Article

Insights from an Evaluation of Nitrate Load Estimation Methods in the Midwestern United States

Daeryong Park [1] , **Myoung-Jin Um** [2,*] , **Momcilo Markus** [3] , **Kichul Jung** [4] , **Laura Keefer** [3]
and Siddhartha Verma [5]

1   Department of Civil and Environmental Engineering, Konkuk University, Seoul 05029, Korea;
     drpark@konkuk.ac.kr
2   Department of Civil Engineering, Kyonggi University, Suwon 16227, Korea
3   Prairie Research Institute, University of Illinois, Champaign, IL 61820, USA; mmarkus@illinois.edu (M.M.);
     lkeefer@illinois.edu (L.K.)
4   Division for Integrated Water Management, Korea Environment Institute, Sejong 30147, Korea;
     kcjung@kei.re.kr
5   Department of Agricultural and Biological Engineering, University of Illinois, Urbana, IL 61801, USA;
     siddhartha.verma20@gmail.com
*   Correspondence: mum@kgu.ac.kr

Abstract: This study investigated the accuracy and suitability of several methods commonly used to estimate riverine nitrate loads at eight watersheds located southwest of Lake Erie in the Midwestern United States. This study applied various regression methods, including a regression estimator with five, six, and seven parameters, an estimator enhanced by composite, triangular, and rectangular error corrections with residual and proportional adjustment methods, the weighted regressions on time, discharge, and season (WRTDS) method, and a simple linear interpolation (SLI) method. Daily discharge and nitrate concentration data were collected by the National Center for Water Quality Research. The methods were compared with subsampling frequencies of 6, 12, and 24 times per year for daily concentrations, daily loads, and annual loads. The results indicate that combinations of the seven-parameter regression method with composite residual and rectangular residual adjustments provided the best estimates under most of the watershed and sampling frequency conditions. On average, WRTDS was more accurate than the regression models alone, but less accurate than those models enhanced by residual adjustments, except for the most urbanized watershed, Cuyahoga. SLI was the most accurate in the Vermilion and Maumee watersheds. The results also provide some information about the effects of rating curve shape and slope, land use, and record length on model performance.

Keywords: five, six and seven parameters; discharge; nitrate concentration; residual correction; SLI; Midwestern United States; WRTDS

1. Introduction

Pollutant loading, from nonpoint sources in particular, can pose significant problems in water bodies. To protect and improve the water quality of the Great Lakes, the Great Lakes Water Quality Agreement was established in 1972 [1]. The National Center for Water Quality Research (NCWQR) at Heidelberg University has been monitoring the nutrient loads of the major watersheds of Lake Erie since the 1970s, and many studies have analyzed the water quality characteristics using monitored water quality data.

Unlike river discharge (streamflow) data, water quality monitoring data typically contain temporal data gaps [2–4]. These gaps produce uncertainties in the estimation of constituent loads. Such uncertainties can be alleviated by filling in the missing water quality monitoring data using regression between water quality and discharge, commonly known as the rating curve approach. Cohn et al. [5] noted the inaccuracy of linear interpolation in estimating nutrient concentration, and attempted to rectify this issue using a multiple

log–linear regression equation based on the rating curve of nutrient concentration and river discharge. Extending Cohn's approach, Runkel et al. [6] developed load estimation (LOADEST), a program for estimating constituent loads in streams. Stenback et al. [7] applied the seven-parameter method to estimate the nitrate and total phosphorus loads at Iowa's river stations; however, the model results fluctuated and showed poor precision.

The authors noted the need for improvements in nutrient estimation methods by identifying the relationships between nutrient concentrations and discharge or other related variables. Aulenbach and Hooper [8] proposed a new composite approach in which they used residual concentrations from the regression model to adjust the model outputs. In an attempt to fine-tune the composite method, Verma et al. [9] applied the composite, triangular, and rectangular residual correction techniques, the seven-parameter regression method, and the ratio- and flow-weighted average methods. The authors evaluated model performance at different stratified sampling frequencies with a root mean square error (RMSE) based on the Upper Sangamon River and the Vermilion River in the Illinois River basin. They also demonstrated that the recommended residual correction techniques produced higher accuracies than the original models did. However, the six years of data for the Upper Sangamon River (1993 to 1999) and the ten years of data for the Vermilion River (1988 to 1999) were much shorter record periods than those available in the NCWQR datasets. To resolve this issue, Park et al. [10] applied copula models by estimating the uncertainty of the nitrate concentration and discharge in the lower Illinois River. As a result, it was deemed necessary to verify the improved error correction techniques for denser sampling and datasets spanning longer periods.

Pellerin et al. [11] determined that nitrate load estimates based on simple linear interpolation (SLI) of discrete concentration data may better characterize anomalies in nitrate concentrations than regression-based approaches, such as the LOADEST, weighted regressions on time, discharge, and season (WRTDS), and composite methods. Lee et al. [12] compared 11 nutrient simulation models, including five- and seven-parameter regression models, with the composite, SLI, and WRTDS methods in estimating the nitrate loads for four watersheds in the United States. For nitrate pollution, the LOADEST seven-parameter model with the composite method encountered smaller errors than the WRTDS and the original LOADEST five- and seven-parameter models using constituent type and sampling frequency in the error analysis. Overall, Lee et al. [12] proved that SLI could provide better simulation results than other methods for nitrate loads, which generally depend on the flow characteristics, sampling frequency, and watershed characteristics. Their study, however, applied different five-parameter regression models and did not consider combined residual correction techniques to improve the nitrate load estimates.

This study attempted to provide a better understanding of the relationship between nitrate concentration and streamflow and nitrate load estimation models through a comparison in the Lake Erie and Ohio River watersheds, which are characterized by a daily record length of approximately 25–50 years, using different sampling frequencies, monitoring record lengths, and watershed characteristics. This study also adopted a subsampling validation approach, sampling at frequencies of 6, 12, and 24 times per year, to assess and compare the performances of the models with observed data, similarly to Guo et al. [13]. Various models have been applied to estimate nitrate loads: the five-, six-, and seven-parameter regression models [5], the regression models combined with residual correction methods [8,9], SLI [12], and the WRTDS method [14]. Model performance was evaluated by estimating the daily concentration, daily load and annual load, and comparing these results with the observed data. The Nash–Sutcliffe efficiency (NSE), RMSE standard deviation ratio (RSR), and coefficient of determination (R^2) were applied as normalized error evaluation methods to analyze the priority ranking of the load estimation models for different stations and data collection frequencies.

2. Methods

2.1. Study Area

The NCWQR monitored various water quality parameters once per day including total suspended solids, total phosphorus, particulate phosphorus, dissolved reactive phosphorus, nitrate, total Kjeldahl nitrogen, and chloride. The present study used a long-period daily time series of water quality data monitored by the NCWQR at eight watersheds in the Midwestern area of the United States, as shown in Figure 1. Six of the watersheds, Cuyahoga (CY), Grand (GD), Maumee (MM), Raisin (RS), Sandusky (SD), and Vermilion (VM), were located around Lake Erie in the Lake Erie basin. Two of the watersheds, Great Miami (GM) and Muskingum (MS), were in the Ohio River basin, as shown in Table 1. The water quality monitoring period differed by watershed. The monitoring of water quality data began in 1975 for MM and SD; in the 1980s for CY, GD, and RS; the 1990s for GM and MS; and in 2001 for VM.
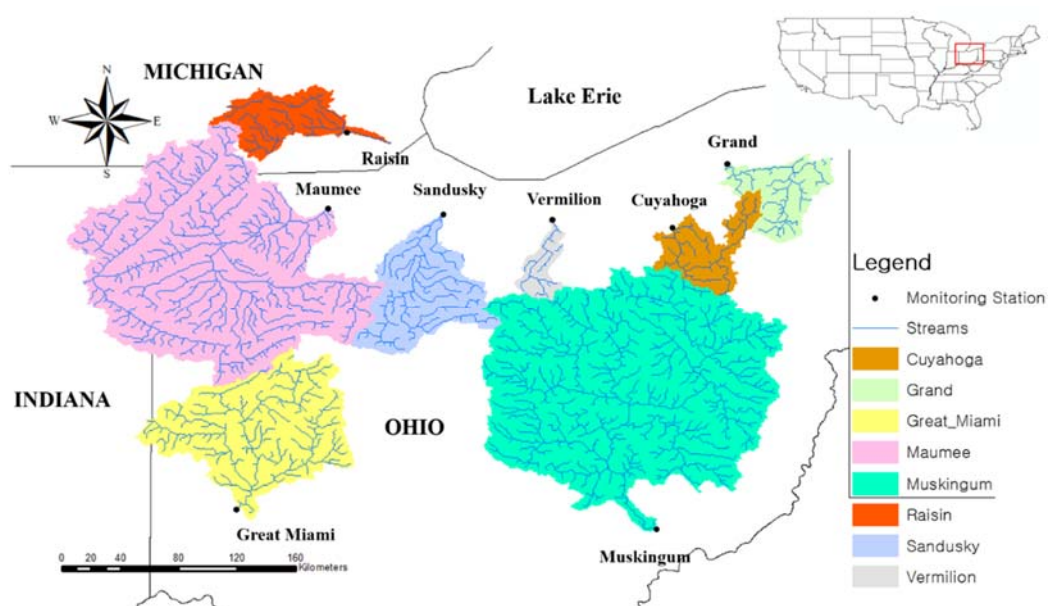

Figure 1. Location of gaging stations and the corresponding watersheds used in this study.

Table 1. Characteristics of applied gaging stations and the corresponding watersheds [15,16].

	USGS Station Number	Monitoring Period	Watershed Size (km^2)	Years Selected	Land Use (%)		
					Agriculture	Urban	Wooded
Lake Erie basin							
Cuyahoga (CY)	04208000	1982–2017	1843	36	17	47	35
Grand (GD)	04212100	1989–2006	1758	18	37	10	52
Maumee (MM)	04193500	1975–2017	16,427	43	81	11	8
Raisin (RS)	04176500	1982–2017	2755	36	72	11	16
Sandusky (SD)	04198000	1975–2017	3285	43	83	9	8
Vermilion (VM)	04199500	2001–2008	697	8	71	1	26
Ohio River basin							
Great Miami (GM)	03271601	1996–2017	6953	22	82	5	10
Muskingum (MS)	03150000	1995–2017	19,208	23	52	2	43

The eight selected watersheds from the NCWQR database are located southwest of Lake Erie in Michigan, Indiana, and Ohio. Land-use data for the CY, GD, MM, RS, and SD basins are shown in Table 1 [15,16]. Other land-use data for the VM, GM, and MS basins were reported by Verma [17]. The shortest monitoring length is 8 years for VM, and the longest is 43 years for MM and SD. According to Table 1, except for CY and GD, six of the eight watersheds report agriculture as the prevailing land use. The primary land uses at

the CY and GD stations are urban and wooded land, respectively [15]. The MM and MS stations each cover a watershed area of over 10,000 km^2, whereas the other six watersheds each cover an area less than 10,000 km^2, and the smallest watershed area, station VM, spans approximately 700 km^2. The relationship between the discharge and nitrate concentration data for the eight stations is shown in Figure 2. The data relations at all stations are positive exponential, except at station CY (Figure 2a), which has a negative exponential relation between discharge and nitrate concentration. The R^2 values of the exponential regressions at all eight stations are between 0.1 and 0.68.
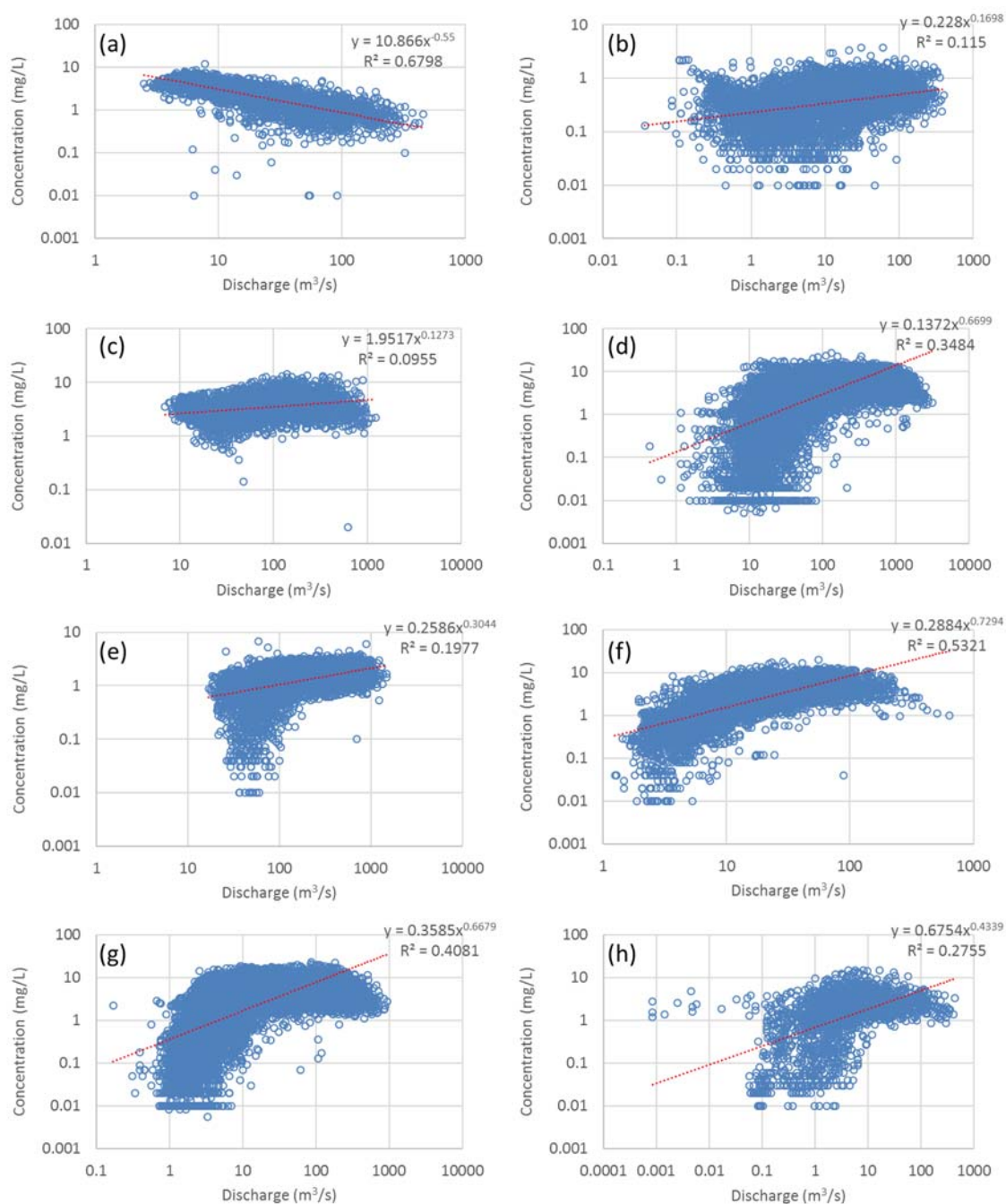

Figure 2. Retrieved nitrate and streamflow data based on station: (**a**) Cuyahoga, (**b**) Grand, (**c**) Great Miami, (**d**) Maumee, (**e**) Muskingum, (**f**) Raisin, (**g**) Sandusky, and (**h**) Vermilion.

2.2. Sampling Frequency

This study applied three sampling frequencies for the observed concentration. A resampling procedure to evaluate the concentrations for models of observed concentration data was used to assess and compare the performances of the models with the observed sampling data, and the accuracies between the model simulation and all observed data are evaluated in Figure 3. The low, medium, and high frequencies were represented by sampling times of 6, 12, and 24 per water year, respectively. These time intervals likely correspond with many water quality monitoring program protocols around the country [18,19]. Sampling months are usually not evenly distributed across the year. To ensure even sampling intervals of 6, 12, and 24 sampling times per year (water year), this study averaged the results of 1000 iterations of each sampling frequency for all the applied estimation methods, as shown in Table 2.
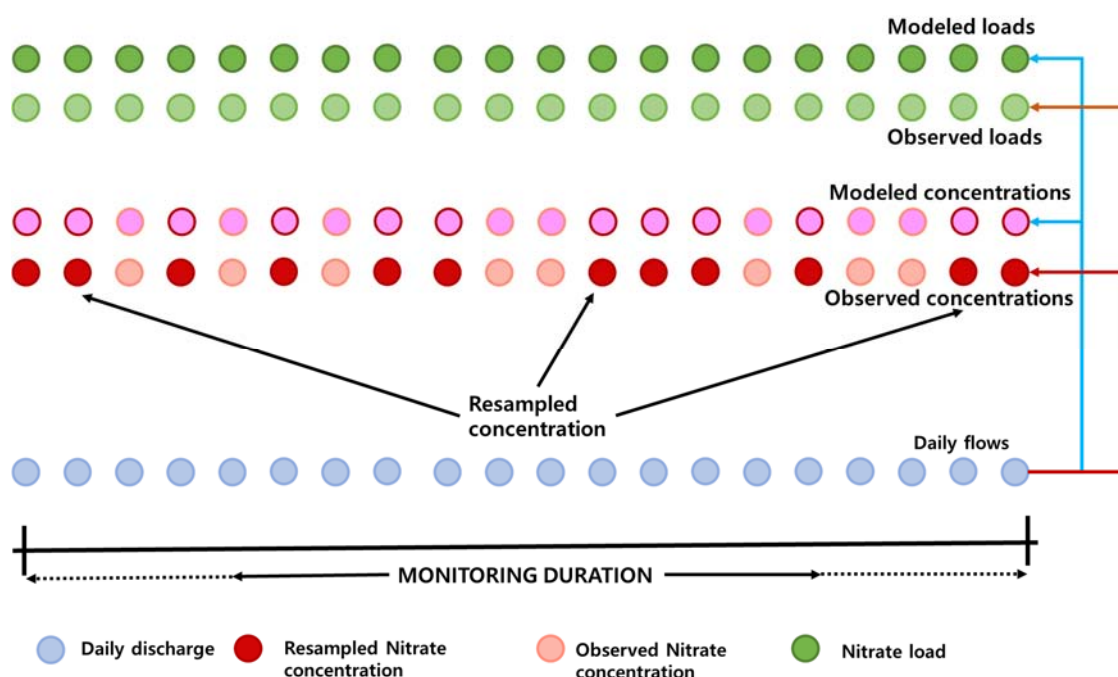
Figure 3. Schematic diagram of comparisons between 23 estimation methods and "TRUE" estimation based on observed data with daily concentrations and loads.

Table 2. Applied daily concentration, daily load and annual load estimation methods in this study and their symbols.

Estimation Methods	Regression (RS)	Composite Residual (CR)	Composite Proportional (CP)	Triangular Residual (TR)	Triangular Proportional (TP)	Rectangular Residual (RR)	Rectangular Proportional (RP)
5-parameter	5RS	5CR	5CP	5TR	5TP	5RR	5RP
6-parameter	6RS	6CR	6CP	6TR	6TP	6RR	6RP
7-parameter	7RS	7CR	7CP	7TR	7TP	7RR	7RP
Simple Linear Interpolation	SLI						
WRTDS	WRT						

2.3. Water Quality Estimation

The Water Quality Laboratory database monitors daily or subdaily pollutant concentrations and flow data for each of the eight watersheds, as shown in Table 1. Based on the monitoring data for each water year, the annual nitrate loads were estimated via a

simple numeric integration of daily loads [8,9,13,20–23]. The daily loads were estimated as follows:

$$Load = k \int C(t)Q(t)\mathrm{d}t \tag{1}$$

where k is the unit conversion factor, C is the pollutant concentration, Q is the flow, and $\mathrm{d}t$ is the duration of the sample. Figure 4 shows a schematic of the load evaluation in this study. This study first applied three regression models (5-, 6-, and 7-parameter). Next, each of the regression models incorporated six residual correction models. Finally, the WRTDS and SLI models were applied to the same dataset. The 23 estimation methods are listed in Table 2 and are explained in the following text.
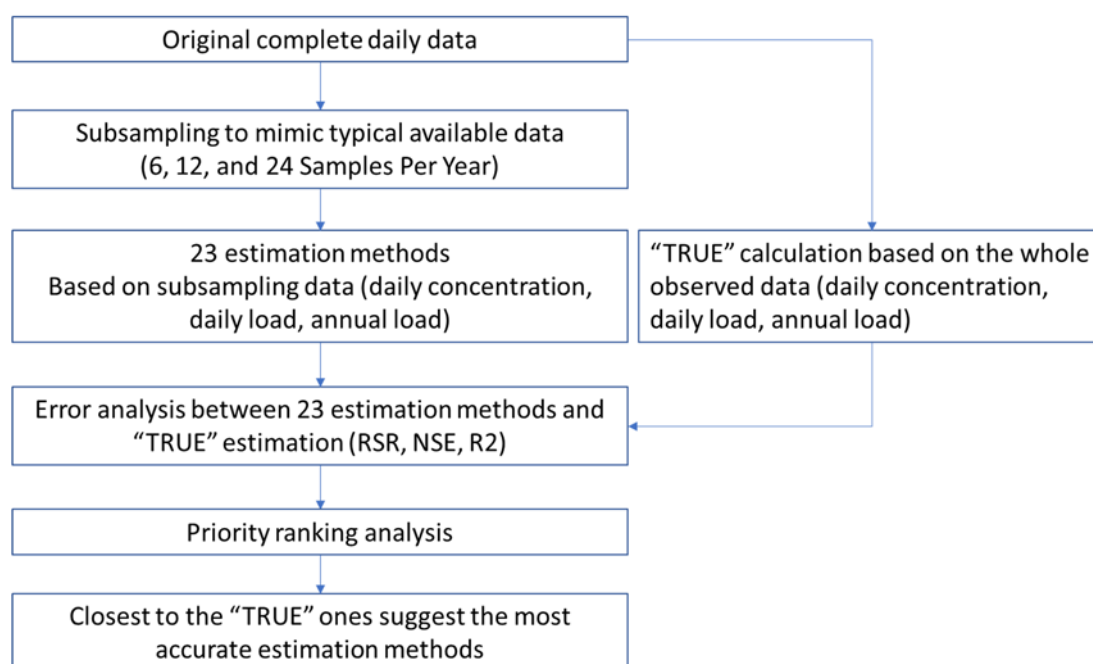

Figure 4. Schematic of the load evaluation process in this study.

2.4. Regression Model with Five, Six, and Seven Parameters

Log-transformed daily flow and decimal time, including seasonal factors with the 7-parameter regression model, were considered to estimate the log-transformed daily concentration, following Cohn et al. [24], as shown below:

$$ln(C) = \beta_0 + \beta_1 ln\left(\frac{Q}{Q'}\right) + \beta_2 \left[ln\left(\frac{Q}{Q'}\right)\right]^2 + \beta_3 (T - T') + \beta_4 (T - T')^2 + \beta_5 \sin(2\pi T) + \beta_6 \cos(2\pi T) + \varepsilon \tag{2}$$

where ln denotes the natural log, β_i refers to the fitting parameters, T denotes the decimal time, Q indicates the streamflow, Q' and T' denote the centering variables used in the model, and ε refers to a model residual, as shown below [1,9]:

$$T\prime = \overline{T} + \frac{\sum_{i=1}^{n}(T_i - \overline{T})^3}{2\sum_{i=1}^{n}(T_i - \overline{T})^2} \tag{3}$$

$$Q\prime = \overline{Q} + \frac{\sum_{i=1}^{n}(Q_i - \overline{Q})^3}{2\sum_{i=1}^{n}(Q_i - \overline{Q})^2} \tag{4}$$

$$\overline{T} = \frac{\sum_{i=1}^{n} T_i}{n} \tag{5}$$

$$\overline{Q} = \frac{\sum_{i=1}^{n} Q_i}{n} \tag{6}$$

where T_i and Q_i refer to T and Q at time i, respectively, and n is the number of observed data points. The fitting parameters β_i are calculated via multiple regression in MATLAB software (Mathworks, Natick, MA, USA).

In the present work, the seven parameters in Equation (2) can be simplified by removing the terms corresponding to the linear (β_3) and quadratic (β_4) terms without considering the effect of time trends on concentration, because time trends are less significant than seasonality and linear and quadratic dependence on flow, and by applying the modified regression equations to the 5- and 6-parameter equations, as follows:

$$ln(C) = \beta_0 + \beta_1 ln\left(\frac{Q}{Q'}\right) + \beta_2 \left[ln\left(\frac{Q}{Q'}\right)\right]^2 + \beta_3(T - T') + \beta_5 \sin(2\pi T) + \beta_6 \cos(2\pi T) + \varepsilon \tag{7}$$

$$ln(C) = \beta_0 + \beta_1 ln\left(\frac{Q}{Q'}\right) + \beta_2 \left[ln\left(\frac{Q}{Q'}\right)\right]^2 + \beta_5 \sin(2\pi T) + \beta_6 \cos(2\pi T) + \varepsilon \tag{8}$$

All fitting parameters (β_i) for each station are different from the resampling data because Equations (3)–(6) are changed by the data.

2.5. Weighted Regressions on Time, Discharge, and Season (WRTDS) and Simple Linear Interpolation (SLI)

The WRTDS method was proposed by Hirsch et al. [14], and one of its important features is that newly collected data have only minimal effects on previous season and water quality trends. The performance of WRTDS is significantly enhanced when there are over 200 observations for at least 20 years [14]. The WRTDS uses the following equation to calculate daily concentration:

$$ln(C) = \beta_0 + \beta_1 T + \beta_2 \ln(Q) + \beta_3 \sin(2\pi T) + \beta_4 \cos(2\pi T) + \varepsilon \tag{9}$$

In the WRTDS method, a unique set of coefficients is estimated for each combination of daily streamflow (Q) and decimal time (T) in the recorded period. For every combination of Q and t, the fitting parameters in Equation (9) are estimated by weighted regression. The weights of each observation in the calibration dataset are based on the distance in time, streamflow, and season between the observed daily streamflow and decimal time. The WRTDS involves a flow normalization method that reduces the effect of random stream variation on the estimated annual flux. The flow normalization estimates of concentration are computed for a given date using weighted regression to estimate the concentration on that date with the streamflow value set to each of the historical streamflow values for that day. The flow normalization concentration on that date is then calculated as the mean of the estimated concentration values from each of these weighted regressions [14].

SLI involves computing pollutant concentrations for nonsampled days via linear interpolation. This method divides the observed daily concentration, daily load (estimated by the product of daily concentration and daily discharge), and annual load data by the sum of the daily load for the SLI of the daily concentration, daily load, and annual load, respectively. SLI is generally sensitive to the timing and number of individual samples.

2.6. Shape of the Residual Adjustments

This study used a method characterized by gradual linear interpolation and triangular- and rectangular-shaped distributions to allocate adjustments around a known residual between the observed and model-simulated values from Figure 5. It also used the composite method defined by Aulenbach and Hooper [8], who demonstrated that this method improved the load estimation accuracy. A schematic of the approach is shown in Figure 5a. Verma et al. [9] developed triangular-shaped (Figure 5b) and rectangular-shaped (Figure 5c) distributions to distribute residual adjustments in the proximity of a known residual. The triangular distribution computes the magnitude of the residuals as the maximum value on the observed days and linearly changes it to zero at midpoints of the temporal intervals

between two continuous observation days. The computed residual value on the nearest observed day in the rectangular distribution is considered the residual for a missing observation day. These residual correction methods account for the temporal correlation in modeling errors by assigning residuals for missed observation days based on known residuals between the model-simulated and observed values, such as daily concentration on proximate sampled days. More detailed descriptions of the triangular and rectangular distributions of the residual values can be found in Verma et al. [9].
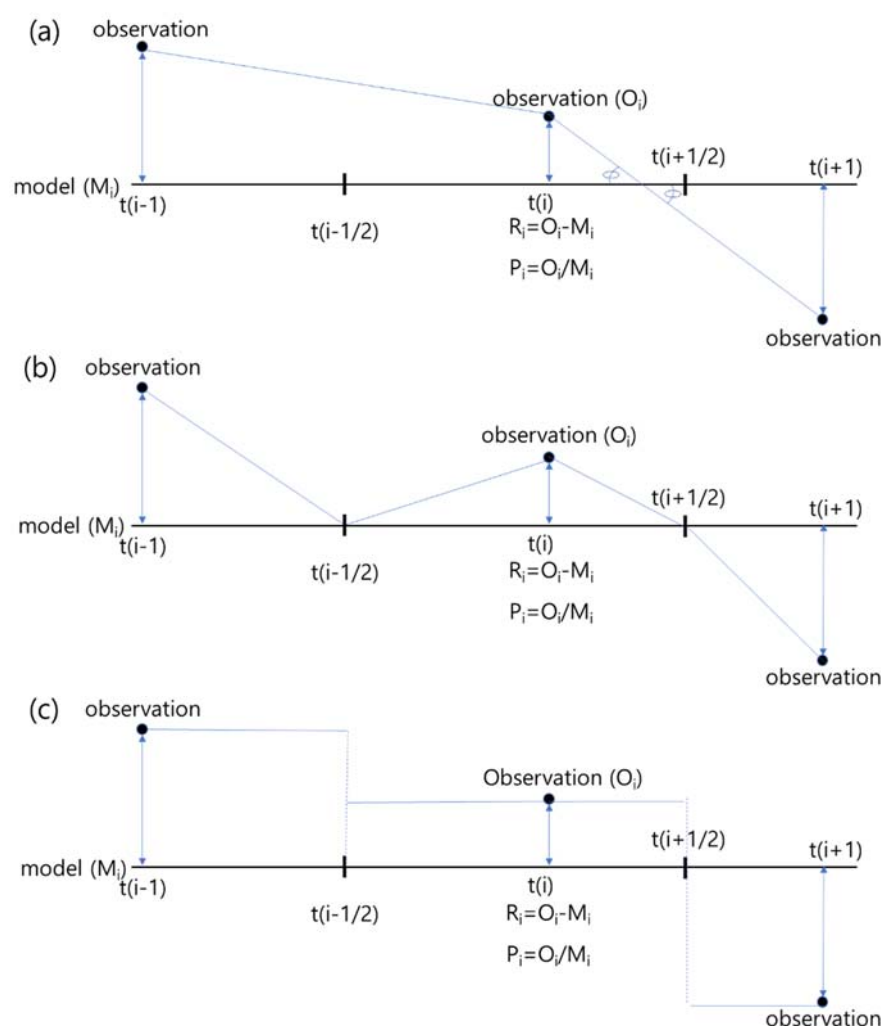
Figure 5. Schematics of the three error-correction distributions for residual and proportional adjustments: (**a**) composite method (piecewise linear interpolation), (**b**) triangular distribution, and (**c**) rectangular distribution.

2.7. Residual and Proportional Adjustment Methods

The present analysis applied residual and proportional adjustment methods. The residual adjustment method proposed by Aulenbach and Hooper [8] was applied on the sampled days by finding the difference between the model-simulated and observed concentrations. The midpoints of the time intervals between all pairs of consecutively sampled days were then determined. These midpoints were then set as vertices in the estimation of the residuals on adjacent unsampled days:

$$R_i = [O_i - M_i] \tag{10}$$

where R_i is the model concentration obtained using the residual adjustment method, O_i is the observed or adjusted observed concentration from the adjusted method, M_i is the

original model concentration, and i is a time step. This adjustment is referred to as the residual adjustment method.

Verma et al. [9] proposed another correction technique called the proportional adjustment method, as follows:

$$P_i = \left[\frac{O_i}{M_i} \right] \qquad (11)$$

where P_i is a concentration obtained using the proportional adjustment method. The proportional adjustment is defined as the ratio between the observed concentration and the measured concentration on a sample day. Once the set of proportional concentrations or load ratios was computed for all sample observations, the composite, rectangular, and triangular distributions were again applied to allocate the concentrations or loads on proximate sample days.

2.8. Accuracy Evaluation

The study used three normalized indicator methods to evaluate the performance of the model [25–30]. The three indicators were calculated using the same dataset. The RSR normalizes the RMSE with the standard deviation of the observation as a normalization factor [31]. The RSR ranges from 0 to any positive value. The optimal value is 0. The equation for the RSR is as follows:

$$\text{RSR} = \frac{\sqrt{\sum_{i=1}^{n} \left(Y_i^{sim} - Y_i^{obs} \right)^2}}{\sqrt{\sum_{i=1}^{n} \left(Y_i^{sim} - \overline{Y^{obs}} \right)^2}} \qquad (12)$$

where Y_i^{sim} is the model value of the ith time point, Y_i^{sim} is the measured value of the ith time point, and Y_i^{sim} is the mean of the measured values.

The NSE, proposed by Nash and Sutcliffe [32], is a normalized indicator of model performance that is advantageous for long-term simulations [31]. Its value ranges from any negative number to 1, and the optimal value is 1. The equation for the NSE is as follows:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n} \left(Y_i^{sim} - Y_i^{obs} \right)^2}{\sum_{i=1}^{n} \left(Y_i^{sim} - \overline{Y^{obs}} \right)^2} \qquad (13)$$

R^2 represents the magnitude of the relationship between the observed and measured data. R^2 has been used in many studies, although it is oversensitive to extreme values [31]. R^2 ranges between 0 and 1, and the optimal value is 1. The equation for R^2 is as follows:

$$R^2 = \left\{ \frac{\sum_{i=1}^{n} \left(Y_i^{sim} - \overline{Y^{sim}} \right) \left(Y_i^{obs} - \overline{Y^{obs}} \right)}{\sqrt{\sum_{i=1}^{n} \left(Y_i^{sim} - \overline{Y^{sim}} \right)^2} \sqrt{\sum_{i=1}^{n} \left(Y_i^{obs} - \overline{Y^{obs}} \right)^2}} \right\}^2 \qquad (14)$$

*2.9. Priority Ranking of 23 Estimation Methods*

In this study, we investigated the priority rankings of the estimation methods. Equation (15) combines all three error analysis methods applied in this study, as follows (modified from Pellerin et al., 2014):

$$E = \sqrt{(R^2 - 1)^2 + (NSE - 1)^2 + RSR^2} \qquad (15)$$

where E is the evaluation combining the three error evaluation methods. The optimal value is 0, and the worst possible value is $+\infty$. Using this equation, we can obtain priority rankings that depend on the stations and sampling frequencies.

3. Results

The five-, six-, and seven-parameter regression methods, their combinations with correction methods, and the SLI and WRTDS models were evaluated to assess their performances, as shown in Table 2. A total of 23 models were considered in the analysis. The performance differed based on various watershed characteristics, sampling frequencies, and target variables, e.g., daily concentrations, daily loads, or annual loads. Figure 6 shows, for example, the daily concentration estimation with seven-parameter regression incorporating residual adjustment methods, as well as the WRTDS method using 12 observed concentration samplings per year. Seven-parameter regression does not match the observed concentration, but seven-parameter regression with the adjustment methods and WRTDS matches the observed concentration. In particular, six seven-parameter regression methods combined with adjustment methods and WRTDS generate different concentrations. This is because the interpolation algorithms between observed data are different for the six adjustment methods and WRTDS.
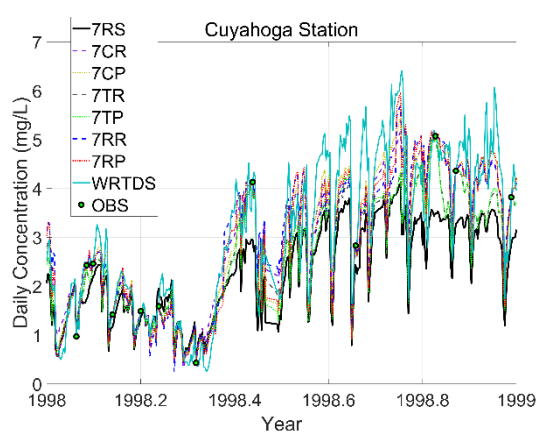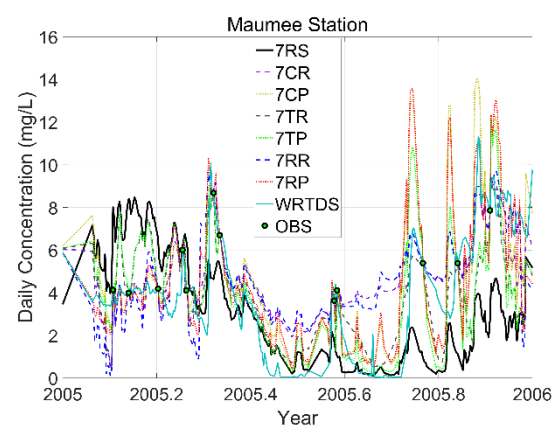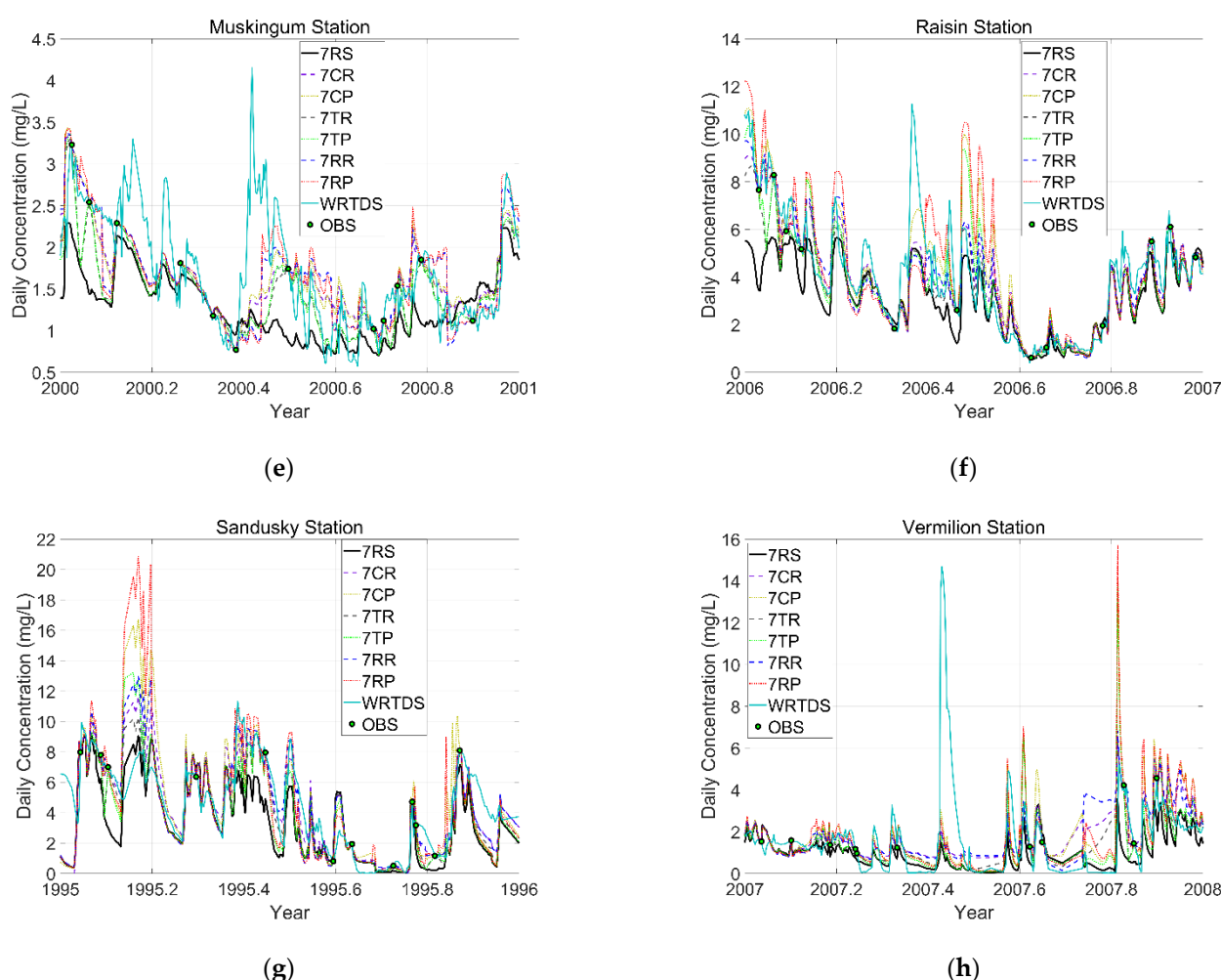
(**a**)
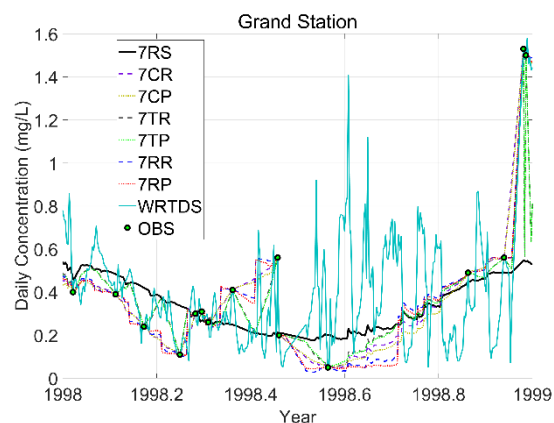
(**b**)
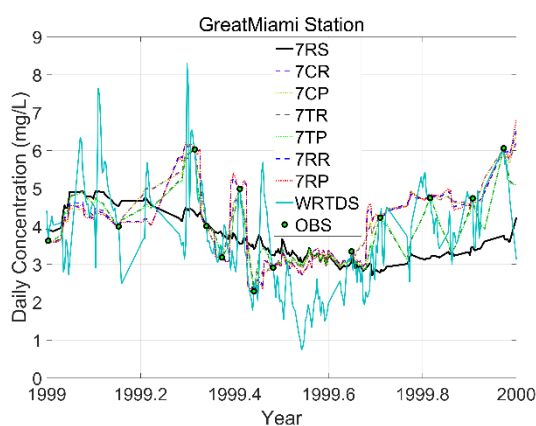
(**c**)

(**d**)

Figure 6. *Cont.*

Figure 6. Daily concentration estimates for 7-parameter estimation, 7-parameter with adjustment methods and WRTDS under the 12 observations per year condition in a particular year: (**a**) Cuyahoga, (**b**) Grand, (**c**) Great Miami, (**d**) Maumee, (**e**) Muskingum, (**f**) Raisin, (**g**) Sandusky, and (**h**) Vermilion.

### 3.1. Uncertainty Analysis of the Model Estimation Performance

Figure 7 presents the error analysis of all daily concentration, daily load and annual load performance results at the eight stations. The error bars for CY, MM, RS, and SD are smaller than those of the other stations. This result indicates that the correlation between the observed discharge and nitrate data is relatively strong, as R2 ranges from 0.35 to 0.68 in Figure 7, with smaller error bar widths. However, GD, GM, MS, and VM show larger error bar widths for the RSR and NSE evaluations. The R2 values for these four stations are as low as 0.01–0.28, as shown in Figure 7.

Figure 8 shows the error bars for the 23 estimation methods in Table 2 after combining the daily concentration, daily load, and annual load estimates for the eight stations. The 5RS, 6RS, 7RS, and WRT approaches produce greater errors than the other approaches for stations GD, GM, and VM, as shown in Figure 8. However, while the means and variances of the five-, six- and seven-parameter regression methods incorporating residual adjustments decrease for GD, GM, and VM compared with those for the 5RS, 6RS, 7RS, and WRT methods, they appear to be similar for the other five stations. In particular, with the exception of the original estimation methods, such as 5RS, 6RS, and 7RS and the WRTDS approaches, the SLIs for all eight stations show error bars similar to those of the five-, six-, and seven-parameter regression methods incorporating the adjustment methods. Error analyses with the RSR and NSE show similar values, but the error bars for R2 illustrate different results because R2 is sensitive to high/extreme values and insensitive to additive

and proportional differences between the observed and model data. On the other hand, the NSE and RSR approaches can be characterized by including normalization factors, and are relatively robust to continuous long-term data [28].
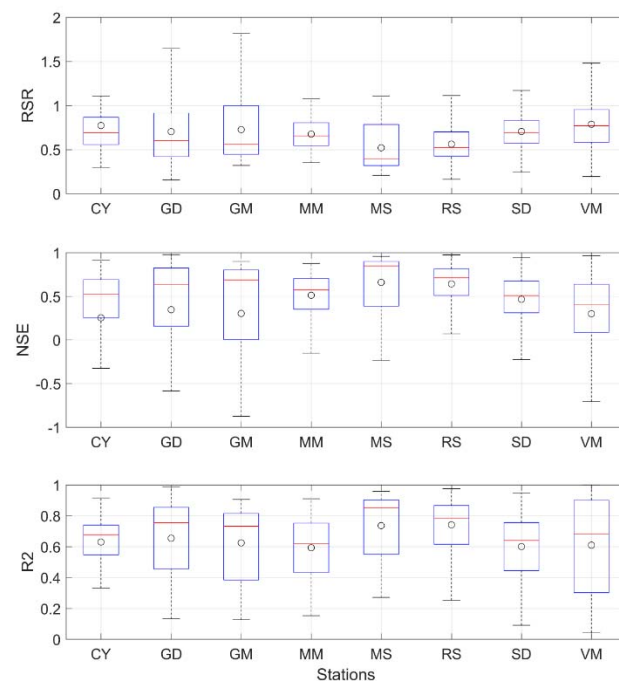
Figure 7. Error analysis of all combining daily concentration, daily load and annual load performance results at eight stations (Note: circles indicate the mean value).
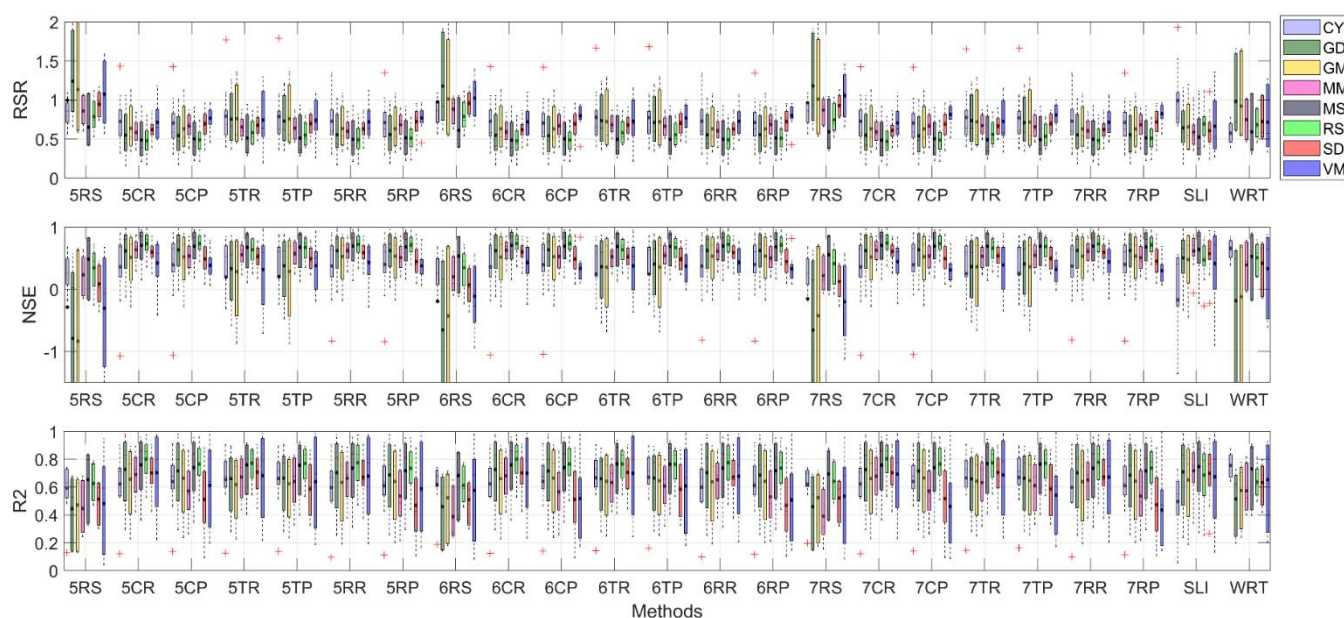
Figure 8. Error analysis results of 23 estimation methods for individual stations (note: black points indicate the mean value).

Figure 9 shows the error bars when combining 23 estimation methods based on three resampling frequencies (6, 12, and 24 resampling per year) for the simulated daily concentration, daily load, and annual load conditions. In general, the estimates of daily concentration by all approaches show greater mean errors than those of the daily load and annual load estimates. However, the daily load and annual load estimates show the similar

means and variances determined by all approaches. This result occurs because the load estimate includes discharge data, and the load is estimated with a lower error variance than that of the observed concentration. In particular, the daily concentrations for 5RS, 6RS, and 7RS have greater mean errors and variances than those of the other approaches. This result also indicates that estimation approaches incorporating adjustment methods are better than the original estimation methods themselves. The SLI approach produces mean errors and variances for daily concentration, daily load, and annual load that are similar to those of the EGR approach.
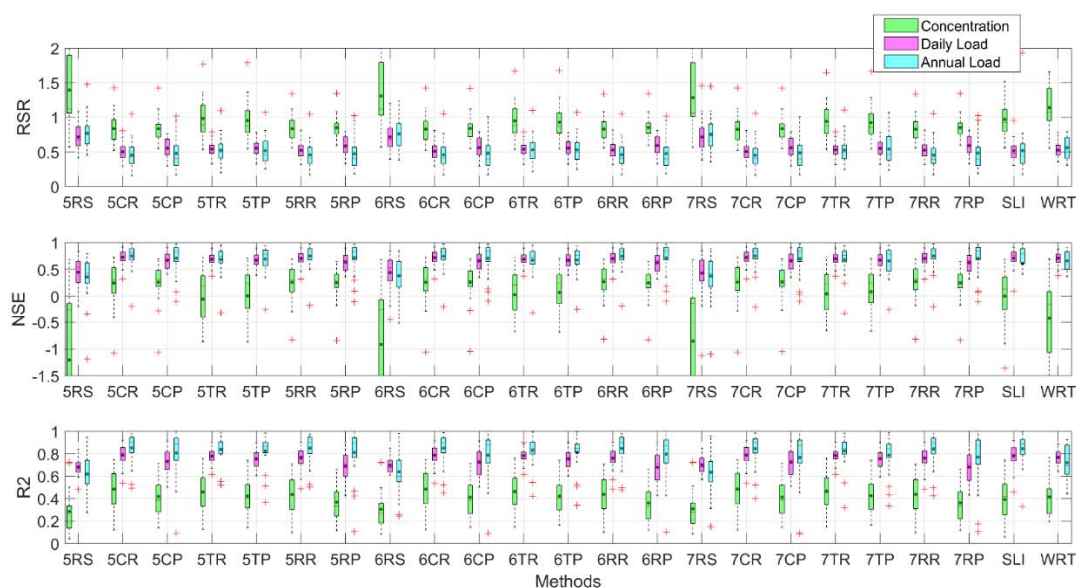
Figure 9. Error analysis results of 23 estimation methods based on concentration, daily load, and annual load (note: black points indicate the mean value).

Figure 10 presents the error analyses of the combination approaches for the daily concentration, daily load, and annual load estimations for all eights stations for the low, medium and high (6, 12, and 24) sampling frequencies using the 23 estimation methods. The variances in the error bars and the mean decrease as the sampling frequency increases. Specifically, the 5RS, 6RS, and 7RS methods have high error bar variances. The remaining 20 estimation methods show small variances in the error bars. Moreover, the means and variances in the error bars for the WRT method are similar, irrespective of the sampling frequency. The SLI approach provides remarkably smaller error bars as the sampling frequency increases, which indicates that SLI is more sensitive to the sampling frequency than the other 22 estimation methods.
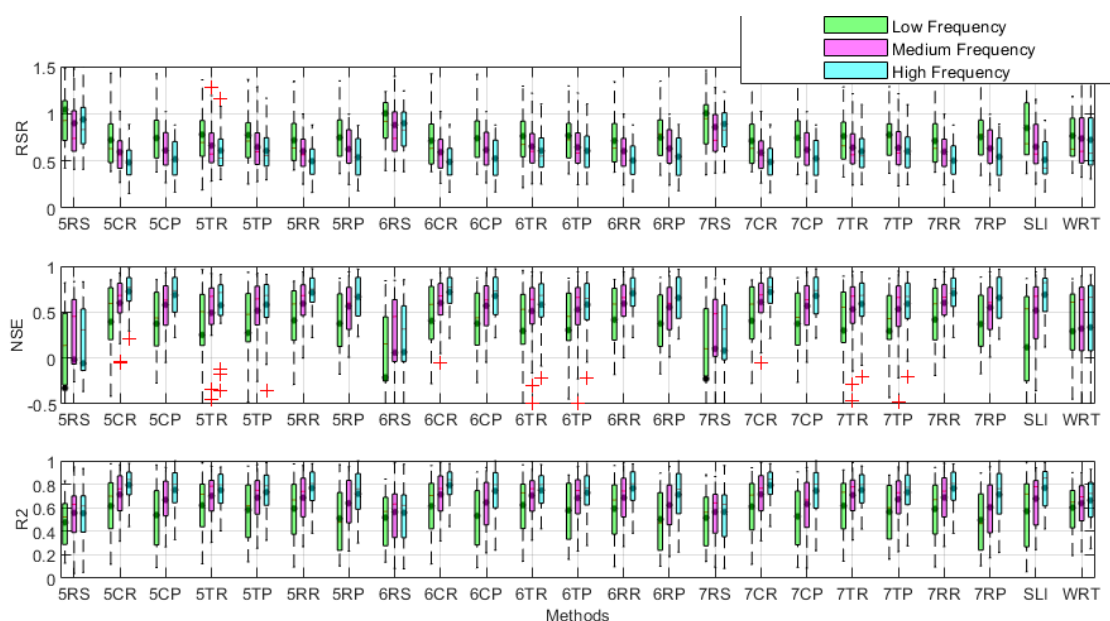
Figure 10. Error analysis results of 23 estimation methods based on sampling frequencies and combined concentration, daily load, and annual load (note: black points indicate the mean value).

3.2. Priority Rankings of Estimation Methods Based on Characteristics of the Stations

The present work investigated combinations of error evaluations for the daily concentration, daily load, and annual load estimation approaches depending on the sampling frequency for each station (Figure 11 and Table 3). In Figure 11, the NSE and R2 are represented via the *x* and *y* axes, respectively, whereas the RSR is expressed by the *z* axis. Table 3 lists the priority rankings of the estimation methods and shows that the 7CP for the GD and GM stations, SLI for the VM station, 7CR for the RS and SD stations, WRT for the CY station, and 6CR for the MS station provide the lowest error. The top ten rankings include the 6CR method for all eight stations and the 5CR, 7CR, and 7RR approaches for seven stations.

Table 3. Priority rankings (Equation (15)) of estimation methods based on the combined daily and annual load estimations in each watershed.

Ranking	Cuyahoga (CY)	Grand (GD)	Great Miami (GM)	Maumee (MM)	Muskingum (MS)	Raisin (RS)	Sandusky (SD)	Vermilion (VM)
1	WRT	7CP	7CP	SLI	6CR	7CR	7CR	SLI
2	6CP	6CP	6CP	5CR	5CR	6CR	6CR	7CR
3	7CP	5CP	7RP	6CR	7CR	5CR	5CR	7RR
4	5CP	7CR	6RP	7CR	7TR	7RR	SLI	5CR
5	5RP	6CR	5CP	5RR	7TP	7CP	7RR	6CR
6	6RP	5CR	7CR	6RR	6TR	6CP	6RR	5RR
7	7RP	7RR	5RP	7RR	6TP	5CP	5RR	6RR
8	5CR	6RR	6CR	5TP	6RR	6RR	7TR	7TR
9	6CR	5RR	7RR	5TR	5RR	5RR	5TR	6TR
10	5RR	6RP	6RR	7TP	7RR	7RP	6TR	5TP

Figure 11. *Cont.*

(**g**)

(**h**)
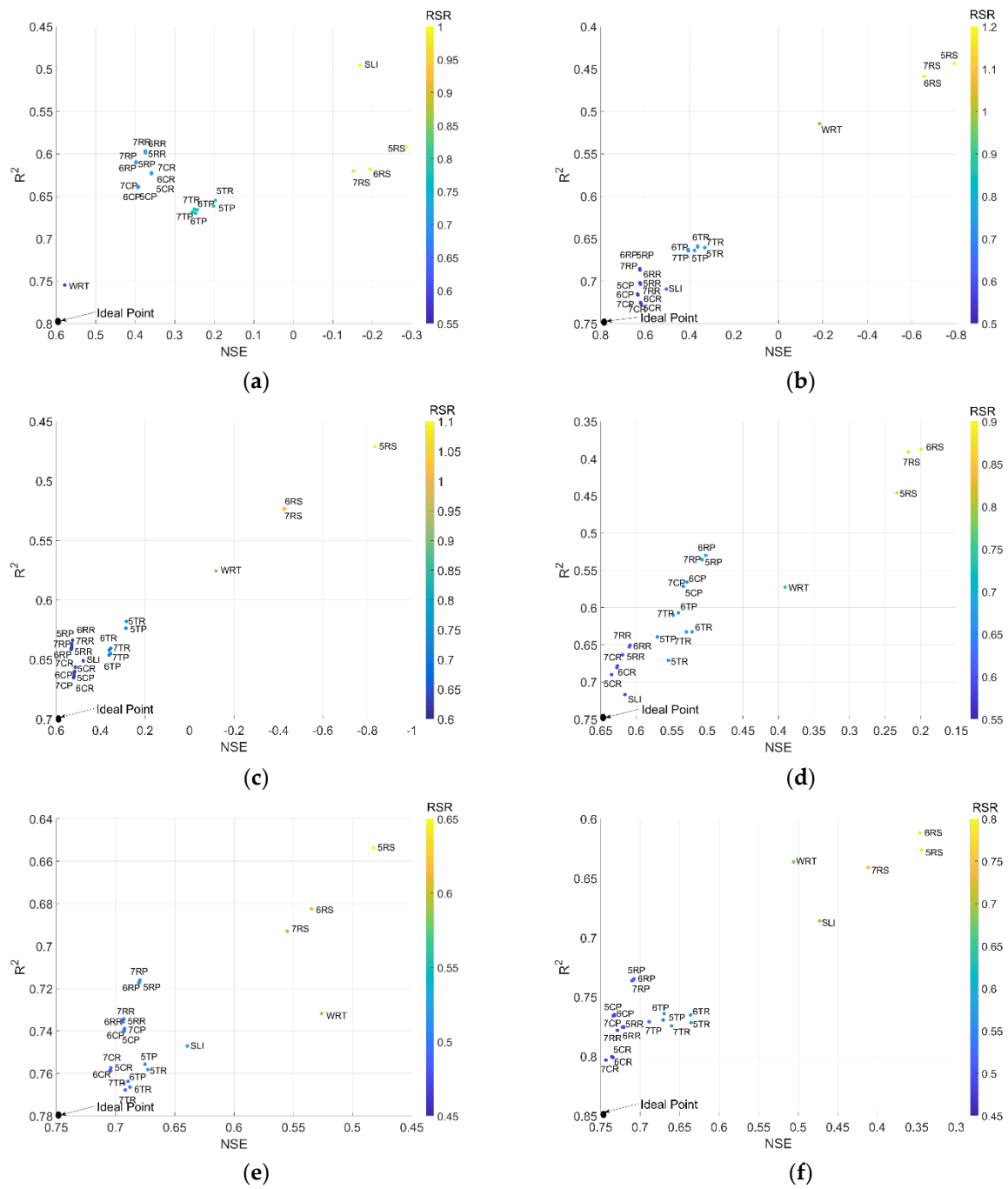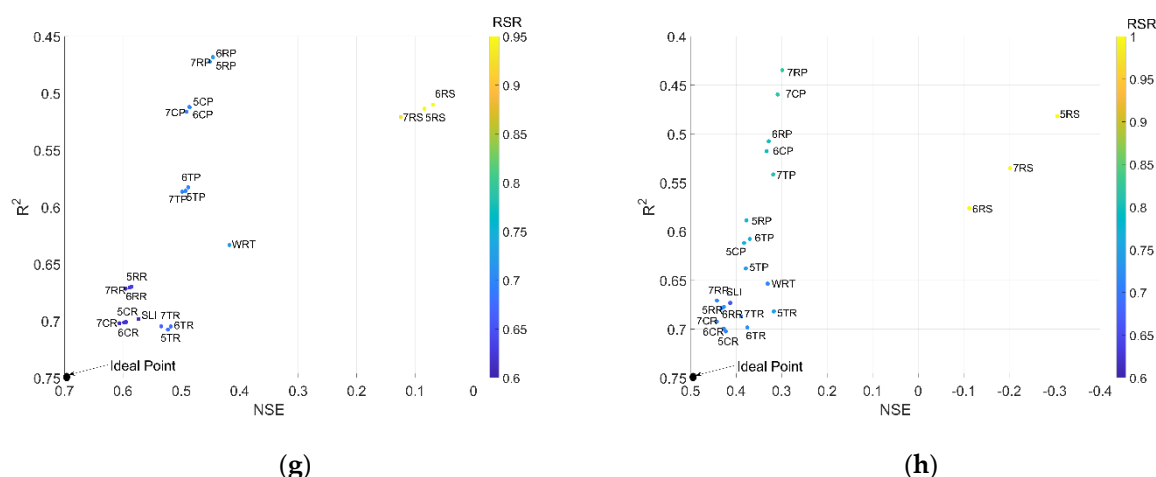
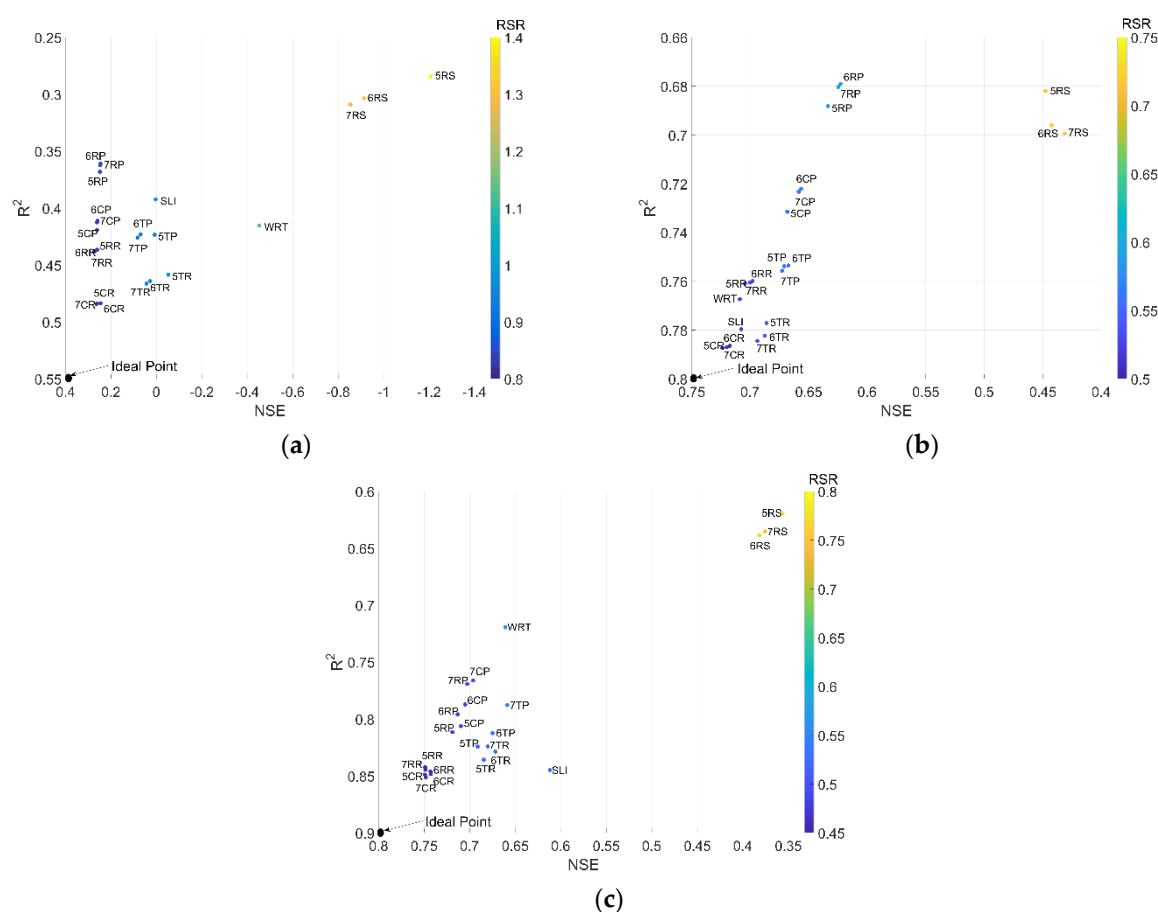**Figure 11.** Evaluation of 23 estimation methods combining three error evaluation methods based on the station: (**a**) Cuyahoga, (**b**) Grand, (**c**) Great Miami, (**d**) Maumee, (**e**) Muskingum, (**f**) Raisin, (**g**) Sandusky, and (**h**) Vermilion.

The most accurate or ideal value would be one, with larger NSE and R2 and smaller RSR values indicating better performance (Figure 11 and Table 3). The most accurate estimation method was the WRT method at CY station (Figure 11a). As noted earlier, the CY station is characterized as a highly impervious area in an urban watershed. As shown in Table 3, the SLI method is the most accurate estimation method for the MM and VM stations, and the fourth most accurate for the SD station. However, the SLI approach is not included among the top ten methods for the other five stations. This result indicates that the SLI method is less reliable for unknown watersheds or monitoring data lengths.

The priority rankings of the estimation methods using unsampled observed and simulated concentrations applied to the daily concentration, daily load, and annual load are evaluated in Figure 12 and Table 4. Overall, the five-, six-, and seven-parameter estimation methods, combined with the CR, RR, and CP methods, are ranked among the top ten, as shown in Table 4. In particular, the 7CR method is ranked first with regard to daily concentration and annual load and second with regard to daily load. The 5CR method is ranked fourth, first, and second with regard to daily concentration, daily load, and annual load, respectively. Moreover, the 6CR method is ranked second, third, and fifth in terms of daily concentration, daily load, and annual load, respectively.

Table 4. Priority rankings (Equation (15)) of estimation methods based on daily concentration, daily load, and annual load estimations.

| Ranking	Daily Concentration	Daily Load	Annual Load
1	7CR	5CR	7CR
2	6CR	7CR	5CR
3	7RR	6CR	7RR
4	5CR	SLI	5RR
5	6RR	EGR	6CR
6	5RR	5RR	6RR
7	5CP	7RR	5RP
8	7CP	7TR	6RP
9	6CP	6RR	5CP
10	5RP	6TR	6CP

Figure 12. Evaluation of 23 estimation methods combining three error evaluation methods based on (**a**) daily concentration, (**b**) daily load, and (**c**) annual load.

Figure 13 and Table 5 show the performance priority rankings for the 23 estimation methods when combining daily concentration, daily load, and annual load based on sampling frequency. For the low and medium sampling frequencies, the top three rankings are assigned to the five-, six-, and seven-parameter load estimation equations incorporating the CR and RR adjustments. For a high sampling frequency, the SLI method provides the most accurate results, followed by the 5-, 6-, and 7RR and CR methods. In contrast, SLI does not appear in the top ten lists for either the low- or medium-frequency conditions. This result indicates that SLI can be used for load estimation when sufficient observed data are available. On the other hand, the CR and RR methods can provide reliable load estimation results even when the quantity of data appears insufficient.

Table 5. Priority rankings (Equation (15)) of estimation methods combining daily concentration, daily load and annual load estimations with eight stations based on sampling frequency.

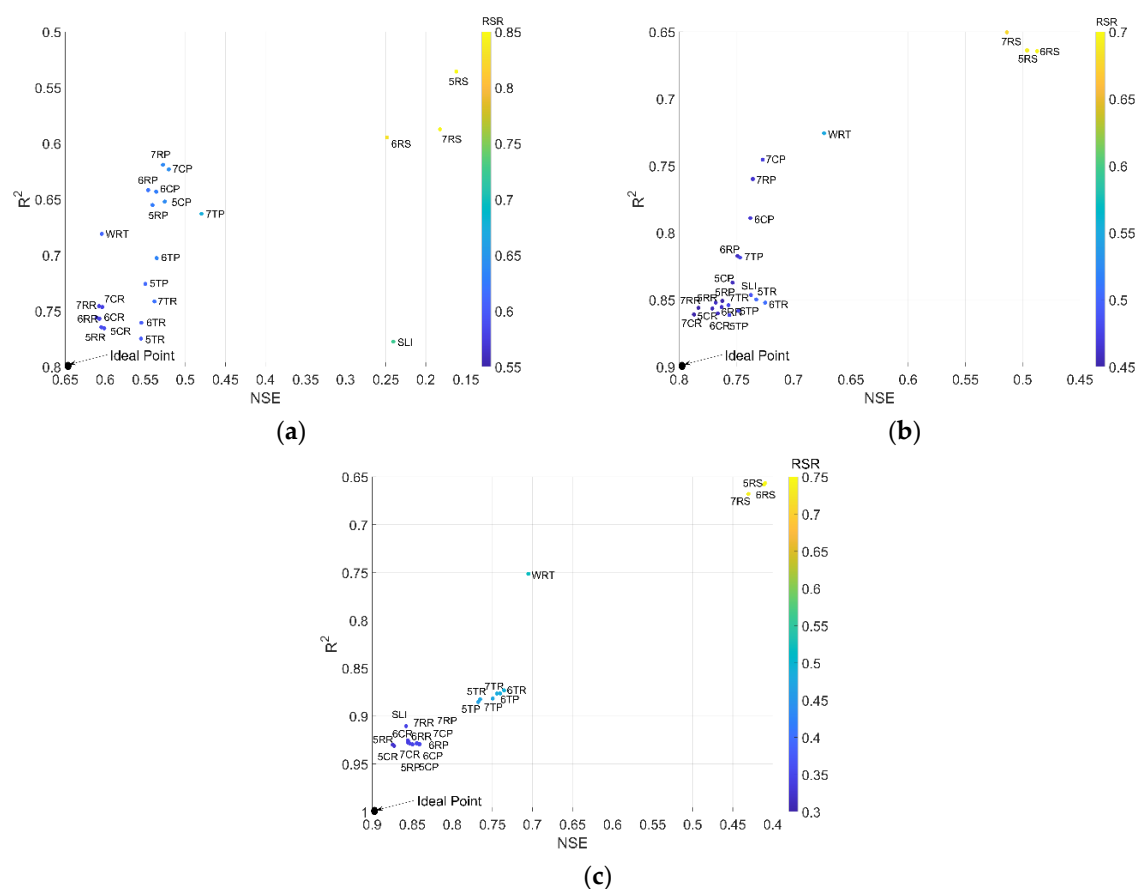| Ranking	Low Frequency	Medium Frequency	High Frequency
1	6CR	7CR	SLI
2	7CR	6CR	5RR
3	5CR	5CR	7RR
4	6RR	7RR	6RR
5	7RR	5RR	5CR
6	5RR	6RR	7CR
7	6TR	7TR	6CR
8	7TR	5CP	5CP
9	5TR	6TR	7CP
10	WRT	6TP	6CP
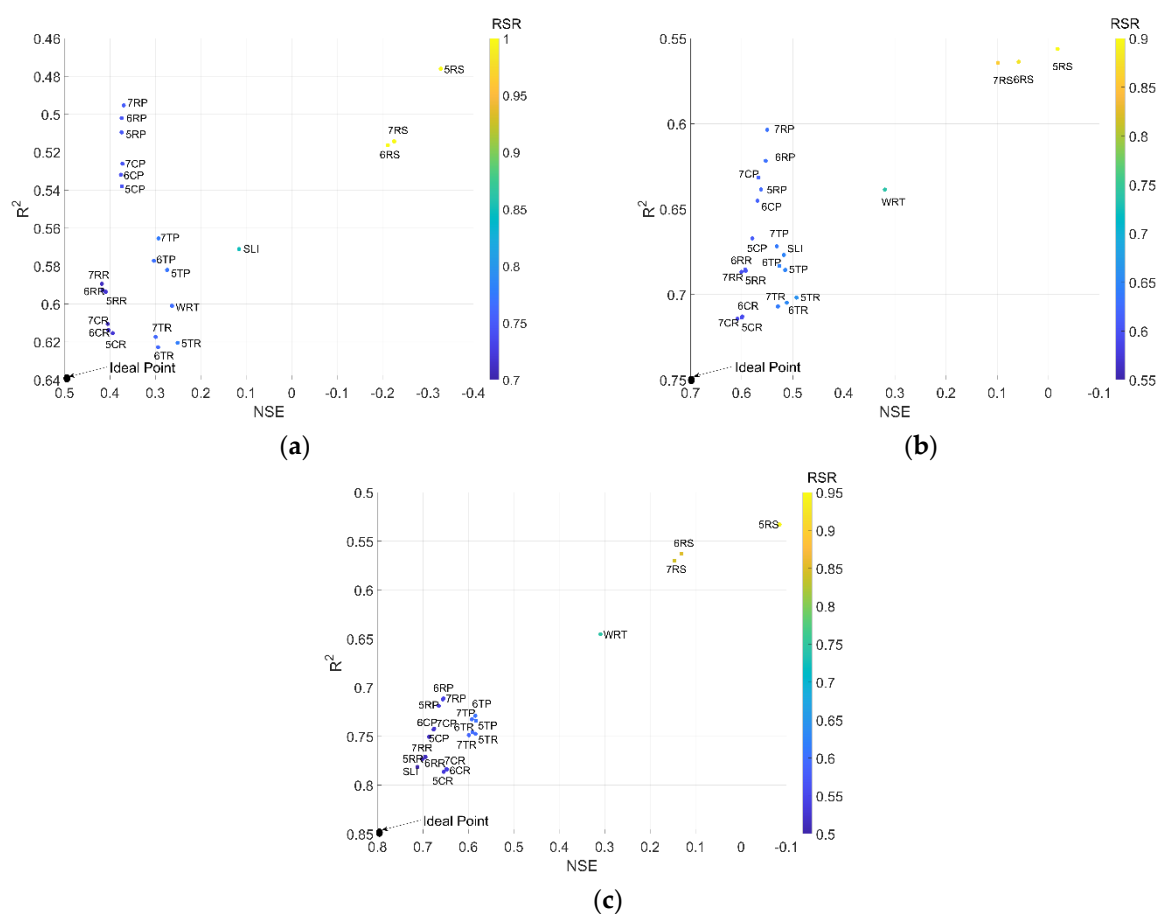
Figure 13. Evaluation of 23 estimation methods combining three mean error evaluation methods based on sampling frequency: (**a**) low frequency, (**b**) medium frequency, and (**c**) high frequency.

In addition, the performance priority rankings of the annual load estimation for the 23 estimation methods were analyzed for different sampling frequencies, as shown in Figure 14 and Table 6. Note that the annual load is one of the most significant factors in nutrient management. The most accurate estimation methods for the low, medium, and high sampling frequencies are the 6RR, 7CR, and 5RR methods, respectively. The 7CR and 7RR methods are ranked within the top four for all three frequencies. Moreover, the 5RR and 5CR methods are ranked within the top six for all three frequencies. The differences in the annual load estimation errors of all the CR and RR methods are small. Based on Figure 13 and Table 6, the 5-, 6-, and 7CR and RR methods are more reliable than the other methods for annual load estimation.

Table 6. The priority rankings of estimation methods combining eight stations for annual load depend on sampling frequency.

Ranking	Low Frequency	Medium Frequency	High Frequency
1	6RR	7CR	5RR
2	6CR	7RR	5CR
3	7RR	5CR	7CR
4	7CR	5RP	7RR
5	5RR	6CR	6RR
6	5CR	5RR	6CR
7	5TR	6RR	5RP
8	6TR	5CP	5CP
9	WRT	5TP	SLI
10	5TP	6RP	7RP

Figure 14. Evaluation of 23 estimation methods combining three error evaluation methods of mean annual load based on sampling frequency: (**a**) low frequency, (**b**) medium frequency, and (**c**) high frequency.

4. Discussion

One of the purposes of this paper was to investigate the results of various models based on long-term nitrate load datasets. Three issues were discussed: (1) the applicability of the SLI method; (2) the most effective estimation method for urban watersheds; and (3) the effects of nitrate load estimation using regression methods with adjustment methods.

4.1. Suitability of the Linear Interpolation Method

The SLI method was the most accurate method for station MM. Figure 15 shows the binning plots of the observed streamflow and nitrate at eight stations. The monitoring data at the MM, SD, and VM stations in Figure 15d,g,h, respectively, show wider ranges of flows (0.37 to 1097 m^3/s, 0.37 to 403 m^3/s, and 0.01 to 148 m^3/s, respectively) and nitrate concentrations (0.01 to 20.10 mg/L) than those of the other stations. Additionally, the relation of logarithmic flow and logarithmic nitrate concentration is positively linear at the MM, SD, and VM stations. The VM station (Figure 15h) shows a high logarithmic nitrate concentration in low-flow classes (0.01 and 0.02 m^3/s) compared with that in higher flow classes (0.05 and 1 m^3/s). However, the quantity of data in low-flow classes is not large, and is negligible. Therefore, the high-concentration data in the low-flow class do not affect the positive linearity of other logarithmic nitrate concentration and logarithmic flow classes in Figure 15h. The SLI method for the SD station is ranked as the fourth most accurate, as shown in Table 3. From Table 3, for the six stations other than the MM and VM stations, we can see that the most accurate methods for combining daily concentration, daily load, and annual load are the WRT, 6CR, 7CR, and 7CP methods. These models are based on logarithmic regression models (Equations (2) and (7)–(9)) incorporating adjustment methods. The accuracy of the logarithmic regression methods decreases because the relationship between nitrate and streamflow data does not follow a

positive linear, quadratic, or cubic relationship [12]. Consequently, for the MM and VM stations, the SLI estimates the nitrate concentration with greater accuracy than that achieved by the other 22 methods. This result suggests that the distribution of nitrate concentration, depending on the streamflow class, is an important factor in the performance of nitrate concentration estimation methods.
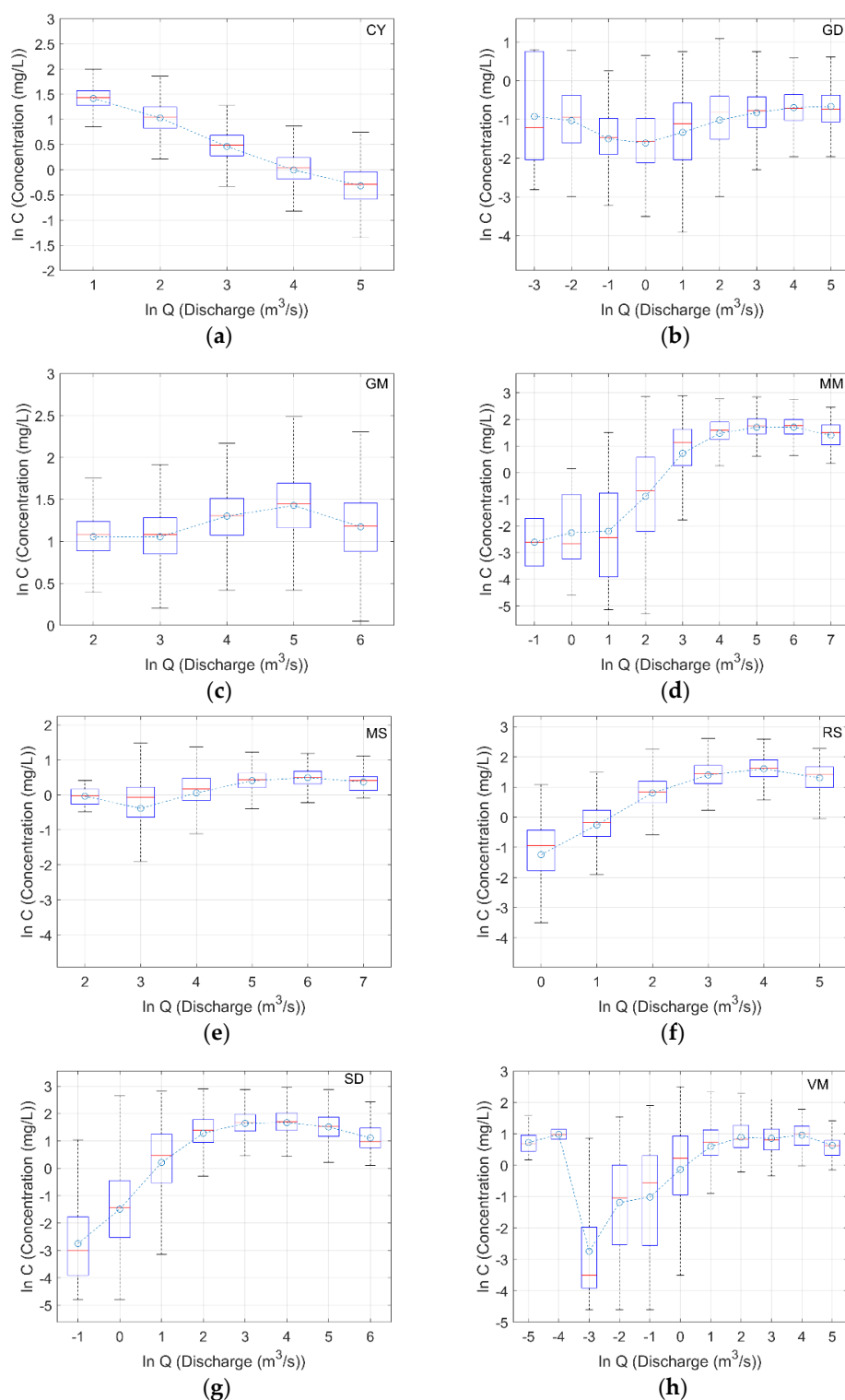
Figure 15. Observed nitrate box plots based on observed streamflow binning: (**a**) Cuyahoga, (**b**) Grand, (**c**) Great Miami, (**d**) Maumee, (**e**) Muskingum, (**f**) Raisin, (**g**) Sandusky, and (**h**) Vermilion (note: circles indicate the mean of binned nitrate concentration).

4.2. Selecting the Best Estimation Method for Urbanized Watershed Stations

As shown in Figure 11a and Table 3, the best fitting method for station CY is the WRTDS method. In total, 47% of the watershed for this station is urbanized, whereas the other stations are urbanized to an extent of 1–11%. In Figure 2a, the relationship between the discharge and nitrate concentration data for station CY is negative and exponential, although this relationship for the other seven stations is positive and exponential. This result implies that the nitrate concentration decreases as the flow increases because high flow dilutes the nitrate concentration in an urban-dominated watershed [33–36]. On the other hand, the nitrate concentration increases as the flow increases because the characteristics of agricultural watersheds are proportional to the nitrate loading with discharge, including effects from both dilution and concentration [37–39]. Additionally, Hirsch [40] noted that while the relationship between the logarithmic discharge and nitrate concentration obtained via the regression models is positive, WRTDS indicates a nonpositive relation. Therefore, the WRTDS method is more suitable than the five-, six-, and seven-parameter regression models as an estimation approach for the CY station given the negative correlation between the log-transformed discharge and the concentration.

4.3. Regression Method with Adjustment Methods

According to Figures 11–14, the original regression methods, namely, 5-, 6-, and 7RS, have low-priority ranks. This result implies that the original regression methods do not provide accurate results, and that they must be modified by adding residual adjustments to improve the model performance. Based on Tables 3–6, the residual adjustment variations of all five-, six-, and seven-parameter methods have higher priority ranks than the five-, six-, and seven-parameter methods alone. This result shows that the residual adjustments are more effective than the proportional adjustment method regardless of the number of parameters in the regression model. In other words, an adjustment method, either residual or proportional, makes the parameter regression methods more effective. Thus, the composite correction method provides more accurate results than the triangular and rectangular correction methods. The variability of the results based on the number of parameters in the regression methods is not significant compared to the magnitude of improvement achieved by the residual adjustment methods because the suitability of the regression methods differs depending on the site conditions. In other words, the results of the regression methods do not differ substantially with the number of parameters, and the errors of the five-, six-, and seven-parameter regression methods can be reduced considerably using residual correction and adjustment techniques. Moreover, Figure 15a and Table 3 show that the regression and WRTDS methods are more suitable for urban watersheds than agricultural or forested watersheds, as indicated by the negative correlation between logarithmic nitrate and logarithmic streamflow.

5. Conclusions

This study applied the five-, six-, and seven-parameter regression methods with combinations of correction/adjustment methods, SLI and WRTDS, and compared the results in terms of the estimation units, including daily concentration, daily load, annual load, gaging stations, and sampling frequencies. In cases in which the relationship between the nitrate concentration and streamflow is strong, the five-, six-, and seven-parameter regression methods incorporating composite, triangular, and rectangular correction with residual adjustments provide a better fit than the original five-, six-, and seven-parameter estimation methods. Additionally, the present work indicates that when the relation between logarithmic concentration and discharge is positively linear, the SLI method provides a more accurate estimate of nitrate loads and concentrations than do the five-, six-, and seven-parameter estimation methods. WRTDS was the most suitable method for the urban-dominated watershed CY, for which the correlation between the logarithmic discharge and nitrate concentration was negatively linear.

The combinations of regression methods with the corrections (composite, triangular, and rectangular) and adjustment methods (residual and proportional) were more accurate than the original regression methods (without corrections) for most conditions in this study. The estimation results were more sensitive to the type of correction and the adjustment methods applied than to the number of parameters in the regression model. While WRTDS provided better estimates than the original regression models (5RS, 6RS, and 7RS), it was typically less accurate than the regression models combined with the residual correction methods. This paper provides recommended priorities to estimate riverine nitrate loads with various methods, but the most appropriate method for a specific watershed should be identified by carefully analyzing the monitoring frequency, watershed scale, and target units.

Author Contributions: Conceptualization, D.P., M.-J.U. and M.M.; methodology, D.P., M.-J.U., L.K. and S.V.; investigation, D.P., M.-J.U. and K.J.; funding acquisition, D.P.; writing, review, and editing, D.P. and K.J. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barbiero, R.P.; Lesht, B.M.; Hinchey, E.K.; Nettesheim, T.G. A brief history of the US EPA Great Lakes National Program Office's water quality survey. *J. Great Lakes Res.* **2018**, *44*, 539–546. [CrossRef]
2. Rodríguez, R.; Pastorini, M.; Etcheverry, L.; Chreties, C.; Fossati, M.; Castro, A.; Gorgoglione, A. Water-quality data imputation with a high percentage of missing values: A machine learning approach. *Sustainability* **2021**, *13*, 6318. [CrossRef]
3. Ratolojanahary, R.; Ngouna, R.H.; Medjaher, K.; Junca-Bourié, J.; Dauriac, F.; Sebilo, M. Model selection to improve multiple imputation for handling high rate missingness in a water quality dataset. *Expert Syst. Appl.* **2019**, *131*, 299–307. [CrossRef]
4. Stets, E.G.; Kelly, V.J.; Crawford, C.G. Regional and temporal differences in nitrate trends discerned from long-term water quality monitoring data. *J. Am. Water Resour. Assoc.* **2015**, *51*, 1394–1407. [CrossRef]
5. Cohn, T.A.; Caulder, D.L.; Gilroy, E.J.; Zynjuk, L.D.; 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 Resour. Res.* **1992**, *28*, 2353–2363. [CrossRef]
6. Runkel, R.L.; Crawford, C.G.; Cohn, T.A. Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. In *Techniques and Methods*; US Geological Survey, U.S. Department of the Interior: Reston, VA, USA, 2004. [CrossRef]
7. Stenback, G.A.; Crumpton, W.G.; Schilling, K.E.; Helmers, M.J. Rating curve estimation of nutrient loads in Iowa rivers. *J. Hydrol.* **2011**, *396*, 158–169. [CrossRef]
8. Aulenbach, B.T.; Hooper, R.P. The composite method: An improved method for stream-water solute load estimation. *Hydrol. Process.* **2006**, *20*, 3029–3047. [CrossRef]
9. Verma, S.; Markus, M.; Cooke, R.A. Development of error correction techniques for nitrate-N load estimation methods. *J. Hydrol.* **2012**, *432*, 12–25. [CrossRef]
10. Park, D.; Markus, M.; Jung, K.; Um, M.-J. Uncertainty analysis of the relationship between discharge and nitrate concentration in the Lower Illinois River using a copula model. *J. Clean. Prod.* **2019**, *222*, 310–323. [CrossRef]
11. Pellerin, B.A.; Bergamaschi, B.A.; Gilliom, R.J.; Crawford, C.G.; Saraceno, J.; Paul Frederick, C.; Downing, B.D.; Murphy, J.C. Mississippi River nitrate loads from high frequency sensor measurements and regression-based load estimation. *Environ. Sci. Technol.* **2014**, *48*, 12612–12619. [CrossRef] [PubMed]
12. Lee, C.J.; Hirsch, R.M.; Schwarz, G.E.; Holtschlag, D.J.; Preston, S.D.; Crawford, C.G.; Vecchia, A.V. An evaluation of methods for estimating decadal stream loads. *J. Hydrol.* **2016**, *542*, 185–203. [CrossRef]
13. Guo, Y.; Markus, M.; Demissie, M. Uncertainty of nitrate-N load computations for agricultural watersheds. *Water Resour. Res.* **2002**, *38*, 3-1–3-12. [CrossRef]

14. Hirsch, R.M.; Moyer, D.L.; Archfield, S.A. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs 1. *JAWRA* **2010**, *46*, 857–880. [PubMed]

15. Verma, S.; Bartosova, A.; Markus, M.; Cooke, R.; Um, M.-J.; Park, D. Quantifying the role of large floods in riverine nutrient loadings using linear regression and analysis of covariance. *Sustainability* **2018**, *10*, 2876. [CrossRef]

16. Jung, K.; Um, M.-J.; Markus, M.; Park, D. Comparison of long short-term memory and weighted regressions on time, discharge, and season models for nitrate-N load estimation. *Sustainability* **2020**, *12*, 5942. [CrossRef]

17. Verma, A.K.; Jha, M.K. Evaluation of a GIS-based watershed model for streamflow and sediment-yield simulation in the upper Baitarani river basin of Eastern India. *J. Hydrol. Eng.* **2015**, *20*, C5015001. [CrossRef]

18. Government of India; Government of The Netherlands. *Protocol for Water Quality Monitoring, Hydrology Project Technical Assistance Version 1.0*; DHV Consultants: New Delhi, India; Delft Hydraulics: Delft, The Netherlands; HALCROW: London, UK; TAHAL: Amsterdam, The Netherlands; CES: Langley, VA, USA; ORG: New York, NY, USA; JPS: Kingston, Jamaica, 1999.

19. Canadian Council of Ministers of the Environment (CCME). *Guidance Manual for Optimizing Water Quality Monitoring Program Design*; Canadian Council of Ministers of the Environment (CCME): Winnipeg, MB, Canada, 2015.

20. El-Shaarawi, A.; Kuntz, K.; Sylvestre, A. Estimation of loading by numerical integration. In *Developments in Water Science*; Elsevier: Amsterdam, The Netherlands, 1986; pp. 469–478.

21. Preston, S.D.; Bierman, V.J., Jr.; Silliman, S.E. An evaluation of methods for the estimation of tributary mass loads. *Water Resour. Res.* **1989**, *25*, 1379–1389. [CrossRef]

22. Richards, R.P.; Baker, D.B.; Kramer, J.W.; Ewing, D.E.; Merryfield, B.J.; Miller, N.L. Storm discharge, loads, and average concentrations in Northwest Ohio rivers, 1975–1995. *J. Am. Water Resour. Assoc.* **2001**, *37*, 423–438. [CrossRef]

23. Toor, G.S.; Harmel, R.D.; Haggard, B.E.; Schmidt, G. Evaluation of regression methodology with low-frequency water quality sampling to estimate constituent loads for ephemeral watersheds in Texas. *J. Environ. Qual.* **2008**, *37*, 1847–1854. [CrossRef] [PubMed]

24. Cohn, T.A.; Delong, L.L.; Gilroy, E.J.; Hirsch, R.M.; Wells, D.K. Estimating constituent loads. *Water Resour. Res.* **1989**, *25*, 937–942. [CrossRef]

25. Jeong, J.; Kannan, N.; Arnold, J.; Glick, R.; Gosselink, L.; Srinivasan, R. Development and integration of sub-hourly rainfall–runoff modeling capability within a watershed model. *Water Resour. Manag.* **2010**, *24*, 4505–4527. [CrossRef]

26. Lamichhane, N.; Sharma, S. Effect of input data in hydraulic modeling for flood warning systems. *Hydrol. Sci. J.* **2018**, *63*, 938–956. [CrossRef]

27. Meng, H.; Sexton, A.; Maddox, M.; Sood, A.; Brown, C.; Ferraro, R.; Murtugudde, R. Modeling Rappahannock River basin using SWAT-pilot for Chesapeake Bay watershed. *Appl. Eng. Agric.* **2010**, *26*, 795–805. [CrossRef]

28. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and water quality models: Performance measures and evaluation criteria. *Trans. ASABE* **2015**, *58*, 1763–1785.

29. Sisay, E.; Halefom, A.; Khare, D.; Singh, L.; Worku, T. Hydrological modelling of ungauged urban watershed using SWAT model. *Model. Earth Syst. Environ.* **2017**, *3*, 693–702. [CrossRef]

30. Jung, K.; Bae, D.H.; Um, M.J.; Kim, S.; Jeon, S.; Park, D. Evaluation of nitrate load estimations using neural networks and canonical correlation analysis with k-fold cross-validation. *Sustainability* **2020**, *12*, 400. [CrossRef]

31. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900. [CrossRef]

32. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual model. Part 1—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290.

33. Burns, D.A.; Boyer, E.W.; Elliott, E.M.; Kendall, C. Sources and transformations of nitrate from streams draining varying land uses: Evidence from dual isotope analysis. *J. Environ. Qual.* **2009**, *38*, 1149–1159. [CrossRef]

34. Jiang, Y.; Frankenberger, J.R.; Bowling, L.C.; Sun, Z. Quantification of uncertainty in estimated nitrate-N loads in agricultural watersheds. *J. Hydrol.* **2014**, *519*, 106–116. [CrossRef]

35. Kaushal, S.S.; Groffman, P.M.; Band, L.E.; Elliott, E.M.; Shields, C.A.; Kendall, C. Tracking nonpoint source nitrogen pollution in human-impacted watersheds. *Environ. Sci. Technol.* **2011**, *45*, 8225–8232. [CrossRef] [PubMed]

36. Poor, C.J.; McDonnell, J.J. The effects of land use on stream nitrate dynamics. *J. Hydrol.* **2007**, *332*, 54–68. [CrossRef]

37. Adeuya, R.; Utt, N.; Frankenberger, J.; Bowling, L.; Kladivko, E.; Brouder, S.; Carter, B. Impacts of drainage water management on subsurface drain flow, nitrate concentration, and nitrate loads in Indiana. *J. Soil. Water Conserv.* **2012**, *67*, 474–484. [CrossRef]

38. David, M.B.; Drinkwater, L.E.; McIsaac, G.F. Sources of nitrate yields in the Mississippi River Basin. *J. Environ. Qual.* **2010**, *39*, 1657–1667. [CrossRef]

39. Lavaire, T.; Gentry, L.E.; David, M.B.; Cooke, R.A. Fate of water and nitrate using drainage water management on tile systems in east-central Illinois. *Agric. Water Manag.* **2017**, *191*, 218–228. [CrossRef]

40. Hirsch, R.M. Large biases in regression-based constituent flux estimates: Causes and diagnostic tools. *J. Am. Water Resour. Assoc.* **2014**, *50*, 1401–1424. [CrossRef]