Article

Comparison of Long Short-Term Memory and Weighted Regressions on Time, Discharge, and Season Models for Nitrate-N Load Estimation

Kichul Jung [1], Myoung-Jin Um [2], Momcilo Markus [3] and Daeryong Park [1],*

[1] Department of Civil and Environmental Engineering, Konkuk University, Seoul 05029, Korea; jkichul11@konkuk.ac.kr

[2] Department of Civil Engineering, Kyonggi University, Suwon 16227, Korea; mum@kyonggi.ac.kr

[3] Prairie Research Institute, University of Illinois, Champaign, IL 61820, USA; mmarkus@illinois.edu

* Correspondence: drpark@konkuk.ac.kr

check for updates

Abstract: The long short-term memory (LSTM) model has been widely used for a broad range of applications entailing the estimation of variables in different fields to improve water quality management in rivers. The main objectives of this study are (1) to develop a novel LSTM-based model for the estimation of nitrate-N loads, which adversely affect water resources, and (2) to evaluate the performance of the model by comparing it with that of Monte Carlo sub-sampling and the weighted regressions on time discharge and season (WRTDS) model. We evaluated the model performance using various numbers of hidden layers, ranging from one to four, in the LSTM model to determine the appropriate number of hidden layers; furthermore, we applied the sampling frequencies of 6, 12, and 24 to assess their impact. Seven polluted river basins in the United States were used for analysis, and the relative root mean squared error (*rRMSE*) and the mean percentage error (*MPE*) metrics were applied for the validation of the model estimates. The proposed model achieved accurate nitrate-N load estimates using three to four hidden layers, and improved model performance was observed when the sampling frequency was increased. The differences among the results obtained using the LSTM model were examined based on a binning technique via a log-log plot of nitrate-N concentration against discharge. The binning analysis showed that the slope obtained from the average rates of discharge and low discharge values apparently influenced the estimates. Furthermore, box plot analyses of the statistical indices such as *rRMSE* and *MPE* demonstrate that the LSTM model seems to exhibit better performance than the WRTDS model. The results of the examination demonstrate that the LSTM model may be a good alternative with regard to estimating nitrate-N loads for the control of water quality constituents.

Keywords: Nitrate-N load estimation; long short-term memory network; weighted regressions on time discharge and season; binning analysis; water quality

1. Introduction

Nitrate-nitrogen (Nitrate-N; NO_3-N) load estimation is crucial to water resource management because excessive nutrients in water increase the degradation of water quality, resulting in water problems in rivers, streams, and receiving water bodies, such as the Laurentian Great Lakes [1,2]. River basins in the Midwestern United States have encountered difficulties with regard to water resource management because of nutrient enrichment and water pollutants derived from crop production using fertilizers and pesticides [3]. Specifically, rivers and streams in the state of Ohio are affected by the presence of high levels of nutrients, and they have been monitored for a long time period to estimate nitrate-N loads, in order to obtain an accurate assessment of water quality [4]. Notably, nonpoint

sources aggravating streams and rivers contributed an average of 79% of the total nitrate-N load in the Ohio River Basin, which should be reduced [5]. Furthermore, the Illinois River watershed is influenced by excess pollutants contributing 17.9% and 12.9% of the total nitrogen and phosphorus therein, respectively [6]. Precise estimates of nitrate-N loads can be gained via the use of appropriate conservation measures by analyzing water pollutant sources and controlling the potential sources of uncertainty in river basins of interest.

Nitrate-N is a common water pollutant derived from point sources, including sewage disposal systems and livestock facilities, and from non-point sources, such as parks, fertilized croplands, and gardens [7]. When pollutants are not adequately treated in drainage systems, nitrates (NO_3) can enter lakes, streams, and rivers, negatively influencing drinking water and ecological health. Nitrate-N loads used for the management of water quality can be expressed as the total mass passing through a river site over a given time period; furthermore, they can be obtained by measuring the sum of nitrate-N concentrations and streamflows. Although measurement of nitrate-N concentration is significant, there are difficulties associated with obtaining these values because of the high cost and difficulty entailed in collecting samples on a daily basis. Therefore, the analysis of nitrate-N load estimation plays a vital role in obtaining accurate estimates and improving water quality in aquatic systems.

Several studies have been performed on nitrate-N load estimation using various approaches. Soil and water assessment tool (SWAT) and hydrological simulation program-Fortran, which are process-based models, were used to calculate nitrate-N loads using the product of discharge and nitrate-N concentration [8–10]. A multiple log-linear regression equation known as LOADEST was used for solving uncertainties in estimating nitrate-N load [11–13]. The LOADEST approach was developed by employing composite, triangular, and rectangular distributions in addition to regression estimators with various numbers of parameters, such as those with five or seven parameters. Weighted regressions on time, discharge, and season (WRTDS) was suggested to produce enhanced load estimates [14]. Artificial neural networks were used as a machine learning technique to predict nitrate-N loads [15,16].

The results of load estimation are dependent on the estimation methods; consequently, the sampling frequency, as well as the observation duration, affects load estimates [17]. Among the sampling-based techniques, the Monte Carlo sampling method is used to provide enhanced estimates [18,19]. Richards and Holloway [20] investigated the use of the Monte Carlo technique to develop sampling strategies to assess the accuracy of solute load estimation. Verma et al. [3] used the Monte Carlo sub-sampling method, which permits reproducing the actual sampling scenarios to achieve nitrate-N load estimation. Monte Carlo simulation was also conducted by Rahman et al. [21] to perform the estimation of a hydrological variable with flood frequency curves. In the analysis, different periodic sampling frequencies, such as one, two, four, six, and eight weeks, were adopted for comparing and evaluating nitrate-N estimates derived from statistical models.

Among machine learning techniques, long short-term memory (LSTM), which is a recurrent neural network (RNN), has been used for a variety of sequential applications based on historical data. Hochreiter and Schmidhuber [22] suggested that the LSTM model can solve the vanishing gradient problem, which occurs when neural networks are trained based on gradient-based learning methods. Yuan et al. [23] performed monthly runoff forecasting for the management of water resources in a river basin in the north of Pakistan. Zhang et al. [24] used the LSTM model for predicting long-term water table depth in agricultural regions in Northwestern China. Tian et al. [25] analyzed four different types of RNNs to forecast discharge in central southern China. Bowes et al. [26] investigated groundwater table forecasting corresponding to storm events in the flood-prone state of Virginia, US. Rainfall-runoff modeling via a LSTM-based sequence-to-sequence model was also carried out by Xiang et al. [27] to estimate the hourly rainfall runoff in Iowa, US. However, there is a paucity of studies based on the LSTM model for the analysis of water pollutants, especially nitrate-N load estimates.

In the present study, we aimed to develop a new model using LSTM to estimate the nitrate-N load in several rivers in Midwestern US. Furthermore, the proposed model was compared with the WRTDS model with regard to the ability to obtain the nitrate-N load to determine the applicability of

the LSTM model. For the validation of the model, we used jackknife resampling techniques based on statistical indices. The reminder of this paper is organized as follows. A description of the datasets used for the study is provided in Section 2. In Section 3, we describe methodologies to estimate and evaluate the nitrate-N load. Sections 4 and 5 present the results and a discussion of the work. Finally, in Section 6, the conclusions are summarized.

2. Data Set

In this study, we estimated the nitrate-N loads for seven river basins, namely, the Cuyahoga (CY), Grand (GD), Great Miami (GM), Maumee (MM), Muskingum (MS), Raisin (RS), and Vermilion (VM) basins, which were chosen to represent the river basins of the Midwest US. Each basin has features that can affect the analysis of the nitrate-N load estimation. The basins consist of urban, wooded, and agricultural areas; particularly, water flow over these areas causes an eventual high nitrate-N concentration that affects the Great Lakes.

CY has a basin area of 1843 km^2. Data regarding this basin are recorded via the United States Geological Survey (USGS) at station number 04208000. The rivers in this basin run through Cleveland city and are heavily influenced by its industrial pollution, which also feeds into Lake Erie. "Urban area" is the most significant land use type of this basin, accounting for 47% of land use. GD is a tributary of Lake Erie; furthermore, the GD basin has a basin area of 1758 km^2. The most significant land use type of this is "woodland," at 52%. Data regarding this basin are gathered at station number 04212100. The GM basin has been assigned the station number 03271601; this basin has an area of 6953 km^2 and is surrounded by the Miami Valley. Furthermore, GM serves as a tributary of the Ohio River. The principal land use type for this basin is agriculture, at 82%. The MM basin has been assigned the station number 04,193,500 with a stream record; this basin has an area of 16,427 km^2. Furthermore, MM flows from northeastern Indiana to northwestern Ohio and Lake Erie. The most significant land use type for this basin is also agriculture, at 81%. The station number for the MS basin is 03150000. Its basin area is 19,208 km^2. It is part of the Ohio River and flows southward via eastern Ohio. Agriculture is the most significant land use type of this basin, at 52%. RS and VM have station numbers 04176500 and 04199500, respectively. Their basins areas are 2755 and 697 km^2, respectively. RS is a river that flows into Lake Erie and VM is a tributary of Lake Erie in northern Ohio. Agriculture is the most significant land use type in both basins, at 72% and 71%, respectively. Figure 1 shows the locations of the seven river basins analyzed in this study.

Table 1 lists the seven river basins according to their outlet, the portion of the land used, and the data collection period for daily discharge and nitrate-N concentration data. The daily discharge and nitrate-N concentration data were applied for load estimation. The daily discharge data were obtained from USGS [28], and the daily nitrate-N concentration data [4] were obtained from the Water Quality Laboratory of the National Center for Water Quality Research at Heidelberg University [29]. The variables used for load estimation were transformed for normality and standardized.

Table 1. Descriptive features for the seven river basins in the USA used for nitrate-N load estimation [30].

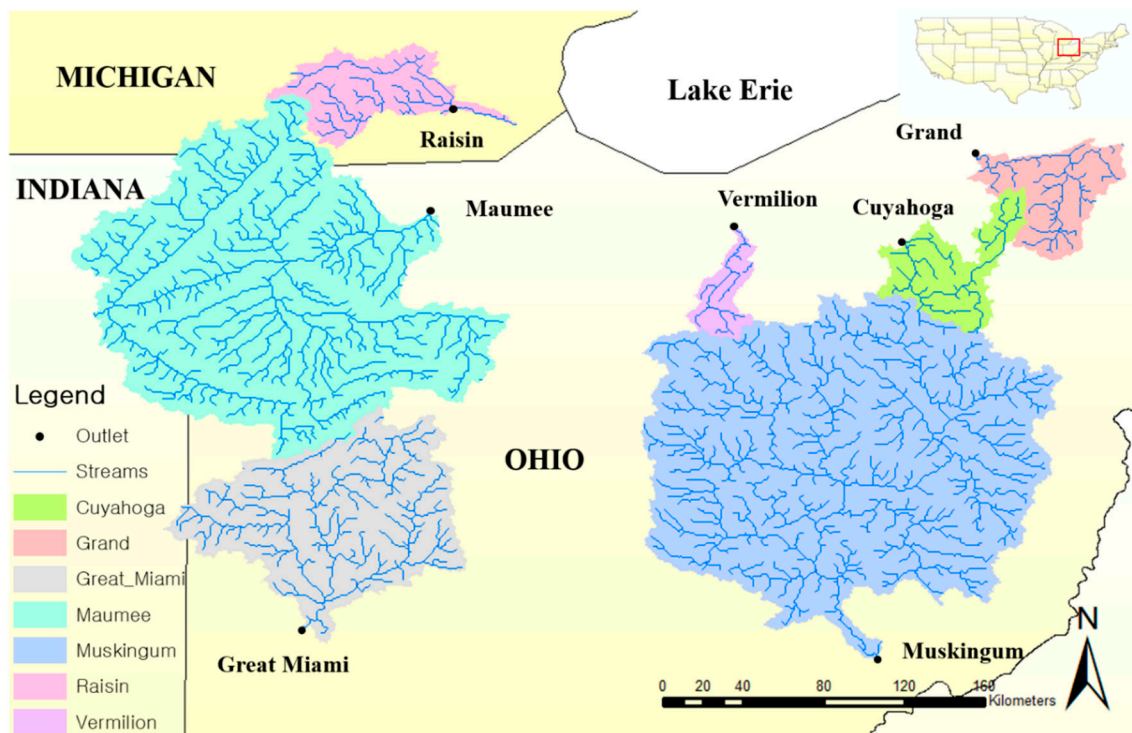| Station Name	Outlet Location		Average Discharge (m^3/s)	Average Nitrate (mg/L)	Year	Drainage Area (km^2)	Land Use (%)		
	Latitude	Longitude					Agriculture	Urban	Wooded
CY	41°23′43″	81°37′48″	1010	2.45	1982–2017	1843	17	47	35
GD	41°43′08″	81°13′41″	1076	0.46	1988–2006	1758	37	10	52
GM	39°36′24″	84°17′13″	3221	3.50	1996–2017	6953	82	5	10
MM	41°30′00″	83°42′46″	6174	4.40	1983–2017	16,427	81	11	8
MS	39°38′42″	81°51′00″	8608	1.40	1994–2017	19,208	52	2	43
RS	41°57′38″	83°31′52″	823	2.95	1982–2010	2755	72	11	16
VM	41°22′55″	82°19′01″	501	2.17	2000–2008	697	71	1	26

Figure 1. Seven river basins, namely, Cuyahoga, Grand, Great Miami, Maumee, Muskingum, Raisin, and Vermilion analyzed in the present work.

The daily nitrate-N concentration was used to develop an appropriate model for estimating the nitrate-N load considering different resampling frequencies and river basins. The data periods used for the CY, GD, GM, MM, MS, RS, and VM basins were 36, 19, 22, 35, 24, 29, and 9 years, respectively. The average discharge for the seven river basins was 3059 m^3/s, with a range of 501 m^3/s to 8608 m^3/s. Furthermore, the average nitrate-N concentration for the basins was 2.48 mg/L, with a range of 0.46 mg/L to 4.40 mg/L. Figure 2 shows the daily and annual rates of discharge for each river basin. This figure uses two different *y*-axis scales, depending on whether the basin is large or small. GM, MM, and MS are all large basins with a large amount of discharge, whereas CY, GD, RS, and VM are small basins with a small amount of discharge. Figure 3 presents the daily and annual nitrate-N concentrations for each basin. In the figure, the average nitrate-N concentrations for CY, GM, MM, and RS are higher than those of GD, MS, and VM.
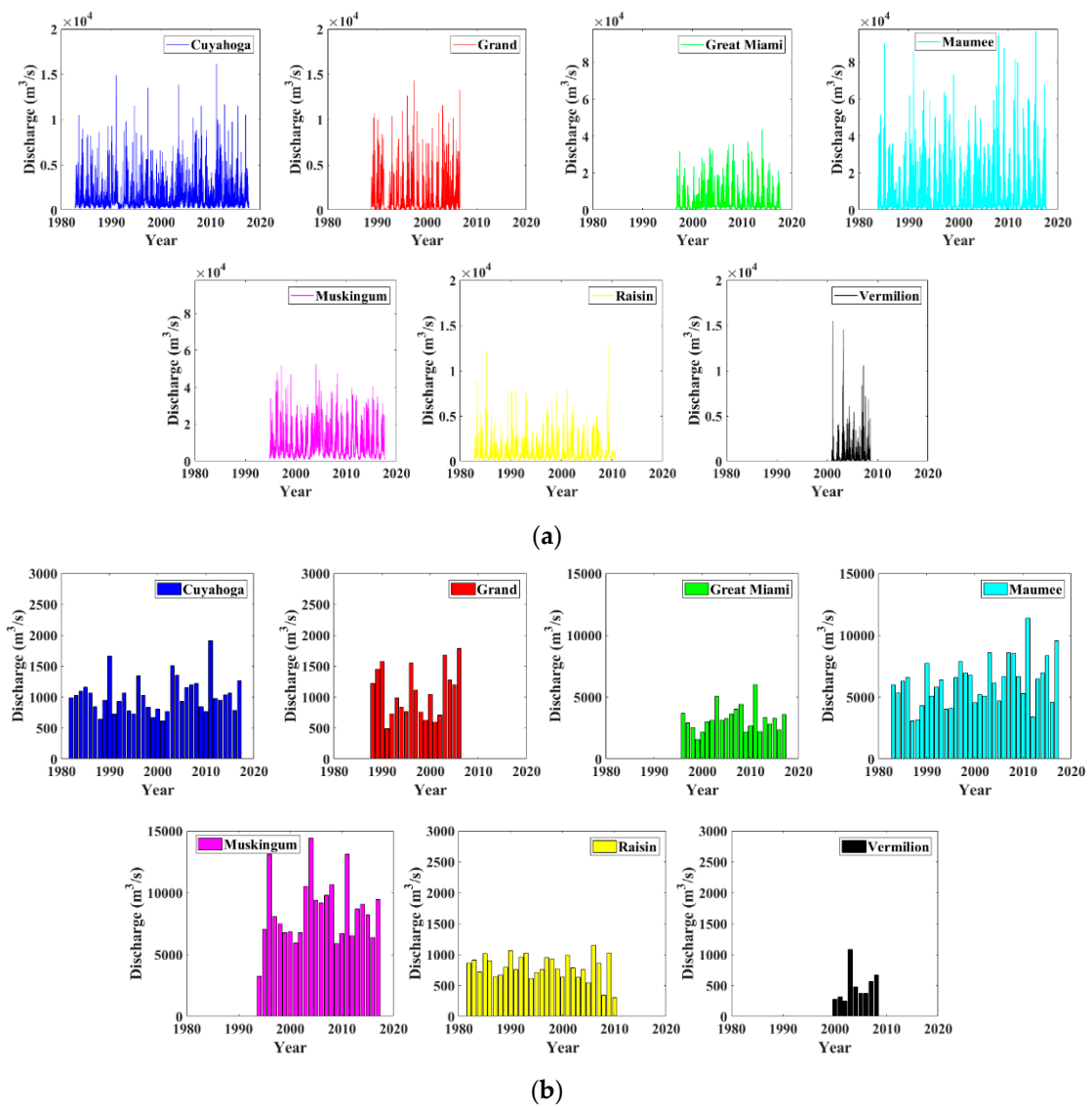
(**a**)

(**b**)

Figure 2. Plot of the (**a**) daily discharge (m^3/s) and (**b**) annual mean discharge (m^3/s) for the seven river basins.
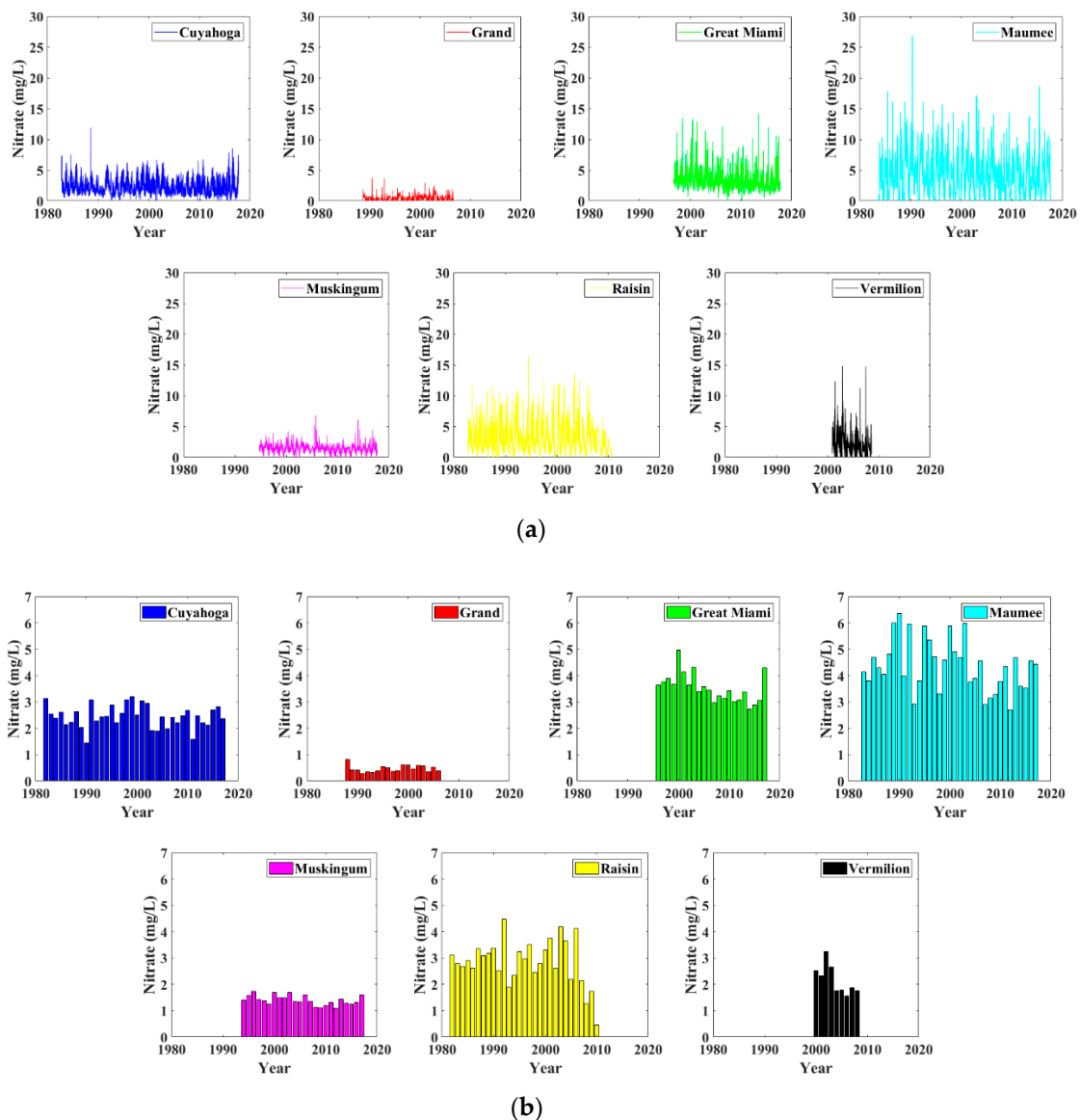
(a)

(b)

Figure 3. Plot of the (**a**) daily nitrate-N concentration (mg/L) and (**b**) annual mean nitrate-N concentration (mg/L) for the seven river basins.

3. Methods

As mentioned previously, the primary objectives of this research are to investigate the use of the LSTM model for obtaining accurate nitrate-N load estimations for river basins in the US. The LSTM model was set up based on the Monte Carlo sub-sampling approach using various sampling frequencies. The discharge and nitrate-N concentrations were used for load estimates by establishing a relationship between the variables. The binning technique was used to examine the relationship, characterize the variables, and verify the results for nitrate-N load estimation. The results obtained from the LSTM model were evaluated using the WRTDS model to validate the performance of the proposed model. In the validation analysis, standard statistical indices were applied to the results.

3.1. LSTM Model Architecture

The LSTM network, which is a type of RNN, was used as an improved model in obtaining load estimation in the current study. The LSTM network was developed to overcome the problem of vanishing gradients [21]. The LSTM model is characterized by a memory cell, C_t, which memorizes

state information over time and permits gradients to flow over sequences. It consists of three gates, including an input gate, i_t, a forget gate, f_t, and an output gate, o_t, from which the information flows into an LSTM cell. The LSTM cell identifies the input derived from the current time, x_t, and the hidden state, h_{t-1}, derived from the previous step by maintaining state information. A diagram of a LSTM cell with the three gates is shown in Figure 4a.
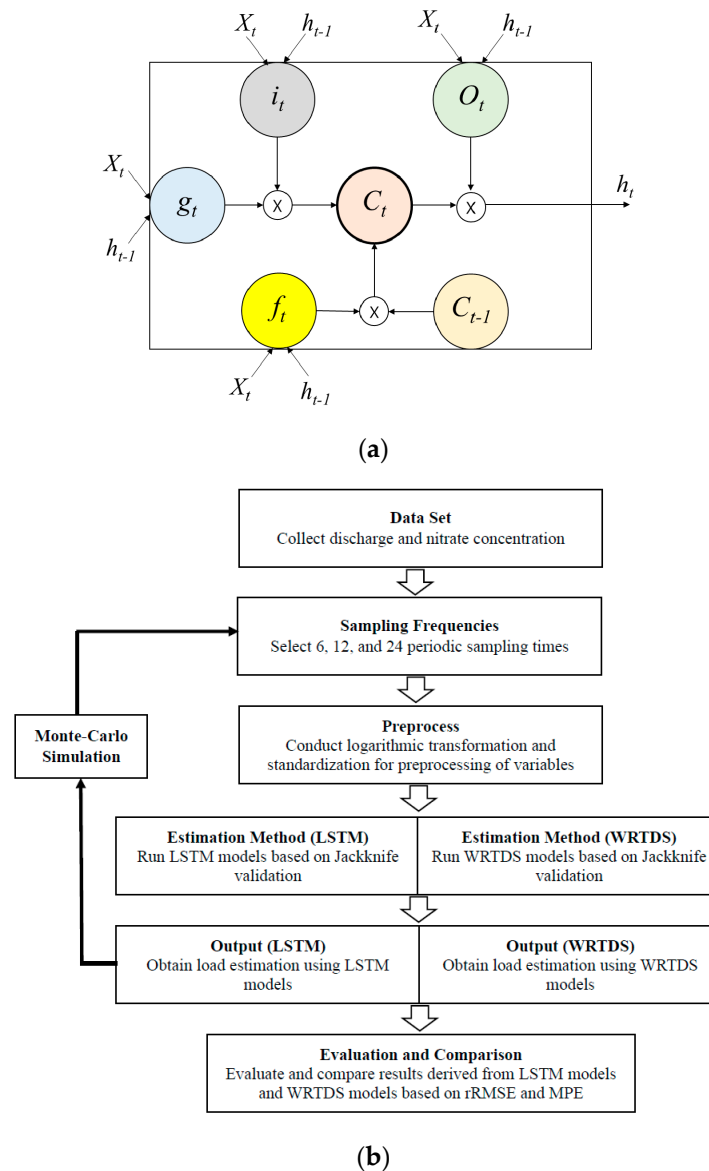

(**a**)

(**b**)

Figure 4. Plot for (**a**) process of the long short-term memory (LSTM) cell over sequences and (**b**) diagram for processes in obtaining nitrate-N load estimation.

With the input and the hidden state, Equation (1) can be defined as the candidate cell state (g_t) based on the tanh function for the LSTM process:

$$g_t = \tanh(W_g x_t + U_g h_{t-1} + b_g) \tag{1}$$

where W_g, U_g, and b_g indicate the weights of the input, recurrent weight, and bias. In the input gate, information that will be stored in the memory cell is identified using an element-wise sigmoid function, as shown in Equation (2). In the forget gate, information that should be eliminated from the cell is

determined using the sigmoid function (Equation (3)). The output gate can be also expressed using this function, as in Equation (4).

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{2}$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{3}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{4}$$

where σ(z) means the element-wise sigmoid function of σ(z) = 1/(1+exp(−z)).

The information in the memory cell is then updated based on the partial forgetting of the information maintained in the previous cell C_{t-1}. Based on the input, forget, and output gates, the memory cell can be denoted as

$$C_t = f_t * C_{t-1} + i_t * g_t \tag{5}$$

where * implies element-wise multiplication. The forget gate is used to determine whether an extent of the past information kept in C_{t-1} will be forgotten. The value of the gate ranges from 0 to 1. If f_t tends to 0, the past information will be forgotten, whereas if it goes to 1, the past information will be stored in the memory cell. Using the analyzed cell state, C_t, the hidden state h_t, as shown in Equation (6), is updated to provide the output of the model.

$$h_t = o_t * \tanh(C_t) \tag{6}$$

The schematic description of the proposed method is presented in Figure 4b, which shows the estimation approaches, sampling frequencies, and assessment techniques used in obtaining load estimations. A brief description of the sampling frequencies and assessment techniques is provided in Sections 3.3 and 3.4, respectively.

Based on the LSTM cell, the LSTM network consists of a sequence input layer, LSTM hidden layers, a fully connected layer, and an output layer. The input layer inputs sequence into the network. The LSTM hidden layers play a significant role in the modeling of correlations between time steps of sequence data. These layers are used to design more complex models that can solve complex issues related to pattern recognition, classification, and estimation. In the present study, the number of hidden layers ranging from one to four, was analyzed to determine the proper values of the hidden layers. The network completes the analysis with a fully connected layer and a regression outer layer, which provides an estimate. To implement the LSTM network, the hidden units are set at 200 for each layer, the maximum number of epochs is set at 250, and the learning rate is set at 0.005. The hidden units of 200 are decided in this study after attempting various numbers of units and by keeping the number of units based on the least forecast error. Adaptive optimization of weights is conducted based on the ADAM (adaptive moment estimation) optimizer algorithm. With the aim of estimating loads, the measured discharge and concentration are used to train the LSTM network. Load estimation is obtained based on the estimated concentration derived from the LSTM model as the output variable. Note that the LSTM network for this analysis can be determined and built using the Deep Network Designer toolbox in MATLAB.

3.2. WRTDS Model Architecture

Weighted regressions on time discharge and season (WRTDS) is an approach used to examine long-term water-quality data by evaluating trends and average nitrate-N concentrations [14]. This method only uses daily flow data, but not necessarily daily concentration data. Notably, daily concentration data are often not present in river water quality monitoring datasets. The model is used to obtain the nitrate-N concentration estimations for each day in the data collection period. The WRTDS model consists of four components, including three deterministic and one random component corresponding to the season, trend, discharge, and random variables. The model chooses samples that are substantially close to estimation points in three dimensions, such as time, season,

and discharge, by prescreening all sampled data for each point [31]. Based on the WRTDS mode with the components, the nitrate-N concentration can be estimated as follows:

$$\ln(c) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \tag{7}$$

where β indicates the fitted coefficient, c is the nitrate-N concentration, Q indicates the discharge, t is the time in the record period, and ε implies the unexplained variation.

This equation is based on weighted regression, in which each observation is weighted using the relevance of the observation to the estimation point. The weight corresponding to each observation can be defined based on a three-dimensional distance metric between the observation point and the estimation point. The form of the weight function determined by Tukey [32] is as follows

$$w = \left\{ \begin{array}{l} \left(1 - (d/h)^3\right)^3 \rightarrow if|d| \leq h \\ \qquad 0 \rightarrow if|d| > h \end{array} \right\} \tag{8}$$

where w indicates the weight, d implies the distance between the observation point and the estimation point, and h means the half window width. Detailed information regarding the processes and characteristics of the WRTDS method is available in a report by Hirsch et al. [14]. Hirsch and De Cicco [33] proposed the exploration and graphics for river trend (EGRET) R package, which includes the WRTDS algorithm. Their study applied an EGRET R package to estimate nitrate-N load concentrations using the WRTDS method.

3.3. Sampling Frequency and Monte Carlo Simulation

To accurately execute the LSTM and WRTDS models proposed in this study for nitrate-N load estimation, data sampling was performed at various frequencies. The accuracy of solute load estimation is significantly dependent on several parameters, including estimation approaches, sampling frequencies, and sampling routines [20,34]. Sampling frequencies of 6, 12, and 24 samples per year, which were used by Verma et al. [3], were employed for load estimation. Periodic sampling, which is the collection of data to represent a continuous daily concentration distribution using a sequence of seasonal and discrete values, was performed using these sampling frequencies. The aforementioned sampling frequencies are equivalent to yearly sampling intervals of 8, 4, and 2 weeks.

Based on the sampling frequencies, Monte Carlo simulation was performed to obtain a uniformly distributed random variable for the models [3]. When an 8-week sampling interval is used, a random day within 8 weeks and another random day within the next 8 weeks are selected as the sampling days on which load estimation will be analyzed. Based on the sampled data, load estimation is carried out by performing 500 iterations for executing the models selected for the present study. After all the 500 iterations are conducted for the eight river basins, the evaluation criteria are computed by averaging the 500 results derived from each simulation. The Monte Carlo sampling method ensures that a broad range of data is used for the evaluation of the proposed models with regard to estimating nitrate-N loads.

3.4. Evaluation Criteria

The performance of the proposed model was evaluated based on its accuracy of estimation of the daily nitrate-N loads for the seven river basins using a leave-one-out cross-validation, jackknife approach. This jackknife validation method has been widely used to assess the performance of estimates derived from neural network models [35–38]. The process of jackknife validation involves the removal of an original sample from the database as a test member, followed by calibration of the network model using the remaining database as training members. The model is calibrated using the training members and assessed using the test members.

The models based on the LSTM and WRTDS approaches were validated using two measures, the relative root mean squared error (*rRMSE*) and the mean percentage error (*MPE*). These statistical indices are commonly used for the evaluation of estimates derived from models [1,3,4,36]. The two measures can be computed as follows:

$$rRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{q_i - \hat{q}_i}{q_i}\right)^2} \tag{9}$$

$$MPE = \frac{100}{n}\sum_{i=1}^{n}\left(\frac{q_i - \hat{q}_i}{q_i}\right) \tag{10}$$

where *n* implies the total number of data points used for the analysis, q_i implies the measured value for day *i*, and $\hat{q}_i$ indicates the estimated value derived from the models for day *i*. The *rRMSE* can range from zero to large positive numbers and the *MPE* can range from large negative to large positive numbers. The optimal value of both *rRMSE* and *MPE* is zero.

To identify the differences in the patterns of the *rRMSE* and *MPE* metrics presented in Section 4, we investigated the characteristics of and relationships among the discharge and nitrate-N concentrations. For the analysis, a binning approach was used to identify the trend of the original data, such as the nitrate-N concentration, by grouping several continuous values into a smaller number of bins. In this approach, the values of nitrate-N concentrations that fall into a given interval are replaced by the value corresponding to that interval. This method has been applied in previous studies to determine and analyze hydrological and environmental phenomena [39,40].

4. Results

4.1. Evaluation of LSTM Models for Load Estimation

In the present study, an LSTM model was developed to perform nitrate-N load estimation for the analysis of nutrient concentrations for water quality management. The model was evaluated using different sampling frequencies of 6, 12, and 24 per year for seven river basins in the US. These sampling frequencies were adopted for the investigation of nitrate-N load estimation based on a study conducted by Lee et al. [1].

Using the LSTM model, the nitrate-N load estimates are obtained by employing various numbers of hidden layers for the three sampling frequencies. Table 2 shows the *rRMSE* and *MPE* metrics averaged over 500 iterations, which are used to evaluate the model estimates, for each number of hidden layers. The blue font indicates the best performing model for each number of hidden layers. The results presented in Table 2 show that the *rRMSE* and *MPE* criteria appear to improve the accuracy in estimating the nitrate-N loads of the seven river basins under the various sampling frequencies when the number of hidden layers increases. The LSTM models with three and four hidden layers show better performance compared to the ones with one and two hidden layers. Thus, a model with a small number of hidden layers seems to have insufficient complexity to fully represent the network system. With regard to the *rRMSE* metric, the model with four hidden layers showed the best performance under the three sampling frequencies for all sevens river basins considered in this study. Notably, with regard to the *MPE* metric, the best performance was exhibited by the model with four hidden layers for the CY, GD, MM, RS, and VM basins under the three sampling frequencies. However, with regard to the *MPE* metric, the model with three hidden layers shows the best performance for the GM basin, with the sampling frequencies of 6 and 12. Moreover, with regard to the statistical index, the model with three hidden layers shows the best performance for the MS basin, with a sampling frequency of 6.

Table 2. Results for relative root mean squared error (rRMSE) and mean percentage error (MPE) based on the LSTM model with various sampling frequencies and hidden layers.

	One Layer						Two Layers					
	6		12		24		6		12		24	
	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE
CY	1.53	−46.19	1.24	−32.78	1.28	−35.79	1.12	−25.77	1.11	−26.30	1.10	−25.60
GD	1.69	−51.78	1.95	−67.33	1.37	−43.77	1.31	−41.32	1.43	−48.97	1.30	−38.66
GM	0.86	−23.35	0.82	−20.49	0.65	−11.40	0.69	−11.86	0.68	−10.58	0.70	−13.22
MM	8.06	−233.19	62.43	−1217.69	59.22	−1143.84	7.35	−210.05	58.88	−1149.37	59.31	−1105.35
MS	1.50	−32.30	1.48	−30.11	1.68	−45.17	1.45	−30.50	1.42	−29.76	1.46	−27.39
RS	4.21	−138.12	19.54	−244.74	20.65	−255.74	3.92	−116.42	18.69	−235.07	18.19	−234.51
VM	5.08	−138.29	6.20	−170.51	41.94	−870.96	3.78	−99.52	4.03	−105.71	31.77	−762.82

(a)

	Three Layers						Four Layers					
	6		12		24		6		12		24	
	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE
CY	1.12	−25.12	1.10	−25.17	1.10	−25.21	0.99	−16.34	0.98	−17.88	0.94	−16.73
GD	1.27	−39.37	1.30	−41.82	1.23	−36.14	1.20	−38.68	1.19	−39.97	1.12	−33.63
GM	0.66	−8.33	0.65	−8.36	0.65	−8.48	0.63	−8.65	0.62	−8.38	0.61	−6.04
MM	6.70	−181.87	58.63	−1081.55	53.49	−1048.46	6.32	−167.85	55.58	−1078.45	49.50	−993.11
MS	1.42	−28.26	1.39	−27.24	1.42	−27.16	1.41	−29.26	1.39	−27.07	1.39	−26.83
RS	3.61	−111.97	18.64	−244.38	17.08	−205.40	3.34	−96.48	15.25	−191.63	15.23	−193.68
VM	3.60	−92.46	3.65	−93.13	26.72	−631.00	3.39	−90.93	3.35	−83.21	20.51	−497.04

(b)

Figure 5a shows the *rRMSEs* for the various numbers of hidden layers, which ranged from one to four, based on the three sampling frequencies for the seven basins. Particularly, Figure 5a shows that the model with four hidden layers exhibits relatively good performance with regard to estimating the nitrate-N loads in CY, GD, and GM compared to the other models. Furthermore, the models with three and four hidden layers exhibit a similar or slightly better performance with regard to the MM, MS, RS, and VM basins. Moreover, Figure 5a shows that the LSTM model with four hidden layers exhibits enhanced performance with regard to the *rRMSE* criterion for all the studied river basins. Furthermore, improved nitrate-N load estimation accuracy is obtained with regard to the CY, GD, GM, and MS basins when the sampling frequency increases. In contrast, a relatively poor performance is observed with regard to the MM and RS basins, even with the use of high sampling frequencies such as 12 and 24. Relatively good performance is obtained for the VM basin when sampling frequencies of 6 and 12 are used. We present a discussion of the results obtained for the river basins, presented in Section 5, by examining the characteristics of discharge and nitrate-N concentrations.

Figure 5b shows the *MPEs* corresponding to the river basins analyzed in this study for the model with four hidden layers with sampling frequencies of 6, 12, and 24. Similar to the *rRMSE* criterion, the *MPE* metric also indicates that the model with four hidden layers exhibits better performance than the models with other numbers of layers, with regard to nitrate-N load estimation for the CY, GD, and GM basins. The models with three or four hidden layers demonstrate a similar or slightly better performance considering the other river basins, including the MM, MS, RS, and VM basins. Based on the results pertaining to the *MPE* criterion, we determined that enhanced performance is obtained for the CY, GD, GM, and MS basins when the sampling frequency of 24 is used, whereas the use of the sampling frequency of 6 or 12 leads to better performance with regard to the MM, RS, and VM basins. These patterns, which may be caused by features of discharge and nitrate-N concentration in the river basins, were also observed via analysis of the *rRMSE* metric.

Figure 6 presents the relationship between discharge and nitrate-N concentration for the seven river basins. In Figure 6, the upper panel shows the relationship between discharge and nitrate-N concentration via a log-log plot, and the lower panel shows a log-log plot of nitrate-N concentration against discharge, based on the binning method, for each basin. As shown in the lower panel, we determined the slope based on the average discharge values that were estimated from non-overlapping bins of nitrate-N concentrations. The regression line was fitted to the plot in the log-log scale; furthermore, it is presented using the dashed line to validate the significance of the observed slopes. The slopes corresponding to the CY, GD, GM, MM, MS, RS, and VM basins are −0.35, 0.01, −0.05, 0.50, 0.44, 0.50, and 0.54, respectively. The CY basin has the largest negative slope of −0.35, demonstrating a different behavior compared to that of the other river basins. The large negative slope implies that nitrate-N concentration increases as the discharge decreases. This result is expected because the CY basin is characterized by a highly impervious watershed entailing an urban area, which affects the nitrate-N concentration [41,42]. In contrast, the MM, RS, and VM basins exhibit large positive slopes, ranging from 0.50 to 0.54. The large positive slope indicates that the nitrate-N concentration increases with an increase in the discharge values. Considering these basins, the performance of the model, with regard to the *rRMSE* and *MPE* indices, decreases when the sampling frequency increases. The large slopes of 0.50 or higher influence nitrate-N load estimation by decreasing the accuracy of the model when large sampling frequencies are employed.
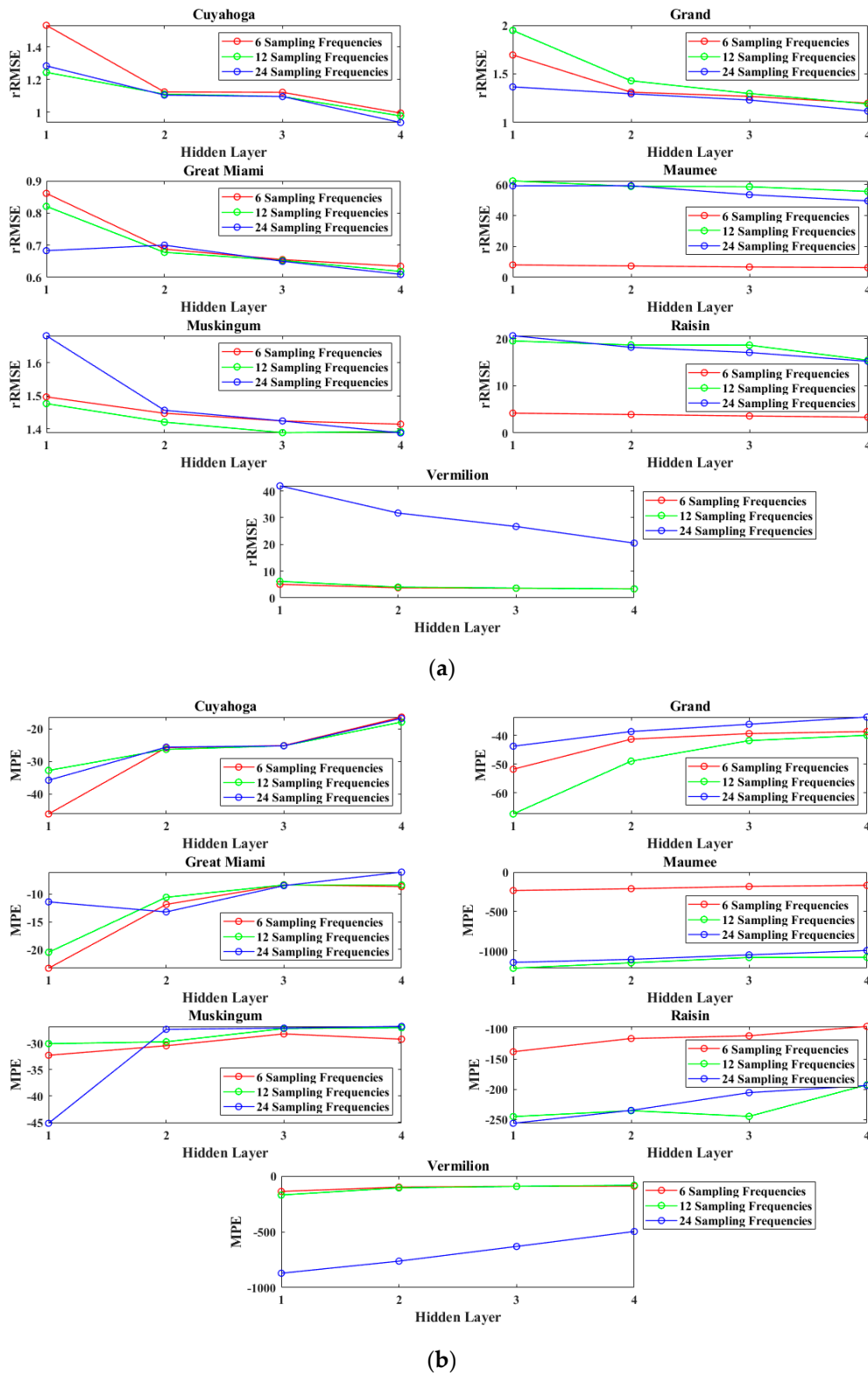
(**a**)

(**b**)

Figure 5. Plot for (**a**) rRMSE and (**b**) MPE of the LSTM model for different hidden layers.
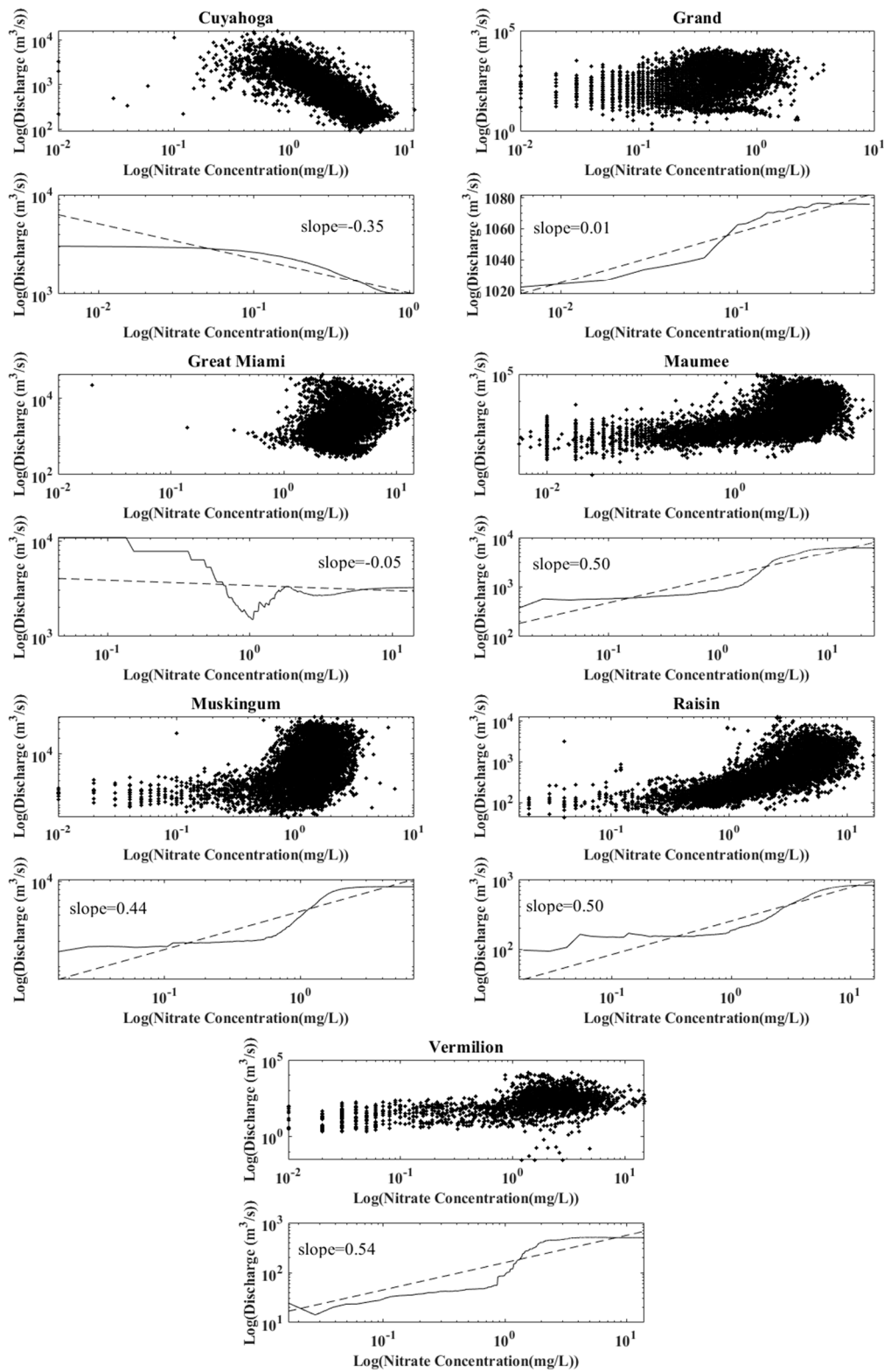
Figure 6. Plot of the upper panel for nitrate-N against discharge in the log–log scale and the lower panel for the variables based on the binning method. The dashed line indicates the slope of the variables.

4.2. Comparison of LSTM and WRTDS Models for Nitrate-N Load Estimation

To validate the proposed model with regard to the estimation of nitrate-N loads, the WRTDS model was applied to estimate the nitrate-N loads of the seven river basins based on the sampling

frequencies of 6, 12, and 24. Hirsch et al. [14] suggested the use of the WRTDS model to estimate daily nitrate-N concentrations, and Kandel and Bhattarai [43] compared various methods, including the WRTDS technique, for predicting nitrate-N loads. Based on the LSTM and WRTDS models, the *rRMSE* and *MPE* metrics for the nitrate-N load estimation were analyzed, as shown in Table 3. For the LSTM model, four hidden layers were used to compare its performance in obtaining nitrate-N load estimations with that of the WRTDS model. The blue font shows the best performing method.

Table 3. Results of daily load estimation based on rRMSE and MPE indices for the LSTM and weighted regressions on time discharge and season (WRTDS) models.

| | LSTM (Four Layers) | | | | | | WRTDS | | | | | |
| | 6 | | 12 | | 24 | | 6 | | 12 | | 24 | |
	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE	rRMSE	MPE
CY	0.99	−16.34	0.98	−17.88	0.94	−16.73	4.27	−16.87	4.16	−16.81	4.16	−16.98
GD	1.20	−38.68	1.19	−39.97	1.12	−33.63	2.80	−85.97	2.69	−79.46	2.64	−77.73
GM	0.63	−8.65	0.62	−8.38	0.61	−6.04	2.45	−14.59	2.10	−13.75	1.90	−13.12
MM	6.32	−167.85	55.58	−1078.45	49.50	−993.11	18.16	−300.70	20.09	−317.02	20.38	−322.22
MS	1.41	−29.26	1.39	−27.07	1.39	−26.83	3.01	−41.34	2.94	−39.05	2.94	−39.30
RS	3.34	−96.48	15.25	−191.63	15.23	−193.68	3.39	−41.57	3.21	−36.54	3.15	−35.74
VM	3.39	−90.93	3.35	−83.21	20.51	−497.04	12.17	−281.05	12.83	−276.61	11.30	−269.52

The average loads estimated using the LSTM and WRTDS models for the sampling frequencies of 6, 12, and 24 are shown in Figure 7. From the figure, we can observe that the estimation error and the bias seem to increase with the load, particularly when the load is substantially large. Regarding CY basin, the LSTM model tends to overestimate the load, whereas the WRTDS model tends to underestimate it. In contrast, in the case of VM basin, the load estimated using the LSTM model seems to be underestimated, whereas that estimated with WRTDS seems to be overestimated. Similar estimation performance is observed for the other basins with regard to the estimation of nitrate-N load using the two models.

Furthermore, box plots were obtained for each river basin based on the *rRMSE* criterion for the analysis of nitrate-N load estimates using the LSTM and WRTDS models. Figure 8 presents the box plots for the *rRMSE* of nitrate-N load estimation with three sampling frequencies. In Figure 8, the centerline of each box plot shows the median value for the estimation, and the top and bottom of each plot represent the 75th and 25th percentiles. The outliers are shown as cross symbols. Figure 8 shows that the LSTM model provides better nitrate-N load estimates compared to the WRTDS model for the sampling frequencies of 6, 12, and 24 in the cases of the CY, GD, GM, and MS basins. Regarding the MM and RS basins, the LSTM model also shows a better performance compared to that of the WRTDS model for a sampling frequency of 6. Considering the VM basin, the LSTM model demonstrates a better performance compared to that of the WRTDS model for the sampling frequencies of 6 and 12. As the results of the LSTM model demonstrate, the results obtained from the WRTDS model show worse performance when the number of sampling frequencies increases in the cases of the MM and VM basins. This observation may result from the slope of the average discharge obtained using the binning method, as discussed in Section 4.1.
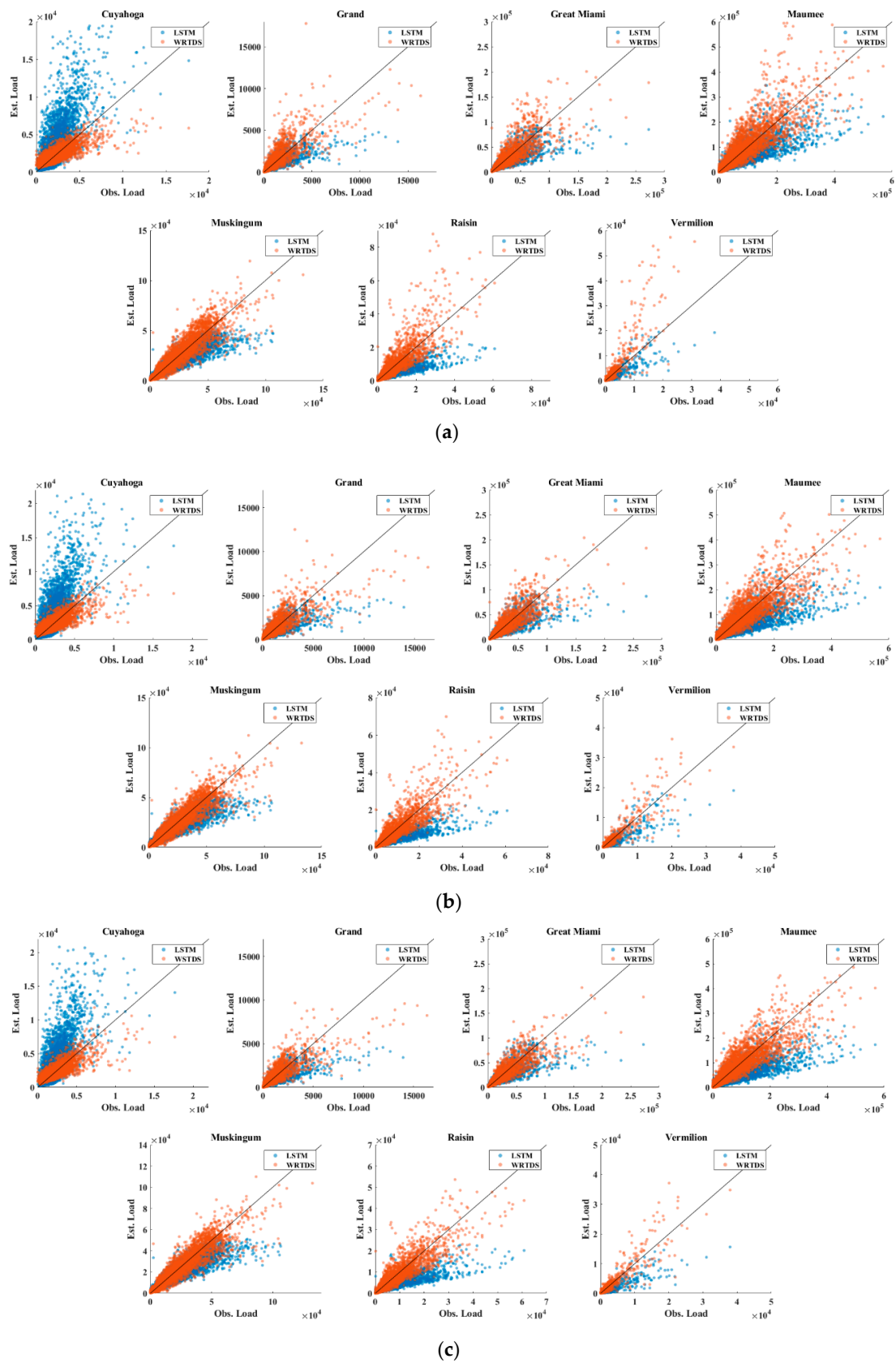
(**a**)

(**b**)

(**c**)

Figure 7. Plot for scatter of the observed load and estimated load derived from LSTM and WRTDS with (**a**) 6 sampling frequencies, (**b**) 12 sampling frequencies, and (**c**) 24 sampling frequencies.
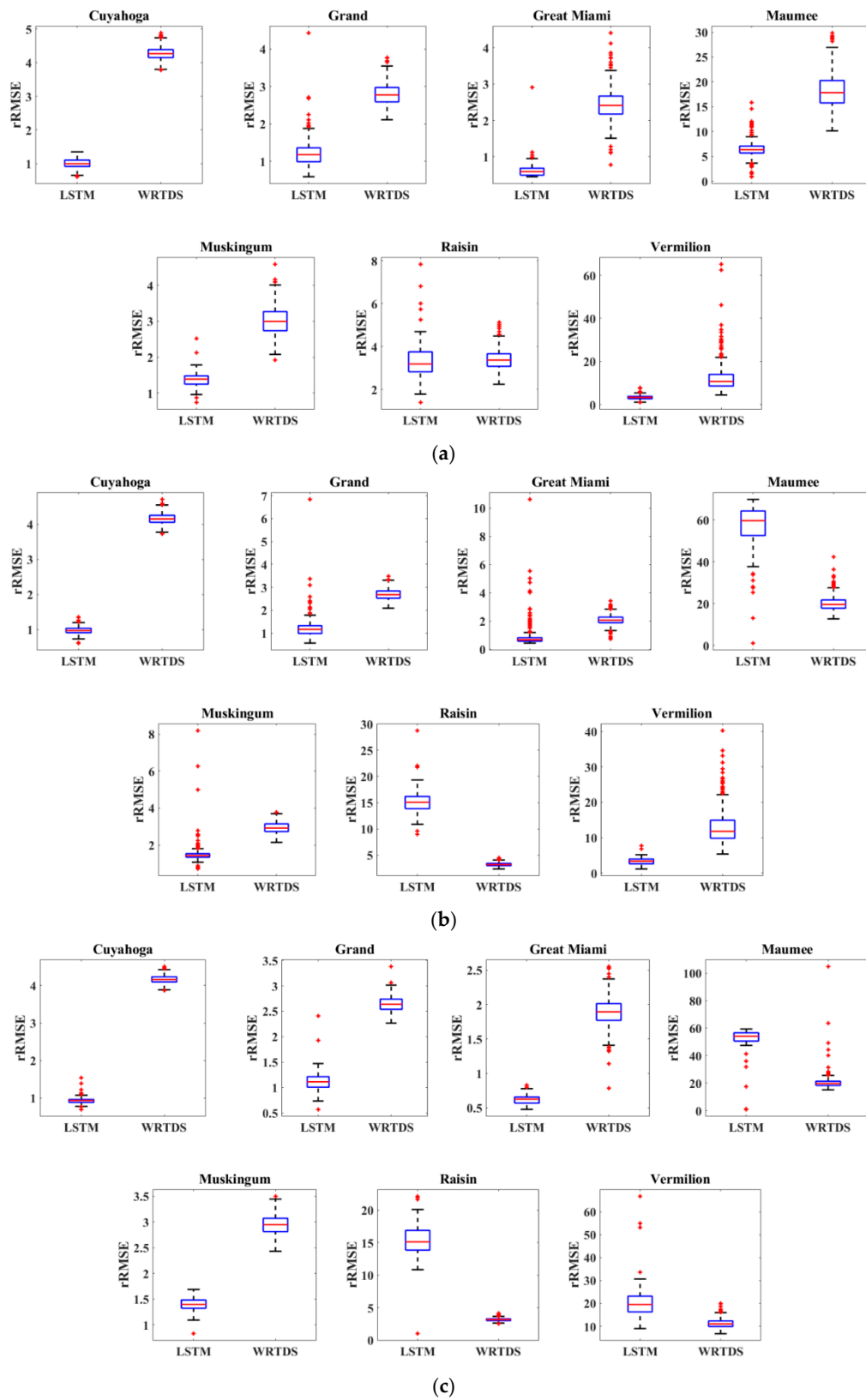
Figure 8. Boxplot of rRMSE based on (**a**) 6 sampling frequencies, (**b**) 12 sampling frequencies, and (**c**) 24 sampling frequencies.

The analysis of the box plots based on the *MPE* was performed for each river basin considering the three sampling frequencies. Figure 9 shows the box plots for the *MPE* criterion of nitration load estimation derived from the LSTM and WRTDS models. The LSTM model performance in the cases of the CY, GD, GM, and MS basins was enhanced when the sampling frequency increased; this was also observed by examining the *rRMSE*. According to the *MPE* metric, the LSTM model with a sampling frequency of 6 showed better performance compared to that of the WRTDS model with regard to the GD, GM, MM, MS, and VM basins. The LSTM model provided better estimates compared to the WRTDS model in the cases of the GD, GM, MS, and VM basins for a sampling frequency of 12, and in the cases of the GD, GM, and MS basins for a sampling frequency of 24. The analysis of the load estimates shows that the LSTM model may provide enhanced load estimates when the sampling frequency is increased in the cases of the CY, GD, GM, and MS basins. Furthermore, the improved estimations can be used for water quality management.
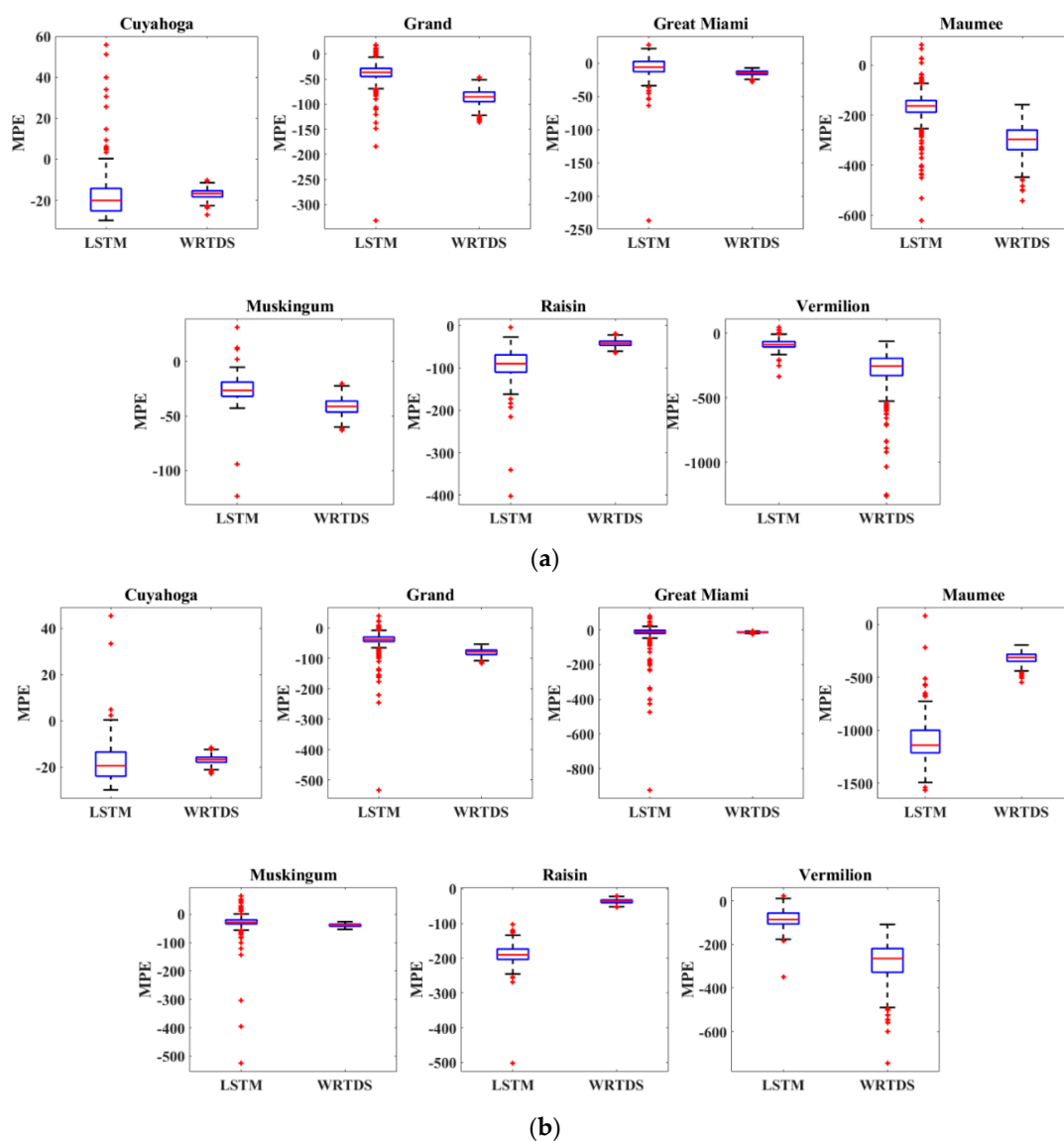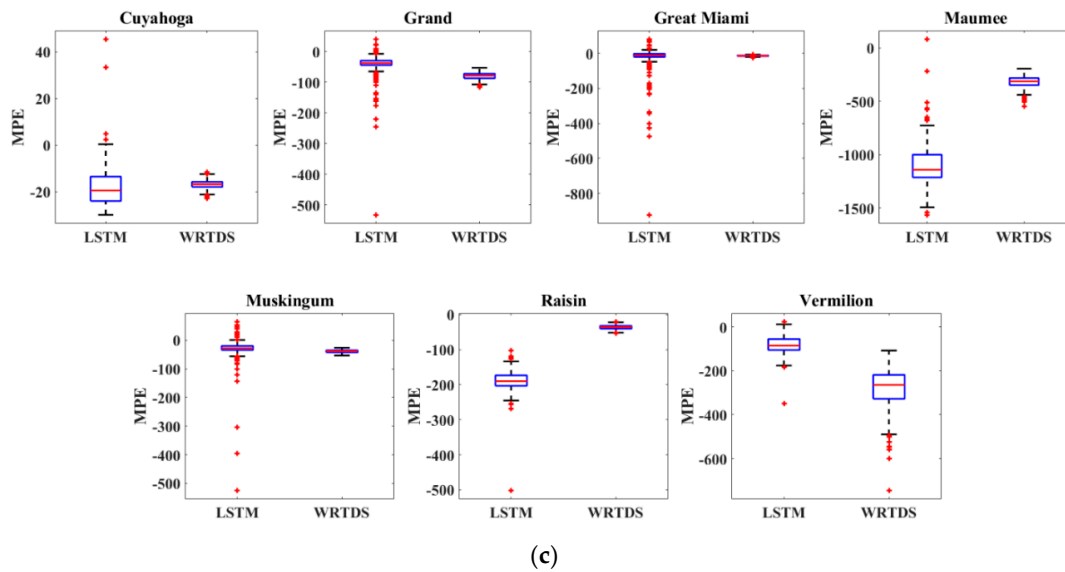

Figure 9. *Cont.*

(**c**)

Figure 9. Boxplot of MPE based on (**a**) 6 sampling frequencies, (**b**) 12 sampling frequencies, and (**c**) 24 sampling frequencies.

5. Discussion

The analysis of nitrate-N load estimation based on the LSTM and WRTDS models was performed for the seven river basins. The models tend to provide improved load estimates when the sampling frequency increases, except for the MM, RS, and VM basins. This may be because hydrological characteristics such as the amount of discharge affect the ability of the model to obtain an accurate load estimation. A correlation study was performed for discharge and nitrate-N concentration and discharge and nitrate-N load for the river basins considered in this study. In Figure 10a, the correlation coefficients between the discharge and nitration concentration range from −0.584 to 0.523, while the R-squared value ranges from 0.007 to 0.341. Figure 10b shows that correlation coefficients between the discharge and nitrate-N load for the seven basins range from 0.776 to 0.919, and the R-squared value ranges from 0.601 to 0844. The red line indicates the linear regression line; it shows the trend of the discharge with regard to the nitrate-N concentration and load.
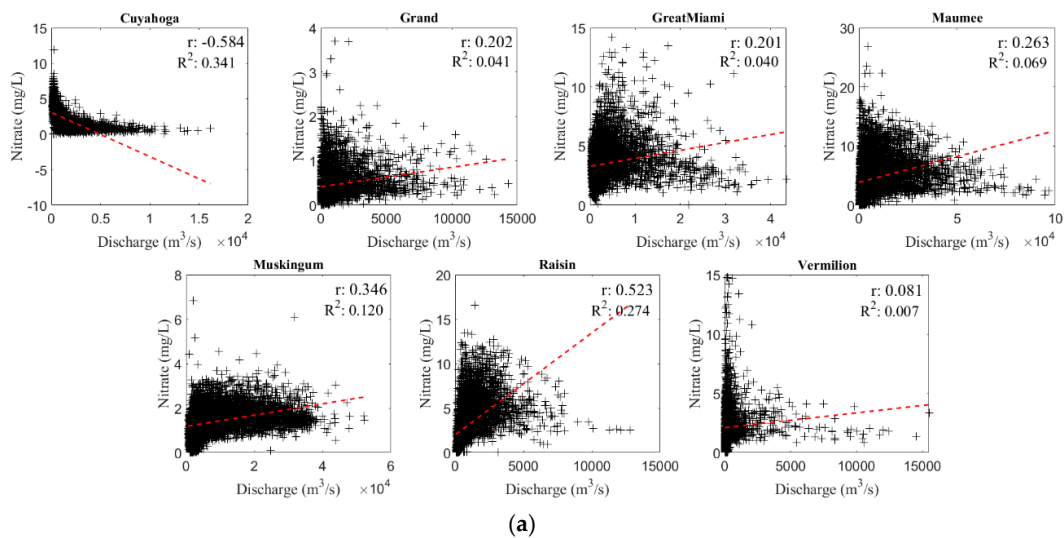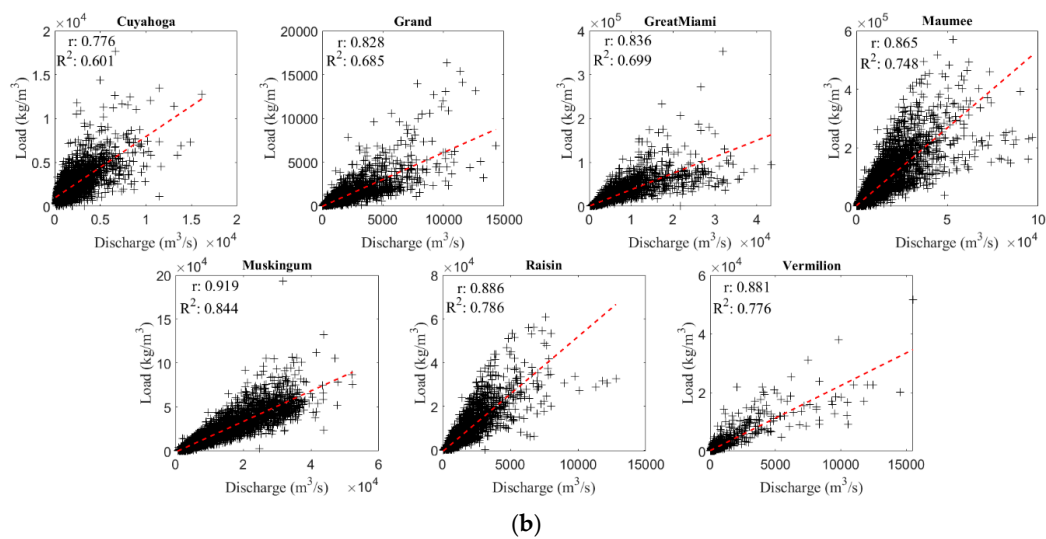

(**a**)

Figure 10. *Cont.*

Figure 10. Relationship (**a**) between daily discharge and daily nitrate-N concentration and (**b**) between annual discharge and daily nitrate-N load for the seven river basins. r implies the correlation coefficient and R^2 means the R-squared. A red line indicates the linear regression line.

Table 4 presents statistical estimates for average discharge values using the binning of nitrate-N concentrations. The maximum values of discharge range from 508.59 to 10,900.00 m^3/s, and the minimum values range from 14.20 to 1507.74 m^3/s. The standard deviation range for the values is from 14.62 to 2254.02 m^3/s. The large value of the slope, the corresponding lowest discharge of which is below 1000 m^3/s, affects the nitrate-N load estimation results of the proposed LSTM model. Relatively worse performance is obtained for the MM, RS, and VM basins when large sampling frequencies are employed, with the related discharge values being lower than 1000 m^3/s. The discharge characteristics and nitrate-N concentration also affect nitrate-N load estimation in the current study. These specific characteristics may limit the applicability of the LSTM approach. Therefore, the model may be improved using a combination of relevant features of the nitrate-N concentration and discharge based on canonical correlation analysis [4]. Several studies have been conducted to evaluate the effects of hydrological and biogeochemical processes on the relationship between nitrate-N concentration and discharge [44–47]. Duncan et al. [45] found that the hydrological variability on a seasonal scale affects nitrate-N concentration in streams, and the slopes of the discharge and nitrate-N concentration are different in a wet year compared to those in a dry year, which is typically characterized by low discharge. In this study, the models with low discharge tend to be sensitive to the nitrate-N load estimation. Similarly, the concentration and discharge relationship analyzed by Duncan et al. [45] was found to be highly dependent on wetness. Future research should focus on sensitivity analysis based on different characteristics of this relationship with regard to nitrate-N load estimation.

Table 4. Statistical estimates for the average discharge based on the binning of the nitrate-N concentration for the seven basins.

	Max	Min	Standard Deviation	Slope
CY	3058	1010	627	−0.35
GD	1077	1022	15	0.01
GM	10,900	1496	1011	−0.05
MM	6178	368	1585	0.50
MS	8608	1508	2254	0.44
RS	822	94	221	0.50
VM	508.59	14.20	127.72	0.54

6. Conclusions

In this study, LSTM was employed for the estimation of nitrate-N loads. The estimated loads can be used to control nutrient enrichment and water pollutants to improve water quality in river basins. The proposed LSTM model was designed based on long-term data records of discharge and nitrate-N concentration in seven river basins in the United States. The Monte Carlo sample method with periodic sampling frequencies of 6, 12, and 24 was applied to the uniformly distributed random variable. The proposed model was evaluated using the *rRMSE* and *MPE* statistical indices and with the WRTDS model for the comparison of model performance.

The appropriate number of hidden layers in the LSTM model was determined to enhance model performance. The statistical metrics showed that the use of three or four hidden layers provided good nitrate-N load estimates. Furthermore, relatively good performance was obtained using four hidden layers in the cases of the CY, GD, and GM basins, whereas a similar or better estimation was obtained with three or four hidden layers in the cases of the MM, MS, RS, and VM basins. Finally, in this study, we used four hidden layers in estimating the nitrate-N loads. When four hidden layers were employed, the LSTM model exhibited an increase in performance with an increase in the sampling frequency, except for the MM, RS, and VM basins. Relatively good nitrate-N load estimates were obtained for the MM, RS, and VM basins using a sampling frequency of 6, compared to those obtained using sampling frequencies of 12 and 24. These differences may be caused by characteristics of discharge and nitrate-N concentration.

To evaluate the proposed LSTM model, the WRTDS model was applied for obtaining nitrate-N load estimates for the seven river basins. The *rRMSE* and *MPE* were represented using box plots for the three sampling frequencies. With regard to the *rRMSE*, the LSTM model in the cases of the CY, GD, GM, and MS basins provided better estimates compared to the WRTDS model for all three sampling frequencies. Considering the MM, RS, and VM basins, the proposed model exhibited a better performance compared to that of WRTDS, based on the *rRMSE*, with sampling frequencies of 6 or 12. With regard to the *MPE*, the LSTM model produced better estimates using a sampling frequency of 6 compared to the WRTDS model in the cases of the GD, GM, MM, MS, and VM basins. Furthermore, the proposed model showed better estimates compared to the WRTDS model in the cases of the GD, GM, MS, and VM basins when using sampling frequencies of 12 or 24. Although the LSTM model employing any of the three sampling frequencies may not be applicable to all river basins with regard to estimating nitrate-N loads, reasonable estimates were obtained for most river basins used in this study. The results of this study show that the LSTM model seems to have excellent potential for being applied to solve environmental issues by reducing water pollutants in river basins.

Future work should focus on the extension of nitrate-N load estimation via LSTM to other river basins in various environments. For example, the nine large tributaries in Chesapeake Bay, analyzed by Hirsch et al. [12], can be explored using the proposed model to reduce nutrient enrichment. Future research should also investigate other aspects related to the relationship between nitrate-N concentration and discharge in various river basins. The extension of the significant variables in the model with regard to estimating nitrate-N loads will be of interest to improve model estimates.

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

References

1. 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]
2. Rabalais, N.N.; Turner, R.E.; Justic, D.; Dortch, Q.; Wiseman, W.J., Jr.; Sen Gupta, B.K. Nutrient changes in the Mississippi River and system responses on the adjacent continental shelf. *Estuaries* **1996**, *19*, 386–407. [CrossRef]
3. 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]
4. 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]
5. Ohio EPA. *Nutrient Mass Balance Study for Ohio's Major Rivers*; Ohio Environmental Protection Agency: Columbus, OH, USA, 2016.
6. Illinois State Water Survey. *An Evaluation of Baseline Nutrient Loadings, Their Trends, and the Effects of Land-Use and Climate Variations in the Illinois River Watershed*; Illinois State Water Survey: Champaign, IL, USA, 2014.
7. Lalonde, V.; Madramootoo, C.A.; Trenholm, L.; Broughton, R.S. Effects of controlled drainage on nitrate concentrations in subsurface drain discharge. *Agr. Water Manag.* **1996**, *29*, 187–199. [CrossRef]
8. Im, S.; Brannan, K.M.; Mostaghimi, S. Simulating hydrologic and water quality impacts in an urbanizing watershed 1. *JAWRA J. Am. Water Resour. Assoc.* **2003**, *39*, 1465–1479. [CrossRef]
9. Schilling, K.E.; Wolter, C.F. Modeling nitrate-nitrogen load reduction strategies for the Des Moines River, Iowa using SWAT. *Environ. Manag.* **2009**, *44*, 671–682. [CrossRef]
10. Ullrich, A.; Volk, M. Influence of different nitrate–N monitoring strategies on load estimation as a base for model calibration and evaluation. *Environ. Monit. Assess.* **2010**, *171*, 513–527. [CrossRef]
11. 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]
12. Runkel, R.L.; Crawford, C.G.; Cohn, T.A. *Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers: U.S. Geological Survey Techniques and Methods Book 4*; US Geological Survey: Reston, VA, USA, 2004.
13. 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]
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 J. Am. Water Resour. Assoc.* **2010**, *46*, 857–880. [CrossRef] [PubMed]
15. Anctil, F.; Filion, M.; Tournebize, J. A neural network experiment on the simulation of daily nitrate-nitrogen and suspended sediment fluxes from a small agricultural catchment. *Ecol. Model.* **2009**, *220*, 879–887. [CrossRef]
16. Yu, C.; Northcott, W.J.; McIsaac, G.F. Development of an artificial neural network model for hydrologic and water quality modeling of agricultural watersheds. *Trans. ASAE* **2004**, *47*, 285–290. [CrossRef]
17. 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]
18. Khu, S.T.; Werner, M.G.F. Reduction of Monte-Carlo simulation runs for uncertainty estimation in hydrological modelling. *Hydrol. Earth Syst. Sci.* **2003**, *7*, 680–692. [CrossRef]
19. Shapiro, A. Monte Carlo sampling methods. *Handb. Oper. Res. Manag. Sci.* **2003**, *10*, 353–425.
20. Richards, R.P.; Holloway, J. Monte Carlo studies of sampling strategies for estimating tributary loads. *Water Resour. Res.* **1987**, *23*, 1939–1948. [CrossRef]
21. Rahman, A.; Weinmann, P.E.; Hoang, T.M.T.; Laurenson, E.M. Monte Carlo simulation of flood frequency curves from rainfall. *J. Hydrol.* **2002**, *256*, 196–210. [CrossRef]
22. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
23. Yuan, X.; Chen, C.; Lei, X.; Yuan, Y.; Adnan, R.M. Monthly runoff forecasting based on LSTM–ALO model. *Stoch. Environ. Res. Risk A* **2018**, *32*, 2199–2212. [CrossRef]

24. Zhang, J.; Zhu, Y.; Zhang, X.; Ye, M.; Yang, J. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* **2018**, *561*, 918–929. [CrossRef]

25. Tian, Y.; Xu, Y.-P.; Yang, Z.; Wang, G.; Zhu, Q. Integration of a parsimonious hydrological model with recurrent neural networks for improved streamflow forecasting. *Water* **2018**, *10*, 1655. [CrossRef]

26. Bowes, B.D.; Sadler, J.M.; Morsy, M.M.; Behl, M.; Goodall, J.L. Forecasting Groundwater Table in a Flood Prone Coastal City with Long Short-term Memory and Recurrent Neural Networks. *Water* **2019**, *11*, 1098. [CrossRef]

27. Xiang, Z.; Yan, J.; Demir, I. A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water Resour. Res.* **2020**, *56*, e2019WR025326. [CrossRef]

28. USGS, USGS Surface-Water Data for the Nation. 2018. Available online: https://waterdata.usgs.gov/nwis/sw (accessed on 10 June 2020).

29. Heidelberg University, Tributary Data Download. 2018. Available online: https://www.heidelberg.edu/tributary-data-download (accessed on 10 June 2020).

30. Verma, S.; Markus, M.; Bartosova, A.; Cooke, R.A. Intra-annual variability of riverine nutrient and sediment loadings using weighted circular statistics. *J. Environ. Eng.* **2018**, *144*, 04018010. [CrossRef]

31. Zhang, Q.; Brady, D.C.; Boynton, W.R.; Ball, W.P. Long-term trends of nutrients and sediment from the nontidal Chesapeake Watershed: An assessment of progress by river and season. *JAWRA J. Am. Water Resour. Assoc.* **2015**, *51*, 1534–1555. [CrossRef]

32. Tukey, J.W. *Exploratory Data Analysis*; Addison-Wesley: Boston, MA, USA, 1977.

33. Hirsch, R.; De Cicco, L. *User Guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R Packages for Hydrologic Data. Technical Report Techniques and Methods Book 4*; US Geological Survey: Reston, VA, USA, 2014.

34. Johnes, P. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *J. Hydrol.* **2007**, *332*, 241–258. [CrossRef]

35. Alobaidi, M.H.; Marpu, P.R.; Ouarda, T.B.M.J.; Chebana, F. Regional frequency analysis at ungauged sites using a two-stage resampling generalized ensemble framework. *Adv. Water Resour.* **2015**, *84*, 103–111. [CrossRef]

36. Jung, K.; Kim, E.; Kang, B. Estimation of Low-Flow in South Korean River Basins Using a Canonical Correlation Analysis and Neural Network (CCA-NN) Based Regional Frequency Analysis. *Atmosphere* **2019**, *10*, 695. [CrossRef]

37. Requena, A.I.; Ouarda, T.B.M.J.; Chebana, F. Low-flow frequency analysis at ungauged sites based on regionally estimated streamflows. *J. Hydrol.* **2018**, *563*, 523–532. [CrossRef]

38. Shu, C.; Ouarda, T.B.M.J. Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resour. Res.* **2007**, *43*. [CrossRef]

39. Jung, K.; Ouarda, T.B.M.J. Classification of drainage network types in the arid and semi-arid regions of Arizona and California. *J. Arid Environ.* **2017**, *144*, 60–73. [CrossRef]

40. Mejía, A.I.; Niemann, J.D. Identification and characterization of dendritic, parallel, pinnate, rectangular, and trellis networks based on deviations from planform self-similarity. *J. Geophys. Res. Earth Surf.* **2008**, *113*. [CrossRef]

41. 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] [PubMed]

42. 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]

43. Kandel, R.; Bhattarai, R. Comparison of various estimation techniques to predict nitrate load in Maumee River. In Proceedings of the 2018 ASABE Annual International Meeting, Detroit, MI, USA, 29 July–1 August 2018.

44. Duncan, J.M.; Band, L.E.; Groffman, P.M. Variable nitrate concentration–discharge relationships in a forested watershed. *Hydrol. Process.* **2017**, *31*, 1817–1824. [CrossRef]

45. Duncan, J.M.; Welty, C.; Kemper, J.T.; Groffman, P.M.; Band, L.E. Dynamics of nitrate concentration-discharge patterns in an urban watershed. *Water Resour. Res.* **2017**, *53*, 7349–7365. [CrossRef]

46. Hagebro, C.; Bang, S.; Somer, E. Nitrate load/discharge relationships and nitrate load trends in Danish rivers. In *Dissolved Loads of Rivers and Surface Water Quantity/Quality Relationships*; International Association of Hydrological Sciences: Wallingford, UK, 1983; pp. 377–386.
47. 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]