

1 Riverine phosphorus gain and loss across the conterminous United 2 States

3 Yiming Wang¹, Xuesong Zhang², Kaiguang Zhao¹, Robert D. Sabo³, Yuxin Miao⁴, Christopher M. Clark³

4
5 1. Ohio Agricultural Research and Development Center, School of Environment and Natural Resources, The Ohio State
6 University, Wooster, OH 44691, USA

7 2. USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD, 20705-2350, USA

8 3. United States Environmental Protection Agency, Office of Research and Development, Center for Public Health and
9 Environmental Assessment, Health and Environmental Effects Assessment Division, Washington, DC, 20460, USA.

10 4. Precision Agriculture Center, Department of Soil, Water and Climate, University of Minnesota, St. Paul, MN 55108, USA

11
12 *Correspondence to:* Yiming Wang (wang.20415@osu.edu); Xuesong Zhang (Xuesong.Zhang@usda.gov)

13 **Abstract.** Excess riverine phosphorus represents a preeminent catalyst for water quality degradation. Spatial mapping and
14 characterization of the net gain and loss of riverine phosphorus help discern the critical source areas. Here, we developed a
15 dataset encompassing phosphate (PO_4^{3-}) and total phosphorus (TP) gain and loss across catchments in the conterminous United
16 States (CONUS). We compiled 51,394 PO_4^{3-} and 285,675 TP concentration measurements and estimated PO_4^{3-} and TP loads
17 at 963 and 2,317 stations, respectively. Next, we leveraged the upstream-downstream topology information from the National
18 Hydrography Dataset Plus (NHDPlus) catchment map at the Hydrologic Unit Catalogue-12 (HUC12) level to derive the net
19 gain and loss of riverine phosphorus across catchments in the CONUS. Such maps can be used to estimate potential
20 contributions of point and non-point sources to riverine phosphorus pollution at refined spatial scales, identify different major
21 factors controlling local riverine P gain and loss compared to P loads, and evaluate watershed model's fidelity for representing
22 riverine P cycling. The resultant dataset is provided in Excel (.xlsx) format, accessible at Figshare
23 (<https://doi.org/10.6084/m9.figshare.28509317>, Wang et al., 2025). Leveraging the HUC12 information for spatialization, the
24 new datasets aim to address the existing gap in regional characterization of riverine phosphorus and support effective
25 management practices across the CONUS.

26 **1 Introduction**

27 Eutrophication in inland waters and estuaries is a widespread water quality challenge across the globe, with significant
28 economic cost (e.g., \$1 billion in Europe and \$2.2 billion annually in the United States (U.S.)) (Wurtsbaugh et al., 2019).
29 Excess phosphorus (P) is a primary contributor to eutrophication in streams and rivers, especially in intensive agricultural
30 regions (Brownlie et al., 2022; Royer et al., 2006). Riverine export of P is also a major contributor to oxygen-depleted dead
31 zones in coastal waters, causing damage to underwater life (Diaz and Rosenberg, 2008). There is an urgent need for global
32 actions to reduce P pollution for the environment and human health (UNEP, 2025).

33 Nonpoint or diffuse sources, particularly nutrients applied to agroecosystems, are often recognized as the primary source of
34 water pollution (Carpenter et al., 1998). P surplus in agricultural soils due to unused fertilization and manure application can
35 be transported to water bodies through surface runoff and groundwater pathways, and cause persistent water pollution
36 (Stackpoole et al., 2019). The diffuse nature of nonpoint source pollution poses challenges for directly identifying and
37 regulating critical source areas. Existing riverine P pollution databases mainly focus on certain agricultural areas (Ringeval et
38 al., 2024). Given the considerable spatial variation in P inputs to rivers (Arheimer and Lidén, 2000; Stackpoole et al., 2019;
39 Zhang et al., 2017), there is a lack of large-scale riverine P datasets at sufficient spatial scales across the conterminous United
40 States (CONUS) to quantify and analyze riverine P gain and loss. Such datasets, in conjunction with other observed and
41 modelled P data (e.g., point source discharge) help identify regions with high non-point source P inputs, thereby supporting
42 more effective targeting of measures for P pollution control. In addition, the datasets can also be used to assess fidelity of
43 distributed watershed models and understand key factors influencing local riverine P cycling.

44 In this study, we aim to develop new datasets for spatial characterization of riverine P gain and loss across the CONUS, to
45 help identify critical source areas and improve prioritization and implementation of nutrient management activities. The
46 subsequent sections of this paper include the method used to generate spatial riverine P gain and loss (Section 2), results
47 detailing the P dataset (Section 3), discussion of influencing factors and potential uncertainties (Section 4), codes and data
48 availability (Section 5), and conclusions (Section 6).

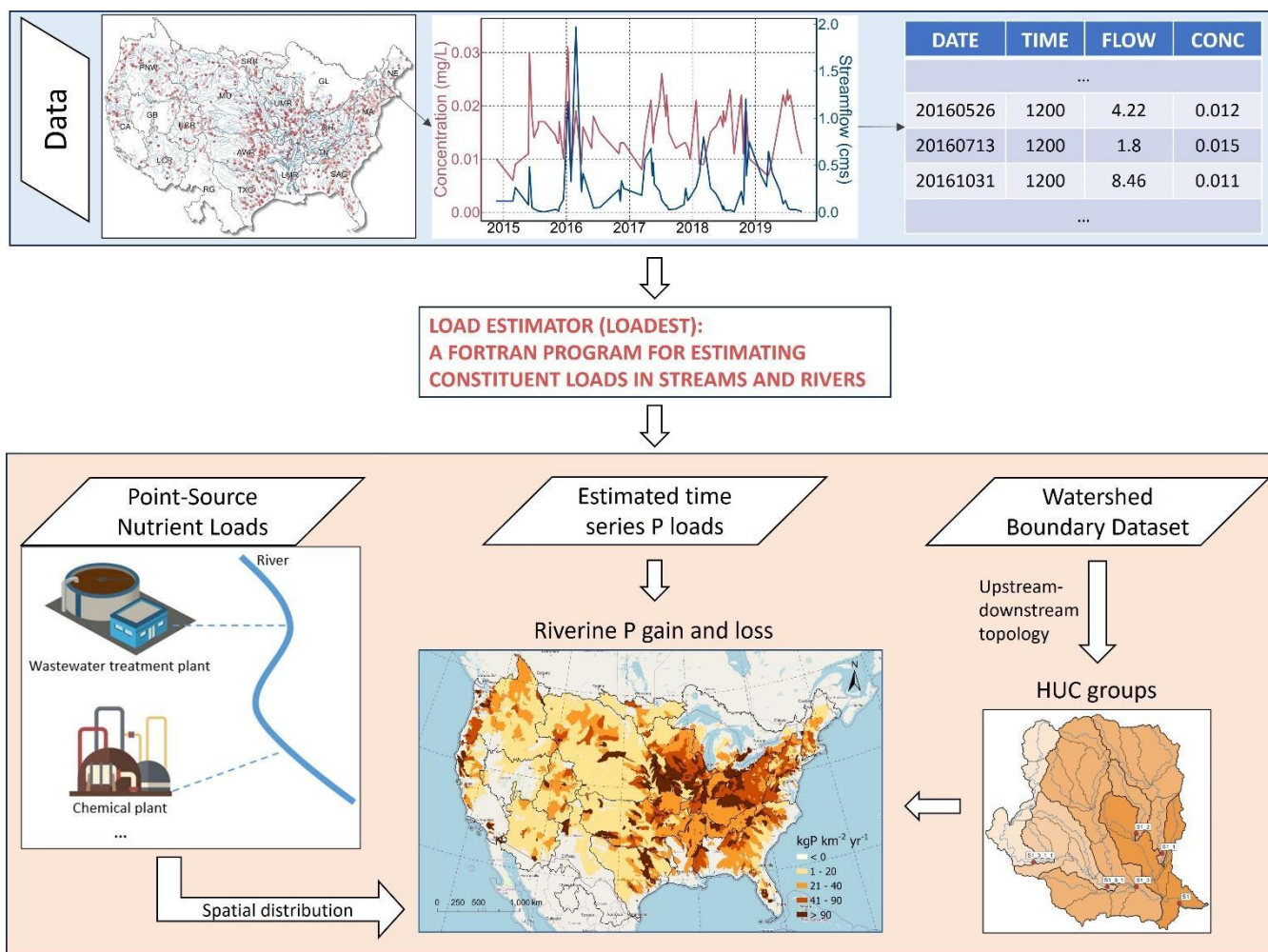
49

50 **2 Materials and Methods**

51 **2.1 Overview**

52 To estimate riverine P gain and loss data across the CONUS, we compiled streamflow and P concentration data (i.e., unfiltered
53 phosphate (PO_4^{3-})) from 963 monitoring stations and total phosphorus (TP) from 2,317 stations, and calculated P loads at these
54 stations using the Load Estimator (LOADEST) program (Runkel et al., 2004) (Fig. 1). Next, we estimated P gain and loss
55 across the catchments measured by one downstream station and its immediate upstream stations using the upstream-
56 downstream connectivity information contained in the National Hydrography Dataset Plus (NHDPlus) catchments

57 (https://nhdplus.com/NHDPlus/NHDPlusV2_data.php), resulting in 547 and 1,225 unique Hydrologic Unit Catalogue (HUC)
 58 groups for PO_4^{3-} and TP, respectively. Each HUC group has a unique pair of downstream and upstream stations that allows us
 59 to calculate the gain and loss of riverine P (as illustrated in Fig. S1 in the Supplemental Information). Note that headwater
 60 HUC groups only have one downstream station without upstream stations draining to them. Due to differences in data
 61 availability, the riverine PO_4^{3-} and TP gain and loss data cover 52,020 and 65,735 Hydrologic Unit Catalogue-12 (HUC12)
 62 catchments, respectively. Then we estimated potential contribution to P pollution from nonpoint sources by subtracting
 63 upstream P inputs and point source P inputs from the riverine P gain and loss in each catchment. Finally, we used land cover
 64 and climate data to evaluate major controls of riverine P gain and loss.



65
 66 **Figure 1: Overview of the generation of TP and PO_4^{3-} gain and loss data across the CONUS** A more detailed description and higher
 67 resolution figure about HUC group generation can be found in Text S2 and Fig. S1.

68 2.2 Study area

69 The CONUS (i.e., the lower 48 states of the U.S.) is located in North America from 46° 20' to 98° 34' W longitude and from
70 39° 49' to 41° 43' N latitude, covering an area of 8,080,464.3 km² with a north-to-south distance of approximately 2,660 km.
71 The terrain features higher elevations in the west and flatter areas in the east. Based on the Watershed Boundary Dataset (WBD;
72 <https://water.usgs.gov/GIS/huc.html>), the CONUS includes 18 major watersheds, encompassing several large rivers such as
73 the Mississippi and Colorado Rivers.

74 2.3 Data compilation

75 We compiled 51,394 PO₄³⁻ (USGS parameter code 00650) concentration data from 963 stations (spanning from 1952 to 2022)
76 and 285,675 TP (USGS parameter code 00665) observations from 2,317 stations (spanning from 1958 to 2023) across the
77 CONUS from the Water Quality Portal (Read et al., 2017). To ensure consistency with streamflow records, only records from
78 the U.S. Geological Survey (USGS) National Water Information System (NWIS) data source were used in this study. For each
79 P observation, we identified co-located monitoring stations (Wang, Zhang, Zhao, et al., 2024) and downloaded and processed
80 daily streamflow data from the USGS NWIS. Only stations with both phosphorus concentration observations and
81 corresponding daily streamflow records were retained. We calculated an average P concentration where there were multiple P
82 concentration observations on the same day. Before calculating the P load at a station with the LOADEST model, we excluded
83 stations with less than 12 measurements. Thus, we finally selected 547 stations for PO₄³⁻ and 1,225 stations for TP. For TP,
84 the "Point-Source Nutrient Loads to Streams of the Conterminous United States" dataset provides estimated annual total point-
85 source inputs during 2012 at the HUC12 level (Skinner and Maupin, 2019). This allowed us to aggregate the total TP input to
86 rivers for the catchments used to calculate riverine TP gain and loss. Note that the point source dataset does not contain PO₄³⁻.
87 Land cover and climatic controls of riverine P gain and loss were also assessed. Land cover data were derived from the National
88 Land Cover Database (NLCD; <https://doi.org/10.5066/P94UXNNTS>), which provided long-term average information on
89 various land cover types: barren land, crops, forest, hay, herbs, impervious surfaces, scrub, water, herbaceous wetlands, and
90 woody wetlands (Homer et al., 2012). Climate data were sourced from the PRISM dataset
91 (<https://www.prism.oregonstate.edu/>), including annual average temperature and total precipitation (Page et al., 2021). Using
92 upstream-downstream topology information, we calculated the total area of each land cover type from the headwater to the
93 current catchment at the HUC12 scale, representing their cumulative impact. For climate data, the local climate within each
94 HUC12 was used. The P surplus was accessed from the National Inventory of Phosphorus (NIP), which provides major inputs
95 and outputs of reactive P at the HUC8 scale across the CONUS (Sabo et al., 2021).
96 To ensure consistency across datasets with different spatial resolutions, all analyses were anchored at the HUC12 group scale.
97 Point-source inputs, originally reported at the HUC12 level, were aggregated to HUC12 groups to match the gain-loss estimates.
98 For datasets available at coarser scales (e.g., HUC8 for NIP inputs and HUC4 for agricultural inputs), gain and loss were
99 upscaled using area-weighted averaging. No downscaling was applied.

100 2.4 Riverine P gain and loss across catchments in the CONUS

101 Riverine P gain and loss was estimated by calculating the difference between P loads at a downstream monitoring station and
102 the sum of P loads from its neighbouring upstream stations. For multiple USGS stations located in the same HUC12 catchment,
103 we kept only one station on the mainstem of the river that is closest to the outlet of the HUC12 catchment, by comparing the
104 drainage area of the gaging station and the HUC12 catchment in which it is located. For headwaters, since there are no upstream
105 gages, the P load was used as the net riverine gain. The upstream-downstream topology relationship between the monitoring
106 stations was derived from the HUC12 catchments from the watershed boundary dataset (WBD;
107 <https://water.usgs.gov/GIS/huc.html>). Such a method has been outlined and tested by Qiu et al. (2023) (Text S2). As explained
108 above, we identified 547 and 1,225 unique HUC groups for PO₄³⁻ and TP, respectively. Note that each HUC group includes
109 multiple HUC12 polygons and these HUC12 catchments share the same gain and loss data.

110 For each HUC group, the balance of riverine P can be expressed as follows:

$$\begin{aligned} 111 \quad P \text{ load at downstream outlet} &= (P \text{ loads from upstream inputs}) + (P \text{ from point sources}) - \\ 112 \quad & (Riverine \text{ removal of } P \text{ from point sources}) + (P \text{ from non-point sources}) - \\ 113 \quad & (Riverine \text{ removal of } P \text{ from non-point sources}) \end{aligned} \quad (1)$$

114 Because in-stream removal terms cannot be directly constrained and may vary across systems (e.g., about 12% globally)
115 (Maavara et al., 2015), we assume the values of these two terms to be negligible for the purpose of deriving a simplified, first-
116 order estimate for nonpoint source inputs. Thus, riverine removal is not explicitly estimated in this study, and the calculated P
117 gain and loss represent the net difference between downstream and upstream loads. Conceptually, this net difference reflects
118 the combined effects of watershed P inputs (from both point and nonpoint sources) and in-stream processes (e.g., retention,
119 transformation, and remobilization) occurring along the flow path, rather than a direct measure of any single process such as
120 in-stream removal. Under this assumption, Eq. (1) reduces to:

$$121 \quad (P \text{ from nonpoint sources}) = P \text{ gain and loss} - (P \text{ from point sources}) \quad (2)$$

122 where $P \text{ gain and loss} = (P \text{ load at downstream outlet}) - (P \text{ loads from upstream inputs})$.

123 Accordingly, negative values of P gain and loss indicate net decreases in load along the flow path, which may reflect a
124 combination of retention, transformation, or other processes, rather than being interpreted solely as riverine removal. This
125 formulation neglects in-stream P removal, and therefore $(P \text{ from nonpoint sources}) = P \text{ gain and loss} -$
126 $(P \text{ from point sources})$ is a lower-end estimate of the nonpoint source contribution to riverine P for each HUC group. Since
127 only TP from point sources is available, we derived nonpoint-source TP loads but not for PO₄³⁻.

128 2.5 Evaluation of estimated riverine load

129 We evaluated the consistency of the PO₄³⁻ and TP loads against another independent dataset derived with the Weighted
130 Regressions on Time, Discharge, and Season (WRTDS) model (Hirsch et al., 2010; Zhang and Hirsch, 2019). First, we
131 compared multi-year average TP loads from 151 monitoring stations. We also evaluated the estimated unfiltered PO₄³⁻ loads,

132 which assess the mass of reactive P susceptible to being released in the water column under various redox conditions.
133 Furthermore, the reliability of the upstream-downstream connectivity information is important for deriving the drainage area
134 of HUC groups that are controlled by pairs of upstream and downstream stations. Here we used a quality-checked and corrected
135 NHDPlus HUC12 catchment map (Wang, Zhang, & Zhao, 2024) that has been verified for reliably deriving the drainage area
136 of each USGS station as compared to the USGS GAGES-II reported values (Falcone and Survey, 2011). These efforts helped
137 ensure the quality of the riverine P gain and loss data developed in this study. In addition, we evaluated long-term trends in
138 riverine P loads using the Sen's slope estimator in combination with the Mann-Kendall test (Hamed and Rao, 1998). To ensure
139 robust trend detection, only monitoring stations with more than 30 years of load estimates were included. This resulted in 405
140 TP stations and 53 PO_4^{3-} stations used for trend analysis. The Sen's slope and corresponding p-values are provided in the
141 dataset.

142 **2.6 Analysis of environmental controls**

143 Recent studies reveal that shifts in land use, agricultural practices, and climatic conditions have introduced a pervasive increase
144 in soluble P concentrations across many different watersheds (Houser and Richardson, 2010; Singh et al., 2023). To assess the
145 spatial factors influencing riverine P gain and loss, we employed random forest modelling to evaluate the relative importance
146 of multiple environmental variables (Breiman, 2001). These factors were categorized into three groups: climatic factors, land
147 cover types, and additional influences such as cumulative agricultural inputs and upstream loads. Given that the NIP dataset is
148 only available at the HUC8 scale and some HUC groups are larger than the HUC8 catchment areas, we calculated the
149 cumulative agricultural inputs at the HUC4 scale. In more detail, we used the ranger package, optimizing the model structure
150 with the caret package in R. Key tuning parameters included the number of variables to use in each split (mtry), the number of
151 trees (n_trees) and the minimum size of data points before splitting a tree (min_n). The tuning process was performed by doing
152 a grid search for mtry (2-6) and min_n (10-20), then a second search was performed to find the optimal n_trees parameter
153 (500-3000). To minimize random effects, the model was run 10 times, and we calculated the average importance value and
154 harmonic mean p-value (Wilson, 2019).

155 **3 Results**

156 **3.1 Riverine phosphorus data**

157 We created two datasets, "Riverine PO_4^{3-} " and "Riverine TP," that encapsulate estimated multi-year average riverine gain and
158 loss and loads, as well as point source and nonpoint source contributions for each HUC group for PO_4^{3-} and TP, respectively
159 (Table 1). Complementing this information, the datasets encompass the location (i.e., longitude and latitude) of the outlet of
160 the HUC group, the area of the HUC group, the count of observations used to calculate P loads, and commencement and
161 termination years of observed data, to facilitate user-defined subsetting of the datasets. Additional information regarding the

162 regression model is also included, such as the form of the regression model selected by LOADEST and the associated
 163 coefficient of determination (r^2) values.

164 Across the board, the average r^2 values for the best-fit model (Table S1) across all sites are 0.76 for PO_4^{3-} loads and 0.83 for
 165 TP loads. Generally, in the load regression, the r^2 values for TP estimation outperformed those for PO_4^{3-} estimation, particularly
 166 in the Mid-Atlantic region (Fig. S2). It is noteworthy, however, that certain monitoring stations exhibited low r^2 values due to
 167 the limited availability of paired P concentration with streamflow data for regression.

168
 169 **Table 1. Data records in the "Riverine PO_4^{3-} " and "Riverine TP" datasets.**

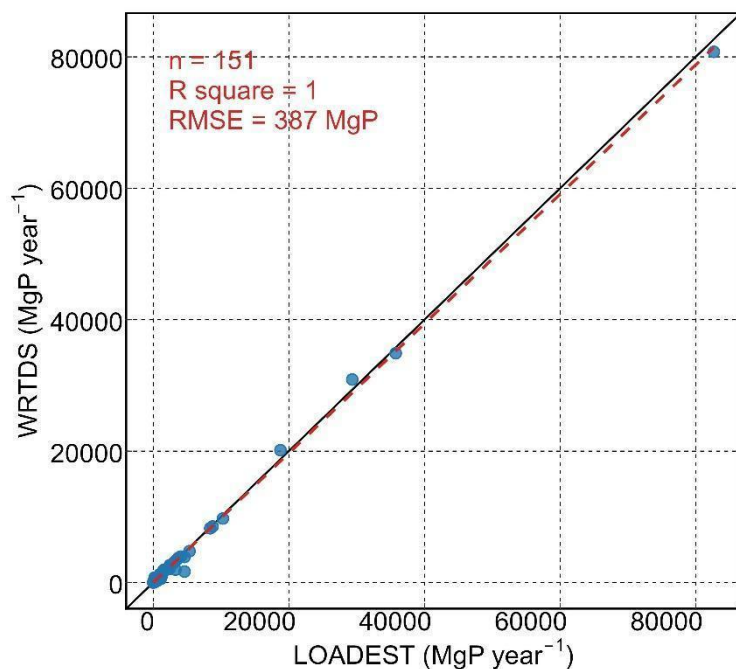
Field name	Description
Station ID	U.S. Geological Survey designated ID; Note that this ID is also used to denote a unique HUC group
Lat	Latitude of the monitoring station at the outlet of a HUC group
Long	Longitude of the monitoring station at the outlet of a HUC group
Area	Area of the HUC group
Load	The amount of PO_4^{3-} or TP loads at the outlet station of a HUC group (kgP yr^{-1})
P gain and loss	The difference between P loads at the outlet of a HUC group and all its immediate upstream stations ($\text{kgP km}^{-2} \text{yr}^{-1}$)
NonPts TP contribution*	Nonpoint-source contribution to riverine phosphorus from a HUC group ($\text{kgP km}^{-2} \text{yr}^{-1}$)
Pts TP load	Point-source TP loads (kg yr^{-1}) from a HUC group
obsNum	The number of phosphorus concentration data with paired streamflow data
startYr	The starting year of the observed data
endYr	The ending year of the observed data
modelID	The ID of regression model used for load estimation
r2	R-Squared (%) for the selected LOADEST regression model used to estimate the P load
Sen's slope	Sen's slope of riverine P loads, representing the long-term rate of change in loads at monitoring stations (kg yr^{-2}).
p-value	The p-value of long-term trends
Areal-normalized slope	Areal-normalized Sen's slope of riverine P loads, representing the long-term rate of change in areal P export ($\text{kg km}^{-2} \text{yr}^{-2}$).

170 Note: * only for TP as only point source TP inputs are available.

171

172 Our TP load estimates are highly consistent with the WRTDS model with high r^2 and low root mean square error (RMSE)
 173 (Fig. 2). The minor disparities observed between these two datasets are likely attributable to variations in temporal coverage

174 and different regression equations. For PO_4^{3-} , we found that only 11 stations with WRTDS estimates matched the stations
175 used here, and all 11 stations are located in small watersheds. Therefore, we leveraged filtered PO_4^{3-} loads estimated by
176 WRTDS to assess if the LOADEST estimated unfiltered PO_4^{3-} loads. Unfiltered PO_4^{3-} measures both the dissolved PO_4^{3-} as
177 well as PO_4^{3-} compounds bound to suspended sediments and organic materials, and thus will have higher load compared to
178 filtered PO_4^{3-} measurements (Fig. S3). Nonetheless, the high correlation indicates our estimates of unfiltered PO_4^{3-} are
179 reasonable.

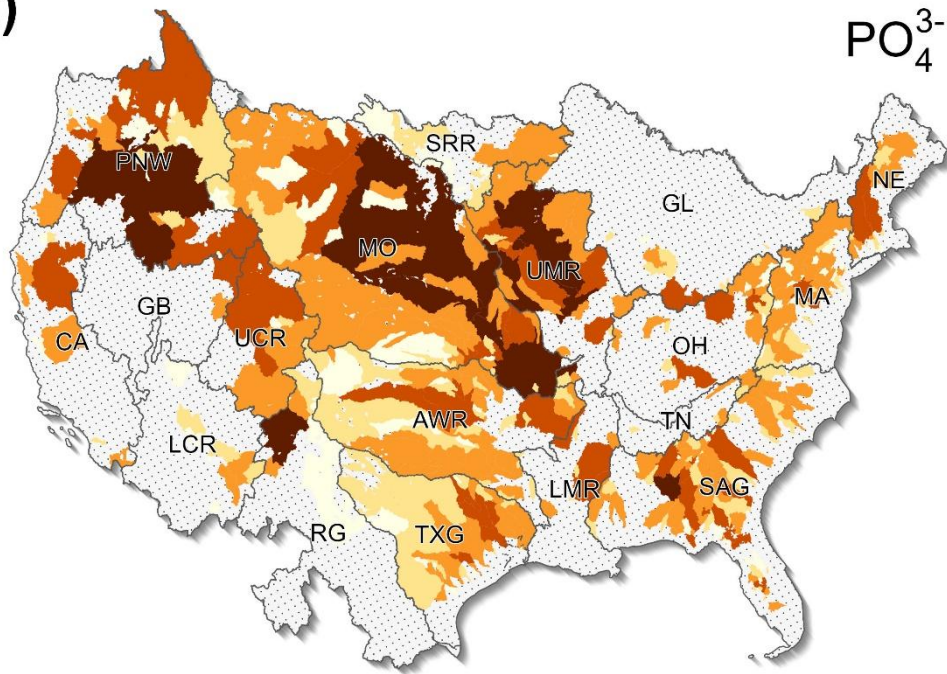


180
181 **Figure 2: Comparison of riverine TP loads between our estimates and previous WRTDS estimated values at 151 monitoring stations.**
182 **Note that the riverine TP loads vary greatly at different locations and over time, with most sites being below 10,000 MgP year⁻¹.**

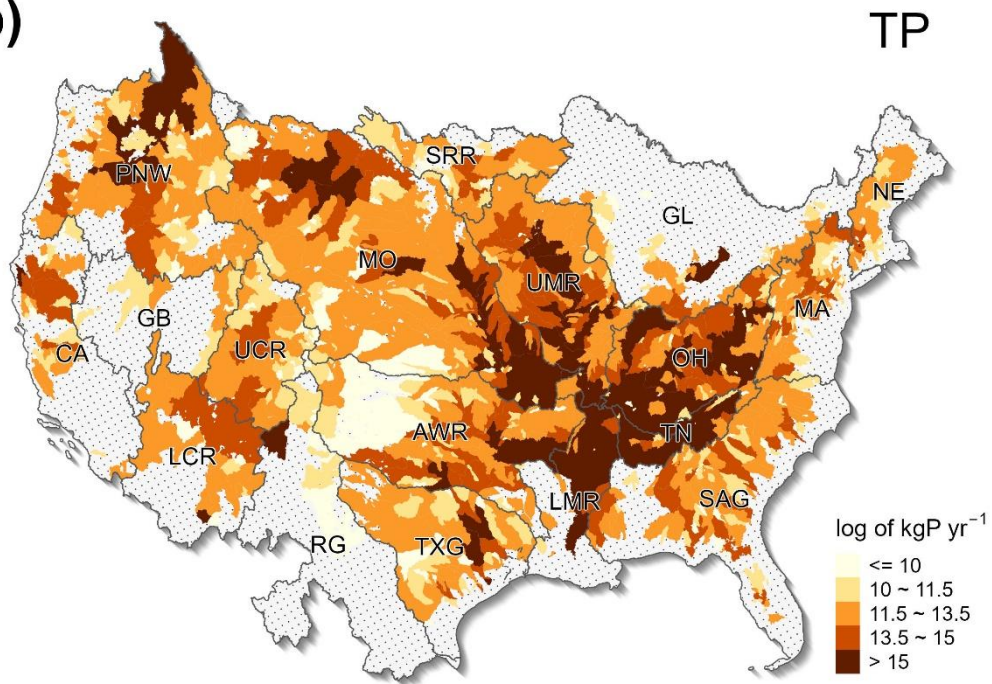
183 Spatial patterns of PO_4^{3-} and TP loads from each HUC group across the CONUS are shown in Fig. 3. Additionally, the location
184 of the station at the outlet of each HUC group and the loads of streams in which it is located are shown in Fig. S4. The datasets
185 encompass 547 stations/HUC groups for PO_4^{3-} and 1,225 stations/HUC groups for TP, covering 4,894,464 and 6,118,360 km²
186 PO_4^{3-} and TP, respectively. At the HUC2 scale, the derived gain and loss estimates cover 21% to 99.9% of basin area for TP
187 and 7% to 96.5% for PO_4^{3-} , with 72% and 44% of HUC2 basins exceeding 50% coverage, respectively (Fig. S5 and Table S2).
188 The difference in spatial coverage is mainly due to the abundance of TP compared to PO_4^{3-} . Stations with high P loads are
189 predominantly situated in the Midwest or proximate to megacities, with a general pattern of higher P loads observed in the
190 eastern U.S. PO_4^{3-} loads range from 111 to 31,671,885 kgP yr⁻¹ and TP loads range from 235 to 336,223,136 kgP yr⁻¹. Median
191 loads are 76,202 and 108,305 kgP yr⁻¹ and average loads are 436,311 and 1,012,363 kgP yr⁻¹, for PO_4^{3-} and TP, respectively.
192 We further examined temporal trends in riverine P loads at stations with long-term records (>30 years). A substantial fraction
193 of stations exhibited statistically significant trends, with 46% of TP stations and 64% of PO_4^{3-} stations showing significant

194 changes. Decreasing trends were more prevalent, accounting for 72% and 58% of TP and PO_4^{3-} stations, respectively. Stations
195 with increasing TP trends were primarily located in the Mississippi River Basin (Fig. S6), suggesting regional differences in
196 nutrient dynamics.

(a)



(b)



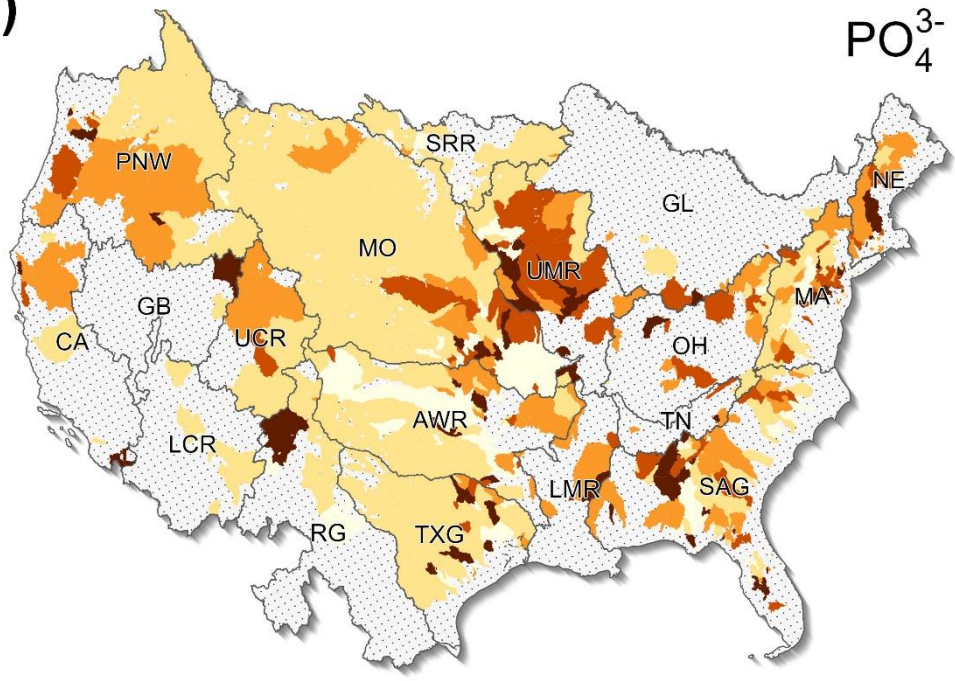
198 **Figure 3: Riverine (a) PO₄³⁻ and (b) TP loads from HUC groups across the CONUS. The boundary lines show the Hydrologic Unit**
199 **Catalogue 2-digit (HUC2) watersheds. Grey areas indicate regions with no data. For visualization purposes, the logarithm was used**
200 **here.**

201

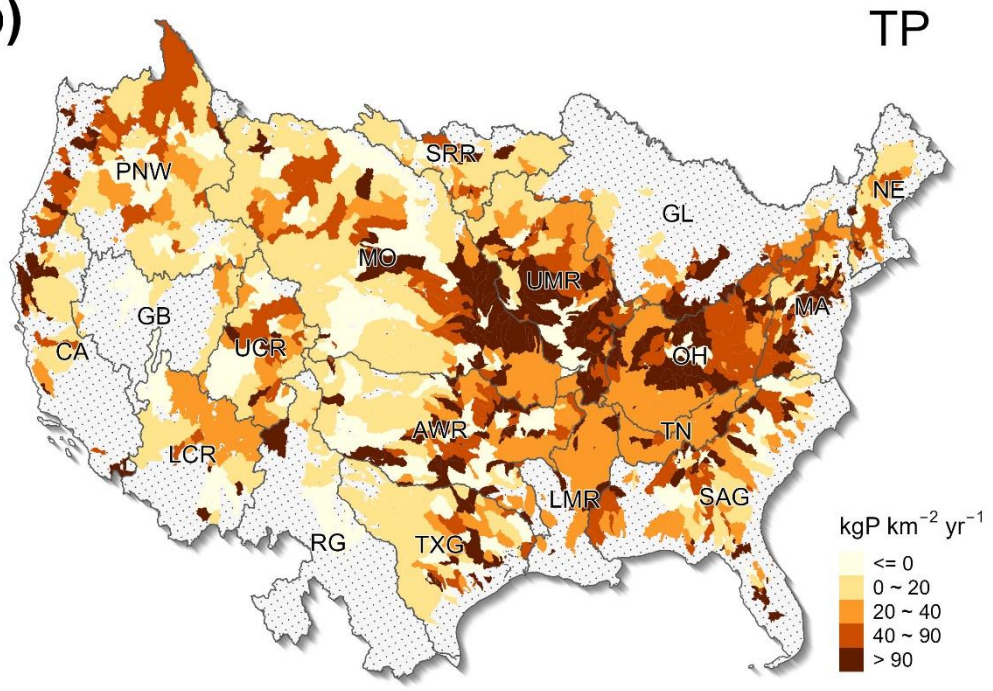
202 The spatial distribution of riverine P gain and loss is shown in Fig. 4. Both PO₄³⁻ and TP gain and loss exhibit similar spatial
203 patterns over the CONUS, with most areas exhibiting riverine P gains. The area-weighted average PO₄³⁻ gain stands at 25.39
204 kgP km⁻² yr⁻¹, and the TP gain is 33.68 kgP km⁻² yr⁻¹. Median PO₄³⁻ and TP gains are lower than averages, standing at 16.75
205 kgP km⁻² yr⁻¹ and 33.57 kgP km⁻² yr⁻¹, respectively. At the HUC group scale, the highest area-weighted PO₄³⁻ gain was
206 identified in the Upper Mississippi Region (UMR), amounting to about 113.96 kgP km⁻² yr⁻¹. The highest TP gain reached
207 186.55 kgP km⁻² yr⁻¹ in the Tennessee Region (TN). Notably, widespread regions in the Midwest exhibit heightened P gains
208 (Fig. S7), particularly in terms of PO₄³⁻, suggesting a discernible impact of human activities (e.g., agricultural fertilization). At
209 the HUC2 level, the lowest area-weighted PO₄³⁻ gain (1.72 kgP km⁻² yr⁻¹) was found in the Rio Grande Region (RG), and the
210 lowest TP gain (0.62 kgP km⁻² yr⁻¹) was found in the Upper Colorado Region (UCR). Refined examination at the HUC group
211 level showed that, over the CONUS, 392,778 km² and 1,468,973 km² areas exhibited riverine PO₄³⁻ and TP losses, respectively.

212

(a)



(b)



214 **Figure 4: The spatial distribution of the gain and loss in riverine (a) PO_4^{3-} and (b) TP over the CONUS. The boundary lines show**
215 **the Hydrologic Unit Catalogue 2-digit (HUC2) watersheds. Grey areas indicate regions with no data.**

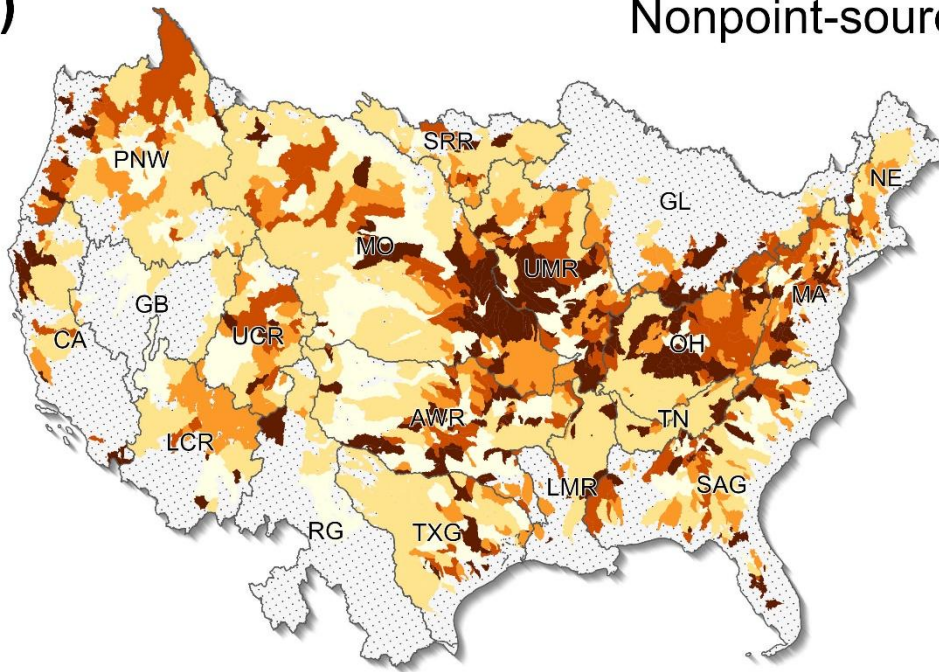
216

217 **3.2 Point and nonpoint source contributions**

218 We also mapped the spatial point source and nonpoint source inputs of TP, as shown in Fig. 5. The nonpoint source
219 contributions are estimated based on Equation (2), which provides a lower-end estimate given that the riverine removal of
220 point and nonpoint source P as shown in Equation (2) was not considered. Point and nonpoint source contributions to riverine
221 TP pollution exhibited large differences in both the magnitude and spatial distribution. Over the CONUS, the area-averaged
222 point source input of TP is $5.44 \text{ kgP km}^{-2} \text{ yr}^{-1}$. By subtracting point source inputs from the calculated TP gain and loss, we
223 obtained an area-averaged nonpoint source TP contribution of $28.24 \text{ kgP km}^{-2} \text{ yr}^{-1}$. Regions characterized by high TP gain with
224 minimal point source pollution were observed in the Midwest. Notably, in most of the agriculturally intensive Missouri and
225 Tennessee-Ohio river basins, total nonpoint source discharge significantly surpassed point source contributions (Fig. S8). Upon
226 the exclusion of point source contributions (Fig. 5b), there is a substantial change, with the areas with riverine TP losses
227 expanding to $1,603,258 \text{ km}^2$, most of them in the Missouri and Arkansas-White-Red River Basins. In general, most watersheds
228 with negative nonpoint sources are concentrated in the western U.S. This does not mean that the nonpoint source P inputs are
229 negative, but indicates that riverine processes likely removed a large fraction of point and nonpoint source P.

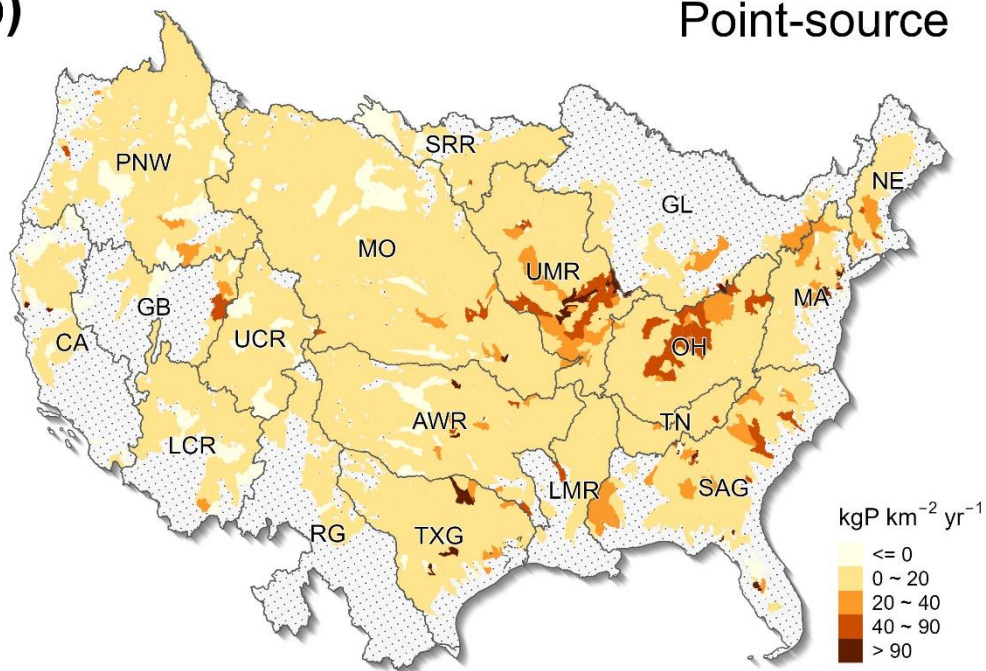
(a)

Nonpoint-source



(b)

Point-source

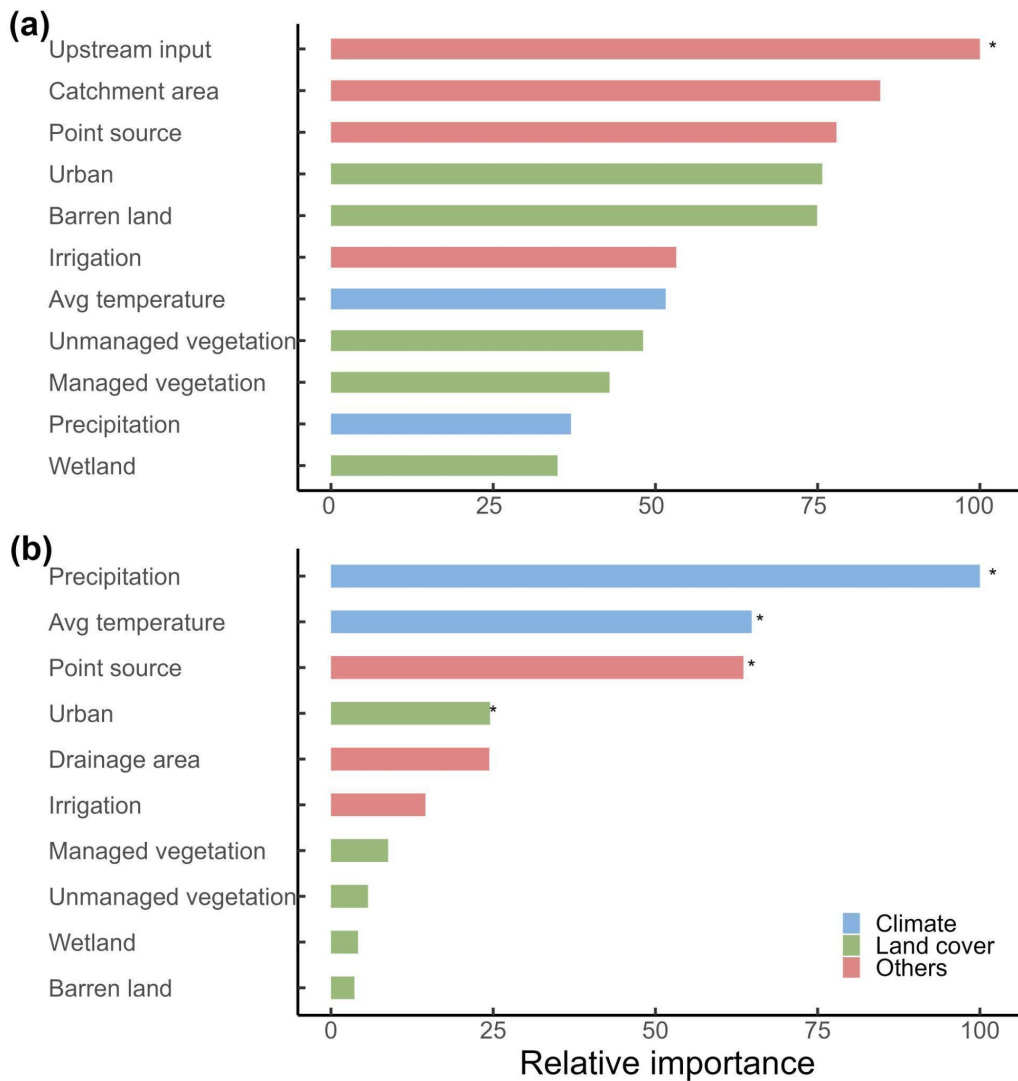


231 **Figure 5: The spatial distribution of (a) nonpoint source and (b) point source contributions to riverine P pollution over the CONUS.**
232 **The boundary lines show the Hydrologic Unit Catalogue 2-digit (HUC2) watersheds. Grey areas indicate regions with no data.**

233

234 **3.3 Factors influencing TP**

235 We employed a random forest model to assess the influence of climate, land use, human activities, and catchment
236 characteristics on riverine TP gain and loss and TP loads (Fig. 6). Note that the calculated P loads represent the outcome of
237 the entire upstream catchment processes, while the riverine P gain and loss data represent both upstream catchment processes
238 (e.g., P inputs from upstream) and local catchment properties (e.g., climate and land use in a HUC group). Such differences
239 lead to the use of different sets of influencing factors (Fig. 6). The land use, climate and point source factors were calculated
240 for the entire upstream area draining to a monitoring station for TP load analysis. In contrast, those factors were averaged over
241 a HUC group for riverine gain and loss analysis. Additionally, for the analysis of riverine P gain and loss data, we included
242 upstream P inputs. Analysis results indicate that upstream input is the sole statistically significant factor affecting TP gain and
243 loss, with climate and land cover showing no notable impact. Conversely, TP loads are predominantly influenced by climatic
244 factors, alongside significant contributions from point source discharges and urban land use.



245

246

247

Figure 6: The importance of factors influencing (a) TP gain and loss and (b) TP loads. Asterisks indicate significance at a level of 0.05.

248

4 Discussion

249

4.1 Important contributions from nonpoint sources to riverine P pollution

250

The estimated area-averaged nonpoint source TP contribution ($28.24 \text{ kgP km}^{-2} \text{ yr}^{-1}$) represents a reduction of 16.2% from the calculated TP gain and loss that includes contributions from both point and nonpoint sources ($33.68 \text{ kgP km}^{-2} \text{ yr}^{-1}$). Given that the TP inputs from point and nonpoint sources are often subject to riverine removal (Withers and Jarvie, 2008), the estimated nonpoint source TP based on Equation (1) should be augmented by the amount of total TP inputs (including both nonpoint and point source) removed through riverine processes. Therefore, the calculated nonpoint source inputs of TP represent an

254

255 underestimate of the contributions from nonpoint sources. If we assume a 12% removal rate for TP inputs (Maavara et al.,
256 2015), then the nonpoint source inputs of TP would increase from 28.24 kgP km⁻² yr⁻¹ to 32.28 kgP km⁻² yr⁻¹. Collectively, the
257 results show that the nonpoint sources likely contribute more than 84% of riverine TP pollution.

258 **4.2 Implications for analysing environmental controls of riverine P**

259 Climatic factors were key drivers of TP loads at the outlet of a watershed, which in general aligns with findings from previous
260 studies (Sabo et al., 2023) and underscores the role of climate in nutrient transport dynamics. Our environmental control
261 analysis using the gain and loss data showed that upstream inputs are leading control of local riverine gain and loss (Fig. 6a),
262 in addition to local inputs of P from point and nonpoint sources and local riverine processes, such as in-stream retention through
263 mechanisms such as P absorption by periphyton via photosynthesis and hydrological processes like reduced streamflow and
264 sedimentation (Dodds, 2003; Withers and Jarvie, 2008). Notably, although accumulated agricultural P inputs (i.e., livestock
265 waste and agricultural fertilizer) positively influenced TP gain and loss (Fig. S9), they were not included in this analysis due
266 to mismatch of spatial scales. In general, using TP loads and riverine P gain and loss can lead to pronounced differences in the
267 analysis of importance of environmental controls. Although climate conditions (i.e., precipitation and temperature) are the
268 major controls of TP loads, which represent the integration of the entire watershed conditions, while riverine P gain and loss
269 indicate that the amount of upstream P inputs entering a local catchment is an important factor influencing the riverine
270 processing of P.

271 Both the differences and the analysis using TP loads and riverine gain and loss data revealed the importance of urban land and
272 agricultural management. For example, both analyses show that irrigation can influence riverine TP, indicating that improving
273 irrigation efficiency and technology holds potential to reduce TP inputs from cropland fertilization (Xia et al., 2020). Though
274 we didn't assess factors influencing PO₄³⁻ due to the lack of point source PO₄³⁻ data, TP hotspots are expected to occur further
275 downstream than PO₄³⁻ hotspots (Fig. 3). This is probably because rivers typically can retain (e.g., periphyton assimilation,
276 adsorption onto suspended or bed sediment) a considerable proportion of incoming soluble-reactive P (e.g., PO₄³⁻) within the
277 upper network, whereas particulate P continue to transport to downstream (Jarvie et al., 2012; Robertson and Saad, 2019;
278 Royer et al., 2006). Overall, the intricate interplay between climate and land use factors underscores the complex nature of P
279 dynamics in riverine systems. These newly developed riverine gain and loss datasets help improve understanding of local
280 controls of riverine P dynamics and identify hotspots of changes in riverine P.

281 **4.3 Limitations and contribution**

282 While the newly developed datasets leverage upstream-downstream topology information at the HUC12 level to help increase
283 the spatial resolution of riverine P gain and loss data, it is essential to acknowledge limitations relevant to understanding and
284 quantifying P cycles and identifying sources of P inputs. First, the accuracy of load estimation via LOADEST is contingent
285 upon the availability of paired P concentration data from monitoring stations. Stations with limited observations may introduce

286 higher uncertainties in load estimations. The robustness of our datasets is partially reflected in the number of observations
287 available. The average and median numbers of PO_4^{3-} observations per site is 54 and 29, respectively; for TP, the average and
288 median numbers of observations per site is 134 and 90, respectively. In addition, the average center year of TP observations is
289 1992, with a median of 1991, while for PO_4^{3-} , the overall observation time is relatively early, with an average of 1981 and a
290 median of 1980. The use of AIC to choose the most parsimonious regression model (average r^2 of 0.76 and 0.83 for PO_4^{3-} and
291 TP, respectively) helped reduce uncertainties in load estimates.

292 Additionally, we calculated the P loads by averaging over different time periods with available data for each monitoring station.
293 The mismatch between observational periods of upstream and downstream stations could introduce uncertainties, given that
294 the available data cover various time periods for different stations. For example, streamflow discharge, which is important for
295 calculating nutrient loads, can vary from year to year. Here, we assumed that multi-year average estimates of P loads are
296 representative of the long-term pattern at a monitoring station. This may not hold for some upstream and downstream stations
297 covering time periods that do not overlap with each other. Therefore, we provided the number and period of observations and
298 model performance information in the datasets, which can help users to refine the calculation of riverine P gain and loss by
299 further screening the P loads data at the monitoring stations. Note that available stations with observed P concentration and
300 streamflow data are relatively sparse in the western vs. eastern U.S., particularly for PO_4^{3-} . This led to large gaps in the spatial
301 coverage of the datasets (Fig. 4). Increasing the number of stations with P observations holds the potential to enhance the
302 accuracy of riverine P estimates in the future.

303 It is also worth noting that the calculated contribution of TP from nonpoint sources represents a conservative estimate, given
304 the unknown rate of TP removal from point and nonpoint sources. Although previous studies showed that TP removal rates
305 were generally small, they could vary substantially across regions, as evidenced by the areas with riverine P loss (Fig. 4). It is
306 reasonable to assume that the estimated TP contribution from nonpoint sources is greater than 84% over the CONUS; however,
307 the local TP contribution from nonpoint sources could be much lower, particularly in regions with high point source inputs.
308 Also, the removal rates of P from point and nonpoint sources are likely different due to differences in the quality of P inputs
309 (e.g., biodegradability and adsorption and desorption to sediments) and flow pathways (Wang et al., 2025a). Therefore, caution
310 should be taken when interpreting the local contribution to P pollution from nonpoint sources.

311 Despite these challenges, our datasets make a unique contribution to the quantification and analysis of riverine P load, gain
312 and loss, and sources across the CONUS. They can support the evaluation and diagnosis of large-scale dynamic watershed
313 models, the examination of environmental controls on riverine P loads, and the estimation of contributions to P gain and loss.
314 For instance, models such as SPARROW could incorporate the spatial patterns of riverine P gain and loss to better constrain
315 nutrient sources and in-stream processing across river networks. The insights derived from our datasets contribute to a more
316 comprehensive understanding of P dynamics under long-term and multi-decadal conditions, providing a foundation for
317 improved water quality management on local, regional, and national scales.

318 **5 Code and data availability**

319 All codes for validating and visualizing PO_4^{3-} and TP gain and loss from the load estimations were run in R version 4.3.1 and
320 are archived at https://github.com/ymwang4924/gain-loss_P. The datasets presented in the paper are available at
321 <https://doi.org/10.6084/m9.figshare.28509317> (Wang et al., 2025b).

322 **6 Conclusions**

323 In this study, we estimated riverine loads of PO_4^{3-} and TP and derived their gain and loss across the CONUS, leveraging the
324 upstream-downstream hydrological connectivity information contained in the NHDPlus catchment map. On average, rivers
325 across the CONUS gain TP at a rate of $33.68 \text{ kgP km}^{-2} \text{ yr}^{-1}$, with notable hotspots in the Midwest. Due to the limitations of
326 data availability, the precision of estimated P gain and loss data could be influenced by the number and periods of observations
327 available at upstream and downstream stations. We provided additional information regarding the number of observations
328 available, temporal coverage of data, the regression model used, and the model's statistical performance, so that users can
329 further subset the datasets to meet certain specific criteria.

330 The riverine P gain and loss datasets allow the estimation of riverine P removal or accrual at a refined spatial resolution to
331 better reflect the impacts of local controls. In contrast, riverine P loads at monitoring stations embody the integrated processes
332 from the entire area upstream of a specific station. Also, by combining point source inputs with the riverine P gain and loss
333 datasets, we derived conservative estimates of the long-term average contribution of nonpoint sources to riverine TP (28.24
334 $\text{kgP km}^{-2} \text{ yr}^{-1}$). The control factor analysis with a random forest model demonstrated that upstream inputs had the greatest
335 influence on the local riverine P gain and loss, while climatic factors dominated riverine P loads at monitoring stations. This
336 suggests that nutrient management practices that prioritize enhancing irrigation efficiency and integrating strategies such as
337 targeted fertilizer application and wetland restoration may more effectively capture and reduce phosphorus mobilization from
338 agricultural lands. The newly developed riverine P datasets in this study extend utility to diverse applications, encompassing,
339 but not limited to, the evaluation of watershed models, identification of critical source areas, and optimization of agricultural
340 management strategies. Future studies may concentrate on filling gaps in the spatial and temporal coverage of the datasets
341 (particularly for PO_4^{3-}).

342 **Author contributions**

343 **Y.W.:** Conceptualization, methodology, investigation, formal analysis, data curation, visualization, and writing—original
344 draft. **X.Z.:** Conceptualization, methodology, investigation, writing—original draft, writing—review and editing, supervision,
345 project administration, and funding acquisition. **K.Z.:** methodology, investigation, writing—review and editing. **R.D.S.:**
346 investigation, writing—review and editing. **Y.M.:** visualization, writing—review and editing. **C.M.C.:** writing—review and
347 editing.

348 **Competing interests**

349 At least one of the (co-)authors is a member of the editorial board of *Earth System Science Data*.

350 **Disclaimer**

351 The views expressed are those of the authors, and do not necessarily represent the views or policies of the U.S. Environmental
352 Protection Agency or any other Federal agency. Any use of trade, firm, or product names is for descriptive purposes only and
353 does not imply endorsement by the U.S. Government.

354 **Financial support**

355 This research is in part supported by the U.S. Department of Agriculture – Agricultural Research Service and the National
356 Aeronautics and Space Administration (NASA 22-CMS22-0027). The mention of trade names or commercial products in this
357 publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by
358 the funding agencies.

359 **Acknowledgement**

360 We thank Sarah Stackpoole, Qian Zhang, and Kate Schofield for constructive comments and editorial suggestions on earlier
361 versions of the manuscript.

362 **References**

- 363 Arheimer, B. and Lidén, R.: Nitrogen and phosphorus concentrations from agricultural catchments—influence of spatial and
364 temporal variables, *J Hydrol (Amst)*, 227, 140–159, [https://doi.org/10.1016/S0022-1694\(99\)00177-8](https://doi.org/10.1016/S0022-1694(99)00177-8), 2000.
- 365 Breiman, L.: Random Forests, *Mach. Learn.*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- 366 Brownlie, W. J., Sutton, M. A., Heal, K. V., Reay, D. S., and Spears, B.: Our phosphorus future: towards global phosphorus
367 sustainability, 2022.
- 368 Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H.: NONPOINT POLLUTION
369 OF SURFACE WATERS WITH PHOSPHORUS AND NITROGEN, *Ecological Applications*, 8, 559–568,
370 <https://doi.org/10.1890/1051-0761>, 1998.
- 371 Diaz, R. J. and Rosenberg, R.: Spreading Dead Zones and Consequences for Marine Ecosystems, *Science* (1979), 321, 926–
372 929, <https://doi.org/10.1126/science.1156401>, 2008.

373 Dodds, W. K.: THE ROLE OF PERIPHYTON IN PHOSPHORUS RETENTION IN SHALLOW FRESHWATER AQUATIC
374 SYSTEMS, *J Phycol*, 39, 840–849, <https://doi.org/10.1046/j.1529-8817.2003.02081.x>, 2003.

375 Falcone, J. A. and Survey, U. S. G.: GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow, Reston, VA,
376 <https://doi.org/10.3133/70046617>, 2011.

377 Hamed, K. H. and Rao, A. R.: A modified Mann-Kendall trend test for autocorrelated data, *J. Hydrol. (Amst.)*, 204, 182–196,
378 1998.

379 Hirsch, R. M., Moyer, D. L., and Archfield, S. A.: Weighted Regressions on Time, Discharge, and Season (WRTDS), with an
380 Application to Chesapeake Bay River Inputs ¹, *JAWRA Journal of the American Water Resources Association*, 46, 857–880,
381 <https://doi.org/10.1111/j.1752-1688.2010.00482.x>, 2010.

382 Homer, C. G., Fry, J. A., and Barnes, C. A.: The national land cover database, 2012.

383 Houser, J. N. and Richardson, W. B.: Nitrogen and phosphorus in the Upper Mississippi River: Transport, processing, and
384 effects on the river ecosystem, *Hydrobiologia*, 640, 71–88, <https://doi.org/10.1007/S10750-009-0067-4/TABLES/2>, 2010.

385 Jarvie, H. P., Sharpley, A. N., Scott, J. T., Haggard, B. E., Bowes, M. J., and Massey, L. B.: Within-River Phosphorus Retention:
386 Accounting for a Missing Piece in the Watershed Phosphorus Puzzle, *Environ. Sci. Technol.*, 46, 13284–13292,
387 <https://doi.org/10.1021/es303562y>, 2012.

388 Maavara, T., Parsons, C. T., Ridenour, C., Stojanovic, S., Dürr, H. H., Powley, H. R., and Van Cappellen, P.: Global
389 phosphorus retention by river damming, *Proceedings of the National Academy of Sciences*, 112, 15603–15608,
390 <https://doi.org/10.1073/pnas.1511797112>, 2015.

391 Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A.,
392 Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson,
393 E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., and McKenzie, J.
394 E.: PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews, *BMJ*, 372,
395 <https://doi.org/10.1136/BMJ.N160>, 2021.

396 Qiu, H., Zhang, X., Yang, A., Wickland, K. P., Stets, E. G., and Chen, M.: Watershed carbon yield derived from gauge
397 observations and river network connectivity in the United States, *Scientific Data* 2023 10:1, 10, 1–13,
398 <https://doi.org/10.1038/s41597-023-02162-7>, 2023.

399 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
400 quality data for national-scale aquatic research: The Water Quality Portal, *Water Resour Res*, 53, 1735–1745,
401 <https://doi.org/10.1002/2016WR019993>, 2017.

402 Ringeval, B., Demay, J., Goll, D. S., He, X., Wang, Y. P., Hou, E., Matej, S., Erb, K. H., Wang, R., Augusto, L., Lun, F.,
403 Nesme, T., Borrelli, P., Helfenstein, J., McDowell, R. W., Pletnyakov, P., and Pellerin, S.: A global dataset on phosphorus in
404 agricultural soils, *Scientific Data* 2024 11:1, 11, 1–34, <https://doi.org/10.1038/s41597-023-02751-6>, 2024.

405 Robertson, D. M. and Saad, D. A.: Spatially referenced models of streamflow and nitrogen, phosphorus, and suspended-
406 sediment loads in streams of the midwestern United States, <https://doi.org/10.3133/sir20195114>, 2019.

407 Royer, T. V., David, M. B., and Gentry, L. E.: Timing of Riverine Export of Nitrate and Phosphorus from Agricultural
408 Watersheds in Illinois: Implications for Reducing Nutrient Loading to the Mississippi River, *Environ. Sci. Technol.*, 40, 4126–
409 4131, <https://doi.org/10.1021/es052573n>, 2006.

410 Runkel, R. L., Crawford, C. G., and Cohn, T. A.: Load Estimator (LOADEST): A FORTRAN program for estimating
411 constituent loads in streams and rivers, <https://doi.org/10.3133/tm4A5>, 2004.

412 Sabo, R. D., Clark, C. M., Gibbs, D. A., Metson, G. S., Todd, M. J., LeDuc, S. D., Greiner, D., Fry, M. M., Polinsky, R., Yang,
413 Q., Tian, H., and Compton, J. E.: Phosphorus Inventory for the Conterminous United States (2002–2012), *J Geophys Res*
414 *Biogeosci.*, 126, <https://doi.org/10.1029/2020JG005684>, 2021.

415 Sabo, R. D., Pickard, B., Lin, J., Washington, B., Clark, C. M., Compton, J. E., Pennino, M., Bierwagen, B., LeDuc, S. D.,
416 Carleton, J. N., and others: Comparing drivers of spatial variability in US lake and stream phosphorus concentrations, *J.*
417 *Geophys. Res. Biogeosci.*, 128, e2022JG007227, 2023.

418 Singh, N. K., Van Meter, K. J., and Basu, N. B.: Widespread increases in soluble phosphorus concentrations in streams across
419 the transboundary Great Lakes Basin, *Nature Geoscience* 2023 16:10, 16, 893–900, [https://doi.org/10.1038/s41561-023-](https://doi.org/10.1038/s41561-023-01257-5)
420 [01257-5](https://doi.org/10.1038/s41561-023-01257-5), 2023.

421 Skinner, K. D. and Maupin, M. A.: Point-source nutrient loads to streams of the conterminous United States, 2012, Data Series,
422 <https://doi.org/10.3133/DS1101>, 2019.

423 Stackpoole, S. M., Stets, E. G., and Sprague, L. A.: Variable impacts of contemporary versus legacy agricultural phosphorus
424 on US river water quality, *Proc Natl Acad Sci U S A*, 116, 20562–20567,
425 https://doi.org/10.1073/PNAS.1903226116/SUPPL_FILE/PNAS.1903226116.SAPP.PDF, 2019.

426 Understanding phosphorus: global challenges and solutions: [https://www.unep.org/news-and-stories/story/what-phosphorus-](https://www.unep.org/news-and-stories/story/what-phosphorus-and-why-are-concerns-mounting-about-its-environmental-impact)
427 [and-why-are-concerns-mounting-about-its-environmental-impact](https://www.unep.org/news-and-stories/story/what-phosphorus-and-why-are-concerns-mounting-about-its-environmental-impact), last access: 26 February 2025.

428 Wang, F., Li, S., Yan, W., Yu, Q., Tian, S., Yan, J., Zhou, D., and Shao, Y.: Dependence of riverine total phosphorus retention
429 and fluxes on hydrology and river size at river network scale, *J. Hydrol. (Amst)*, 652, 132676,
430 <https://doi.org/10.1016/j.jhydrol.2025.132676>, 2025a.

431 Wang, Y., Zhang, X., and Zhao, K.: A dataset of riverine nitrogen yield across watersheds in the Conterminous United States,
432 *Scientific Data* 2024 11:1, 11, 1–8, <https://doi.org/10.1038/s41597-024-03552-1>, 2024a.

433 Wang, Y., Zhang, X., Zhao, K., and Singh, D.: Streamflow in the United States: Characteristics, trends, regime shifts, and
434 extremes, *Sci Data*, 11, 788, <https://doi.org/10.1038/s41597-024-03618-0>, 2024b.

435 Wang, Y., Zhang, X., Zhao, K., Sabo, R. D., Miao, Y., and Clark, C. M.: Riverine Phosphorus Gain and Loss across the
436 conterminous United States, <https://doi.org/10.6084/m9.figshare.28509317>, 2025b.

437 Wilson, D. J.: The harmonic mean p -value for combining dependent tests, *Proceedings of the National Academy of Sciences*,
438 116, 1195–1200, <https://doi.org/10.1073/pnas.1814092116>, 2019.

439 Withers, P. J. A. and Jarvie, H. P.: Delivery and cycling of phosphorus in rivers: A review, *Science of The Total Environment*,
440 400, 379–395, <https://doi.org/10.1016/j.scitotenv.2008.08.002>, 2008.

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

443 Xia, Y., Zhang, M., Tsang, D. C. W., Geng, N., Lu, D., Zhu, L., Igalavithana, A. D., Dissanayake, P. D., Rinklebe, J., Yang,
444 X., and Ok, Y. S.: Recent advances in control technologies for non-point source pollution with nitrogen and phosphorous from
445 agricultural runoff: current practices and future prospects, *Appl. Biol. Chem.*, 63, 8, [https://doi.org/10.1186/s13765-020-0493-](https://doi.org/10.1186/s13765-020-0493-6)
446 6, 2020.

447 Zhang, Q. and Hirsch, R. M.: River Water-Quality Concentration and Flux Estimation Can be Improved by Accounting for
448 Serial Correlation Through an Autoregressive Model, *Water Resour. Res.*, 55, 9705–9723,
449 <https://doi.org/10.1029/2019WR025338>, 2019.

450 Zhang, W., Jin, X., Liu, D., Lang, C., and Shan, B.: Temporal and spatial variation of nitrogen and phosphorus and
451 eutrophication assessment for a typical arid river — Fuyang River in northern China, *Journal of Environmental Sciences*, 55,
452 41–48, <https://doi.org/10.1016/J.JES.2016.07.004>, 2017.

453