Article

Assessing linkages between watershed nutrient loading and estuary water quality in Lavaca Bay, Texas

Michael Schramm[1]

[1] Texas A&M AgriLife Research - Texas Water Resources Institute 1001 Holleman Dr. E. College Station, TX 77840-2118; michael.schramm@ag.tamu.edu

Simple Summary: A Simple summary goes here.

Abstract: A single paragraph of about 200 words maximum. For research articles, abstracts should give a pertinent overview of the work. We strongly encourage authors to use the following style of structured abstracts, but without headings: 1) Background: Place the question addressed in a broad context and highlight the purpose of the study; 2) Methods: Describe briefly the main methods or treatments applied; 3) Results: Summarize the article's main findings; and 4) Conclusion: Indicate the main conclusions or interpretations. The abstract should be an objective representation of the article, it must not contain results which are not presented and substantiated in the main text and should not exaggerate the main conclusions.

Keywords: keyword 1; keyword 2; keyword 3 (list three to ten pertinent keywords specific to the article, yet reasonably common within the subject discipline.).

Citation: Schramm, M. Assessing linkages between watershed nutrient loading and estuary water quality in Lavaca Bay, Texas. *Water* **2023**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

1. Introduction

Like many estuaries globally, estuaries along the Texas Gulf coast are facing pressures from increasing population, increases in point source and non-point source pollution and alterations to freshwater inflows leading to increases in the occurrences and risks of algal blooms and eutrophication [1–3]. Recent studies indicate that estuary water quality dynamics in both agriculturally dominated and urban watersheds within Texas are displaying signals of conditions increasingly conducive to eutrophication [3–6].

2. Materials and Methods

2.1. Study Area and Data

Lavaca Bay is a secondary bay in the Matagorda Bay system located on the Texas Gulf coast, roughly halfway between the cities of Houston and Corpus Christi (Figure **??**). Lavaca Bay is 190 km^2 with the majority of freshwater inflow provided by the Lavaca and Navidad River systems. The Garcitas-Arenosa, Placedo Creek, and Cox Bay watersheds provide additional freshwater inflows. The entire watershed land area for Lavaca Bay is 8,149 km^2. The Lavaca and Navidad River watersheds are a combined 5,966 km^2, or approximately 73% of the entire Lavaca Bay watershed area. Discharge from the Navidad River is regulated by Lake Texana which has been in operation since 1980. Lake Texana provides 170,000 acre-feet of water storage and discharges into the tidal section of the Navidad River which ultimately joins the tidal section of the Lavaca River 15 km upstream of the confluence with the Bay.

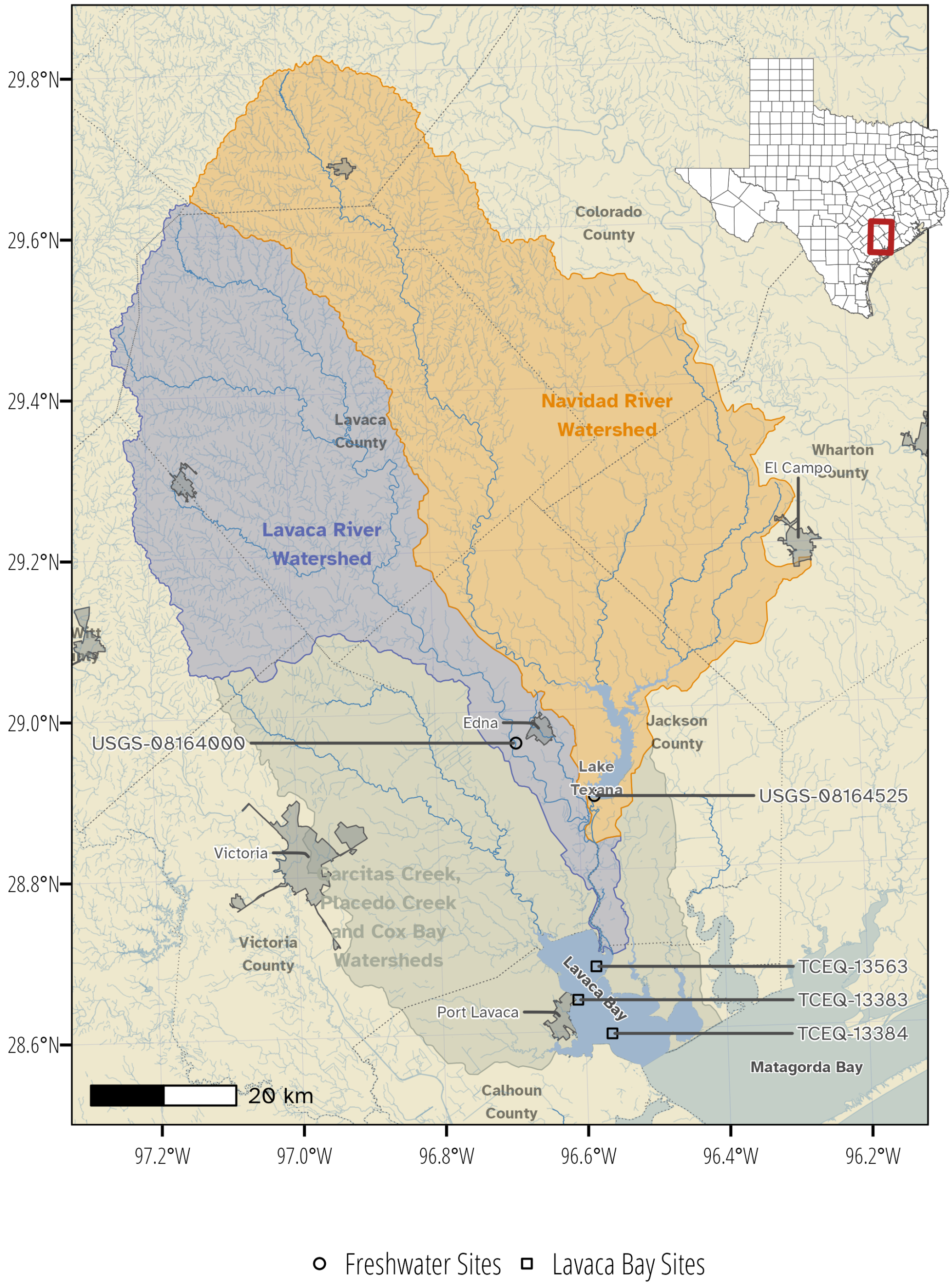

Figure 1. Map of the Lavaca Bay watershed, location of USGS gages where nutrient loads were calculated, and location of estuary water quality sampling sites.

Daily discharges for the Lavaca River (USGS-08164000) were obtained from the United States Geologic Survey (USGS) National Water Information System using the *dataRetrieval* R package [7]. Gaged dailly discharges from Lake Texana (USGS-0816425) were provided by the Texas Water Development Board (TWDB) (April 21, 2022 email from R. Neupane, TWDB).

Water quality sample data for both freshwater and estuary locations were o btained from the Texas Commission on Environmental Quality (TCEQ) Surface Water Quality

Monitoring Information System. Data submitted through the system are required to be collected under Quality Assurance Project Plans and lab method procedures outlined by the TCEQ's procedures manual. The QAPP and procedures manuals ensure the consistent collection and laboratory methods are applied between samples collected by different entities and under different projects. For freshwater locations, total phosphorus (TP) and nitrate-nitrogen (NO_3) data were downloaded (Table 1). Unfortunately, insufficient data was available for assessment of total nitrogen (TN) and total Kjeldahl nitrogen (TKN) loadings, so analysis was restricted to TP and NO_3 loads. For estuary locations, we obtained data for TP, Nitrite+Nitrate (NO_x), TKN, chlorophyll-*a*, and dissolved oxygen (Table 2).

Table 1. Summary of gauged streamflow and freshwater water quality samples between January 1, 2000 and December 31, 2020.

Station ID		Mean	SD	N
USGS-08164000	TP (mg/L)	0.21	0.09	80
	NO_3 (mg/L)	0.18	0.24	74
	Mean Daily Streamflow (cfs)	332.78	1667.47	7671
USGS-08164525	TP (mg/L)	0.20	0.08	81
	NO_3 (mg/L)	0.29	0.26	62
	Mean Daily Streamflow (cfs)	666.14	2957.79	7671

Table 2. Summary of estuary water quality samples collected between January 1, 2005 and December 31, 2020.

Station ID		Mean	SD	N
TCEQ-13383	TP (mg/L)	0.11	0.05	47
	NO_x (mg/L)	0.07	0.15	51
	TKN (mg/L)	0.94	0.49	45
	Chlorophyll-*a* (μg/L)	9.43	5.31	47
	DO (mg/L)	7.22	1.35	55
TCEQ-13384	TP (mg/L)	0.08	0.03	51
	NO_x (mg/L)	0.06	0.08	52
	TKN (mg/L)	0.76	0.40	48
	Chlorophyll-*a* (μg/L)	8.22	6.44	46
	DO (mg/L)	7.51	1.32	54
TCEQ-13563	TP (mg/L)	0.13	0.06	50
	NO_x (mg/L)	0.09	0.13	53
	TKN (mg/L)	0.94	0.37	49
	Chlorophyll-*a* (μg/L)	9.67	5.33	49
	DO (mg/L)	7.91	1.34	56

2.2. Estimating Watershed Based Nutrient Loads

Estimates of nutrient loads were developed using Generalized Additive Models (GAMs) relating nutrient concentration to river discharge, season, and time. Separate models were fit at each station for each parameter and used to predict nutrient concentrations for each day in the study period. GAMs can be specified in a functionally similar manner to the commonly used LOADEST [8] or WRTDS [9] regression models and have been shown to produce reliable estimates of nutrient and sediment loadings [10–16]. GAMs are a semiparametric extension of generalized linear models where the linear predictor is represented as the sum of multiple unknown smooth functions and parametric linear predictors [17]. Although the underlying parameter estimation procedure of GAMs is substantially different than WRTDS, both the functional form and results are demonstrated to be similar [18]. The use of GAMs over other regression-based approaches was (1) the ability to easily explore and incorporate different model terms, (2) the ability to incorporate non-linear smooth function without explicit apriori knowledge of the expect shape, and (3) the ability to specify a link function that relates the expected value of the response to the linear predictors and allows use to avoid data transformations as much as possible.

GAMs were fit using the *mgcv* package in R which makes available multiple types of smooth functions with automatic smoothness selection [17]. The general form of the model relating NO_3 and TP concentration to streamflow, season, and time was:

$$\begin{aligned} g(\mu) &= \alpha + f_1(ddate) + f_2(yday) + f_3(log1p(Q)) + f_4(ma) + f_5(fa) \\ y &\sim \mathcal{N}(\mu, \sigma^2), \end{aligned} \tag{1}$$

where μ is the conditional expected NO_3-N or TP concentration, *g()* is the log-link, α is the intercept, $f_n()$ are smoothing functions. y is the response variable (NO_3 or TP concentration) modeled as normally distributed with mean μ and standard deviation σ. *ddate* is the date converted to decimal notation, *yday* is numeric day of year (1-366), and *log1p(Q)* is the natural log of mean daily streamflow plus 1.

Moving average (*ma*) is an exponentially smoothed moving average that attempts to incorporate the influence of prior streamflow events on concentration at the current time period. Wang *et al.* [10], Kuhnert *et al.* [12] and Zhang and Ball [19] refer to this as averaged or smoothed discounted flow and demonstrated improvements in nutrient loading models by including the term. Kuhnert *et al.* [12] expresses MA as:

$$ma(\delta) = d\kappa_{i-1} + (1-\delta)\hat{q}_{i-1} \quad \text{and} \quad \kappa_i = \sum_{m=1}^{i} \hat{Q}_m, \tag{2}$$

where δ is the discount factor (here, set equal to 0.95), κ_i is the cumulative flow (Q) up to the ith day.

Flow anomaly (*fa*) is a unitless term that represents how wet or dry the current time period is from a previous time period [19,20]. Long-term flow anomaly (*ltfa*) is the streamflow over the previous year relative to the entire period and calculated as described by Zhang and Ball [19]:

$$ltfa(t) = \bar{x}_{1\,year}(t) - \bar{x}_{entire\,period} \tag{3}$$

and the short-term flow anomaly (*stfa*) calculated as the current day flow compared to the preceding 1-month streamflow:

$$stfa(t) = x_{current\,day}(t) - \bar{x}_{1\,month}(t) \tag{4}$$

where x are the averages of log-transformed streamflow over the antecedent period (*1-year*, *1-month*, etc.) for time t. We used *ltfa* in NO~3 models and *stfa* in TP models based on results from Zhang and Ball [19] demonstrating major improvements in NO_x regression models that incorporated *ltfa* and moderate improvements in TP regression models that incorporated *stfa*.

The calculation of model terms for the Lake Texana site were slightly modified because daily loads are not a function of natural stream flow processes alone, but of dam releases and nutrient concentrations at the discharge point of the lake. Q, *ma*, and *fa* terms were calculated based on total gaged inflow from the 4 major tributaries to the lake. Thin-plate regression splines were used for *ddate*, *log1p(Q)*, *fa*, and *ma*. A cyclic cubic regression spline was used for *yday* to ensure the ends of the spline match (day 1 and day 366 are expected to match). First order penalties were applied to the smooths of flow-based variables which penalize departures from a flat function to help constrain extrapolations for high flow measurements.

Left-censored data were not uncommon in this dataset. Several methods are available to account for censored data. We transformed left-censored nutrient concentrations to one-half the detection limit. Although this simple approach can introduce bias [21], we considered it acceptable because high concentrations and loadings are associated with high-flow events and low-flow/low-concentration events will account for a small proportion of total loadings [15].

Daily loads were estimated as the predicted concentration multiplied by the daily streamflow. For the Lake Texana site, model terms were slightly modified because daily loads are a function of dam releases and nutrient concentration, but concentration will be a function of lake inflows and or other lake processes. *Q*, *ma*, and *fa* terms were calculated based on total gaged inflow from the 4 major tributaries to the lake and laily loads at the dam were calculated from the discrete daily concentration at the discharge point of the lake and corresponding reported daily discharge from the dam. Flow-normalized loads were estimated similar to WRTDS by setting flow-based covariates on each day of the year equal to each of the historical values for that day of the year over the study period [9]. The flow-normalized estimate was calculated as the mean of all the predictions for each day considering all possible flow values. Standard deviations and credible intervals were obtained by drawing samples from the multivariate normal posterior distribution of the fitted GAM [15,22,23]. Uncertainty in loads were reported as 90% credible intervals developed by drawing 1000 realizations of parameter estimates from the multivariate normal posterior distribution of the model parameters. GAM performance was evaluated using repeated 5-fold cross validation [24] and average Nash-Suttclifee Efficiency (NSE), r^2 and percent bias (PBIAS) metrics across folds were calculated for each model.

2.3. Linking Estuary Water Quality to Hydrology and Nutrient Loads

To test if changes in freshwater inflow and nutrient loading had explanatory effect on changes in estuary water quality a series of GAM models were fit at each site relating parameter concentration to temporal trends, inflow, and nutrient loads [25]:

$$\begin{aligned} g(\mu) &= \alpha + f_1(ddate) + f_2(yday) \\ y &\sim \Gamma(\mu, \lambda), \end{aligned} \tag{5}$$

$$\begin{aligned} g(\mu) &= \alpha + f_1(ddate) + f_2(yday) + f_3(Q) \\ y &\sim \Gamma(\mu, \lambda), \end{aligned} \tag{6}$$

$$\begin{aligned} g(\mu) &= \alpha + f_1(ddate) + f_2(yday) + f_3(Q) + f_4(Load) \\ y &\sim \Gamma(\mu, \lambda), \end{aligned} \tag{7}$$

where μ is the conditional expected response (nutrient concentration), *g()* is the log link, and response variable was modeled as Gamma distributed with mean μ and scale λ. $f_1(ddate)$ is decimal date smoothed with a thin-plate regression spline, $f_2(yday)$ is the numeric day of year smoothed with a cyclic cubic regression spline, $f_3(Q)$ is mean daily inflow (the combined measurements from Lavaca River and Lake Texana) and $f_4(Load)$ is the total NO_3 or TP watershed load. The set of models specified for each water quality response are in Table **??**.

Table 3. Set of GAM models specified for each water quality parameter response.

Water Quality Response Parameter	Model	Model Terms
TP	Temporal	s(ddate) + s(yday)
	Flow	s(ddate) + s(yday) + s(Q)
	Flow+Load	s(ddate) + s(yday) + s(Q) + s(TP Load)
NO_x	Temporal	s(ddate) + s(yday)
	Flow	s(ddate) + s(yday) + s(Q)
	Flow+Load	s(ddate) + s(yday) + s(Q) + s(NO_3 Load)
Chlorophyll-*a*	Temporal	s(ddate) + s(yday)
	Flow	s(ddate) + s(yday) + s(Q)
	Flow+Load	s(ddate) + s(yday) + s(Q) + s(TP Load) + s(NO_3 Load)
Dissolved Oxygen	Temporal	s(ddate) + s(yday)
	Flow	s(ddate) + s(yday) + s(Q)
	Flow+Load	s(ddate) + s(yday) + s(Q) + s(TP Load) + s(NO_3 Load)
TKN	Temporal	s(ddate) + s(yday)
	Flow	s(ddate) + s(yday) + s(Q)

Because streamflow and nutrient loads are tightly correlated, freshwater inflow can mask signals from nutrient loads alone. Following the methodology implemented by Murphy *et al.* [25], both streamflow and nutrient loads were prepossessed to account for season and flow. Instead of using raw freshwater inflow and nutrient loading values, these values were replaced by seasonally adjusted inflow and flow-adjusted nutrient loads by fitting a GAM relating season (day of year) to log transformed daily freshwater inflow values:

$$g(\mu) = \alpha + f_1(yday), \tag{8}$$

and a GAM relating log transformed NO_3 or TP loads to log transformed daily inflow:

$$g(\mu) = \alpha + f_1(log(Q)), \tag{9}$$

where the response variables were modeled as normally distributed with an identity link function. Response residuals from the respective GAM models were used as Q and *Load* in Equation 6 and Equation 7.

3. Results

3.1. Watershed Nutrient Loads

Based on criteria provided by Moriasi *et al.* [26], GAMs ranged from "satisfactory" to "very good" based on median NSE, r^2 and PBIAS calculated from repeated 5-fold cross validation on predicted and measured nutrient loads. NO_3 GAM models had median NSE values of 0.34 and 0.48 at USGS-08164000 and USGS-08164390 respectively. Median r^2 values were 0.70 (USGS-08164000) and 0.87 (USGS-08164525) and PBIAS values were 2.00 (USGS-08164000) and 10.90 (USGS-08164525) for NO_3 loads. Median NSE values were 0.80 (USGS-08164000) and 0.91 (USGS-08164525) for TP loads. Median r^2 values for TP loads were 0.93 (USGS-08164000) and 0.99 (USGS-08164525) and median PBIAS values were -7.20 (USGS-08164000) and -3.30 (USGS-08164525). Density plots of metrics show similar distribution of values between sites for the same parameter, with the exception r^2 values for NO_3 loads where USGS-08164000 showed a much larger variance in values compared to USGS-08164525 (Figure 2).In addition to higher average NSE and r^2 values, GAMs had smaller variance in metric values for TP compared to NO_3.

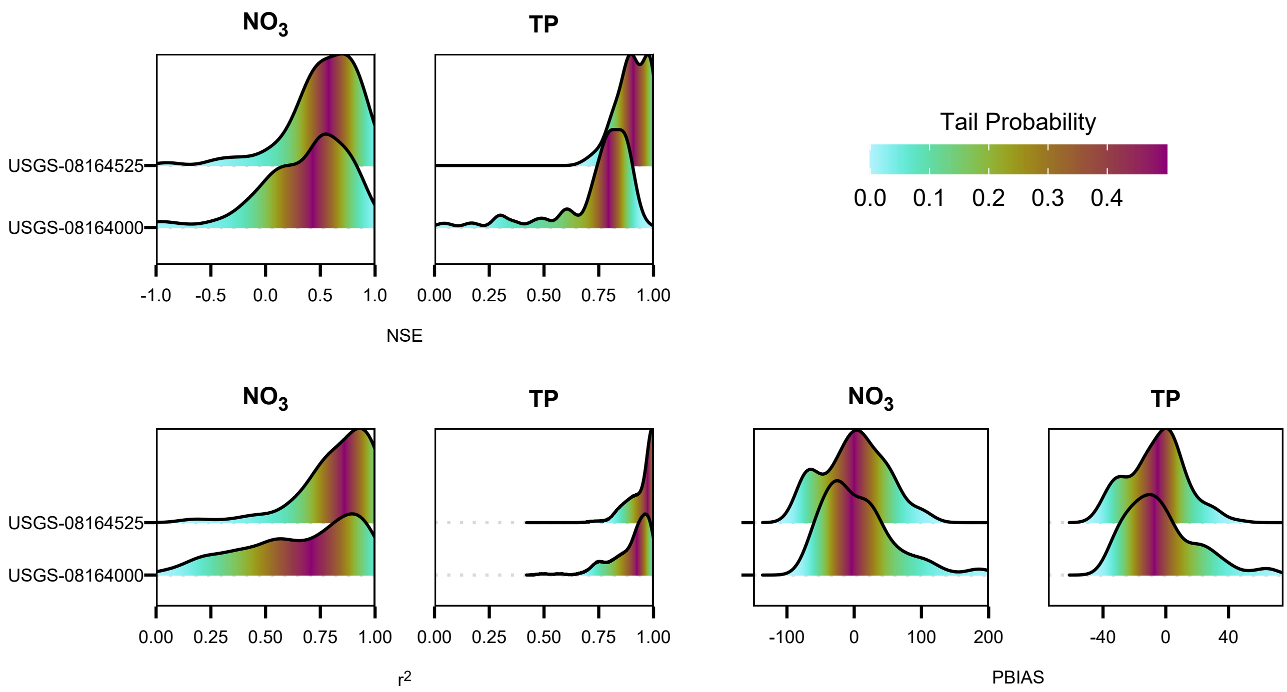

Figure 2. Density plots of goodness-of-fit metrics (NSE, r^2, and PBIAS) from repeated 5-fold cross validation between predicted nutrient loads from GAM models and measured nutrient loads. Color indicates the tail probability calcualted from the empirical cumulative distribution of the goodness-of-fit metrics.

Annual NO_3 and TP loads show considerable variation, generally following patterns in discharge (Figure 3, Figure 4). Flow-normalized TP loads at both sites and flow-normalized loads at USGS-08164000 indicate watershed based loads have not changed much over time when accounting for changes in streamflow (Figure 3). Flow-normalized loads at USGS-08164000 show small variation over time with some decreases in NO_3 loads since 2013. Aggregated across both sites, the mean annual NO_3 load 2005 through 2020 was 205,405 kg (126,867 kg - 341,569 kg, 90% CI). Annual NO_3 loads ranged from 12,574 kg in 2011 to 794,510 kg in 2007.

Total annual TP loads ranged from 7,839 kg in 2011 to 595,075 kg in 2007. Mean annual TP loading from 2005 through 2020 was 182,673 kg (152,227 kg - 219,310 kg, 90% CI). On average, USGS-08164525 accounted for 68% of NO_3 loads and 59% of TP loads from 2005 through 2020. However, during periods of extreme drought the Lavaca River (USGS-08164000) became the primary source of nutrient loading in the watershed with the Navidad River only accounting for 15% and 25% of NO_3 and TP loads in 2011 (Figure 4).

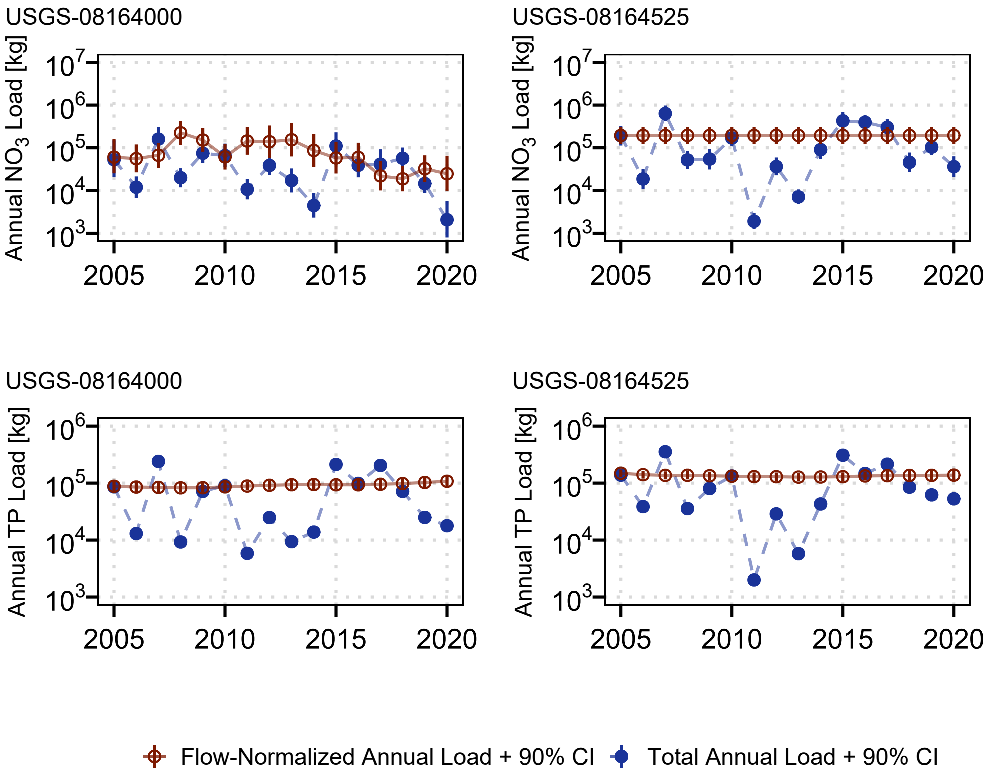

Figure 3. Aggregated estimated annual and flow-normalized annual NO extsubscript3 and TP loads for USGS-08164000 and USGS-08164525.

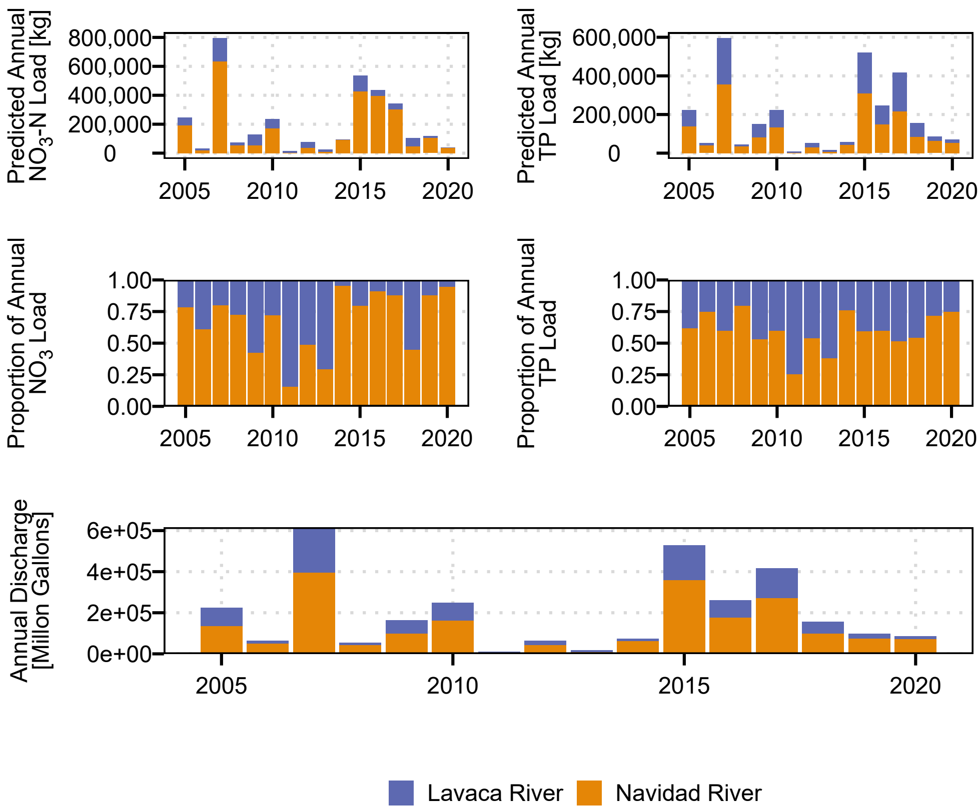

Figure 4. Comparison of delivered annual loads at USGS-08164000 and USGS-08164525.

3.2. Linkages Between Water Quality and Watershed Flows and Loads

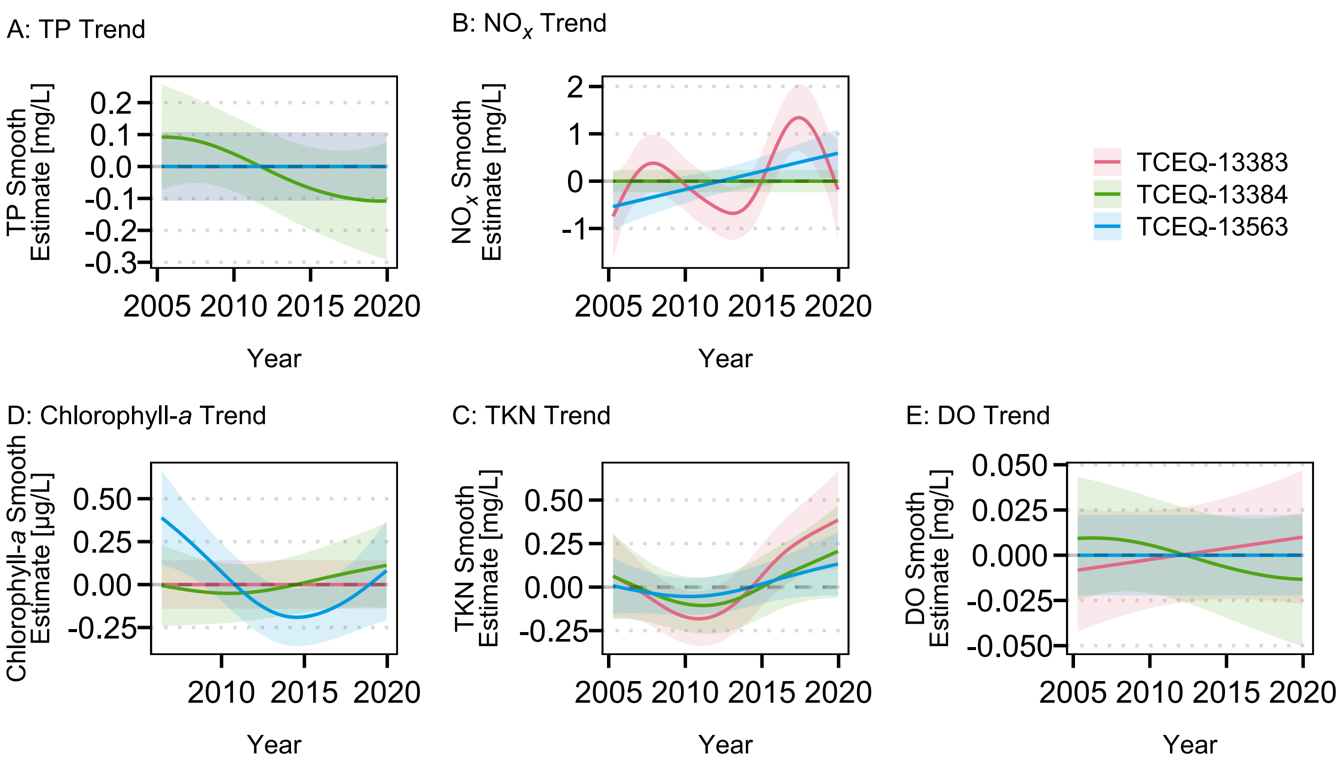

Figure 5. Smoothed temporal trend component for water quality paramaters obtained from temporal estuary GAMs.

GAMs did not identify significant changes in TP or DO concentrations at any of the Lavaca Bay sites from 2005 through 2020 (Figure 5). The upper-bay site, TCEQ-13563, had a linear increase in NO_x concentration and and decrease in chlorophyll-*a* from 2005 through 2014. The mid-bay site, TCEQ-13383, showed a periodic pattern in NO_x concentration that appeared similar to precipitation/inflow patterns, as well as a post 2011 increase in TKN concentrations. No significant long-term trends in concentrations were identified by GAMs for the lower-bay TCEQ-13384 site.

Freshwater inflow provided additional explanation for changes in TP and NO_x concentration at all of the Lavaca Bay sites according to AIC_c and model probability values (Table 4). TCEQ-13563, the site closest to the river outlet, was the only site that had improvements in the explanations of DO and TKN concentration with the inclusion of inflow. Both TCEQ-13563 and TCEQ-13383, the mid-bay site, saw improvements in explanations for variations in chlorophyll-*a* with the inclusion of freshwater inflow. The addition of nutrient loads (both TP and NO_3) terms did not provide additional explanation for changes in chlorophyll-*a* or DO concentrations. Inclusion of TP loads provided additional explanation of TP concentrations at the upper- and mid-bay sites, TCEQ-13563 and TCEQ-13383. Inclusion of NO_3 loads only provided marginal improvements in the explanation of NO_X concentration at the lower-bay TCEQ-13384 site.

Table 4. Estuary GAM AIC_c values and associated model probabilities (in parenthesis). Models with the highest probability for each site and water quality parameter combination are bolded and italicized for emphasis.

Parameter	Site	Temporal	Flow	Flow + Load
TP	TCEQ-13383	-152.1 (0.03)	-156.1 (0.24)	***-158.2 (0.72)***
	TCEQ-13384	-194.4 (0.03)	***-200.2 (0.49)***	-200.2 (0.49)
	TCEQ-13563	-145.3 (0)	-156.6 (0.41)	***-157.3 (0.59)***
NO_x	TCEQ-13383	-218.9 (0)	***-244.8 (0.5)***	-244.8 (0.5)
	TCEQ-13384	-263.4 (0)	-311.7 (0.48)	***-311.9 (0.52)***
	TCEQ-13563	-175.1 (0)	***-190.2 (0.5)***	-190.2 (0.5)
Chlorophyll-*a*	TCEQ-13383	279.7 (0.18)	***278.1 (0.41)***	278.1 (0.41)
	TCEQ-13384	***268.2 (0.33)***	268.2 (0.33)	268.2 (0.33)
	TCEQ-13563	289.5 (0.08)	***286.1 (0.46)***	286.1 (0.46)
TKN	TCEQ-13383	***42.2 (0.66)***	43.5 (0.34)	-
	TCEQ-13384	***34.3 (0.57)***	34.8 (0.43)	-
	TCEQ-13563	31.1 (0.22)	***28.7 (0.78)***	-
DO	TCEQ-13383	***146.4 (0.34)***	146.4 (0.34)	146.5 (0.32)
	TCEQ-13384	***135.9 (0.47)***	137 (0.27)	137 (0.27)
	TCEQ-13563	138.3 (0.25)	***137.2 (0.43)***	137.8 (0.32)

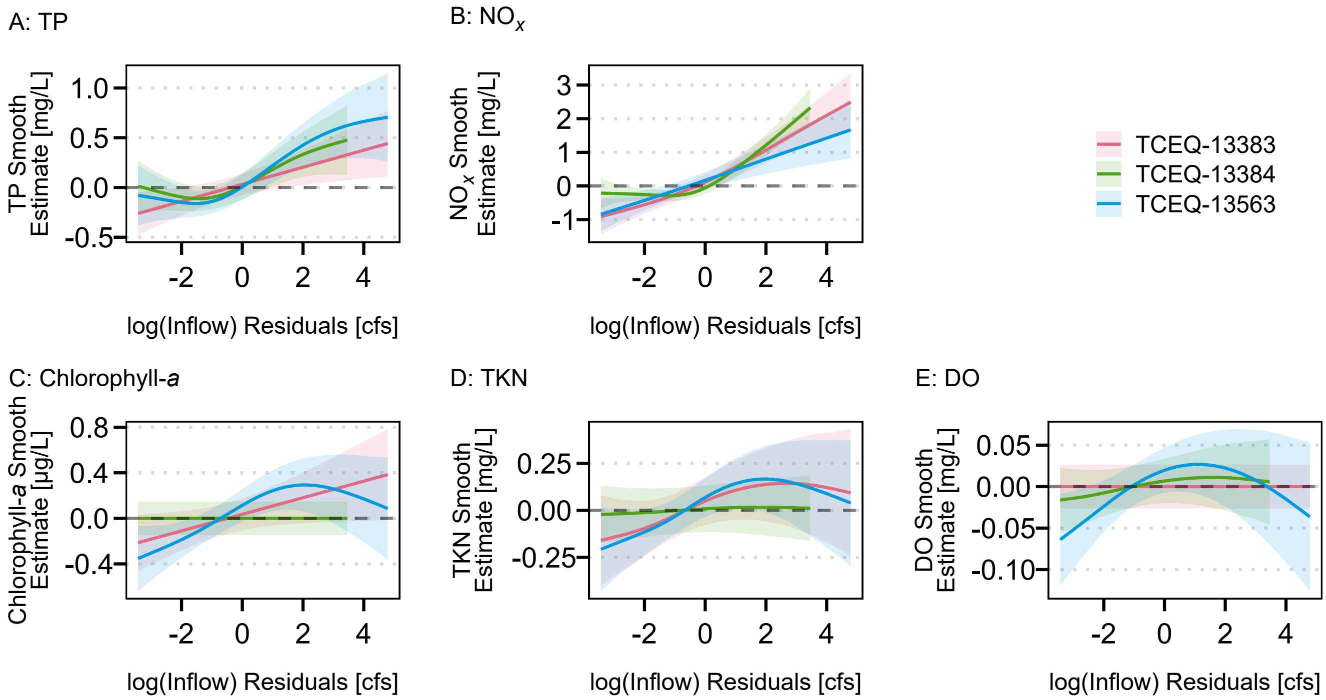

Figure 6. Estimated effects of mean daily inflow residuals on mean TP, NO_x, chlorophyll-*a*, TKN, and DO concentrations in Lavaca Bay obtained from flow estuary GAMs.

GAMs show that increases in freshwater inflow result in nearly linear increases in TP and NO_x at all three sites (Figure **??**). At the upper-bay TCEQ-13563 site, GAMs showed increases in freshwater inflow initially increased chlorophyll-*a* and DO concentration but concentrations leveled and potentially decreased at higher flows. The mid-bay TCEQ-13383 site showed a nearly linear increased in chlorophyll-*a* concentration with increases in freshwater inflow. Freshwater flow did not have significant effects on chlorophyll-*a*, TKN, or DO at the lower-bay TCEQ-13384 site.

Increases in TP loads resulted in nearly linear increases of TP concentration at the upper- and mid-bay sites, TCEQ-13563 and TCEQ-13383 respectively (Figure 7). The relative effect size appeared to much smaller than the effect of freshwater inflow alone. Increases in NO_3 loads only showed an effect at the lower-bay TCEQ-13384 site. The effect was quite small compared to streamflow and provided only small improvements to the model (Table 4). As noted above, nutrient loadings did not provide any explanation in changes in the remaining assessed water quality parameters.

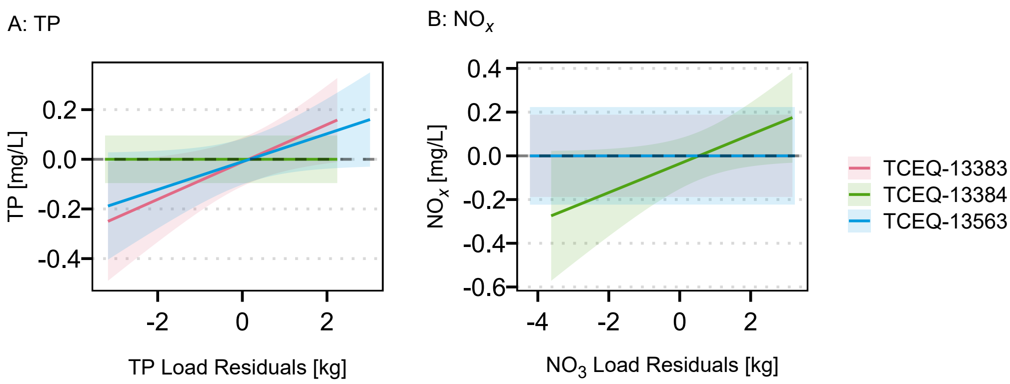

Figure 7. Estimated effects of nutrient load residuals on TP and NO_x concentrations in Lavaca Bay obtained from flow+load estuary GAMs.

4. Discussion

Table 5. Mean estimates of annual TP yield in the Lavaca River watershed in published studies.

Parameter	Reported Yield ($kg \cdot km^{-2} \cdot year^{-1}$)	Approach	Time Period	Reference
TP	35.2 (28.8, 43.3)*	GAM	2005-2020	This work
TP	45.2	SPARROW	2000-2014	Wise *et al.* [27]
TP	42	SWAT	1977-2005	Omani *et al.* [28]
TP	20.81-91.58†	SPARROW	1980-2002	Rebich *et al.* [29]
TP	28.9	LOADEST	1972-1993	Dunn [30]

* Values represent the mean of annual point estimates, lower and upper 95% credible intervals.
† A single point estimate was not reported, these value represent the range depicted on the choropleth map provided in the report.

Authors should discuss the results and how they can be interpreted in perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

5. Conclusion

This section is not mandatory, but can be added to the manuscript if the discussion is unusually long or complex.

Funding: This paper was funded by financial assistance provided by the Coastal Zone Management Act of 1972, as amended, administered by the National Oceanic and Atmospheric Administration (NOAA), Office for Coastal Management, pursuant to NOAA Award No. NA21NOS4190136. The

views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA, the U.S. Department of Commerce, or any of their subagencies.

Data Availability Statement: Data and code are openly available in Zenodo at https://doi.org/10.5281/zenodo.7330754.

Acknowledgments: All sources of funding of the study should be disclosed. Please clearly indicate grants that you have received in support of your research work. Clearly state if you received funds for covering the costs to publish in open access.

Conflicts of Interest: The author declares no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, an in the decision to publish the results.

References

1. Bricker, S.; Longstaff, B.; Dennison, W.; Jones, A.; Boicourt, K.; Wicks, C.; Woerner, J. Effects of Nutrient Enrichment in the Nation's Estuaries: A Decade of Change. *Harmful Algae* **2008**, *8*, 21–32. https://doi.org/10.1016/j.hal.2008.08.028.
2. Kennicutt, M.C. Water Quality of the Gulf of Mexico. In *Habitats and Biota of the Gulf of Mexico: Before the Deepwater Horizon Oil Spill*; Ward, C.H., Ed.; Springer New York: New York, NY, 2017; pp. 55–164. https://doi.org/10.1007/978-1-4939-3447-8_2.
3. Bugica, K.; Sterba-Boatwright, B.; Wetz, M.S. Water Quality Trends in Texas Estuaries. *Marine Pollution Bulletin* **2020**, *152*, 110903. https://doi.org/10.1016/j.marpolbul.2020.110903.
4. Wetz, M.S.; Hayes, K.C.; Fisher, K.V.; Price, L.; Sterba-Boatwright, B. Water Quality Dynamics in an Urbanizing Subtropical Estuary(Oso Bay, Texas). *Marine Pollution Bulletin* **2016**, *104*, 44–53. https://doi.org/10.1016/j.marpolbul.2016.02.013.
5. Wetz, M.S.; Cira, E.K.; Sterba-Boatwright, B.; Montagna, P.A.; Palmer, T.A.; Hayes, K.C. Exceptionally High Organic Nitrogen Concentrations in a Semi-Arid South Texas Estuary Susceptible to Brown Tide Blooms. *Estuarine, Coastal and Shelf Science* **2017**, *188*, 27–37. https://doi.org/10.1016/j.ecss.2017.02.001.
6. Chin, T.; Beecraft, L.; Wetz, M.S. Phytoplankton Biomass and Community Composition in Three Texas Estuaries Differing in Freshwater Inflow Regime. *Estuarine, Coastal and Shelf Science* **2022**, *277*, 108059. https://doi.org/10.1016/j.ecss.2022.108059.
7. De Cicco, L.A.; Hirsch, R.M.; Lorenz, D.L.; Watkins, W.; Johnson, M. dataRetrieval: R Packages for Discovering and Retrieving Water Data Available from Federal Hydrologic Web Services. U.S. Geological Survey, 2022. https://doi.org/10.5066/P9X4L3GE.
8. 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 Resources Research* **1992**, *28*, 2353–2363. https://doi.org/10.1029/92WR01008.
9. Hirsch, R.M.; Moyer, D.L.; Archfield, S.A. Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs1: Weighted Regressions on Time, Discharge, and Season (WRTDS), With an Application to Chesapeake Bay River Inputs. *JAWRA Journal of the American Water Resources Association* **2010**, *46*, 857–880. https://doi.org/10.1111/j.1752-1688.2010.00482.x.
10. Wang, Y.G.; Kuhnert, P.; Henderson, B. Load Estimation with Uncertainties from Opportunistic Sampling Data – A Semiparametric Approach. *Journal of Hydrology* **2011**, *396*, 148–157. https://doi.org/10.1016/j.jhydrol.2010.11.003.
11. Kroon, F.J.; Kuhnert, P.M.; Henderson, B.L.; Wilkinson, S.N.; Kinsey-Henderson, A.; Abbott, B.; Brodie, J.E.; Turner, R.D. River Loads of Suspended Solids, Nitrogen, Phosphorus and Herbicides Delivered to the Great Barrier Reef Lagoon. *Marine Pollution Bulletin* **2012**, *65*, 167–181. https://doi.org/10.1016/j.marpolbul.2011.10.018.
12. Kuhnert, P.M.; Henderson, B.L.; Lewis, S.E.; Bainbridge, Z.T.; Wilkinson, S.N.; Brodie, J.E. Quantifying Total Suspended Sediment Export from the Burdekin River Catchment Using the Loads Regression Estimator Tool: REGRESSION ESTIMATOR TOOL FOR POLLUTANT LOADS. *Water Resources Research* **2012**, *48*. https://doi.org/10.1029/2011WR011080.
13. Robson, B.J.; Dourdet, V. Prediction of Sediment, Particulate Nutrient and Dissolved Nutrient Concentrations in a Dry Tropical River to Provide Input to a Mechanistic Coastal Water Quality Model. *Environmental Modelling & Software* **2015**, *63*, 97–108. https://doi.org/10.1016/j.envsoft.2014.08.009.
14. Hagemann, M.; Asce, S.M.; Kim, D.; Park, M.H. Estimating Nutrient and Organic Carbon Loads to Water-Supply Reservoir Using Semiparametric Models. *J. Environ. Eng.* **2016**, p. 9. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001077.
15. McDowell, R.W.; Simpson, Z.P.; Ausseil, A.G.; Etheridge, Z.; Law, R. The Implications of Lag Times between Nitrate Leaching Losses and Riverine Loads for Water Quality Policy. *Scientific Reports* **2021**, *11*, 16450. https://doi.org/10.1038/s41598-021-95302-1.
16. Biagi, K.; Ross, C.; Oswald, C.; Sorichetti, R.; Thomas, J.; Wellen, C. Novel Predictors Related to Hysteresis and Baseflow Improve Predictions of Watershed Nutrient Loads: An Example from Ontario's Lower Great Lakes Basin. *Science of The Total Environment* **2022**, *826*, 154023. https://doi.org/10.1016/j.scitotenv.2022.154023.
17. Wood, S.N. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models: Estimation of Semiparametric Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2011**, *73*, 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x.

18. Beck, M.W.; Murphy, R.R. Numerical and Qualitative Contrasts of Two Statistical Models for Water Quality Change in Tidal Waters. *JAWRA Journal of the American Water Resources Association* **2017**, *53*, 197–219. https://doi.org/10.1111/1752-1688.12489.
19. Zhang, Q.; Ball, W.P. Improving Riverine Constituent Concentration and Flux Estimation by Accounting for Antecedent Discharge Conditions. *Journal of Hydrology* **2017**, *547*, 387–402. https://doi.org/10.1016/j.jhydrol.2016.12.052.
20. Vecchia, A.V.; Gilliom, R.J.; Sullivan, D.J.; Lorenz, D.L.; Martin, J.D. Trends in Concentrations and Use of Agricultural Herbicides for Corn Belt Rivers, 1996-2006. *Environmental Science & Technology* **2009**, *43*, 9096–9102. https://doi.org/10.1021/es902122j.
21. Hornung, R.W.; Reed, L.D. Estimation of Average Concentration in the Presence of Nondetectable Values. *Applied Occupational and Environmental Hygiene* **1990**, *5*, 46–51. https://doi.org/10.1080/1047322X.1990.10389587.
22. Wood, S.N. On Confidence Intervals for Generalized Additive Models Based on Penalized Regression Splines. *Australian & New Zealand Journal of Statistics* **2006**, *48*, 445–464. https://doi.org/10.1111/j.1467-842X.2006.00450.x.
23. Marra, G.; Wood, S.N. Coverage Properties of Confidence Intervals for Generalized Additive Model Components: Coverage Properties of GAM Intervals. *Scandinavian Journal of Statistics* **2012**, *39*, 53–74. https://doi.org/10.1111/j.1467-9469.2011.00760.x.
24. Burman, P. A Comparative Study of Ordinary Cross-Validation, v-Fold Cross-Validation and the Repeated Learning-Testing Methods. *Biometrika* **1989**, *76*, 503–514. https://doi.org/10.1093/biomet/76.3.503.
25. Murphy, R.R.; Keisman, J.; Harcum, J.; Karrh, R.R.; Lane, M.; Perry, E.S.; Zhang, Q. Nutrient Improvements in Chesapeake Bay: Direct Effect of Load Reductions and Implications for Coastal Management. *Environmental Science & Technology* **2022**, *56*, 260–270. https://doi.org/10.1021/acs.est.1c05388.
26. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. *Transactions of the ASABE* **2015**, *58*, 1763–1785. https://doi.org/10.13031/trans.58.10715.
27. Wise, D.R.; Anning, D.W.; Miller, O.W. Spatially Referenced Models of Streamflow and Nitrogen, Phosphorus, and Suspended-Sediment Transport in Streams of the Southwestern United States. Scientific Investigations Report 2019-5106, U.S. Geological Survey, Reston, Virginia, 2019. https://doi.org/10.3133/sir20195106.
28. Omani, N.; Srinivasan, R.; Lee, T. Estimation of Sediment and Nutrient Loads to Bays from Gauged and Ungauged Watersheds. *Applied Engineering in Agriculture* **2014**, pp. 869–887. https://doi.org/10.13031/aea.30.10162.
29. Rebich, R.A.; Houston, N.A.; Mize, S.V.; Pearson, D.K.; Ging, P.B.; Evan Hornig, C. Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico from Streams in the South-Central United States1: Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico From Streams in the South-Central United States. *JAWRA Journal of the American Water Resources Association* **2011**, *47*, 1061–1086. https://doi.org/10.1111/j.1752-1688.2011.00583.x.
30. Dunn, D. Trends in Nutrient Inflows to the Gulf of Mexico from Streams Draining the Conterminous United States, 1972-93. Water-Resources Investigations Report 96-4113, USGS, Austin, Texas, 1996. https://doi.org/10.3133/wri964113.