https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

Application of Sparse Sensing for Stream Nutrient Monitoring Programs

Wasif Bin Mamoon[1], Kun Zhang[2], Mitul Luhar[3] and Anthony J. Parolari[1]

[1]Department of Civil, Construction, and Environmental Engineering, Marquette University, Milwaukee, WI, USA.
[2]Department of Civil and Environmental Engineering, University of Minnesota Duluth, Duluth, MN, USA.
[3]Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA, USA.

Correspondence to: Anthony J. Parolari (anthony.parolari@marquette.edu)

Abstract. Existing stream nutrient (i.e., nitrogen or phosphorus) monitoring approaches often exhibit low sampling frequency and associated high uncertainty due to financial and maintenance constraints. Data-driven Sparse Sensing (DSS) offers an alternative approach to estimate concentration data at high resolution using fewer measurements. DSS transforms stream data from training locations into a reduced-dimension space and identifies optimal sampling times. These optimal measurements are used to reconstruct high-resolution concentration data at target locations. In this study, we used the DSS framework to estimate stream nutrient concentrations and loads (nitrate nitrite as NOx and total phosphorus as TP) across the US Midwest region, with additional analyses in other hydrologic regions to examine regional variability in optimal sampling times. The modeling approach was designed to address key issues in nutrient monitoring programs, including determining the required size and length of training data and establishing data selection procedures. The base model predicted NOx and TP concentrations and loads with good accuracy (NSE> 0.5, load error < ±4% for NOx; NSE>0.45, load error < ±10% for TP) using only 40-60 samples per year (~10-15% of total measurements). Optimal sampling times were concentrated in spring and early summer in the Midwest but varied across regions. Findings indicated that DSS requires only 2-3 years of training data from either 10-15 regional monitoring locations (NOx and TP) or the target location itself (NOx). This study demonstrates that DSS can be integrated into nutrient monitoring programs to generate high-resolution stream data, estimate loads, and support resource-efficient management decisions.

1 Introduction

While nutrients (i.e., nitrogen and phosphorus) are necessary for plant and microbial growth, excessive amounts in surface waters can trigger eutrophication, adversely affecting human health and the environment. Nutrient pollution results in economic losses arising from social, ecological, and policy-related issues (Pretty et al., 2003). In the US, nutrient pollution-related costs are estimated to total more than \$2B per year in freshwater systems, mostly due to reduced property values and increased expenses in drinking water treatment processes (Dodds et al., 2009). In addition, coastal areas face an annual economic loss of approximately \$100M in the US (Davidson et al., 2014). While attempts to lessen nutrient loading from

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

point sources have been successful, managing pollution from nonpoint sources, especially from pasture and croplands, remains a challenge (Carpenter et al., 1998; Stets et al., 2020; Wurtsbaugh et al., 2019).

To manage nutrient pollution effectively, reliable nutrient load estimates are needed. These can be obtained via high-frequency water quality sampling, which minimizes the error or uncertainty in load estimations (Gao et al., 2020; Wang et al., 2024) and captures changes that occur over short time scales (Pellerin et al., 2016). As a result, hourly or daily monitoring is often recommended for nutrients (Cassidy & Jordan, 2011; Gao et al., 2020; Jones et al., 2012; Minaudo et al., 2017; Reynolds et al., 2016; Valkama & Ruth, 2017). However, most monitoring programs sample at biweekly or monthly intervals because of the costs associated with regular grab sampling and the high expense of installing and maintaining in-situ sensors (Blaen et al., 2016; Pellerin et al., 2016; Read et al., 2017; Sprague et al., 2017). Monitoring phosphorus is particularly challenging as it requires highly selective analytical methods to differentiate forms and detect small concentrations (Worsfold et al., 2016).

Nutrient models (i.e., statistical, physical, or deep learning approaches) have been widely used to estimate continuous nutrient loads and concentrations. However, their efficacy is often compromised by substantial data requirements, site-specific parameter calibration, inappropriate scalability, and inability to fully represent underlying physical mechanisms (Blöschl & Sivapalan, 1995; Castrillo & García, 2020; Y. Chen et al., 2020; Harrison et al., 2021; Holmberg et al., 2006; Hu et al., 2021; Khullar & Singh, 2021; Nour et al., 2006; Schärer et al., 2006; Varadharajan et al., 2022). Consequently, there is an increasing demand for alternative approaches that can accurately estimate nutrient loads and capture short-term changes while easing the requirement of continuous high-frequency sampling.

Many prior studies have reported environmental signals or time-series (i.e., stream flow, temperature, nutrient concentrations, dissolved oxygen, soil moisture) to be "sparse" (Katul et al., 2007; Kirchner & Neal, 2013; Parolari et al., 2021; Zhang et al., 2023b). Sparse signals can be represented with fewer coefficients than the number of individual samples in an appropriately transformed domain (e.g., Fourier). The sparsity of stream signals, where only a few active modes can capture most of the variance (Katul et al., 2007), can be leveraged to estimate continuous stream data, easing sampling requirements (Bin Mamoon et al., 2025; Zhang et al., 2023a; Zhang et al., 2023b). This eventually promotes the use of dimensionality-reduction techniques such as compressed sensing and data-driven sparse sensing to generate reduced-order models. These dimension reduction techniques rely on sparsity and low dimensionality to reconstruct high frequency data from minimal measurements (Donoho, 2006; Manohar et al., 2018).

Compressed sensing (CS) has been effective in compression and recovery of MRI images, radar data, soil moisture data, and land cover data (Lustig et al., 2008; Wei et al., 2017; L. Zhang et al., 2010). CS transforms the signals into a universal basis, the Fourier domain, where only a few active Fourier modes can represent the signals. For example, the variances of daily stream flow and concentration signals were captured with only 0.3% to 36% most energetic coefficients in the Fourier domain (Zhang et al., 2023a). Based on the sparsity of the signals, CS can reconstruct the signal with very few sparse measurements. With only 5-10% of total measurements, 15-minute stream flow and concentration time-series (from the Midwest region in the US) were successfully reconstructed using CS (Zhang et al., 2023a). Further, nitrogen loads were

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

estimated with an average error of 6.6% at an effective sampling frequency of 10 days. However, CS has some limitations for practical applications. First, the Fourier space is on a universal basis, which may not be optimal for all signals. Second, CS requires random sampling in time, which is difficult to implement in existing monitoring programs. To overcome these drawbacks of CS, alternative dimension reduction techniques, such as Data-driven Sparse Sensing (DSS), have been proposed (Manohar et al., 2018).

Data-driven sparse sensing (DSS) has been widely used in image reconstruction, reconstruction of two-dimensional flow and temperature fields, as well as optimal sensor placement (Manohar et al., 2022; Ohmer et al., 2022). DSS transforms the signal into a tailored reduced-dimension space, identified via a singular value decomposition, where the signals can be sparsely represented. Furthermore, DSS can identify optimal sampling times and locations based on their information content (Manohar et al., 2018; Zhang et al., 2023b). Using DSS, daily streamflow data across the US were reconstructed with only 5-50 samples per year (Zhang et al., 2023b). The optimal sampling times were distinct across the regions and showed seasonal patterns (Zhang et al., 2023b). The DSS framework was also applied to estimate continuous nutrient concentration data using high-frequency surrogate water quality data and sparse concentration data (Bin Mamoon et al., 2025). Results showed that with 20-80 samples per year, the DSS model (trained with flow and concentration data) estimated continuous and reliable concentration data at target locations.

Bin Mamoon et al. (2025) established that DSS can reconstruct high-resolution concentration data and loads from minimal measurements. This study builds on this prior work with a focus on addressing the practical challenges of incorporating DSS into stream monitoring. First, this study evaluates two training strategies to leverage existing water quality data to design a monitoring program, training on either regional data (i.e., from nearby locations) or local, historical data (i.e., from the same location). Within these training strategies, we identify the minimum training data required to achieve performant models in terms of the number of locations and time-series length. Secondly, this study investigates whether DSS performance depends on climate and watershed properties, toward identifying locations with potential for sampling improvements. Lastly, our prior work was restricted to a limited number of training data sets with multiple surrogate water quality variables. We found that DSS performed well with only flow and concentration data. Therefore, the objectives of the present study are addressed with an expanded data set including more locations with longer time-series, enabling analysis of DSS data requirements and the historical data training strategy. By addressing these challenges, this study informs the strategic implementation of the DSS framework into nutrient monitoring programs to ensure more efficient use of monitoring resources.

2 Data and Methodology

2.1 Data-driven Sparse Sensing (DSS) Framework

The DSS framework uses the singular value decomposition (SVD) and QR factorization to represent the signal in an appropriate reduced-dimension space and identify the most informative sampling times and locations (Manohar et al., 2018). SVD is used in the DSS framework to reduce dimensionality by detecting dominant spatial or temporal modes that capture

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

key features of the system dynamics (Kutz, 2013; Taira et al., 2017). These singular modes represent a tailored lower-dimensional subspace onto which higher-dimensional data can be projected, preserving important information while reducing data requirements (Brunton & Kutz, 2022). Once the dimensionality of the data is reduced, DSS uses QR factorization with column pivoting to rearrange the data matrix based on information content. This identifies the most informative rows or columns, which represent the best sampling times or sensor locations for estimation or reconstruction (Brunton & Kutz, 2022). Thus, the DSS framework ensures that sparse measurements offer enough information to minimize the number of samples needed for accurate reconstruction (Manohar et al., 2018; Zhang et al., 2023b). The DSS framework is illustrated in Fig. 1.

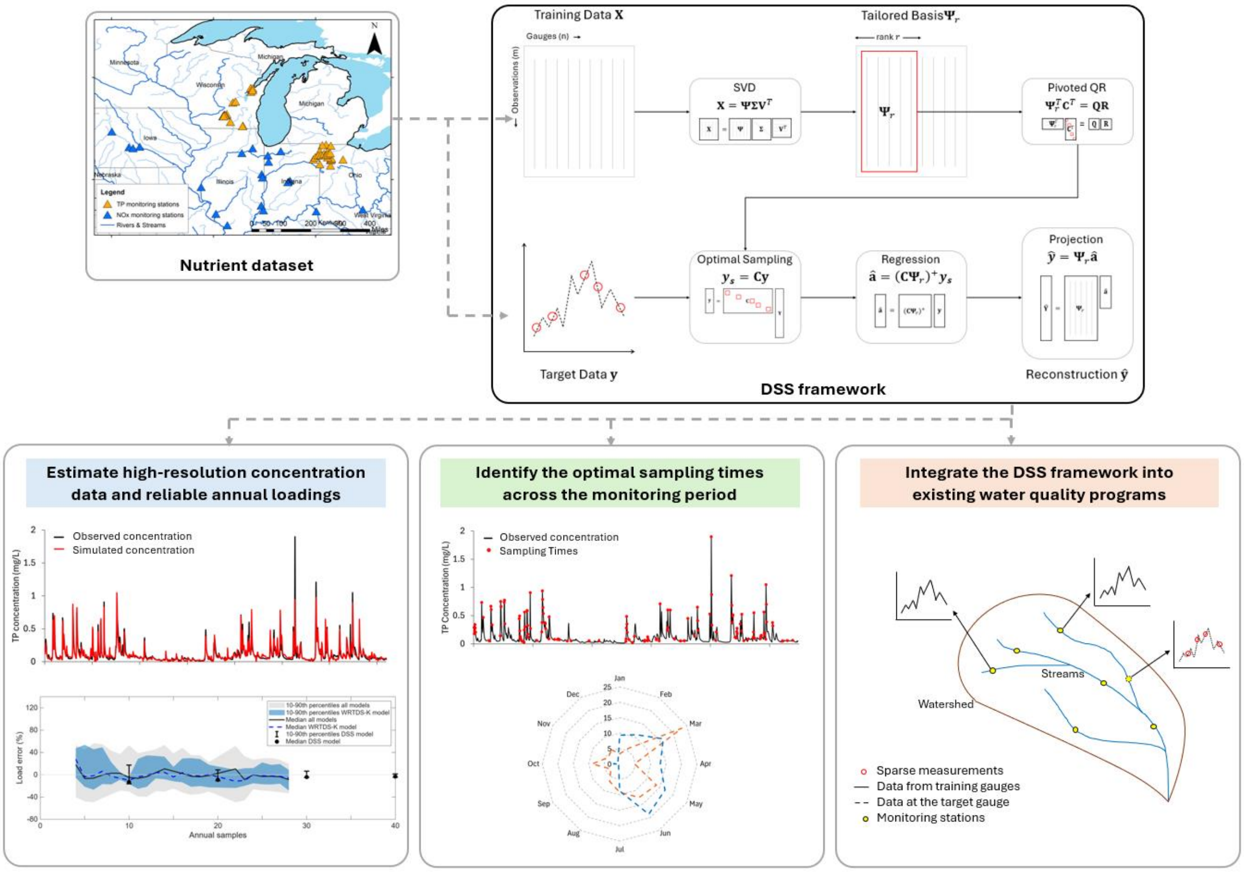


Figure 1: Data-driven sparse sensing (DSS) framework for reconstructing high-resolution stream nutrient concentration time series, identifying optimal sampling times, and improving monitoring efficiency in watershed systems. This figure was adapted from Bin Mamoon et al. (2025).

Reconstruction of high-frequency time series using DSS followed the procedure described below. Available flow and concentration data were organized into a matrix $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2 \dots \mathbf{a}_n] \in \mathbb{R}^{m\times n}$ where each of the $\boldsymbol{n}$ columns represented different stream monitoring locations, and each of the $\boldsymbol{m}$ row elements represented daily samples from those locations. For

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

each model iteration, one of the concentration time-series was selected as the target data $\mathbf{y}$ while 80% of the remaining flow and concentration time-series were randomly selected as the training data $\mathbf{X}$. To ensure that the predictive performance remained consistent despite variations in the training set, we repeated the selection of training gauges 5 times for each target gauge prediction. An SVD of training data $\mathbf{X}$:

$$\mathbf{X} = \mathbf{\Psi\Sigma V}^{\mathrm{T}}, \tag{1}$$

yields a set of orthonormal temporal basis functions ($\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2 \ldots \boldsymbol{\psi}_{\mathrm{n}}] \in \mathbb{R}^{\mathrm{m}\times\mathrm{m}}$) and spatiotemporal correlational functions ($\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_{\mathrm{n}}] \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$). These modes are ordered by the cumulative variance in the data captured by each mode pair. The diagonal matrix $\mathbf{\Sigma} \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$ contains the singular values $\boldsymbol{\sigma}$, such that $\sigma_1 > \sigma_2 > .. > \sigma_{\mathrm{n}}$, where σ_1 represents the variance captured by mode pair $(\boldsymbol{\psi}_1, \mathbf{v}_1)$ and so forth. The optimal rank-r approximation of the dataset $\mathbf{X}$ is

$$\mathbf{X} \approx \mathbf{\Psi}_{\mathrm{r}}\mathbf{\Sigma}_{\mathrm{r}}\mathbf{V}_{\mathrm{r}}^{\mathrm{T}}\ , \tag{2}$$

where $\mathbf{\Psi_r} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2 \ldots \boldsymbol{\psi}_{\mathrm{r}}]$ is the tailored temporal basis that explains 85% of the cumulative variability in the training data. To obtain the optimal sampling points from the r basis modes $\mathbf{\Psi_r}$, QR factorization with column pivoting is performed, yielding

$$\mathbf{\Psi}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} = \mathbf{QR}\ , \tag{3}$$

where $\mathbf{C}$ is the measurement matrix that yields p column vectors ordered by their informational value. The pivot points in $\mathbf{C}$ signify either the best times to sample or the best locations to sample, if $\mathbf{X}^{\mathrm{T}}$ is used as the training data.

The target data $\mathbf{y}$ can be represented in terms of the tailored basis $\mathbf{\Psi_r}$ as

$$\mathbf{y} = \mathbf{\Psi}_{\mathrm{r}}\mathbf{a}\ , \tag{4}$$

where $\mathbf{a}$ is a vector that yields coefficients for the different basis functions in $\mathbf{\Psi_r}$. This study assumes the target watersheds are poorly gauged, and so reconstruction must proceed from limited measurements. Hence, the values of the coefficients in $\mathbf{a}$ are from sparse measurements obtained at the optimal sampling times at the target location, $\mathbf{y}_{\mathrm{s}}$, as

$$\mathbf{y}_{\mathrm{s}} = \mathbf{Cy} \approx \mathbf{C\Psi}_{\mathrm{r}}\hat{\mathbf{a}}\ , \tag{5}$$

which is further used to estimate high-resolution time series at the target gauge:

$$\hat{\mathbf{y}} = \mathbf{\Psi}_{\mathrm{r}}\hat{\mathbf{a}}\ , \tag{6}$$

In summary, the DSS framework incorporates two common dimension-reduction techniques, SVD and QR factorization, to minimize dimensionality while preserving the key features of the system. The SVD provides a tailored basis of the dataset while QR factorization identifies the informative sampling times and locations from this basis. Subsequent measurements at the target locations can be mapped onto the appropriate bases to produce high-resolution time series data.

This study implemented the DSS framework to reconstruct concentration time series (NOx and TP) by modifying the MATLAB code for DSS retrieved from the following GitHub repository: https://github.com/kmanohar/SSPOR_pub.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

2.2 Data and Study Area

Nutrient concentration (NOx and TP) and corresponding flow (Q) time-series across different regions in the US were obtained from the United States Geological Survey (USGS) National Water Information System (https://waterdata.usgs.gov/nwis) and used in this study. The compiled dataset included 21 NOx gauges and 39 TP gauges (along with their corresponding flow gauges) from the Midwest region and 5 NOx gauges (along with their corresponding flow gauges) from both the South Atlantic Gulf region and the Mid-Atlantic region to compare the regional DSS models (Fig. 2a, 2b). For specific scenario-based analysis (discussed in the following section), subsets of these gauges were used as per the data requirements of each experimental setup. The USGS gauges had daily measurements spanning at least 3 years with less than 10% missing values. To fill in the missing data, cubic spline interpolation was implemented, and the data were normalized to z-scores before analysis. A summary of the collected dataset (station name, location, hydrologic region, corresponding catchment properties, i.e., basin area, mean daily flow, mean concentrations, dominant land-class) for these USGS monitoring gauges is provided (Table S1). Land use distribution for the respective basins was obtained from the National Land Cover Database (NLCD 2021) (https://www.mrlc.gov/data) and categorized into seven broad groups: Barren land, Developed, Forest, Grassland/Shrub, Pasture/Cropland, Rivers/Lakes, and Wetlands (Fig. S1).

2.3 Model Scenarios

he DSS framework was tested under a base model scenario and various real-world scenarios to evaluate predictive performance. The number of gauges and the duration of their records varied across simulation scenarios (Table 1).
The base model was simulated to assess its performance relative to benchmark load estimation models. A comparative assessment of data and sampling requirements was also conducted between DSS and another similar sparse modeling technique (i.e., compressed sensing). In addition to the base NOx model from the Midwest region, separate DSS models were developed in different hydrologic regions to examine the variation in optimal sampling times.

Table 1: Summary of datasets used under base and scenario analyses.

Analysis Type	Nutrient	Region	No of gauges	Monitoring period
Base Models	NOx	Midwest	21	May 2021 – April 2024
	TP		39	October 2019 – April 2022
Regional Models	NOx	Midwest	21	May 2021 – April 2024
	NOx	Mid-Atlantic	5	
	NOx	South Atlantic	5	
Scenario 1 (Historic)	NOx	Midwest	4	April 2014 -April 2024
	TP		3	October 2008 - September 2022
Scenario 2 (Regional)	NOx	Midwest	22*	April 2014 to April 2024

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

	TP		40*	October 2010 - September 2022

* Additional gauges with shorter records (≤ 2 years) were included to maximize the dataset

Two scenarios were simulated as prototypes to improve the sampling efficiency of water quality monitoring programs. These include: (1) High-resolution data estimated at the target location using the historic stream data from the same location, and (2) High-resolution data estimated at the target location using concurrent stream data from regional locations (Fig. 2c).

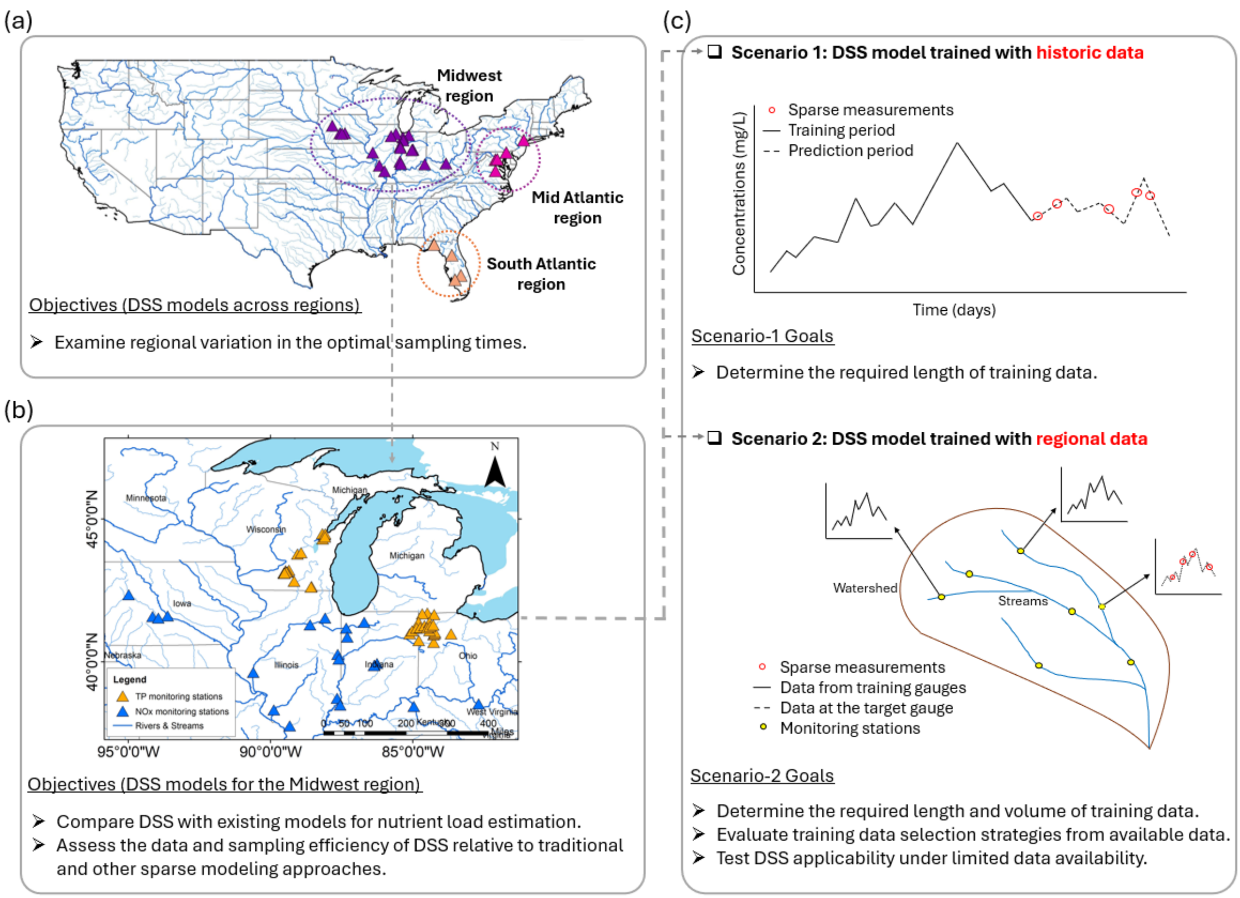


Figure 2: USGS monitoring locations (NOx) across the US considered in this study (a), NOx and TP monitoring stations in the Midwest region (b), and schematic outline for scenarios 1 and 2, using historical and regional data for training, respectively (c).

Both Scenario 1 (model trained with historic data) and Scenario 2 (model trained with regional data) were used to determine the number of monitoring years required to train the DSS model. For Scenario 1, historic data from the target location was used as training data with lengths ranging from 1-8 years for NOx and 1-12 years for TP. For Scenario 2, regional data was used as training data with ranges of 1-10 years for NOx and 1-12 years for TP.

In addition to the number of monitoring years, Scenario 2 was used to determine the number of monitoring gauges required to train the model with regional data. In this case, DSS model performance was compared across training data sizes of 5, 10,

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

and 15 pairs of flow and concentration time series for NOx reconstruction, and 5, 15, and 25 pairs of time series for TP reconstruction.

Scenario 2 was further evaluated to examine the impact of different training data selection processes. To simulate a scenario in which limited concentration data are available for training, models trained with both flow and concentration data were compared to models trained with only flow data. Model performance was examined in two additional cases: one in which concentration and flow data were paired and drawn from the same monitoring sites, and another in which concentration and flow data were selected independently from different sites within the region.

2.4 Model Performance Evaluation

Model performance was evaluated using the Nash-Sutcliffe Efficiency (NSE) and the Mean Absolute Error (MAE). NSE reflects the model's capacity to explain the observed variability (Nash & Sutcliffe, 1970) and MAE denotes the absolute magnitude of deviation between simulated values and observed values (Willmott & Matsuura, 2005). NSE and MAE can be written as

$$\mathbf{NSE} = \mathbf{1} - \frac{\sum_{i=1}^{N}(y_{obs}-\hat{y}_{sim})^2}{\sum_{i=1}^{N}(y_{obs}-y_{obs,mean})^2}\ , \tag{7}$$

$$\mathbf{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{\boldsymbol{y}}_{sim} - \boldsymbol{y}_{obs}|\ , \tag{8}$$

where $\boldsymbol{y}_{obs}$ is the observed concentration, $\hat{\boldsymbol{y}}_{sim}$ is the simulated concentration and N is the number of total samples in the time-series. For daily NOx time-series, an NSE greater than 0.35 signifies satisfactory performance, whereas an NSE over 0.5 denotes good performance (Pandit et al., 2025). A lower NSE threshold was set for TP, with over 0.3 signifying satisfactory, while over 0.45 signifying good performance (Pandit et al., 2025).

Annual nutrient loads were calculated using either observed or predicted daily flow and concentrations (Moatar & Meybeck, 2005) as

$$\mathbf{L} = \mathbf{K}\sum_{j=1}^{n}\mathbf{X_j Q_j}\ , \tag{9}$$

where $\boldsymbol{L}$ is nutrient loads expressed in tons per year, $\boldsymbol{K}$ represents a unit conversion factor, $\boldsymbol{X_j}$ denotes the daily nutrient concentrations in mg/L, $\boldsymbol{Q_j}$ indicates the daily streamflow in m³/sec, and $\boldsymbol{n}$ signifies the number of days. The percentage error in the annual nutrient load was compared to benchmark results obtained using state-of-the-practice load estimation models (Lee et al., 2019). The benchmark results were obtained from a USGS dataset that estimated annual nutrient loads (at 23 locations) using 13 load-estimation models: interpolation, ratio estimator methods, LOADEST regression models, and weighted regression models. The models also assessed 6 sampling strategies: biweekly (~26 samples/year), monthly (12 samples/year), bimonthly (6 samples/year), high-flow targeted (18 samples/year), early high-flow targeted (18 samples/year), and monthly plus high-flow samples (18 samples/year). Among these models, WRTDS-K (Q. Zhang & Hirsch, 2019)

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

models provided the best results for bi-weekly sampling or flow-targeted sampling, with error rates of ~10% for nitrate-nitrite, and ~20% for total phosphorus.

3 Results

3.1 Overview of Base Model Performance

3.1.1 Informative Sampling Times and Locations

QR factorization with column pivoting allows the identification of the most informative sampling times from the tailored basis. The months from Spring and Summer were the most informative for sampling NOx and TP, as per the DSS model. While reconstructing the NOx time series, the majority of the informative samples (almost 65%) were distributed across the months from March to June (Fig. 3a). For the TP time series, March, May, June, and July were the months with the majority of the informative samples (almost 56%) (Fig. 3a). In terms of seasons, Spring and early Summer were found to be the most suitable times for sampling NOx and TP (Fig. 3b). When the optimal sampling times were projected onto the concentration time series, they generally coincided with the time of peak values or near-peak values of concentration (Fig. 3c).

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

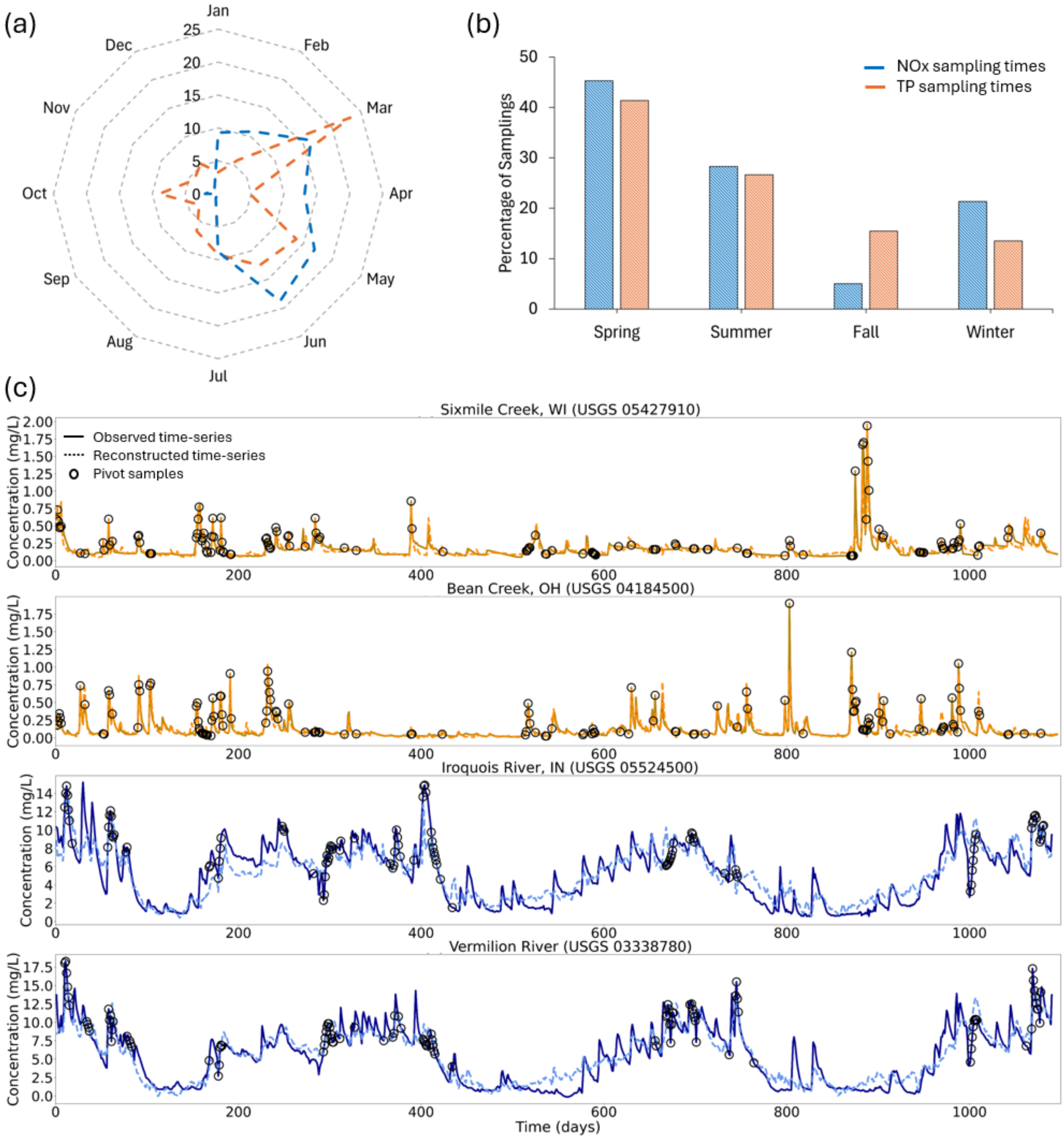


Figure 3: Monthly (a) and seasonal (b) distribution of informative sampling times. Lower panels show examples of TP and NOx time series reconstruction from optimal samples (c). This figure shows results corresponding to the top 100 informative sampling times.

QR pivoting can also identify the most informative sampling locations. For the current dataset, the DSS model does not suggest any identifiable spatial patterns in the most informative sampling locations for reconstructing the NOx or TP time series. The most informative gauges, for both NOx and TP, were distributed across the watersheds (Fig. S2). Most of the top

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

15 sampling locations were nutrient monitoring gauges (11 out of 15 for NOx and 12 out of 15 for TP) (Tables S2 and S3). For the top 5 gauges, the most informative flow time series were from smaller watersheds, and concentration time series were from larger watersheds. For NOx, the only flow time-series consistently ranked in the top 5 had a basin area of 1,813 km^2 (8.78 – 29,299 km^2 across all Q time-series with a median of 3,891 km^2) and a mean daily flow of 4.67 m3/s (0.072 – 2,271 m^3/s across all Q time-series with a median of 39.60 m^3/s) (Table S2). For the 5 NOx time series that were consistently ranked in the top 5 informative gauges, the ranges of basin area and mean daily concentrations were 679 – 13,435 km^2 (8.78 – 69,571 km^2 across all NOx time-series with a median of 4,247 km^2) and 0.86 – 5.54 mg/L (0.68 – 7.74 mg/L across all NOx time-series with a median of 2.24 mg/L), respectively (Table S2). For TP, the only flow time-series consistently ranked in the top 5 had a basin area of 40.7 km^2 (12.97 – 14,360 km^2 across all Q time-series with a median of 183.4 km^2) and a mean daily flow of 0.535 m^3/s (0.064 – 141.6 m^3/s across all Q time-series with a median of 4.32 m^3/s) (Table S3). For the 4 TP time-series that were consistently ranked in the top 5 informative gauges, the ranges of basin area and mean daily concentrations were 15.85 – 1,341 km^2 (12.97 – 14,360 km^2 across all TP time-series with a median of 138.6 km^2) and 0.13 – 0.27 mg/L (0.07 – 0.47 mg/L across all TP time-series with a median of 0.182 mg/L), respectively (Table S3). In terms of land-use classes, agricultural lands were more prominently featured among the top NOx and TP gauges. Among the top 15 gauges, 12 NOx gauges and 14 TP gauges were located in watersheds dominated by pasture or croplands. This may be because of the dataset's composition, which is primarily composed of agricultural watersheds in the Midwest.

3.1.2 Model Performance across Monitoring Gauges

The DSS model demonstrated good performance for most of the NOx and TP concentration time series in terms of NSE values with 50% sample reduction (Figure S3). For NOx, 13 out of the 21 gauges had mean NSE values greater than 0.5, while 7 gauges had a mean NSE value less than 0.35. For TP, the DSS model yielded a mean NSE value greater than 0.45 at 34 out of 39 TP gauges, with only 3 gauges yielding mean NSE values less than 0.3. Examples of reconstructed concentration time series at selected nutrient monitoring gauges show that reconstructed TP concentrations were very similar to the observed concentrations, while the reconstructed NOx concentrations underestimated some of the observed peak magnitudes (Fig. 3c).

The DSS model achieved good accuracy for predicting NOx and TP concentrations with 40-60 samples annually. With only 40 samples per year (~10% of total measurements) at the target gauge, daily NOx time-series were reconstructed with a mean NSE of 0.36 ± 0.27 and a mean MAE of 0.84 ± 0.55 mg/L using data from the top 15 gauges as the training set (Fig. 4a, 4c). Daily TP time-series reconstructed utilizing 40 samples annually at the target gauge yielded a mean NSE of 0.43 ± 0.28 and a mean MAE of 0.08 ± 0.07 mg/L, using data from the top 20 gauges as the training set (Fig. 4b, 4d).

As more measurements were taken, reconstruction performance increased. The mean NSE value increased to 0.49 ± 0.26 (NOx) and 0.51 ± 0.28 (TP) when 60 samples (~15% of total measurements) instead of 40 samples were taken at the target gauge (Fig. 4a, 4b). As more training data were incorporated into the DSS model, reconstruction performance improved significantly with the same number of samples being used. When the training gauges increased from 15 to 20, the mean NSE

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

value for NOx reconstructions increased from 0.36 ± 0.27 to 0.51 ± 0.40 using 40 samples per year at the target location (Fig. 4a). For TP, the mean NSE value increased from 0.43 ± 0.28 to 0.54 ± 0.28 when the training gauges increased from 20 to 30. The mean NSE value further increased to 0.59 ± 0.31 when all gauges were used (Fig. 4b).

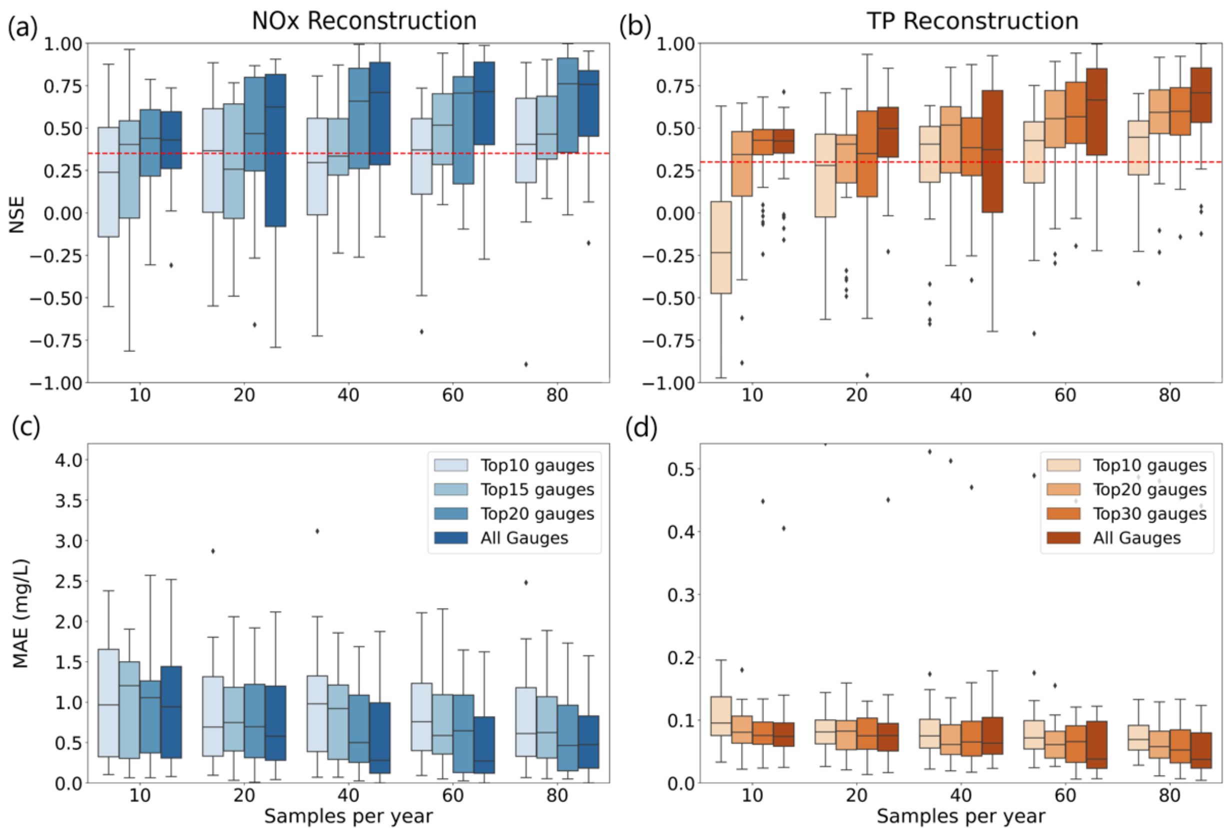


Figure 4: Model performance (in terms of NSE and MAE) for NOx (a, c) and TP (b, d) reconstructions considering different numbers of top informative samples (different groups) and top informative gauges included in the training set (different bars within each group). The length of each bar/box represents the range of performance metrics among all the gauges. The red dashed line (upper panels) represents the threshold NSE value for the model to be satisfactory.

3.1.3 Estimation of Annual Nutrient Loadings

Accurate annual nutrient loads can be obtained from fewer measurements via the DSS model (Fig. 5). The annual NOx load was estimated with a mean error of -3.16 ± 6.11% using only 20 samples per year (~5% of total yearly measurements), whereas the mean error for annual TP loads was -9.20 ± 10.85% with the same number of samples being used. As the number of samples taken increased to 40 samples per year (~10% of total yearly measurements), the mean error reduced to -2.28 ± 4.42% and -3.81 ± 5.93% for annual NOx and TP loads, respectively.

The annual nutrient load errors obtained from the DSS model (using 10, 20, and 30 samples per year) were compared against the benchmark USGS dataset, where the WRTDS-K model performed comparatively better than the other load estimation

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

models (Lee et al., 2019; Zhang and Hirsch, 2019). The DSS model's performance was identical to, or somewhat better than, the benchmark results (Fig. 5). The results were comparable to the WRTDS-K model results (median annual load error within ±3% and ±10% for NOx and TP loads, respectively) when the annual loads were estimated using a concentration time-series that was reconstructed with more than 10 samples per year.

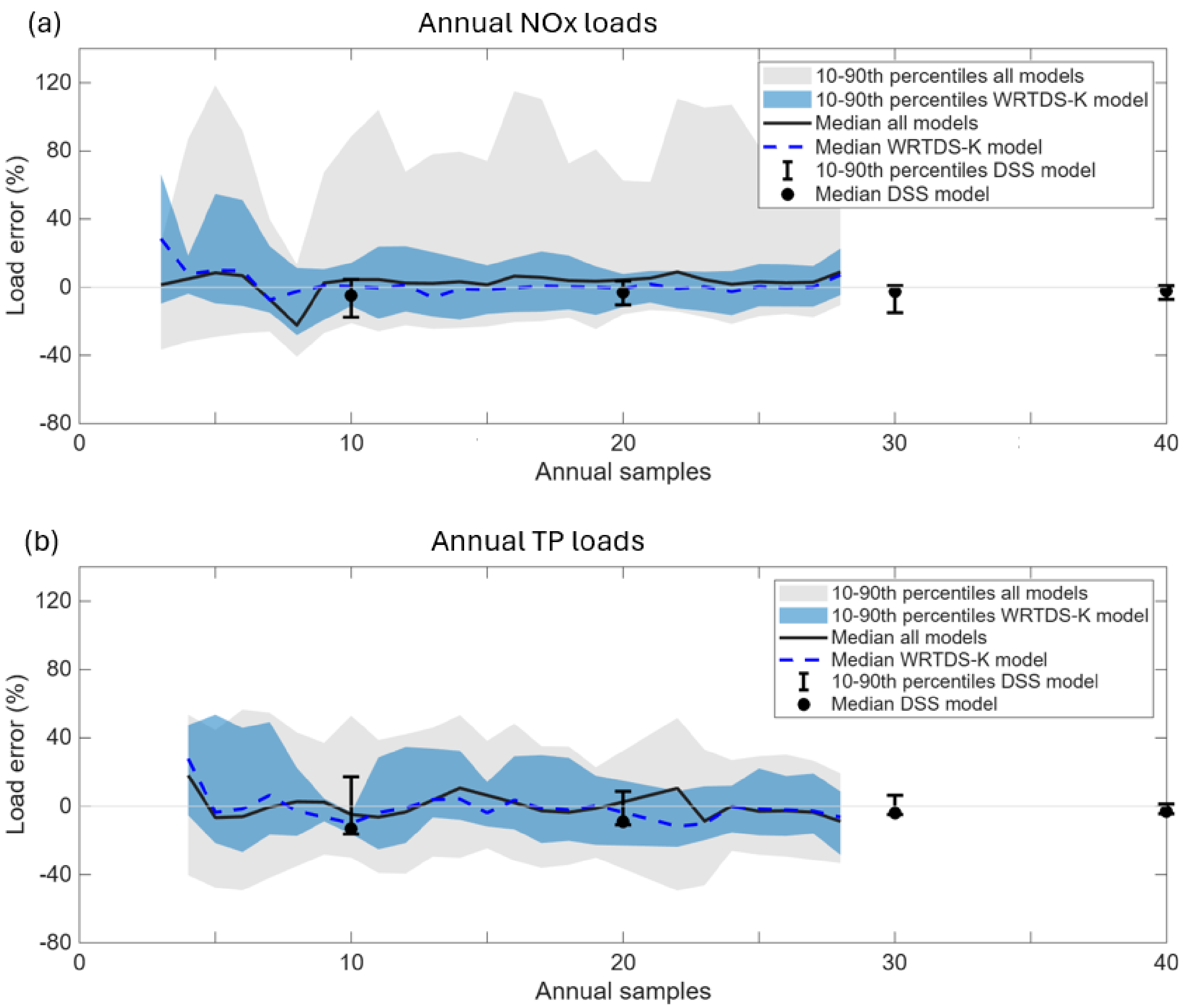


Figure 5: Comparison of model performance (in terms of percentage error in annual loads) between the WRTDS-K model and the DSS model for NOx (a) and TP (b). The DSS models were trained with all the available flow and concentration gauges.

3.1.4 Relation between Model Performance and Catchment Properties

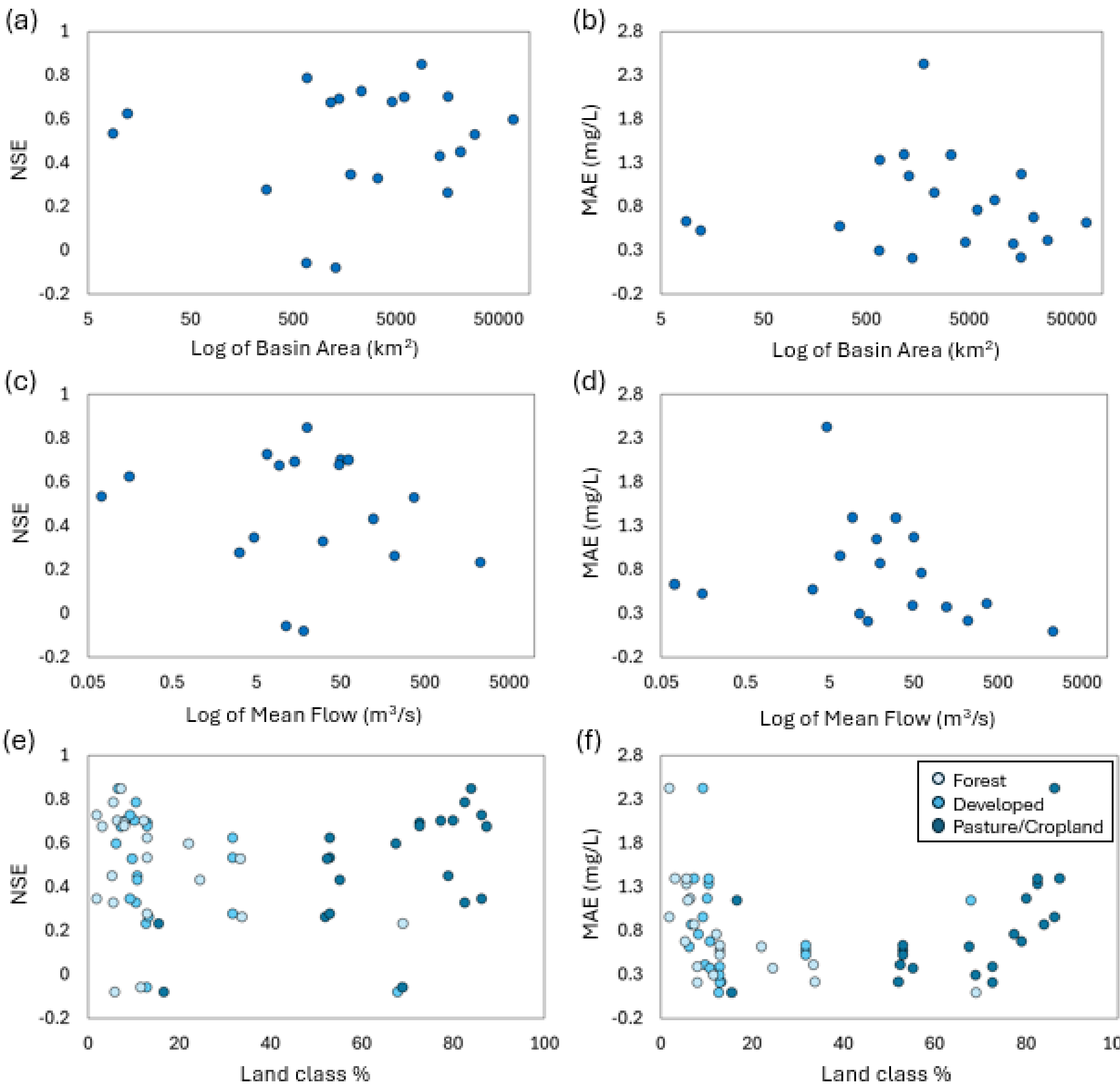
An attempt was made to understand the impacts of key catchment properties, such as watershed area, mean flow, and land class distributions, on reconstruction accuracy using the DSS framework. In general, higher NSE and lower MAE values

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

aligned with larger basin area and higher mean flow, specifically for TP (Fig. 6, 7). For example, the mean NSE value for the NOx model increased from 0.58 ± 0.36 (across all the gauges) to 0.64 ± 0.19 for the gauges with a basin area greater than 2,500 km2 and to 0.62 ± 0.22 for gauges having a mean flow greater than 15 m3/s (Fig. 6a, 6c). Similarly, the TP model yielded a mean NSE value of 0.59 ± 0.31 across all the gauges. But for the gauges with a basin area greater than 250 km2, the mean NSE value was 0.67 ± 0.15. The mean NSE value increased to 0.66 ± 0.16 when the mean flow was greater than 3 m3/s at the target gauge (Fig. 7a, 7c).

Model performance did not show any consistent trends with individual land-class percentages when segregated by dominant land-use combination (aerial extent of at least 10% in most of the basins) (Fig. 6e, 7e).

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026



300 **Figure 6: Model performance for NOx as a function of basin area (a, b), mean daily stream flow (c, d), and land-use composition (e, f). Model performance is quantified by Nash-Sutcliffe Efficiency (NSE) and the mean absolute error (MAE, mg/L).**

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

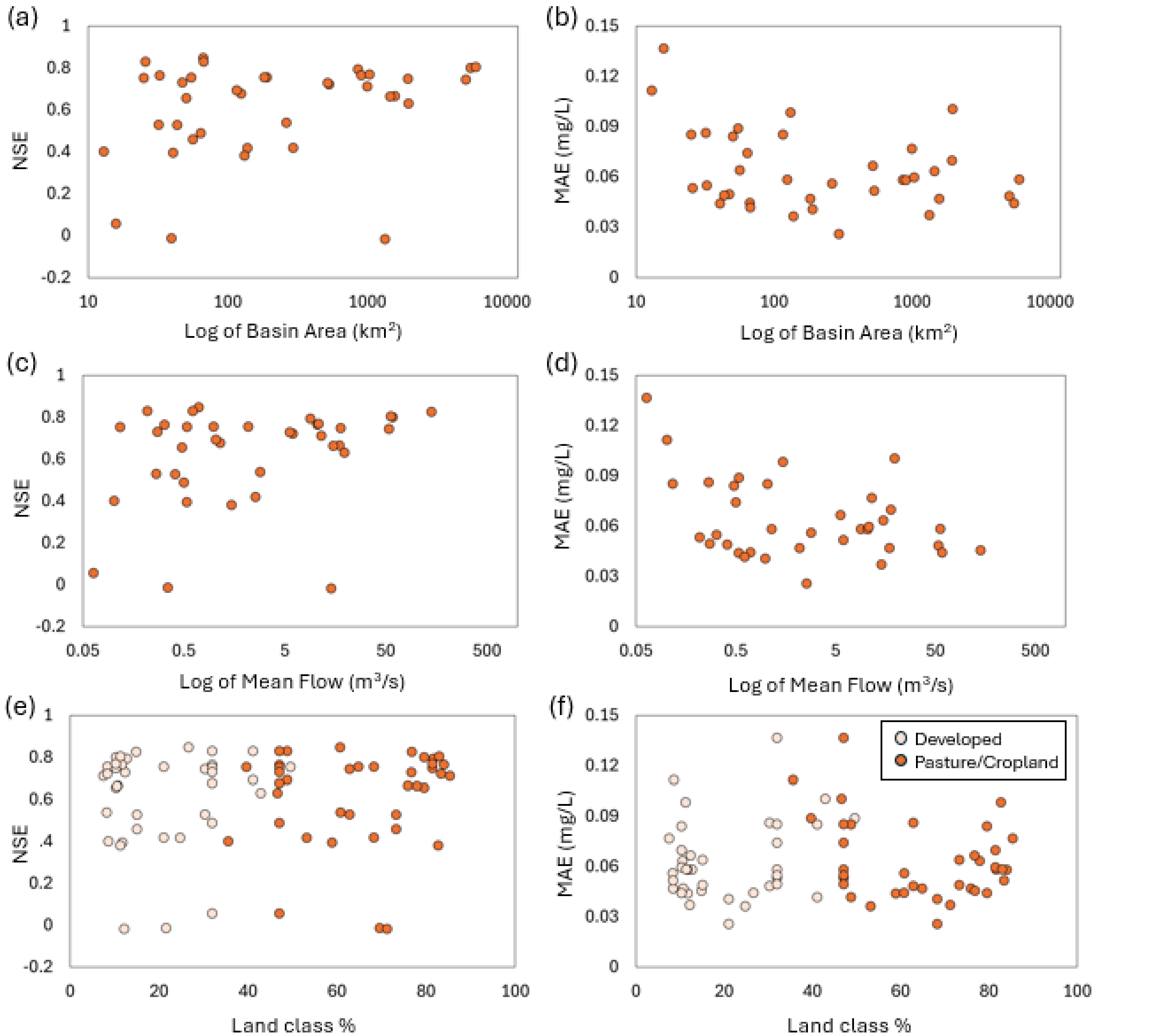


Figure 7: Model performance for TP as a function of basin area (a, b), mean daily stream flow (c, d), and land-use composition (e, f). Model performance is quantified by Nash-Sutcliffe Efficiency (NSE) and the mean absolute error (MAE, mg/L).

3.1.5 Models from Different Hydrologic Regions

305 In addition to the base DSS model for the Midwest region, two new DSS models (only for NOx) were developed for the South Atlantic and the Mid-Atlantic regions (Fig. 2a). The new DSS models showed less satisfactory model performance compared to the model for the Midwest region. Specifically, the models for the Mid-Atlantic and South Atlantic Gulf regions

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

yielded mean NSE values lower than 0.35 (Fig. 8a). This deterioration in performance can be attributed to the smaller size of the datasets available for training (5 gauges with 3 years of monitored data), as discussed in the following section.

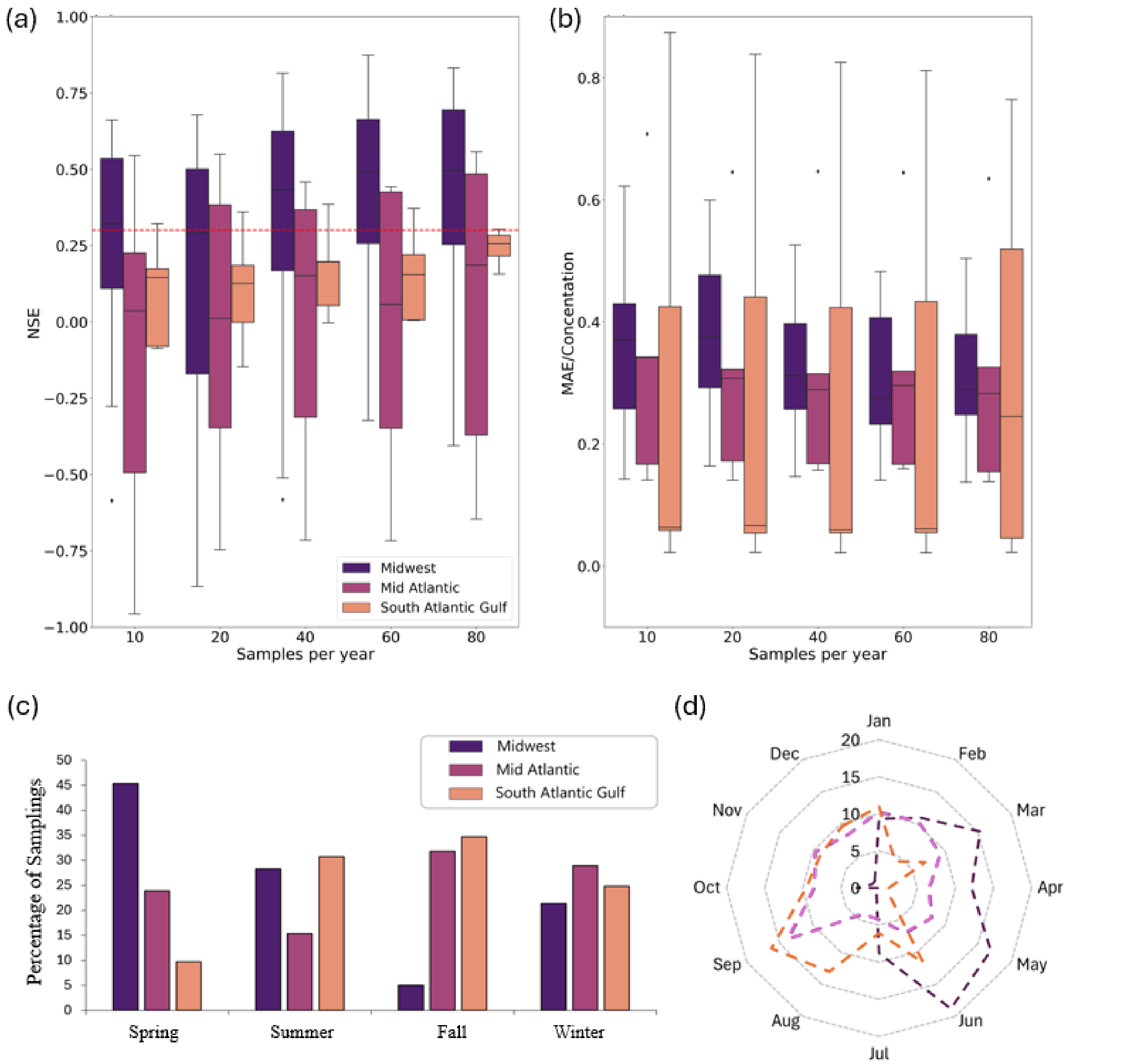


Figure 8: Model performance for NOx reconstructions at different regions (a, b) as a function of different numbers of samples (different groups) and distribution of informative sampling times by season (c) and month (d) for different regions. The red dashed line (a) represents the threshold NSE value for the model to be satisfactory.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

While the DSS model in the Midwest region identified the months from Spring and Summer as the most informative for sampling NOx, the optimal sampling times for the Mid Atlantic and South Atlantic Gulf regions were distributed across Fall, Winter, and/or Summer (Fig. 8c). In the Mid-Atlantic region, almost 60% of the pivot sampling times occurred in Fall and Winter (31.81% in September through November, 28.91% in December to February) (Fig. 8d). In the South Atlantic Gulf region, almost 65% of the informative sampling times were in Summer and Fall (30.73% in June to August, 34.72% in September to November) (Fig. 8d). These differences may reflect variations in hydrology and land use practices across these regions.

3.2 Model Performance across Selected Scenarios

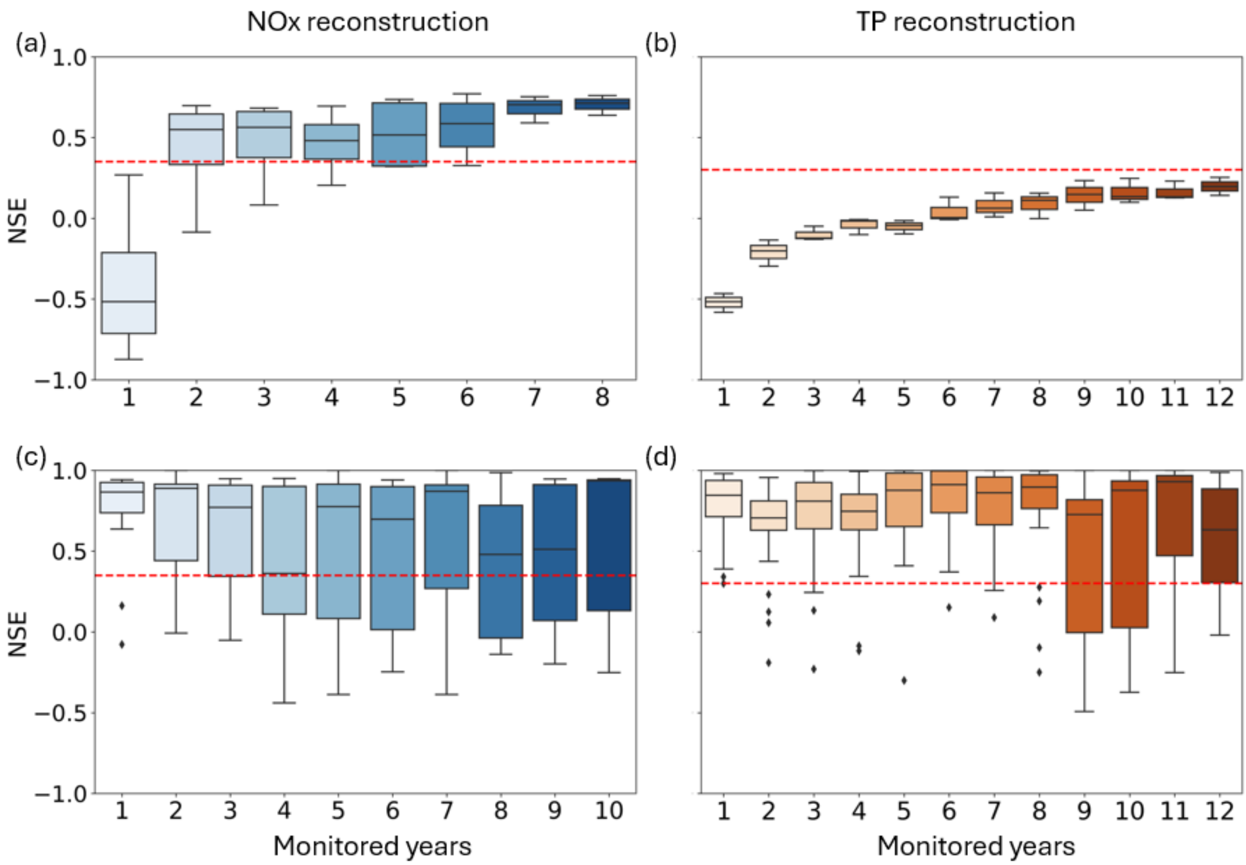
The DSS model performance was evaluated using 2 training data scenarios (i.e., historical and regional) to examine the impacts of variations in data size, data span, and training data selection methods.

3.2.1 Length of Training Data

In scenario 1 (model trained with historic flow and concentration data from the target location), the NOx model performance improved gradually as the amount of training data increased (from 2 to 8 years of monitored data). Specifically, using 60 samples per year, the mean NSE value increased from 0.47 ± 0.36 with 3 years of training data to 0.61 ± 0.29 with training data spanning 8 years (Fig. 9a). In contrast, TP model performance remained below the satisfactory level (< 0.30) throughout the scenario. With 60 samples per year, the TP model yielded a mean NSE value of 0.18 ± 0.06, even with 12 years of training data (Fig. 9b).

In scenario 2 (model trained with regional flow and concentration data), model performance (for both NOx and TP) showed low sensitivity but performed above a satisfactory level regardless of training data extent (Fig. 9c, 9d). With training data spanning 3 years, the NOx model yielded a mean NSE value of 0.64 ± 0.37 using 60 samples at the target gauge. For the same number of samples, the mean NSE value was 0.55 ± 0.57 when the training data spanned 10 years (Fig. 9c). Similarly, with 60 samples per year, the TP model yielded a mean NSE value of 0.66 ± 0.24 and 0.56 ± 0.47 for training data lengths of 2 years and 12 years, respectively (Fig. 9d).

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026



340 **Figure 9: Model performance (in terms of NSE) for NOx (a, c) and TP (b, d) reconstructions across different model configurations (different bars), i.e., variations in length of training data taken from either historic data at the target site (a, b) or regional stream data (c, d). The results shown here are based on the model that used 60 samples per year. The length of each bar/box represents the range of performance metrics among all the gauges. The red dashed line represents the threshold NSE value for the model to be satisfactory.**

345 ### 3.2.2 Size of Training Data

DSS model performance was sensitive, particularly for NOx, to the number of gauges included in the training set. With 60 samples obtained annually at the target gauge, the NOx model yielded a mean NSE value of 0.36 ± 0.22 using 10 pairs of flow and concentration time series for training, which increased to 0.45 ± 0.24 using 15 pairs of time series for training (Fig. 10a). In contrast, the TP model showed less sensitivity to the training data size; it achieved a mean NSE value of 0.45 ± 0.33
350 and 0.42 ± 0.24, respectively, with 15 and 25 pairs of flow and concentration time series used as training data (Fig. 10b).

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

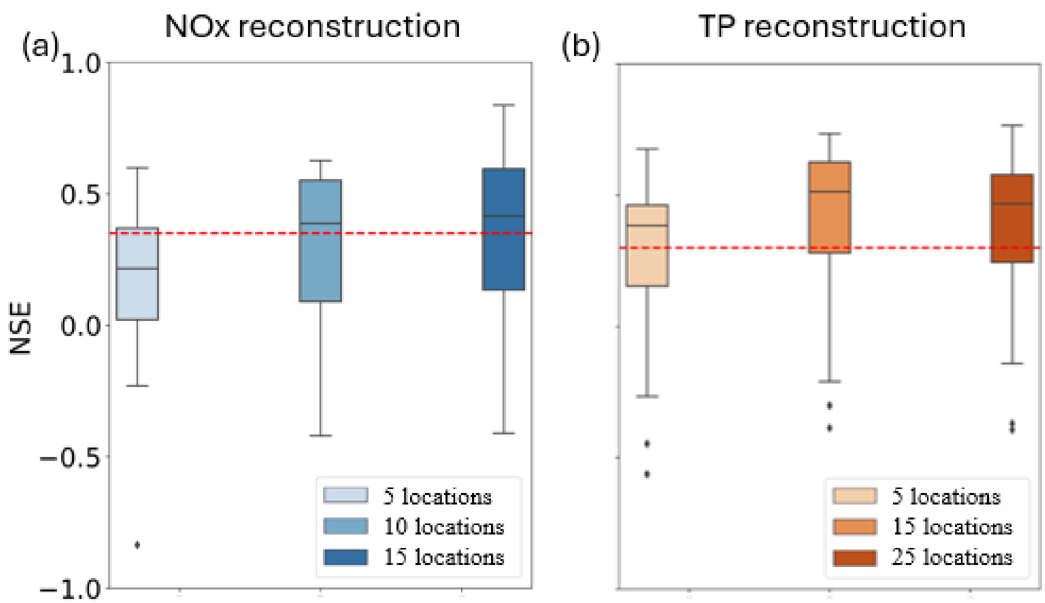


Figure 10: Model performance (in terms of NSE) for NOx (a) and TP (b) reconstructions across different numbers of training gauges. The results shown here are based on the model that used 60 samples per year. The length of each bar/box represents the range of performance metrics among all the gauges. The red dashed line represents the threshold NSE value for the model to be satisfactory.

3.2.3 Selection Processes of Training Data

Model performance improved for both NOx and TP when concentration data was included in the training data (Fig. 11a, 11b). When trained with both flow and concentration and using 60 samples per year, the NOx model yielded a mean NSE value of 0.40 ± 0.33 (Fig. 11a), and the TP model yielded a mean NSE value of 0.37 ± 0.31 (Fig. 11b). However, the mean NSE values were below the satisfactory levels (0.35 for NOx and 0.3 for TP) when trained with flow only. Some sites did show acceptable performance when trained on flow only, suggesting that DSS sampling strategies can be developed to initiate a water quality sampling program in some cases. Model results were similar for both NOx and TP, regardless of how the training data were selected (Fig. 11c, 11d). With 60 samples per year at the target gauge, the NOx model yielded a mean NSE value of 0.44 ± 0.25 when the flow and concentration data were selected independently. With paired flow and concentration data, the mean NSE was 0.39 ± 0.28 (Fig. 11c). For TP, the model yielded a mean NSE value of 0.38 ± 0.30 for independently selected flow and concentration data and a mean NSE value of 0.41 ± 0.27 with paired flow and concentration data (Fig. 11d).

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

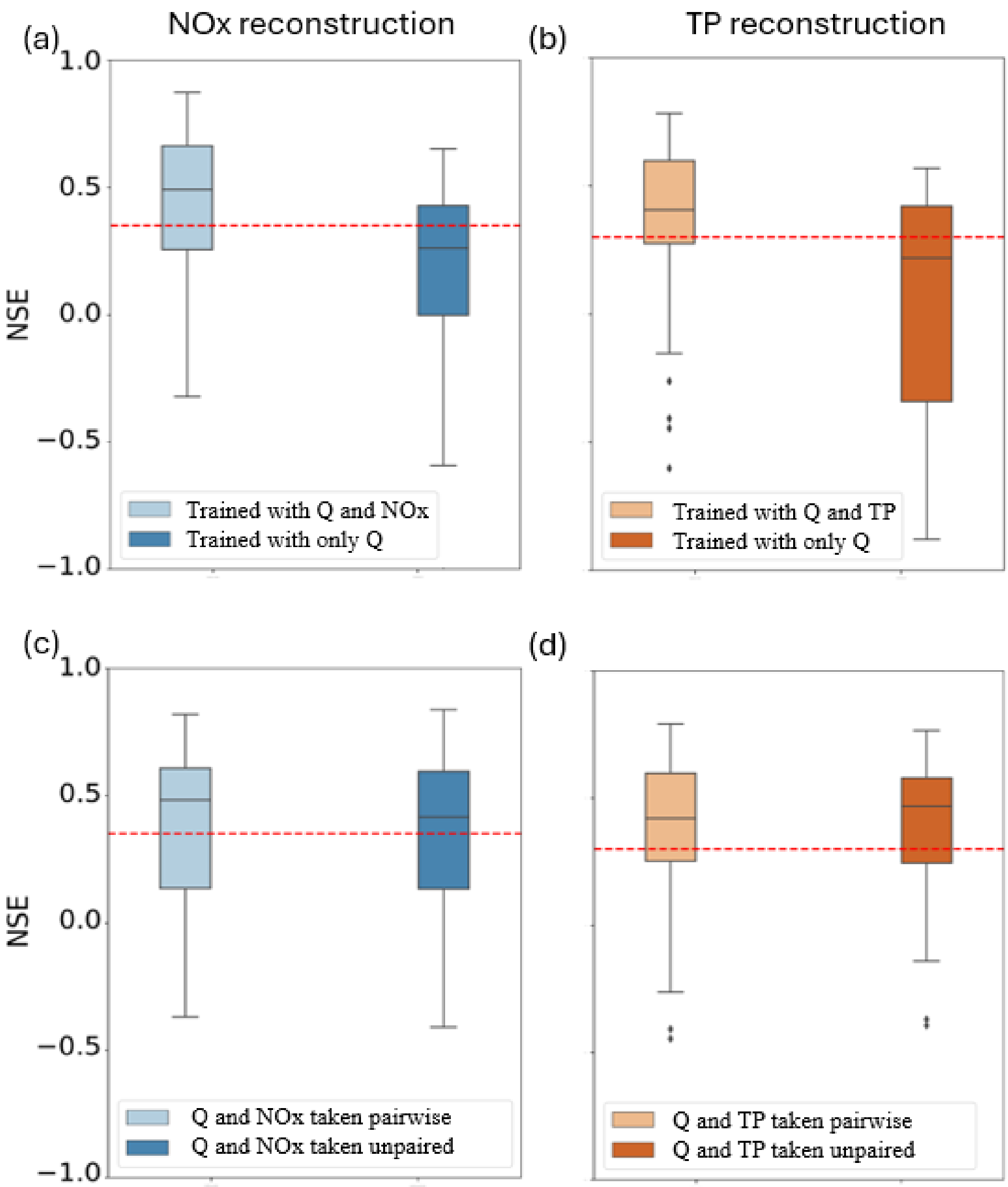


Figure 11: Model performance (in terms of NSE) for NOx (a, c) and TP (b, d) reconstructions across different training data selection methods, i.e., model trained with flow and concentration versus only flow (a, b) and random versus paired selection of flow and concentration data (c, d). The results shown here are based on the model that used 60 samples per year. The length of each bar/box represents the range of performance metrics among all the gauges. The red dashed line represents the threshold NSE value for the model to be satisfactory.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

4 Discussion

4.1 Optimizing Sampling Frequency and Times with DSS

The DSS model was able to predict high-resolution NOx and TP concentrations and annual loads with 40-60 samples per year (approximately equivalent to weekly sampling). This aligns with the results from prior research that employed a DSS modeling framework (Bin Mamoon et al., 2025). The DSS model not only uses a small amount of data to generate continuous concentration time series, but it can also identify this subset of data based on their information content. In conventional monitoring practices, sampling is typically performed at intervals of fixed frequency, i.e., weekly, bi-weekly, or monthly (Bieroza et al., 2018; Bowes et al., 2009; Gao et al., 2020; Minaudo et al., 2017; Reynolds et al., 2016; Robertson et al., 2018; Villa et al., 2019), and in some cases additional sampling is carried out around storm events (Minaudo et al., 2017; Robertson et al., 2018). National and regional programs also opt for similar fixed sampling intervals (14-52 samples per year for USGS; 20 samples per year for Chesapeake; 20-35 samples per year for Minnesota) (Egge et al., 2021; U.S. EPA, 2010). Unlike these conventional approaches, one can pinpoint the most informative sampling times and locations before sampling based on historical data using the DSS model framework.

For the Midwest region, optimal sampling times occurred in spring and early summer, which coincides with peak values and most variability in nutrient concentrations. Similar findings were also reported in prior studies that implemented the DSS framework (Bin Mamoon et al., 2025; Zhang et al., 2023b). However, the optimal sampling times were different for the Mid Atlantic region (fall and early winter) and the South Atlantic region (fall and summer). These differences may reflect variations in hydrology (i.e., snowmelt, storm events, subsurface transport) and land use practices (i.e., fertilizer application) across these regions (Duan et al., 2012; Kaushal et al., 2014; Mulholland & Hill, 1997; Royer et al., 2006; Van Meter et al., 2020). For example, sudden spikes in streamflow due to late-season storms and snowmelt events are common in the Mid-Atlantic region, which can elevate nutrient concentrations in fall to early winter (Inamdar et al., 2018). In the South Atlantic region, row crops (i.e., corn and cotton) usually receive fertilizer during early spring (at planting) and late spring (as side-dress), causing peaks in nutrient levels during the following months. These differences in sampling times imply that a “one-size-fits-all” sampling approach may not work across different regions. Rather, the DSS framework can identify the sampling times that are relevant to regional watershed processes and characteristics.

In terms of optimal sampling locations for NOx and TP, an evenly distributed pattern (with respect to basin size, mean daily flow, land use classes, and mean daily concentrations) was observed. Indeed, the availability of diverse flow and concentration data that can capture the entire spectrum of environmental variability has the potential to improve model-independent (data driven) methods (Kratzert et al., 2019). The present study highlights the potential of the DSS framework as a "regional" model that does not require explicit prior knowledge of the hydrological system.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

4.2 Sampling Efficiency of the DSS Framework

The DSS framework demonstrated a substantial reduction (~85-90%) in sampling requirements while estimating daily concentrations or annual loadings. Like all modeling methods, the DSS framework needs concentration data at the target location for model training. However, it can generate high-frequency concentration time-series and reliable annual load estimates from a limited number of observations. This highlights the feasibility of the DSS framework for monitoring programs with limited resources or watersheds with sparse observational data.

For comparison, the DSS model outperformed the results obtained through a similar dimension reduction technique, Compressed Sensing (CS), that was applied to reconstruct streamflow and nutrient signals (Zhang et al., 2023a). With only 20 to 80 samples annually (equivalent to a sampling frequency of 0.05-0.22 days), the DSS model could estimate TP load with an error rate < ±10%. To get similar results using CS, sampling was needed every 9-10 hours (sampling frequency of 0.4 days). With a sampling frequency of 10 days (35 to 40 samples annually), CS was able to estimate annual NOx loads with an error rate of −6.6% ± 3.8%, whereas the error rate was < ±3% using the DSS model. Another major advantage of using DSS over CS is that it identifies the optimal sampling times and locations rather than relying on random sampling.

4.3 DSS Modeling and Catchment Properties

The DSS model for TP performed better for gauges with larger basin areas and higher flow regimes, whereas the model for NOx showed less consistent sensitivity to these factors. Large basins possess a smaller number of active degrees of freedom or reduced dimensionality (Bin Mamoon et al., 2026; Ombadi et al., 2021; Vignesh et al., 2015) as they filter out the randomness driven by stochastic rainfall signals and basin heterogeneities via the "smoothing effect" (Müller et al., 2021; Smith et al., 2004; Nester et al., 2011). While high-flow regimes usually minimize concentration spikes, concentration levels are elevated due to reduced dilution capacity during low-flow periods (Macintosh et al., 2011; Withers et al., 2011). Overall, larger basins and high-flow regimes provide stability and buffering capacity, thereby enhancing model performance.

4.4 Nutrient Monitoring and Modeling with DSS

The findings from the scenario-based analysis provide insights into the minimum data requirements, model transferability, and the flexibility of data selection strategies. The DSS framework can guide strategic planning for existing or proposed monitoring programs by identifying minimum data requirements and optimal sampling strategies. For example, the results indicate that 2-3 years of monitored flow and concentration data from 10-15 nearby locations can yield effective high-resolution reconstructions at target locations. Results suggest that using long term historical data from a single location provided limited benefit (specifically for TP), which may reflect the effects of land-use or climate change (Kaushal et al., 2014). On the other hand, increasing the number of regional data captures a broad range of hydrologic and biogeochemical variability that improves model generalization (Kratzert et al., 2019).

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

This framework offers a viable starting point for monitoring or modeling efforts in ungauged or poorly gauged locations. The results presented in this study imply that programs with limited resources should prioritize short-term, spatially distributed monitoring at nearby locations and targeted sparse measurements from the target locations. This contradicts the suggestion of long-term, high-frequency monitoring at the target location, as proposed in prior studies (Kamrath et al., 2023; Kaushal et al., 2014). The findings also demonstrate the ability of the DSS framework to learn transferable patterns across locations with minimal data, aligning with recent studies that leverage transfer learning across sites or catchment heterogeneity to improve the performance of water quality models (X. Chen et al., 2024; Ma et al., 2024)..

5 Conclusion

High-frequency stream nutrient data are essential for accurately identifying the sources of pollutants and assessing the effectiveness of water quality control measures. However, monitoring stream nutrients at high frequency is challenging due to financial and operational constraints. This study utilized the DSS model framework to assess training data and sampling requirements for reliable prediction of NOx and TP concentrations and annual loads. Results showed that NOx and TP concentrations and loads can be estimated at target locations with good accuracy using only 40-60 samples per year (~10-15% of total measurements) obtained at the optimal sampling times. However, DSS model performance in the Mid-Atlantic and South Atlantic Gulf regions was challenged by limited data.

A major goal of this study was to evaluate the DSS model performance under different scenarios to inform future monitoring programs. Our scenario-based analysis revealed that daily NOx and TP concentrations can be accurately predicted with only 2-3 years of historic (only NOx) or regional (both NOx and TP) flow and concentration data, regardless of how they are selected (randomly or paired by location), using sparse measurements from at least 15 of these monitoring locations. This suggests that monitoring programs can use in-situ high-frequency sensors across locations and supplement with occasional grab sampling. The results also imply that high resolution concentration data can be generated from regional flow and concentration data (need not to be from same location) as long as these gauges span different catchment sizes, flow rates, and concentrations. Overall, this DSS model framework showed the potential to ensure the efficient use of the resources associated with nutrient monitoring programs, ultimately resulting in improved water quality management efforts.

Code and data availability

All streamflow and concentration data were collected by the United States Geological Survey (USGS) and are available through NWIS (https://waterdata.usgs.gov/nwis). Gauge numbers and web links to data sources are listed in Table S1 in the Supporting Information. The DSS algorithm was implemented by modifying MATLAB code downloaded from the GitHub repository: https://github.com/kmanohar/SSPOR_pub.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

Supplement link

The supplementary material is published as a separate file with this article.

Acknowledgements

This work relates to Department of Army award ("Novel Technologies to Mitigate Water Contamination for Resilient Infrastructure"; Federal Award Identification Number: W9132T2220001) issued by the U. S. Army Corps of Engineers (USACE) Engineer Research and Development Center (ERDC). The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of USACE-ERDC.

Author contributions

WBM: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. KZ: Conceptualization, Supervision, Visualization, Writing – review and editing. ML: Visualization, Writing – review and editing. AJP: Conceptualization, Funding acquisition, Project administration, Supervision.

Conflict of Interest Disclosure

The authors declare there are no conflicts of interest for this manuscript.

References

Bieroza, M. Z., Heathwaite, A. L., Bechmann, M., Kyllmar, K., and Jordan, P.: The concentration-discharge slope as a tool for water quality management, Sci. Total Environ., 630, 738–749, https://doi.org/10.1016/j.scitotenv.2018.02.256, 2018.

Bin Mamoon, W., Povinelli, R. J., and Parolari, A. J.: Nonlinear dynamics of stream nutrient concentrations: Low dimensionality, structured complexity, and seasonal predictability horizon, J. Hydrol., 673, 135432, https://doi.org/10.1016/j.jhydrol.2026.135432, 2026.

Bin Mamoon, W., Zhang, K., Luhar, M., and Parolari, A. J.: Stream nutrient load and concentration estimation from minimal measurements, Geophys. Res. Lett., 52, e2025GL114935, https://doi.org/10.1029/2025GL114935, 2025.

Blaen, P. J., Khamis, K., Lloyd, C. E. M., Bradley, C., Hannah, D., and Krause, S.: Real-time monitoring of nutrients and dissolved organic matter in rivers: Capturing event dynamics, technological opportunities and future directions, Sci. Total Environ., 569–570, 647–660, https://doi.org/10.1016/j.scitotenv.2016.06.116, 2016.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: A review, Hydrol. Process., 9, 251–290, https://doi.org/10.1002/hyp.3360090305, 1995.

Bowes, M. J., Smith, J. T., and Neal, C.: The value of high-resolution nutrient monitoring: A case study of the River Frome, Dorset, UK, J. Hydrol., 378, 82–96, https://doi.org/10.1016/j.jhydrol.2009.09.015, 2009.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H.: Nonpoint pollution of surface waters with phosphorus and nitrogen, Ecol. Appl., 8, 559–568, https://doi.org/10.1890/1051-0761(1998)008%255B0559:NPOSWW%255D2.0.CO;2, 1998.

Cassidy, R. and Jordan, P.: Limitations of instantaneous water quality sampling in surface-water catchments: Comparison with near-continuous phosphorus time-series data, J. Hydrol., 405, 182–193, https://doi.org/10.1016/j.jhydrol.2011.05.020, 2011.

Castrillo, M. and García, Á. L.: Estimation of high frequency nutrient concentrations from water quality surrogates using machine learning methods, Water Res., 172, 115490, https://doi.org/10.1016/j.watres.2020.115490, 2020.

Chen, X., Sun, W., Jiang, T., and Ju, H.: Enhanced prediction of river dissolved oxygen through feature- and model-based transfer learning, J. Environ. Manage., 372, 123310, https://doi.org/10.1016/j.jenvman.2024.123310, 2024.

Chen, Y., Song, L., Liu, Y., Yang, L., and Li, D.: A review of the artificial neural network models for water quality prediction, Appl. Sci., 10, 5776, https://doi.org/10.3390/app10175776, 2020.

Davidson, K., Gowen, R. J., Harrison, P. J., Fleming, L. E., Hoagland, P., and Moschonas, G.: Anthropogenic nutrients and harmful algae in coastal waters, J. Environ. Manage., 146, 206–216, https://doi.org/10.1016/j.jenvman.2014.07.002, 2014.

Dodds, W. K., Bouska, W. W., Eitzmann, J. L., Pilger, T. J., Pitts, K. L., Riley, A. J., Schloesser, J. T., and Thornbrugh, D. J.: Eutrophication of U.S. freshwaters: Analysis of potential economic damages, Environ. Sci. Technol., 43, 12–19, https://doi.org/10.1021/es801217q, 2009.

Donoho, D. L.: Compressed sensing, IEEE Trans. Inf. Theory, 52, 1289–1306, https://doi.org/10.1109/TIT.2006.871582, 2006.

Duan, S., Kaushal, S. S., Groffman, P. M., Band, L. E., and Belt, K. T.: Phosphorus export across an urban to rural gradient in the Chesapeake Bay watershed, J. Geophys. Res.-Biogeo., 117, G01025, https://doi.org/10.1029/2011JG001782, 2012.

Egge, L., Anderson, P., and Ganske, L.: Minnesota's Water Quality Monitoring Strategy 2021–2031, Minnesota Pollution Control Agency, St. Paul, Minnesota, USA, 2021.

Gao, L., Qi, J., Li, S., Benoy, G., Xing, Z., and Meng, F.: Effects of sampling frequency on estimation accuracies of annual loadings for water quality parameters in different sized watersheds, Water Qual. Res. J. Can., 55, 261–277, https://doi.org/10.2166/wqrj.2020.012, 2020.

Harrison, J. W., Lucius, M. A., Farrell, J. L., Eichler, L. W., and Relyea, R. A.: Prediction of stream nitrogen and phosphorus concentrations from high-frequency sensors using Random Forests regression, Sci. Total Environ., 763, 143005, https://doi.org/10.1016/j.scitotenv.2020.143005, 2021.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

Holmberg, M., Forsius, M., Starr, M., and Huttunen, M.: An application of artificial neural networks to carbon, nitrogen and phosphorus concentrations in three boreal streams and impacts of climate change, Ecol. Model., 195, 51–60, https://doi.org/10.1016/j.ecolmodel.2005.11.009, 2006.

Hu, X., Chu, L., Pei, J., Liu, W., and Bian, J.: Model complexity of deep learning: A survey, Knowl. Inf. Syst., 63, 2585–2619, https://doi.org/10.1007/s10115-021-01605-0, 2021.

Inamdar, S., Johnson, E., Rowland, R., Warner, D., Walter, R., and Merritts, D.: Freeze–thaw processes and intense rainfall: The one-two punch for high sediment and nutrient loads from mid-Atlantic watersheds, Biogeochemistry, 141, 333–349, https://doi.org/10.1007/s10533-017-0417-7, 2018.

Jones, A. S., Horsburgh, J. S., Mesner, N. O., Ryel, R. J., and Stevens, D. K.: Influence of sampling frequency on estimation of annual total phosphorus and total suspended solids loads, J. Am. Water Resour. Assoc., 48, 1258–1275, https://doi.org/10.1111/j.1752-1688.2012.00684.x, 2012.

Kamrath, B., Yuan, Y., Manning, N., and Johnson, L.: Influence of sampling frequency and estimation method on phosphorus load uncertainty in the Western Lake Erie Basin, Ohio, USA, J. Hydrol., 617, 128906, https://doi.org/10.1016/j.jhydrol.2022.128906, 2023.

Katul, G. G., Porporato, A., Daly, E., Oishi, A. C., Kim, H., Stoy, P. C., Juang, J., and Siqueira, M. B.: On the spectrum of soil moisture from hourly to interannual scales, Water Resour. Res., 43, W05428, https://doi.org/10.1029/2006WR005356, 2007.

Kaushal, S. S., Mayer, P. M., Vidon, P. G., Smith, R. M., Pennino, M. J., Newcomer, T. A., Duan, S., Welty, C., and Belt, K. T.: Land use and climate variability amplify carbon, nutrient, and contaminant pulses: A review with management implications, J. Am. Water Resour. Assoc., 50, 585–614, https://doi.org/10.1111/jawr.12204, 2014.

Khullar, S. and Singh, N.: Machine learning techniques in river water quality modelling: A research travelogue, Water Supply, 21, 1–13, https://doi.org/10.2166/ws.2020.277, 2021.

Kirchner, J. W. and Neal, C.: Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection, Proc. Natl. Acad. Sci. USA, 110, 12213–12218, https://doi.org/10.1073/pnas.1304328110, 2013.

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., and Nearing, G.: Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets, Hydrol. Earth Syst. Sci., 23, 5089–5110, https://doi.org/10.5194/hess-23-5089-2019, 2019.

Lee, C. J., Hirsch, R. M., and Crawford, C. G.: An evaluation of methods for computing annual water-quality loads, U.S. Geological Survey Scientific Investigations Report 2019–5084, U.S. Geological Survey, Reston, Virginia, USA, 2019.

Lustig, M., Donoho, D. L., Santos, J. M., and Pauly, J. M.: Compressed sensing MRI, IEEE Signal Process. Mag., 25, 72–82, https://doi.org/10.1109/MSP.2007.914728, 2008.

Ma, K., Shen, C., Xu, Z., and He, D.: Transfer learning framework for streamflow prediction in large-scale transboundary catchments: Sensitivity analysis and applicability in data-scarce basins, J. Geogr. Sci., 34, 963–984, https://doi.org/10.1007/s11442-024-2235-x, 2024.

Macintosh, K. A., Jordan, P., Cassidy, R., Arnscheidt, J., and Ward, C.: Low flow water quality in rivers; septic tank systems and high-resolution phosphorus signals, Sci. Total Environ., 412–413, 58–65, https://doi.org/10.1016/j.scitotenv.2011.10.012, 2011.

Manohar, K., Brunton, B. W., Kutz, J. N., and Brunton, S. L.: Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns, IEEE Control Syst. Mag., 38, 63–86, https://doi.org/10.1109/MCS.2018.2810460, 2018.

Manohar, K., Kutz, J. N., and Brunton, S. L.: Optimal sensor and actuator selection using balanced model reduction, IEEE Trans. Automat. Contr., 67, 2108–2115, https://doi.org/10.1109/TAC.2021.3082502, 2022.

Minaudo, C., Dupas, R., Gascuel-Odoux, C., Fovet, O., Mellander, P.-E., Jordan, P., Shore, M., and Moatar, F.: Nonlinear empirical modeling to estimate phosphorus exports using continuous records of turbidity and discharge, Water Resour. Res., 53, 7590–7606, https://doi.org/10.1002/2017WR020590, 2017.

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

Mulholland, P. J. and Hill, W. R.: Seasonal patterns in streamwater nutrient and dissolved organic carbon concentrations: Separating catchment flow path and in-stream effects, Water Resour. Res., 33, 1297–1306, https://doi.org/10.1029/97WR00490, 1997.

Müller, M. F., Roche, K. R., and Dralle, D. N.: Catchment processes can amplify the effect of increasing rainfall variability, Environ. Res. Lett., 16, 084032, https://doi.org/10.1088/1748-9326/ac153e, 2021.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nour, M. H., Smith, D. W., El-Din, M. G., and Prepas, E. E.: The application of artificial neural networks to flow and phosphorus dynamics in small streams on the Boreal Plain, with emphasis on the role of wetlands, Ecol. Model., 191, 19–32, https://doi.org/10.1016/j.ecolmodel.2005.08.010, 2006.

Ohmer, M., Liesch, T., and Wunsch, A.: Spatiotemporal optimization of groundwater monitoring networks using data-driven sparse sensing methods, Hydrol. Earth Syst. Sci., 26, 4033–4053, https://doi.org/10.5194/hess-26-4033-2022, 2022.

Ombadi, M., Nguyen, P., Sorooshian, S., and Hsu, K.: Complexity of hydrologic basins: A chaotic dynamics perspective, J. Hydrol., 597, 126222, https://doi.org/10.1016/j.jhydrol.2021.126222, 2021.

Pandit, A., Hogan, S., Mahoney, D. T., Ford, W. I., Fox, J. F., Wellen, C., and Husic, A.: Establishing performance criteria for evaluating watershed-scale sediment and nutrient models at fine temporal scales, Water Res., 274, 123156, https://doi.org/10.1016/j.watres.2025.123156, 2025.

Parolari, A. J., Sizemore, J., and Katul, G. G.: Multiscale legacy responses of soil gas concentrations to soil moisture and temperature fluctuations, J. Geophys. Res.-Biogeo., 126, e2020JG005865, https://doi.org/10.1029/2020JG005865, 2021.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

Pellerin, B. A., Stauffer, B. A., Young, D. A., Sullivan, D. J., Bricker, S. B., Walbridge, M. R., Clyde, G. A., and Shaw, D. M.: Emerging tools for continuous nutrient monitoring networks: Sensors advancing science and water resources protection, J. Am. Water Resour. Assoc., 52, 993–1008, https://doi.org/10.1111/1752-1688.12386, 2016.

Pretty, J. N., Mason, C. F., Nedwell, D. B., Hine, R. E., Leaf, S., and Dils, R.: Environmental costs of freshwater eutrophication in England and Wales, Environ. Sci. Technol., 37, 201–208, https://doi.org/10.1021/es020793k, 2003.

Read, E. K., Carr, L., De Cicco, L., Dugan, H. A., Hanson, P. C., Hart, J. A., Kreft, J., Read, J. S., and Winslow, L. A.: Water quality data for national-scale aquatic research: The Water Quality Portal, Water Resour. Res., 53, 1735–1745, https://doi.org/10.1002/2016WR019993, 2017.

Reynolds, K. N., Loecke, T. D., Burgin, A. J., Davis, C. A., Riveros-Iregui, D., Thomas, S. A., St. Clair, M. A., and Ward, A. S.: Optimizing sampling strategies for riverine nitrate using high-frequency data in agricultural watersheds, Environ. Sci. Technol., 50, 6406–6414, https://doi.org/10.1021/acs.est.5b05423, 2016.

Robertson, D. M., Hubbard, L. E., Lorenz, D. L., and Sullivan, D. J.: A surrogate regression approach for computing continuous loads for the tributary nutrient and sediment monitoring program on the Great Lakes, J. Great Lakes Res., 44, 26–42, https://doi.org/10.1016/j.jglr.2017.10.003, 2018.

Royer, T. V., David, M. B., and Gentry, L. E.: Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: Implications for reducing nutrient loading to the Mississippi River, Environ. Sci. Technol., 40, 4126–4131, https://doi.org/10.1021/es052573n, 2006.

Schärer, M., Page, T., and Beven, K.: A fuzzy decision tree to predict phosphorus export at the catchment scale, J. Hydrol., 331, 484–494, https://doi.org/10.1016/j.jhydrol.2006.05.034, 2006.

Smith, M. B., Koren, V. I., Zhang, Z., Reed, S. M., Pan, J.-J., and Moreda, F.: Runoff response to spatial variability in precipitation: An analysis of observed data, J. Hydrol., 298, 267–286, https://doi.org/10.1016/j.jhydrol.2004.03.039, 2004.

Sprague, L. A., Oelsner, G. P., and Argue, D. M.: Challenges with secondary use of multi-source water-quality data in the United States, Water Res., 110, 252–261, https://doi.org/10.1016/j.watres.2016.12.024, 2017.

Stets, E. G., Sprague, L. A., Oelsner, G. P., Johnson, H. M., Murphy, J. C., Ryberg, K., Vecchia, A. V., Zuellig, R. E., Falcone, J. A., and Riskin, M. L.: Landscape drivers of dynamic change in water quality of U.S. rivers, Environ. Sci. Technol., 54, 4336–4343, https://doi.org/10.1021/acs.est.9b05344, 2020.

U.S. EPA: Chesapeake Bay Total Maximum Daily Load for Nitrogen, Phosphorus, and Sediment, U.S. Environmental Protection Agency, Washington, DC, USA, 2010.

Valkama, P. and Ruth, O.: Impact of calculation method, sampling frequency and hysteresis on suspended solids and total phosphorus load estimations in cold climate, Hydrol. Res., 48, 1594–1610, https://doi.org/10.2166/nh.2017.199, 2017.

Van Meter, K. J., Chowdhury, S., Byrnes, D. K., and Basu, N. B.: Biogeochemical asynchrony: Ecosystem drivers of seasonal concentration regimes across the Great Lakes Basin, Limnol. Oceanogr., 65, 848–862, https://doi.org/10.1002/lno.11353, 2020.

https://doi.org/10.5194/egusphere-2026-3230
Preprint. Discussion started: 18 August 2026

Varadharajan, C., Appling, A. P., Arora, B., Christianson, D. S., Hendrix, V. C., Kumar, V., Lima, A. R., Müller, J., Oliver, S., Ombadi, M., Perciano, T., Sadler, J. M., Weierbach, H., Willard, J. D., Xu, Z., and Zwart, J.: Can machine learning accelerate process understanding and decision-relevant predictions of river water quality?, Hydrol. Process., 36, e14565, https://doi.org/10.1002/hyp.14565, 2022.

Vignesh, R., Jothiprakash, V., and Sivakumar, B.: Streamflow variability and classification using false nearest neighbor method, J. Hydrol., 531, 706–715, https://doi.org/10.1016/j.jhydrol.2015.10.056, 2015.

Villa, A., Fölster, J., and Kyllmar, K.: Determining suspended solids and total phosphorus from turbidity: Comparison of high-frequency sampling with conventional monitoring methods, Environ. Monit. Assess., 191, 617, https://doi.org/10.1007/s10661-019-7775-7, 2019.

Wang, J., Bouchez, J., Dolant, A., Floury, P., Stumpf, A., Bauer, E., Keefer, L., Gaillardet, J., Kumar, P., and Druhan, J.: Sampling frequency, load estimation and the disproportionate effect of storms on solute mass flux in rivers, Sci. Total Environ., 906, 167379, https://doi.org/10.1016/j.scitotenv.2023.167379, 2024.

Wei, J., Wang, L., Liu, P., Chen, X., Li, W., and Zomaya, A. Y.: Spatiotemporal fusion of MODIS and Landsat-7 reflectance images via compressed sensing, IEEE Trans. Geosci. Remote Sens., 55, 7126–7139, https://doi.org/10.1109/TGRS.2017.2742529, 2017.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 30, 79–82, 2005.

Withers, P. J. A., Jarvie, H. P., and Stoate, C.: Quantifying the impact of septic tank systems on eutrophication risk in rural headwaters, Environ. Int., 37, 644–653, https://doi.org/10.1016/j.envint.2011.01.002, 2011.

Worsfold, P., McKelvie, I., and Monbet, P.: Determination of phosphorus in natural waters: A historical review, Anal. Chim. Acta, 918, 8–20, https://doi.org/10.1016/j.aca.2016.02.047, 2016.

Wurtsbaugh, W. A., Paerl, H. W., and Dodds, W. K.: Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum, WIREs Water, 6, e1373, https://doi.org/10.1002/wat2.1373, 2019.

Zhang, K., Bin Mamoon, W., Schwartz, E., and Parolari, A. J.: Reconstruction of sparse stream flow and concentration time-series through compressed sensing, Geophys. Res. Lett., 50, e2022GL101177, https://doi.org/10.1029/2022GL101177, 2023.

Zhang, K., Luhar, M., Brunner, M. I., and Parolari, A. J.: Streamflow prediction in poorly gauged watersheds in the United States through data-driven sparse sensing, Water Resour. Res., 59, e2022WR034092, https://doi.org/10.1029/2022WR034092, 2023.

Zhang, L., Xing, M., Qiu, C.-W., Li, J., Sheng, J., Li, Y., and Bao, Z.: Resolution enhancement for inversed synthetic aperture radar imaging under low SNR via improved compressive sensing, IEEE Trans. Geosci. Remote Sens., 48, 3824–3838, https://doi.org/10.1109/TGRS.2010.2048575, 2010.

Zhang, Q. and Hirsch, R. M.: River water-quality concentration and flux estimation can be improved by accounting for serial correlation through an autoregressive model, Water Resour. Res., 55, 9705–9723, https://doi.org/10.1029/2019WR025338, 2019.