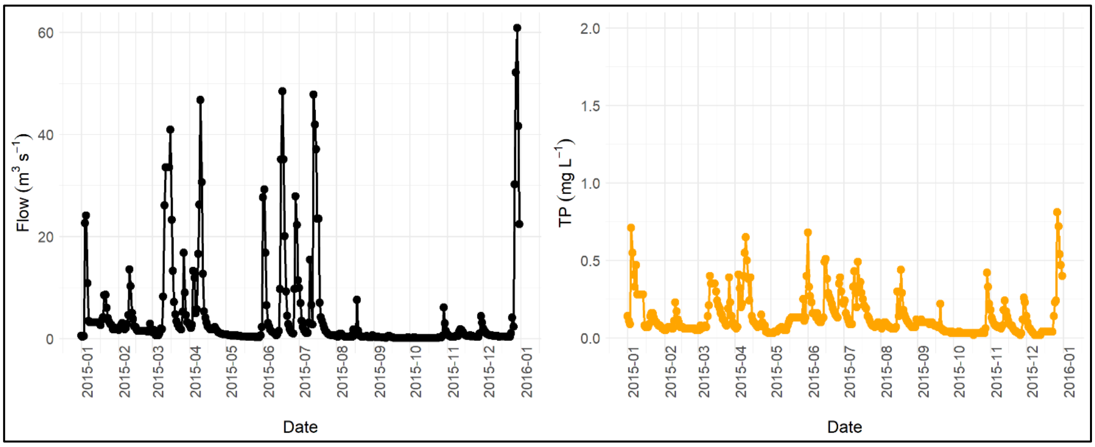

Fig. 3. Time series for Honey Creek dataset included in `RiverLoad` package showing variation in (a) discharge (m^3 s^{-1}) and (b) total phosphorus concentration (mg L^{-1}) from 2013-01-01 to 2013-12-31.

4.2 Bootstrap results

The true load from the daily flow and concentration data of Maumee River and Honey Creek is equal to 2494.12 t year^{-1} and 50.75 t year^{-1}, respectively. While these loads are derived from measurements of concentration and discharge that cannot be considered free of all error, they approximate the actual or true load at these sites during the study period. Accordingly, these estimates were used as the reference values against which the error of the load estimates, produced for the Maumee River and Honey Creek under several sampling scenarios, were assessed. The error in load estimates is a combination of bias (accuracy) and precision, whose values are reported in Table 3–4. The accuracy of each subsampling scenario is calculated as the mean of the relative errors (ε); the precision is estimated as the standard deviation of the percentage errors (σ). The values are reported in percentage to easily compare the results between the two different datasets. As reported by Moatar and Meybeck (2005), many studies have highlighted an inverse relationship between accuracy and precision of the different estimation methods and suggested the root-mean-square error ($RMSE = \sqrt{\bar{\varepsilon}^2 + \bar{\sigma}^2}$), which combines bias and precision, as a suitable evaluation criterion (Table 3–4).

Table 3. The percent bias 'ε (%)' and standard deviation of the percentage errors 'σ (%)' of the weekly, fortnightly and monthly estimated load of total phosphorus for Maumee River.

Method no.	Statistics	WEEKLY	FORTNIGHTLY	MONTHLY
1	**ε (%)**	-91	-88	-85
	σ (%)	17	38	55
	RMSE (%)	93	96	101
2	**ε (%)**	-1	0	-1
	σ (%)	24	49	70
	RMSE (%)	24	49	70
3	**ε (%)**	-18	-51	-118
	σ (%)	13	31	23
	RMSE (%)	23	60	120
4	**ε (%)**	-92	-93	-94
	σ (%)	5	12	16
	RMSE (%)	92	94	95
5	**ε (%)**	-2	-5	-11
	σ (%)	14	26	37
	RMSE (%)	14	26	38
6	**ε (%)**	-19	-39	-56
	σ (%)	10	19	24
	RMSE (%)	21	43	61
7	**ε (%)**	1	-1	-6
	σ (%)	14	27	39
	RMSE (%)	14	27	39
9	**ε (%)**	-34	-36	-37
	σ (%)	6	13	20
	RMSE (%)	34	38	42
11	**ε (%)**	-24	-26	-27
	σ (%)	6	13	20
	RMSE (%)	25	29	33

Table 4. The percent bias 'ε (%)' and standard deviation of the percentage errors 'σ (%)' of the weekly, fortnightly and monthly estimated load of total phosphorus for Honey Creek.

Method no.	Statistics	WEEKLY	FORTNIGHTLY	MONTHLY
1	ε (%)	-125	-143	-127
	σ (%)	23	52	76
	RMSE (%)	127	152	148
2	ε (%)	11	-1	0
	σ (%)	25	55	79
	RMSE (%)	27	55	79
3	ε (%)	-38	-90	-142
	σ (%)	16	39	36
	RMSE (%)	41	98	147
4	ε (%)	-144	-152	-152
	σ (%)	8	17	25
	RMSE (%)	144	153	154
5	ε (%)	4	-7	-15
	σ (%)	12	26	38
	RMSE (%)	13	27	41
6	ε (%)	-27	-71	-98
	σ (%)	11	24	29
	RMSE (%)	29	75	102
7	ε (%)	6	-3	-10
	σ (%)	12	26	39
	RMSE (%)	14	27	40
9	ε (%)	-13	-16	-16
	σ (%)	8	17	26
	RMSE (%)	15	23	30
11	ε (%)	2	-1	0
	σ (%)	9	17	26
	RMSE (%)	9	17	26

The various methods show different reliability compared to the true load, as it can be seen from the RMSE values and from the boxplots, showing the percentage relative error of the estimation of the different replicated estimates (Figs. 4–5). Given the different subsampling scenarios, the procedures that provide the most accurate and precise estimation are method 11 (Ferguson rating curve), method 7 (Beale ratio estimator) and method 5. On the contrary, method 4 and method 1 produce the greatest error in terms of RMSE. A common feature of these two algorithms is that concentration and discharge are separately averaged and then multiplied, and consequently the relationship between flow and concentration, if any, is lost. It is likely that this kind of methods could give a reliable estimation for other rivers that display a low inter-annual variability, without marked seasonal variations.

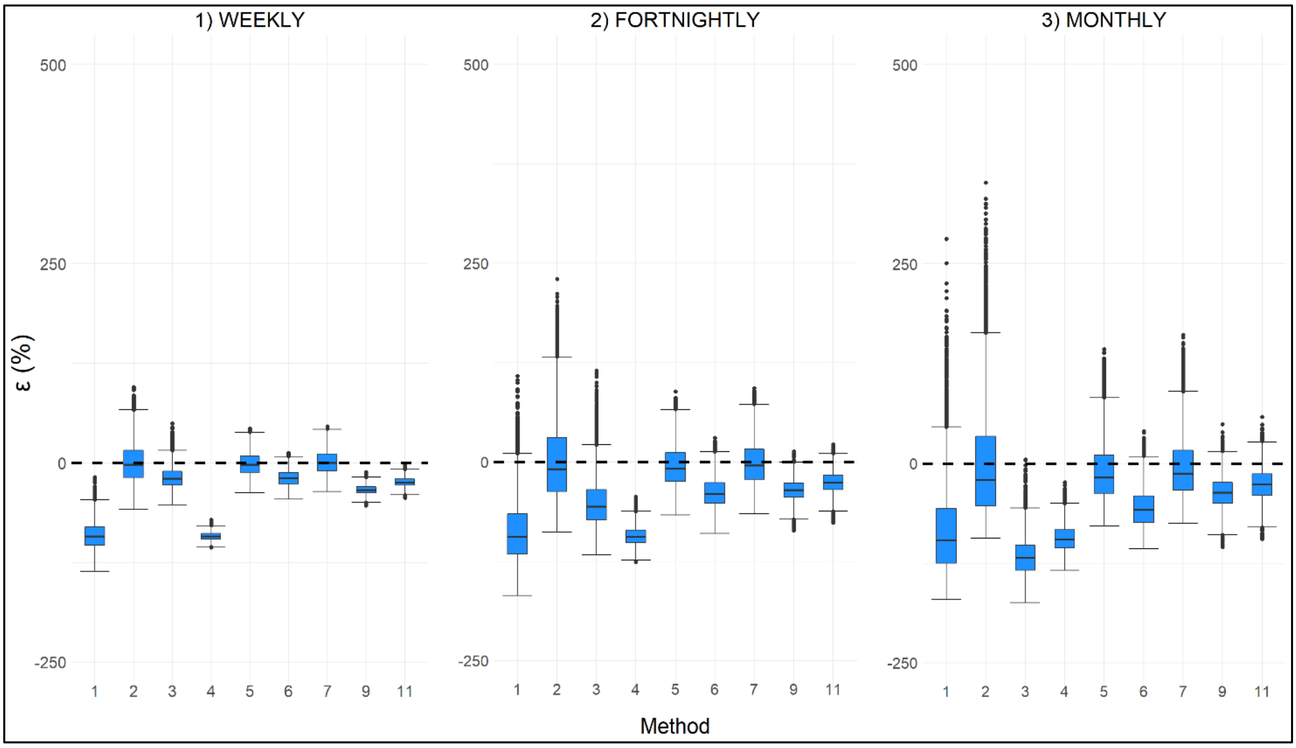

Fig. 4. Boxplot of the percentage relative error 'ε (%)' obtained from the bootstrap (n=10000) estimation of the load of total phosphorus (TP) of Maumee River (Ohio), based on the nine estimation methods implemented (see Table 1 for the acronym of the method on the x-axis). The results are reported for the three different sampling frequencies tested: 1) weekly; 2) fortnightly; 3) monthly.

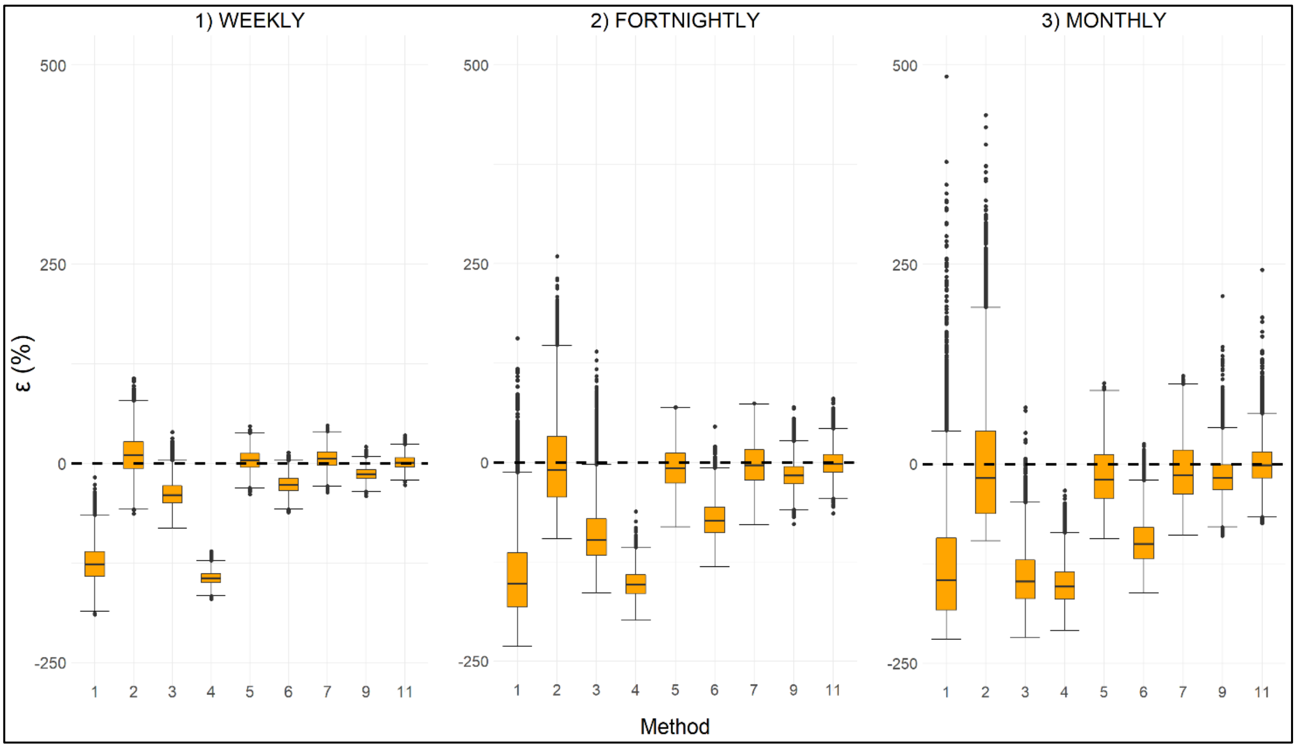

Fig. 5. Boxplot of the percentage relative error 'ε (%)' obtained from the bootstrap (n=10000) estimation of the load of total phosphorus (TP) of Honey Creek (Ohio), based on the nine load calculation methods implemented (see Table 1 for the acronym of the method on the x-axis). The results are reported for the three different sampling frequencies tested: 1) weekly; 2) fortnightly; 3) monthly.

Many procedures seem to underestimate the reference load, and this is especially verified for method 4. For Honey Creek, the absolute value of the median is between 1.1% and 144% of the reference load at a weekly frequency, 1.6–168% at a fortnightly frequency and 1.8–179% at a monthly frequency, where the best estimation is given by the method 11 (Ferguson rating curve) and the worst by methods 1 and 4. Whereas, for the Maumee River, the equivalent range is between 0.1% and 93% at a weekly frequency, 5–101% at a fortnightly frequency, and 14%–119% at a monthly frequency, where the best estimation is given by method 7 (Beale ratio estimator) and the worst by the methods 1 and 3. The reduction in sampling frequency results in a reduction of precision and, to a lesser extent, to an enhancement of bias. Indeed, the results obtained through a monthly sampling show a larger difference in load estimates, which indicates that there was insufficient sampling and the estimated true load might be biased and imprecise. Therefore, great attention must be paid to the selection of the most suitable methods, as a not careful consideration could lead to a large error in the estimation, especially when concentration data are sparse (*e.g.* less often than twice a month).

It is important to underline as the bootstrap procedure was performed on a random basis and the data were not specifically selected. In real cases, the knowledge of the studied site and the researcher expertise play an important role, and accurate evaluation of the data before employing them in the package must be done, as the quality of input data largely affects the quality of the outcome obtained. An effective selection of the sampling dates, due to the prior knowledge of the studied site, and the evaluation of "outlier" conditions is recommended.

5 Conclusion and future improvements

`RiverLoad` is a useful tool to facilitate load estimation, information of primary importance for water quality assessment and pollution source identification. Load data constitute prior information in a wide range of ecological studies; however, their estimation can represent a complex challenge. Indeed, different algorithms have been proposed by various authors and the application of whole of them can be a very time-consuming activity. This package allows different methods to be easily compared and we also provided documented references to allow the user to learn about the development of the algorithms and previous applications. `RiverLoad` is suitable for both extended and limited dataset and allows the contemporary estimation for different compounds.

The package provides twelve different functions, but the quality of their estimation changes depending on different parameters, e.g., the field study characteristics and extension, the frequency and distribution of the sampling, the variability in flow, and the user must take this into account. Indeed, the outcomes obtained may be wrong if the method is not properly chosen. A general rule for the selection of the right procedure could not be provided, because the validity of the results is largely dependent on the site features. Moreover, the validation of estimates, is in general, a perplexing task as it would require a complete sequence of data that, if any, makes the use of an estimation method unnecessary. However, we provided some guidance to load estimation methods as a certain procedure may be more appropriate over another in some situations.

We believe that `RiverLoad` is highly useful and it benefits from being a completely open source tool, open to further examination and extension.

Acknowledgements

This work was supported by University of Milano-Bicocca (FA). The authors declare no conflict of interest. We thank the anonymous reviewers for their comments and feedbacks, which helped to improve this article.

Author's contribution

V.N. developed the package and wrote the manuscript; V.N., M.P., B.L., M.R. tested all the codes and the examples; B.L., M.R. revised and improved the manuscript. All authors have approved the final version to be published.

References

Aulenbach, B.T., Burns, D.A., Shanley, J.B., Yanai, R.D., Bae, K., Wild, A.D., Yang, Y., Yi, D., 2016. Approaches to stream solute load estimation for solutes with varying dynamics from five diverse small watersheds. Ecosphere 7, 1–22. https://doi.org/10.1002/ecs2.1298

Aulenbach, B.T., Hooper, R.P., 2006. The composite method: An improved method for stream-water solute load estimation. Hydrol. Process. 20, 3029–3047. https://doi.org/10.1002/hyp.6147

Beale, E.M.L., 1962. Some uses of computers in operational research. Ind. Organ. 31, 27–28.

Canty, A., Ripley, B., 2017. boot: Bootstrap R (S-plus) Functions.

Chu, Y., Salles, C., Cernesson, F., Perrin, J.L., Tournoud, M.G., 2008. Nutrient load modelling during floods in intermittent rivers: An operational approach. Environ. Model. Softw. 23, 768–781. https://doi.org/10.1016/j.envsoft.2007.09.008

Cooper, D.M., Watts, C.D., 2002. A comparison of river load estimation techniques: Application to dissolved organic carbon. Environmetrics 13, 733–750. https://doi.org/10.1002/env.525

Craven, J., Angarita, H., Corzo Perez, G.A., Vasquez, D., 2017. Development and testing of a river basin management simulation game for integrated management of the Magdalena-Cauca river basin. Environ. Model. Softw. 90, 78–88. https://doi.org/10.1016/j.envsoft.2017.01.002

Dolan, D.M., Yui, A.K., Geist, R.D., 1981. Evaluation of River Load Estimation Methods for Total Phosphorus. J. Great Lakes Res. 7, 207–214. https://doi.org/10.1016/S0380-1330(81)72047-1

Elwan, A., Singh, R., Patterson, M., Roygard, J., Horne, D., Clothier, B., Jones, G., 2018. Influence of sampling frequency and load calculation methods on quantification of annual river nutrient and suspended solids loads. Environ. Monit. Assess. 190, 78. https://doi.org/10.1007/s10661-017-6444-y

Ferguson, R.I., 1987. Accuracy and precision of methods for estimating river loads. Earth Surf. Process.

Landforms 12, 95–104. https://doi.org/10.1002/esp.3290120111

Ferguson, R.I., 1986. River Loads Underestimated by Rating Curves. Water Resour. Res. 22, 74–76. https://doi.org/10.1029/WR022i001p00074

Heidelberg University, 2019. Heidelberg Tributary Loading Program [WWW Document].

Lee, C.J., Hirsch, R.M., Schwarz, G.E., Holtschlag, D.J., Preston, S.D., Crawford, C.G., Vecchia, A. V., 2016. An evaluation of methods for estimating decadal stream loads. J. Hydrol. 542, 185–203. https://doi.org/10.1016/j.jhydrol.2016.08.059

Letcher, R. a, Jakeman, a J., Calfas, M., Linforth, S., Baginska, B., Lawrence, I., 2002. A comparison of catchment water quality models and direct estimation techniques. Environ. Model. Softw. 17, 77–85. https://doi.org/http://dx.doi.org/10.1016/S1364-8152(01)00054-8

Littlewood, I.G., 1995. Hydrological regimes, sampling strategies, and assessment of errors in mass load estimates for United Kingdom rivers. Environ. Int. 21, 211–220. https://doi.org/10.1016/0160-4120(95)00011-9

Loftus, T.T., Baker, D., V Setzler, J., Crumrine, J., 2006. Honey Creek Watershed Action Plan.

Moatar, F., Meybeck, M., 2005. Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire. Hydrol. Process. 19, 429–444. https://doi.org/10.1002/hyp.5541

Moog, D.B., Whiting, P.J., 2002. Climatic and Agricultural Factors in Nutrient Exports from Two Watersheds in Ohio. J. Environ. Qual. 31, 72–83. https://doi.org/10.2134/jeq2002.7200

Newham, L.T.H., Letcher, R.A., Jakeman, A.J., Kobayashi, T., 2004. A framework for integrated hydrologic, sediment and nutrient export modelling for catchment-scale management. Environ. Model. Softw. 19, 1029–1038. https://doi.org/10.1016/j.envsoft.2003.11.006

Ongley, E.D., 1973. Sediment discharge from Canadian basins into Lake Ontario. Can. J. Earth Sci. 10, 146–156.

Phillips, J.M., Webb, B.W., Walling, D.E., Leeks, G.J.L., 1999. Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. Hydrol. Process. 13, 1035–1050. https://doi.org/10.1002/(SICI)1099-1085(199905)13:7<1035::AID-HYP788>3.0.CO;2-K

Polyakov, V., Fares, A., Kubo, D., Jacobi, J., Smith, C., 2007. Evaluation of a non-point source pollution model, AnnAGNPS, in a tropical watershed. Environ. Model. Softw. 22, 1617–1627. https://doi.org/10.1016/j.envsoft.2006.12.001

Preston, S.D., Bierman, V.J., Silliman, S.E., 1989. An evaluation of methods for the estimation of tributary mass loads. Water Resour. Res. 25, 1379–1389. https://doi.org/10.1029/WR025i006p01379

Quilbé, R., Rousseau, A.N., Duchemin, M., Poulin, A., Gangbazo, G., Villeneuve, J.P., 2006. Selecting a calculation method to estimate sediment and nutrient loads in streams: Application to the Beaurivage River (Québec, Canada). J. Hydrol. 326, 295–310. https://doi.org/10.1016/j.jhydrol.2005.11.008

Richards, R.P., 1998. Estimation of pollutant loads in rivers and streams: A guidance document for NPS programs, Prepared under Grant X998397-01-0 U.S. Environmental Protection Agency, Region VIII David Rathke, Project Officer.

Rodda, J.C., Jones, G.N., 1983. Preliminary estimates of loads carried by rivers to estuaries and coastal waters around Great Britain derived from the Harmonized Monitoring Scheme. J. Inst. Water Eng. Sci. 37.

Shrestha, S., Kazama, F., 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environ. Model. Softw. 22, 464–475. https://doi.org/10.1016/j.envsoft.2006.02.001

Smith, V., Wood, S., McBride, C., Atalah, J., Hamilton, D., 2016. Phosphorus and nitrogen loading restraints are essential for successful eutrophication control of Lake Rotorua, New Zealand. Inl. Waters 6, 273–283. https://doi.org/10.5268/IW-6.2.998

Verhoff, F.H., Melfi, D.A., Yaksich, S.M., 1980. River nutrient and chemical transport estimation. J. Environ. Eng. Div. 106, 591–608.

Walling, D.E., Webb, B.W., 1985. Estimating the discharge of contaminants to coastal waters by rivers: Some cautionary comments. Mar. Pollut. Bull. 16, 488–492. https://doi.org/10.1016/0025-326X(85)90382-0

Walling, D.E., Webb, B.W., 1981a. The reliability of suspended sediment load data. Eros. sediment Transp. Meas. (Proceedings Florence Symp. 177–194.

Walling, D.E., Webb, B.W., 1981b. The reliability of suspended sediment load data. Eros. sediment Transp. Meas. (Proceedings Florence Symp. 177–194.

Webb, B.W., Phillips, J.M., Walling, D.E., 2000. A new approach to deriving “best-estimate” chemical fluxes for rivers draining the LOIS study area. Sci. Total Environ. 251–252, 45–54. https://doi.org/10.1016/S0048-9697(00)00413-7

Williams, M.R., King, K.W., Macrae, M.L., Ford, W., Van Esbroeck, C., Brunke, R.I., English, M.C., Schiff, S.L., 2015. Uncertainty in nutrient loads from tile-drained landscapes: Effect of sampling frequency, calculation algorithm, and compositing strategy. J. Hydrol. 530, 306–316. https://doi.org/10.1016/j.jhydrol.2015.09.060

Winslow, L.A., Zwart, J.A., Batt, R.D., Dugan, H.A., Iestyn Woolway, R., Corman, J.R., Hanson, P.C., Read, J.S., 2016. LakeMetabolizer: An R package for estimating lake metabolism from free-water oxygen using diverse statistical models. Inl. Waters 6, 622–636. https://doi.org/10.5268/IW-6.4.883

Wolfs, V., Meert, P., Willems, P., 2015. Modular conceptual modelling approach and software for river hydraulic simulations. Environ. Model. Softw. 71, 60–77. https://doi.org/10.1016/j.envsoft.2015.05.010

Worrall, F., Howden, N.J.K., Burt, T.P., 2013. Assessment of sample frequency bias and precision in fluvial flux calculations - An improved low bias estimation method. J. Hydrol. 503, 101–110. https://doi.org/10.1016/j.jhydrol.2013.08.048