



Predicting salinity and alkalinity fluxes of U.S. freshwater in a changing climate: Integrating anthropogenic and natural influences using data-driven models

Beibei E^{a,1}, Shuang Zhang^b, Elizabeth Carter^{a,c}, Tasmeem Jahan Meem^c, Tao Wen^{a,*}

^a Department of Earth and Environmental Sciences, Syracuse University, NY, 13244, United States

^b Department of Oceanography, Texas A&M University, College Station, TX, 77843, United States

^c Department of Civil and Environmental Engineering, Syracuse University, NY, 13244, United States

ARTICLE INFO

Editorial handling by Dr. Zimeng Wang

Keywords:

Salinization
Alkalinization
Climate change
Rock weathering
Machine learning

ABSTRACT

Climate change is an ongoing and intensifying threat. Previous studies indicate that U.S. rivers are undergoing salinization and alkalinization driven by both natural (e.g., temperature and precipitation) and anthropogenic (e.g., population density) factors. In this study, random forest models were developed to predict how the salinity (i.e., sodium) and alkalinity fluxes from 226 U.S. rivers will vary with changing population density and climatic forcings (i.e., temperature and precipitation) from 2040 to 2100 for three socioeconomic development pathways. The models predicted a lower future sodium flux in the northern U.S., likely due to reduced winter road salting under projected warmer winter. In the southern and western U.S., where road salting is uncommon, the models predicted little or no change in future sodium flux, however, a projected warmer and drier climate might exacerbate soil salinization in these regions. The models also indicated that carbonate weathering rates are inhibited when temperatures exceed 10 °C, leading to a lower future alkalinity flux in carbonate-rich watersheds at high temperatures. In siliciclastic-dominated or organic carbon-rich watersheds, rising temperatures are associated with increased riverine alkalinity flux, likely through the acceleration of silicate weathering and decomposition of soil organic carbon. Higher precipitation and enhanced transport capacity were generally deemed to contribute to higher solute fluxes before reaching a plateau. This study underscores the urgency for policymakers and scientists to adapt strategies for managing rivers, focusing on mitigating the impacts of river salinization and shifts in riverine alkalinity driven by global warming.

1. Introduction

Over the past decades, global and U.S. river has been experiencing salinization and alkalinization due to anthropogenic salt input (e.g., road salt, mining waste), human-accelerated mineral weathering (Kaushal et al. 2017, 2018; Haq et al., 2018) and associated feedbacks. In particular, Kaushal et al. (2018, 2017) and Haq et al. (2018) suggested that river salinization enhanced the ion exchange on the surface of building concrete and rock, leading to more alkalinity released into river within human-dominated watersheds, which can threaten drinking water supplies, impair river biodiversity, accelerate corrosion of infrastructure, and mobilize inorganic and organic contaminants (DeVilbiss et al., 2021; Duan and Kaushal, 2015; Kaushal, 2016; Hintz and Relyea,

2019). For example, elevated salinity level harms sensitive aquatic species, e.g., fishes and amphibians by reducing their populations and disrupting food webs. In addition, high salinity can degrade the quality of drinking water sources, potentially increasing costs and energy use for desalination and water treatment. Utilizing the same datasets as Kaushal et al. (2018), E et al. (2023) built a machine learning model – Random Forest (RF) to predict historical salinity and alkalinity fluxes in 226 U.S. watersheds from 1942 to 2021 using an array of watershed properties (i.e., hydrology, climate, geomorphology, soil chemistry, geology, land use and cover), which indicated that anthropogenic activities (e.g., urbanization) and natural processes (e.g., runoff, bedrock geology) were the primary drivers explaining the spatial variability in salinity and alkalinity fluxes across U.S. rivers over the past 80 years, respectively.

* Corresponding author.

E-mail address: twen08@syr.edu (T. Wen).

¹ Current address: Department of Oceanography, Texas A[amp]M University, College station, TX, 770843, United States.

Mediated by rock weathering, alkalinity in terrestrial water systems dominates net carbon fluxes from land to ocean and, as such, has important implications for climate change: terrestrial riverine alkalinity acts as the primary natural mechanism regulating carbon dioxide levels in the atmosphere (Meybeck, 1987; Berner, 2004).

Human activities, e.g., fossil fuel combustion and urbanization, are predicted to drive up the average global temperature at the end of 21st century by 1.5–2.7 °C above the pre-industrial level (Lenton et al., 2023). It is also anticipated that regional and global climate patterns will be largely reshaped, such as the snowfall line moving to higher latitudes (Quante et al., 2021) and shifts in global precipitation patterns and regional hydroclimate (Dore, 2005). Changing climate patterns will both directly and indirectly impact river salinity/alkalinity. For example, northward migration of snowfall will likely lead to less use of road salt in mid-latitude regions, and may therefore mitigate the trend of river salinization (Stirpe et al., 2017), while extremely high temperatures and frequent droughts due to less precipitation might result in severe soil salinization and alkalinization in certain regions (Perri et al., 2020; Perri, 2022; Stirpe et al., 2017). In the regions with increased precipitation, higher erosion and river transport capability might deliver more alkalinity to rivers and expose more fresh rock to erosion and weathering (Penman et al., 2020; Isson et al., 2020; Berner, 2004). Population density is also projected to vary across space and time throughout the 21st century (Wang et al., 2022b; Chen and Mueller, 2018). Changing population density will not only alter the direct salt input to rivers through urbanization (E et al., 2023), but also impact river water chemistry through land use changes and agricultural activities, such as fertilizer usage and soil plowing (Hansen et al., 2018; Thorslund et al., 2021).

Future projections of climatic forcing (i.e., temperature and precipitation) and population can vary significantly depending on assumptions about how human society will develop. Duan and Kaushal (2015) emphasized a warming climate will increase carbon and nutrient fluxes from sediments in streams across various land uses. Lehmann et al. (2023) revealed that weathering flux to the ocean could significantly increase under new climate patterns as early as 2100, by up to 68%, depending on the environmental conditions. Welsch et al. (2006) suggested that soil CO₂ concentrations have the potential to increase as climate gets warmer and wetter leading to enhanced soil weathering. Previous studies highlight the critical need to further understand of how climate change will impact alkalinity variation in the future under different Shared Socioeconomic Pathways (SSPs). SSPs are widely used in climate change modeling [e.g., Coupled Model Intercomparison Project Phase 6 (CMIP6)] to describe potential pathways of future socioeconomic development (O'Neill et al., 2010; Riahi et al., 2017). Each SSP is defined based on socioeconomic development indices such as demographics, economic development, and urbanization as provided by the International Institute for Applied Systems Analysis (Kc and Lutz, 2017; Riahi et al., 2017). In particular, SSP126, SSP370, and SSP585 represent development pathways of sustainability, regional rivalry, and fossil-fueled development, respectively (Riahi et al., 2017). Population density for SSP126, SSP370, and SSP585 will show a slow increase, a rapid decrease (due to increased regional conflict and/or war), and a rapid increase, respectively (Kc and Lutz, 2017). Therefore, different development scenarios reflect varying levels of human activity and greenhouse gas emissions, which modulate future climate change, subsequently impacting future salinity and alkalinity fluxes in U.S. rivers. Given the critical role of river salinity and alkalinity in deciphering the global carbon cycle and managing watershed ecosystem services, it is essential to accurately predict how salinity and alkalinity will vary with changing climate and population density in the future.

In this study, we used machine learning models to predict future salinity and alkalinity fluxes from U.S. rivers from 2040 to 2100 under different projected socioeconomic pathways. This study builds on the published preliminary assessment by E et al. (2023), which examined the relationship between watershed properties and salinity and

alkalinity fluxes across 226 U.S. watersheds. We compiled sodium and alkalinity concentrations from 226 U.S. Geological Survey (USGS) river monitoring sites across the U.S. (Kaushal et al., 2018; E et al., 2023). Sodium, rather than chloride, was selected as the salinity proxy to maintain consistency with previous studies. Various watershed properties including climate, geomorphology, geology, soil chemistry, land use, and land cover were compiled/calculated for each watershed and used as predictor features in the machine learning model. The objective of this analysis is to gain further understanding of how human activities and climatic forcings interact to regulate future river salinization and alkalinization in U.S. rivers. Findings from this study can help guide the design of mitigation and adaptation policies to address the challenges posed by inevitable climate change and global warming.

2. Materials and methods

2.1. Random forest model and SHapley Additive exPlanation (SHAP) method

In this study, we built Random Forest (RF) models to predict sodium and alkalinity fluxes from U.S. rivers using watershed properties. The RF model, a popular machine learning technique introduced in the 1990s (Tin, 1995), offers several advantages and strong performance in addressing complex research questions. These advantages include its capability to handle non-linear relationships between variables, provide insight into variable importance, avoid overfitting, and require minimal dataset preprocessing (e.g., Tyralis et al., 2019; Mimeau et al., 2024; Pazola et al., 2024). Detailed descriptions of the RF model can be found in the literature (e.g., E et al., 2023). An RF regressor makes predictions by combining multiple decision trees. Each decision tree makes an independent prediction on the target variable (e.g., sodium or alkalinity flux) based on a random subset of input data. In each tree, a random set of predictor variables (i.e., watershed properties) is selected at each split to divide the training data (Breiman, 2001). The RF model must be carefully trained and optimized to ensure robust predictions on new data (i.e., unseen dataset). Here, the RF model was trained using 10-fold cross-validation, during which three hyperparameters were fine-tuned, i.e., num.trees (i.e., number of sub-models or trees), min.node.size (i.e., minimal size of the tree branch in each sub-model), and mtry (i.e., number of predictor variables selected at each split). The optimal hyperparameters (Table S1) were determined by minimizing the mean squared error (MSE) in the training phase with respect to num.trees (100–1500), min.node.size (1–15), and mtry (1–15). Model performance was assessed by comparing prediction values with the ground-truth values in hold-out datasets. The development of the RF model was implemented in R using the “mlr” package (Bischl et al., 2016).

Partial Dependence Plots (PDP) were used to visualize the relationship between sodium or alkalinity flux and each important predictor variable (Le et al., 2019). In addition, SHapley Additive exPlanations (SHAP) values (Lundberg and Lee, 2017) were calculated for each predictor feature using R packages, including "fastshap", "mlr", and "ranger". SHAP values, based on game theory, assign an importance value to each feature in a model, allowing the assessment of how individual features and their interactions impact the prediction of the target variable. SHAP Dependence Plots (SDPs) (Lundberg et al., 2020) were generated using the R package "SHAPforxgboost" to visualize SHAP values.

2.2. Datasets

Except for future climate forcings and population density, all other datasets discussed in this study are from E et al. (2023), which compiled salinity and alkalinity concentrations from the Water Quality Portal (Read et al., 2017) for a total of 226 USGS river monitoring sites across the U.S.: Northeast (n = 43), Southeast (n = 31), Midwest (n = 62), Northwest (n = 21), Southwest (n = 65), and Pacific (n = 4) (Fig. 1).

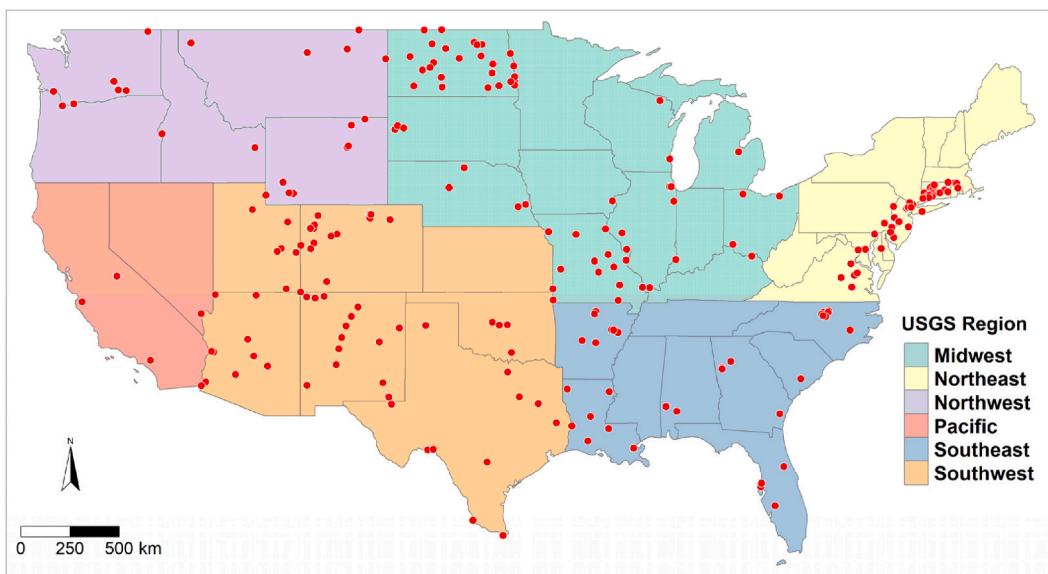


Fig. 1. Location of 226 USGS river monitoring sites in this study.

These sites were selected because they reported at least 30 years of continuous river chemistry data between 1942 and 2021. Before feeding the data in the RF model, E et al. (2023) aggregated daily sodium and alkalinity measurements by year and month before calculating the annual-averaged monthly water chemistry data, i.e., each site had at most twelve data points. After aggregation, a total of 2685 sodium and 2691 alkalinity measurements were retained. These calculated monthly-aggregated concentrations (in unit of mg/L) were then converted to monthly flux in unit of mg per km² per month following Equation (1).

$$F = \frac{IC \times D \times 28.32 \frac{\text{liter}}{\text{ft}^3} \times 3600 \frac{\text{second}}{\text{hour}} \times 24 \frac{\text{hour}}{\text{day}} \times 30 \frac{\text{day}}{\text{month}}}{WA} \quad (1)$$

where F represents sodium or alkalinity flux (mg/km²-mo), IC is the monthly chemical concentration (mg/L), D denotes river discharge (ft³/sec) while WA is the watershed area (km²). For alkalinity flux, it is expressed as mg CaCO₃/km²-mo. The constants of 28.32, 3600, 24, and 30 are unit conversion factors. For each site, E et al. (2023) delineated their corresponding watersheds. Then the average value of each of the 32 watershed properties (Table S2) was calculated for each watershed based on publicly accessible datasets. E et al. (2023) proposed a workflow to select a subset of 18 features by removing redundant features and retaining relevant features based on statistical analysis and domain knowledge, among which three predictor features (runoff, soil moisture, and impervious surface area) were not included due to the lack of high-resolution (1 km) future data. The omission of these three features did not substantially reduce RF model performance in predicting sodium and alkalinity fluxes (refer to section 3.1 for details). The 15 remaining features include climate ($n = 2$, Temperature and Precipitation), geomorphology ($n = 2$, Erosion Rate and Elevation), soil chemistry ($n = 2$, Soil Organic Carbon and Soil pH), geology ($n = 5$, Carbonate Sediment, Siliciclastic Sediment, Unconsolidated Sediment, Igneous Basic, and Metamorphic), land use ($n = 2$, Population Density and Cultivated Vegetation), and land cover ($n = 2$, Trees and Flooded Vegetation) (Table S2, marked by asterisks). For future (2040–2100) alkalinity/salinity projections, high-resolution future precipitation and air temperature were based on the GFDL-ESM4 model from the CHELSA data repository (Karger et al., 2017) for three SSPs: SSP126 (Sustainability), SSP370 (Regional Rivalry), and SSP585 (Fossil-fueled Development) (Boke-Olén et al., 2017) (Table S3). Global population density data (1 km resolution; spanning from 2020 to 2100) for all three SSPs

were derived from Wang et al. (2022). Future values of the other twelve predictor features were assumed to remain constant, and watershed-scale local mean historical values were used.

Based on the compiled historic datasets of watershed properties and water chemistry, we trained and validated RF models before applying them to predict future sodium and alkalinity fluxes using future population density, temperature, and precipitation plus the other 12 predictor variables under three SSP scenarios (SSP126, SSP370, and SSP585) for two time periods: 2040 to 2070 and 2070 to 2100. Comparisons of population density, temperature, and precipitation across the three SSP scenarios and the historical average are presented in Figs. S1, S2, S3, and S4.

3. Results and discussion

3.1. Performance evaluation of sodium and alkalinity fluxes prediction models

We used 15 features to build RF models to predict future salinity and alkalinity fluxes in U.S. rivers. The optimized hyperparameters for both models are listed in Table S1. Calculated mean squared errors (MSE) are plotted as a function of hyperparameters in Fig. S5. Sodium and alkalinity prediction models explained 81% and 78% variation in the target variable in the hold-out dataset, respectively (Figs. S6a and S7a). Calculated residual values showed no statistically significant correlation with predicted flux, suggesting strong generalizability in the trained models (Figs. S6b and S7b). In addition, we investigated spatiotemporal biases in the models predictions by plotting model residuals over space and time. Neither model showed spatiotemporal clustering of residuals (Figs. S8a and S9a). Furthermore, no statistically significant correlation was found between the percentage of monthly data points with >30% relative errors and the total number of monthly measurements in each state (Figs. S8b and S9b). These results indicate an absence of spatial bias in the trained RF models (Carter et al., 2023). As to temporal aspect, no specific months exhibited significantly higher erroneous predictions for riverine sodium/alkalinity flux, especially when a higher relative error threshold of 100% was used to evaluate model performance. To sum up, the RF models with 15 features show strong performance and generalizability in predicting future salinity and alkalinity fluxes in U.S. rivers.

In the model development phase, the averaged absolute SHAP values were calculated for each predictor feature. Predictors were ranked based

on their SHAP values illustrating the importance of each feature in flux prediction (Fig. 2). For the sodium (salinity) flux model, population density emerged as the most important feature. For the alkalinity flux model, bedrock geology, temperature, and precipitation were among the most important features, suggesting that alkalinity flux is primarily governed by natural processes (Fig. 2). These findings align with the previous study (E et al., 2023), which indicated that salinity and alkalinity fluxes are mainly governed by anthropogenic and natural processes, respectively. Furthermore, air temperature, a key climatic variable, emerged as the second most influential factor in both RF models (Fig. 2). This underscores the substantial impact that climate change may have on riverine sodium and alkalinity fluxes. In the alkalinity prediction model, the percentage of watershed underlain by carbonate sediment (rather than siliciclastic sediment) was ranked as the most important feature. This result is likely due to the kinetic differences in mineral weathering rates. In particular, the dissolution of calcite minerals of 1 mm of diameter typically takes 10^{-1} years, while silicate mineral dissolution can take 10^2 years (Lasaga, 1984). Carbonate weathering rates are highly sensitive to environmental conditions, including soil moisture, temperature, vegetation, and soil respiration (i.e., soil pCO_2) (Gaillardet et al., 2019).

Partial Dependence Plot (PDP) and SHAP Dependence Plot (SDP) were used to assess how analyte flux varied with key predictor features, focusing on population density, temperature, and precipitation

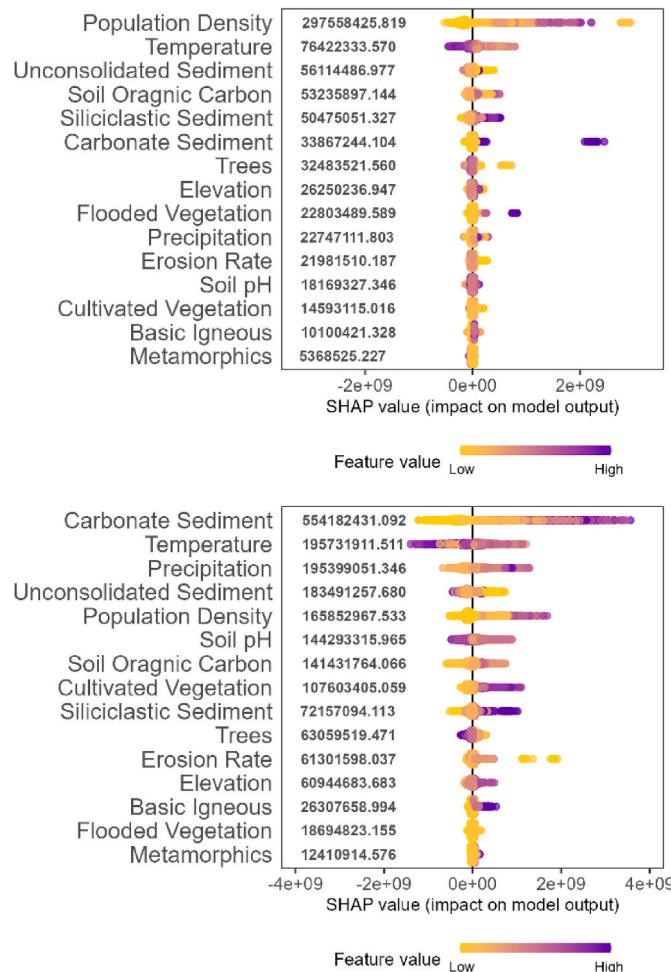


Fig. 2. Predictor feature ranking based on SHAP values from the riverine alkalinity flux prediction model based on 15 features (excluding runoff, soil moisture, and impervious surface percentage) for (a) sodium flux and (b) alkalinity flux. A higher SHAP value indicates the corresponding predictor has a higher impact on prediction of the flux of interest.

(Figs. S10 and 3). The PDP shows that increasing population density significantly augment riverine sodium flux (Fig. S10). This relationship highlights the contribution of anthropogenic activities, such as urban runoff. Furthermore, the PDP reveals that sodium flux is substantially higher at lower temperatures (below 0 °C) compared to higher temperatures (above 10 °C), with the peak sodium flux occurring around 2 °C (Fig. S10c). The SDP supports this finding, showing a positive correlation between sodium flux and temperatures below 8 °C (Fig. S10d). These patterns suggest that road salts applied during winter (i.e., lower temperature) represent a significant driver of riverine salinity. However, it is important to note that these PDPs and SDPs may not fully capture the contributions of other human activities, such as effluent discharge, on riverine sodium flux. Such activities could influence sodium levels independently of temperature. Finally, precipitation generally positively correlates with sodium flux (Fig. S10). Higher precipitation leads to greater runoff, which enhances solute transport from land surfaces to rivers.

SDP and PDP plots show that alkalinity flux is suppressed at colder temperatures (<0 °C) or much warmer temperatures (>10 °C), with the optimum temperature being 4 °C–10 °C (Fig. 3). These observations are consistent with previous studies. In particular, Lehmann et al. (2023) investigated global riverine alkalinity across six continents, from 44°S to 51°N, and suggested that mean annual temperature acts as a first-order control on riverine alkalinity concentration. They reported that 5 °C–15 °C correspond to the most extensive carbonate weathering due to abundant water supply and soil acidity. Romero-Mujalli et al. (2019) proposed that the dependence of the dissolution rate of calcium carbonate on land surface temperature can be described by a Gaussian function, i.e., a “boomerang-shaped” curve, with peak alkalinity concentrations at approximately 11 °C. Furthermore, Gaillardet et al. (2019) also observed a boomerang-shaped relationship between carbonate weathering and temperature, with the optimum temperature being 10 °C–15 °C. Higher temperature will reduce the solubility of carbonate and cause more CO_2 degassing in carbonate systems. Alkalinity flux increases with precipitation before reaching a plateau at about 200 mm/month (Fig. 3). Note that most USGS sites used in this study report precipitation values of < 200 mm/month (Fig. 3d). The precipitation SDP shows that river alkalinity flux is suppressed when the precipitation is < 60 mm/month, above which alkalinity generation is promoted as precipitation increases. We hypothesize that this threshold of 60 mm/month reflects the average minimum precipitation required to generate effective runoff to deliver weathering products to rivers. It is important to note that alkalinity can still be generated and transported to the river through other mechanisms when precipitation is below 60 mm/month. Here, we treat precipitation as a proxy of runoff. Runoff is controlled by multiple factors including precipitation, temperature, soil type, topography, land use, and human activities (Ahmed et al., 2022; Grigorev et al., 2022; Hansen et al., 2018; Jiang et al., 2021; Wang et al., 2022a; Zhou et al., 2023) with precipitation being the primary driver (Hewlett, 2009, McDonnell et al., 1967). Similar to alkalinity flux, salinity flux also exhibits a “boomerang-shaped” relationship with temperature (Fig. S10), with an optimal temperature (maximum salinity flux) of approximately 8 °C. Multiple watershed processes could contribute to this relationship. Among them, we hypothesize that as temperature increases, salinity flux initially rises due to enhanced weathering processes, but beyond a certain point, the contribution of anthropogenic salt inputs (e.g., road salting) may decline as temperatures continue to increase. Further research is needed to better understand these “boomerang-shaped” relationships.

3.2. Human and climate impacts on future riverine sodium flux

Trained RF models were used to predict future sodium and alkalinity fluxes using future projections of the 15 predictor features, among which three features (air temperature, precipitation, and population density) were varying with time or SSP, while the other 12 features were held

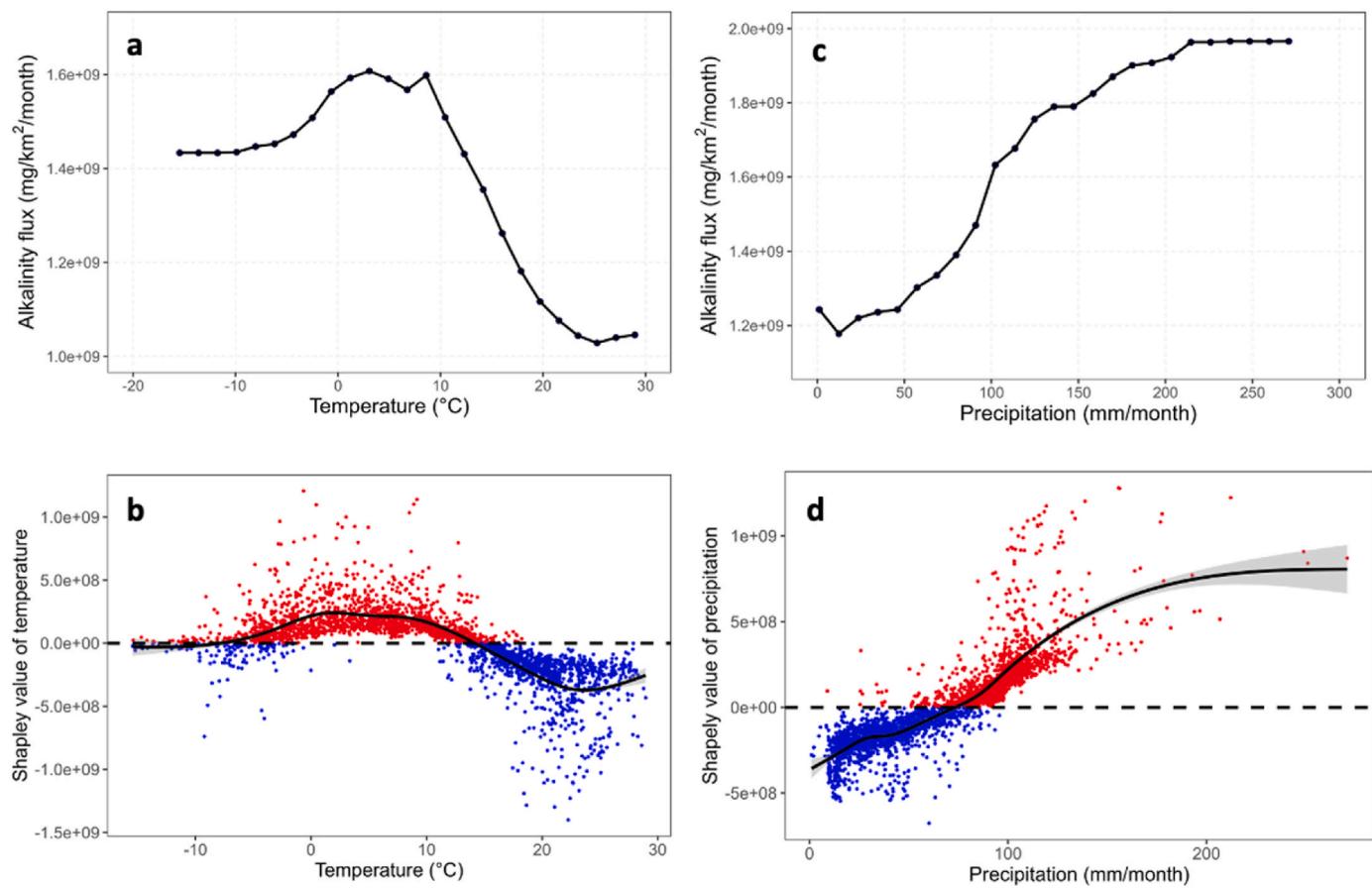


Fig. 3. Partial Dependence Plot (a and c) and SHAP Dependence Plot (b and d) derived from the 15 features based random forest model that predict alkalinity flux. In SDP figures (b and d), SHAP values >0 are colored in red, representing a positive contribution to riverine alkalinity fluxes from the corresponding predictor feature. Samples in blue denote negative SHAP values, for which predictor feature makes a negative contribution to riverine alkalinity fluxes. Both PDP and SDP visualize the relationship between alkalinity flux and predictor features including temperature (a and d) and precipitation (c and d). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

constant at their historic values. A detailed description of projected future population density, temperature, and precipitation is available in the supplementary information. To ensure accuracy, four USGS sites (07374525 and 07373420 in Louisiana, and 06934500 and 06818000 in Missouri) were excluded when assessing the spatial variability of predictor features and predicted results. These four sites have watersheds that span multiple regions (e.g., northern and southern U.S.), making their predicted fluxes less representative of region-specific situations.

By the end of 2100, the model predicts that climate change will lead to a change in the U.S. riverine salinity fluxes by 3.28% (SSP126), -11.74% (SSP370), and 7.95% (SSP585). An average warming of 1 °C across the U.S. has the potential to alter river salinity fluxes by 1.49% (SSP126), -3.33% (SSP370), and 2.02% (SSP585). Over time, both future sodium flux and population density show an increase from SSP370 to SSP126 then SSP585, suggesting increased anthropogenic sodium inputs into rivers through activities, such as irrigation, fertilizer,

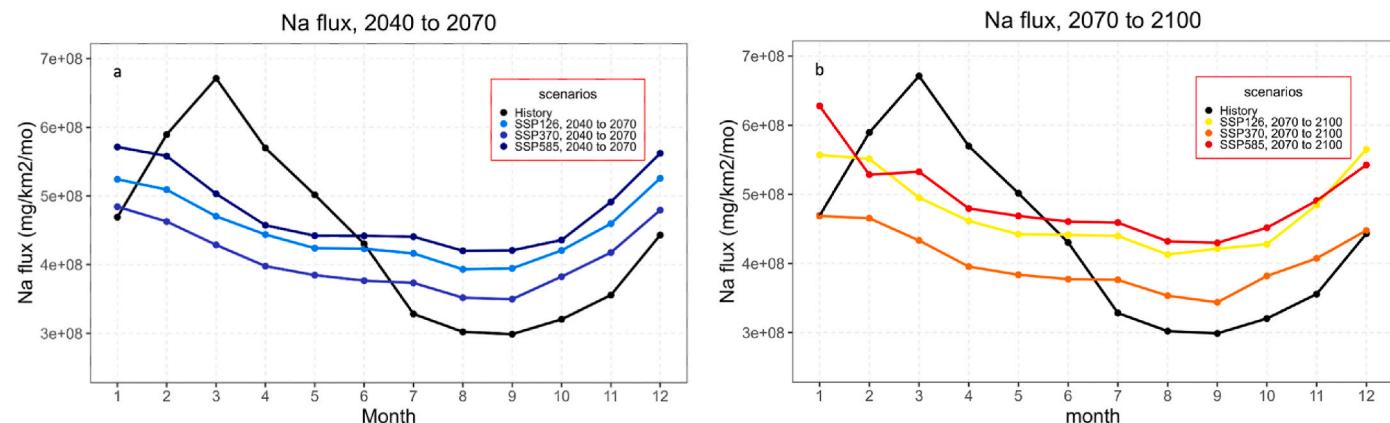


Fig. 4. Predicted future monthly averaged sodium flux for all three scenarios (SSP126, SSP370, and SSP585) and for both periods of time: (a) 2040 to 2070 and (b) 2070 to 2100.

sewage, and industry (Figs. S1a and 4). In the first half of the calendar year, future riverine sodium flux is generally lower than historical levels (Fig. 4). This reduction is attributed to rising future temperatures during winter months (December to February) compared to historical levels (Figs. S1b and 4), which may lead to reduced road salt application. Lower road salt application not only results in lower salinity level in U.S. rivers during winter months but also leads to less storage of road salt in soils, riverbanks, and riverbeds. This might explain the observed smaller declines in future riverine sodium flux following the winter months (i.e., January–June) compared to historical values (Fig. 4). In the second half of the calendar year, both future precipitation and sodium flux are projected to exceed historical levels (Figs. S1 and 4). Since climatological precipitation is typically higher during these months, the increased runoff leads to higher salinity transport capacity to rivers. As such, high precipitation rates (e.g., >60 mm/month) may amplify salinity flux in U.S. rivers during these months.

The change in population density from the present to the future differs by scenario. For most states, population density increases under SSP126 and SSP585 but declines under SSP370 (Fig. S2). As sodium flux is positively correlated with population density, such trends in population density can explain the projected increase in riverine sodium fluxes in some northern states (e.g., Connecticut, Iowa, Indiana, Washington) (Fig. 5). Unlike the northern U.S., the southern and western U.S. show little to no increase in future sodium flux, even under scenarios with projected increases in population density. This is likely due to the limited use of road salts in the southern U.S., where current and future winter temperatures rarely fall below freezing. Furthermore, the relative increase in future temperature in most southern states is much larger than that in future precipitation, which will lead to a projected warmer and drier climate (Figs. S3 and S4). While such projected climate patterns in southern states could exacerbate soil salinization, river salinization might not necessarily worsen significantly due to reduced discharge, which limits the transport of salts from the watershed to the river.

3.3. Human and climate impacts on future riverine alkalinity flux

Over the first half of the year (January–June), future river alkalinity flux peaks in March (SSP370 and SSP585) or in April (SSP126) (Fig. 6), coinciding with increasing monthly average temperature (Fig. S1b). However, projected future alkalinity peak flux is lower than historical values during these six months, with the offset beginning to increase in February and reaching its maximum in May (Fig. 6). This observation likely reflect the temperature control on riverine alkalinity flux. From February to June, monthly temperatures increase under all SSPs, and move further away from the optimum temperature of 4 °C for carbonate weathering (Fig. 3). This seasonal deviation from the optimum temperature for carbonate weathering is less pronounced in historical temperature. Therefore, future increases in temperature might limit carbonate weathering, reducing alkalinity fluxes in U.S. rivers during the first half of the year. In the second half of the year (July–December), under all three SSP scenarios, future temperatures decrease with month, moving towards the optimum temperature of 4 °C (Fig. 6). Combined with increased future precipitation, this creates conditions favorable for carbonate weathering and the transport of alkalinity into rivers. Enhanced precipitation during these months increase the watershed capacity for flushing and transporting alkalinity in the future. The interplay between precipitation and temperature leads to a higher future alkalinity flux in rivers during the second half of the year (Fig. 4).

Our model predicts that the U.S. riverine alkalinity flux will vary by 4.28% (SSP126), -1.64% (SSP370), and 4% (SSP585) by 2100 due to climate change, while an average warming of 1 °C across the U.S. has the potential to change river alkalinity fluxes by 1.94% (SSP126), -0.47% (SSP370), and 1.02% (SSP585). From 2040 to 2100 under all three scenarios, most states reporting a decrease or less than 5% increase in riverine alkalinity flux are located in the Midwest, Southeast, and Southwest U.S. (Fig. 7), particularly along both sides of the Mississippi River. River watersheds in these regions are predominantly underlain by carbonate sediments (Fig. S11), which are highly sensitive to

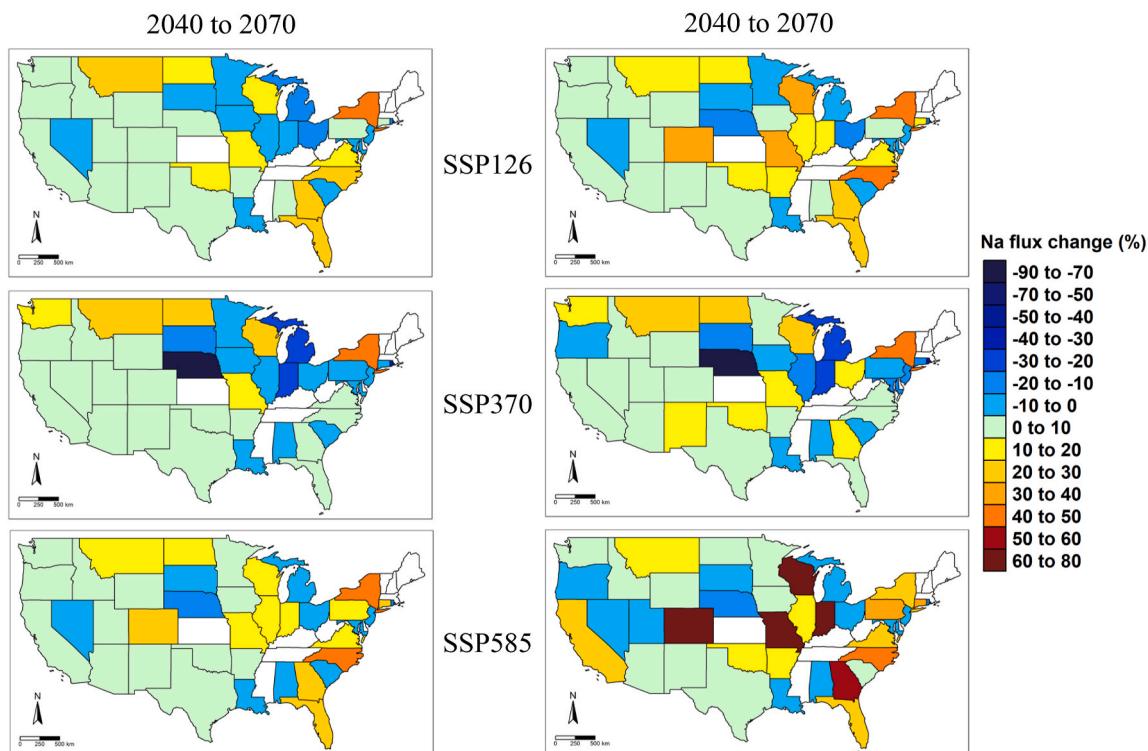


Fig. 5. U.S. states color coded by relative change in monthly averaged riverine sodium flux in each state [(future value – historical value)/historical value] for both periods of time (2040–2070 and 2070 to 2100) and three scenarios (SSP126, SSP370, and SSP585). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

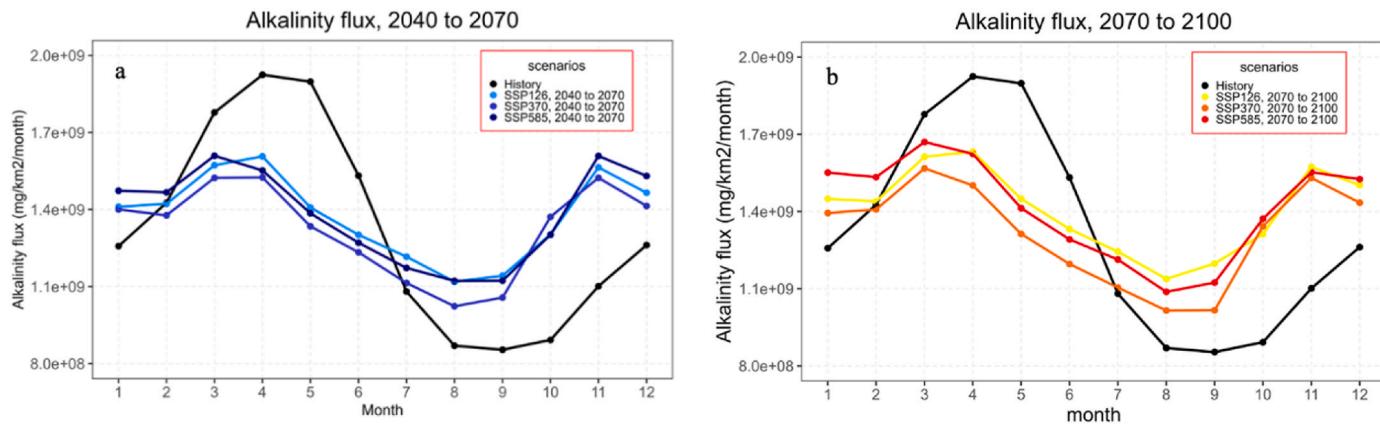


Fig. 6. Predicted future monthly averaged alkalinity flux for all three scenarios (SSP126, SSP370, and SSP585) and for both periods of time: (a) 2040 to 2070 and (b) 2070 to 2100.

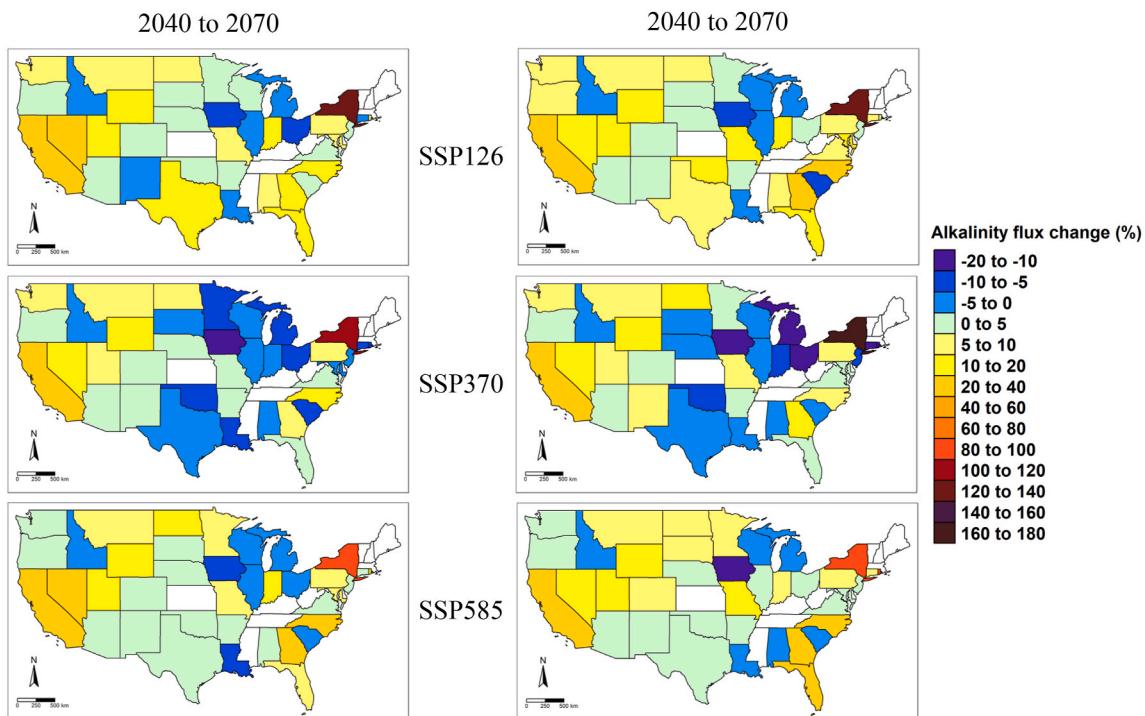


Fig. 7. U.S. states color coded by relative change in monthly averaged riverine alkalinity flux in each state [(future value – historical value)/historical value] for both periods of time (2040–2070 and 2070 to 2100) and three scenarios (SSP126, SSP370, and SSP585). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

temperature increases. In these states, projected future increases in temperature (Fig. 3) often exceed optimal range for carbonate weathering, i.e., 4°C–10 °C. These temperatures might suppress carbonate weathering, lowering the supply of alkalinity into rivers. In contrast, states in the Northeast, Northwest, and Pacific regions are mainly underlain by a combination of siliciclastic and unconsolidated sediments (Fig. S11). Temperature increases enhance weathering of silicates (Banwart et al., 2009; West et al., 2005) and unconsolidated sediments (Berner and Berner, 2012). Therefore, projected increases in alkalinity flux in these regions are consistent with enhanced weathering of siliciclastic and unconsolidated sediments.

In the Northeast U.S., soil organic carbon (SOC) is also relatively enriched compared to other regions (Fig. S11). The decomposition of SOC by microorganisms can lead to the formation of carbonic acid, a weak acid which can have varied effects on soil pH. In soils with low cation supply capacity, carbonic acid lowers soil pH. In soils with higher

cation supply capacity, carbonic acid can react with base cations in soil, producing alkalinity to buffer the soil exchangeable pH (Cotrufo et al., 2015; Dong et al., 2022). If a catchment has a higher base cation supply than the carbonic acid input flux, the accelerated decomposition of SOC associated with rising temperatures could potentially contribute to the increase in alkalinity flux in the northeast region (Fig. 7). However, if carbonic acid supply from the SOC decomposition overpasses the soil buffering capability, soil pH and exchangeable alkalinity can decline.

As presented above, it is critical to predict future salinity and alkalinity fluxes under projected climate patterns and human activities. To address this knowledge gap, this study demonstrates a preliminary application of machine learning models to predict future riverine fluxes and their interactions with natural and human factors. However, it is important to acknowledge several limitations of the dataset and models used in this study. For instance, while population density can infer impervious surface area, it does not fully capture other critical land-use

changes. Incorporating future high-resolution land-use data (e.g., Hou et al., 2022; Zhang et al., 2023) is highly recommended to enhance the performance and interpretability of the RF model. Additionally, the RF model is still considered “gray box” model, as it cannot provide explicit equations to describe the relevant physical processes that dictate the nonlinear relationships between salinity or alkalinity fluxes and climatic and hydrological conditions. Future research should prioritize improving the availability of high-resolution datasets that describe critical predictor variables, such as land use and anthropogenic inputs. These enhanced datasets would allow for more accurate and representative modeling of riverine salinity and alkalinity fluxes. In addition, efforts should focus on developing more explainable machine learning models to better elucidate the nonlinear interactions between predictor variables and riverine fluxes.

4. Conclusion

The objective of this study was to identify important drivers of salinity and alkalinity fluxes in U.S. rivers and use projected changes in these drivers to evaluate how riverine salinity/alkalinity will respond to future climate change. The 15 most physically and/or empirically important variables were used to build two Random Forest (RF) models to predict future (2040–2100) monthly sodium and alkalinity fluxes at 226 river monitoring sites across the U.S. The RF results showed that population density and the percentage of watershed underlain by carbonate sediment were the most important features for predicting riverine sodium and alkalinity flux, respectively. This suggested that variability in riverine salinity is primarily associated with anthropogenic forcings (e.g., road salting), whereas variability in riverine alkalinity is primarily associated with natural forcings. In addition, air temperature ranked second in both models, indicating strong climatic control on both riverine sodium and alkalinity fluxes.

Trained RF models were applied to predict future sodium and alkalinity fluxes in U.S. rivers from 2040 to 2100 under various socioeconomic pathways. Model results showed lower riverine sodium fluxes in rivers in the northern U.S., likely due to higher winter temperatures and associated reductions in road salting. The southern and western U.S. generally showed little to no wintertime increase in sodium flux due to little or no road salting. It is important to note that the relatively small increases in future precipitation, compared to large increases in future temperature in the southern U.S., will lead to a warmer and drier climate pattern, which is projected to exacerbate soil and river salinization in southern states. Our models also suggested carbonate weathering is limited by temperatures above 10 °C, which led to lowered riverine alkalinity flux in carbonate-dominated watersheds under future climate scenarios. This is likely associated with the lower solubility of carbonate and increases in CO₂ degassing rates at higher temperatures. However, in watersheds dominated by siliciclastic and unconsolidated sediments or with high soil organic carbon, rising temperatures could accelerate silicate weathering or the decomposition of organic carbon, resulting in an increase in riverine alkalinity flux.

This work highlights the potential consequences of rising temperatures on river salinization and provides insights into the future challenges posed by climate-induced changes in riverine alkalinity. The prediction results under different SSP scenarios indicate that future riverine salinity and alkalinity fluxes are projected to increase across most U.S. states. However, under the sustainable development scenario (SSP126), the increase in salinity flux is smaller compared to the fossil-fuel development scenario (SSP585). For both salinity and alkalinity fluxes, climatic and hydrologic variables, such as temperature and precipitation, emerge as key predictors. This suggests that river water quality, with respect to salinity and alkalinity, can be managed by mitigating variability in climatic and hydrologic conditions. Overall, this study provides critical insights for policymakers and geoscientists working towards the sustainable management of river ecosystems in the context of ongoing global change.

CRediT authorship contribution statement

Beibei E: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shuang Zhang:** Writing – original draft, Resources, Methodology, Investigation. **Elizabeth Carter:** Writing – review & editing, Resources, Investigation. **Tasmeem Jahan Meem:** Writing – original draft, Resources, Investigation. **Tao Wen:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Open Research

The data and codes discussed in this article are deposited in an online data repository and are publicly and freely available via this DOI: <https://doi.org/10.4211/hs.dcf1135ac217455d90f62c709ceabc6>.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. OAC-2209864 to TW.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apgeochem.2025.106285>.

Data availability

The data and codes discussed in this article are deposited in an online data repository and are publicly and freely available via this DOI: <https://doi.org/10.4211/hs.dcf1135ac217455d90f62c709ceabc6>.

References

- Ahmed, N., Wang, G., Booij, M.J., Xiangyang, S., Hussain, F., Nabi, G., 2022. Separation of the impact of landuse/landcover change and climate change on runoff in the upstream area of the yangtze river, China. *water resour. Manag.* 36, 181–201.
- Banwart, S.A., Berg, A., Beerling, D.J., 2009. Process-based modeling of silicate mineral weathering responses to increasing atmospheric CO₂ and climate change. *Global Biogeochem. Cycles* 23. Online. <https://onlinelibrary.wiley.com/doi/abs/10.1029/2008GB003243>.
- Berner, E.K., Berner, R.A., 2012. *Global Environment*. Princeton University Press. Online: <http://www.jstor.org/stable/j.ctv30pnvjd>.
- Berner, R.A., 2004. *The Phanerozoic Carbon Cycle: CO₂ and O₂*. Oxford University Press, Oxford ; New York.
- Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G., Jones, Z.M., 2016. mlr: Machine Learning in R. *Journal of Machine Learning Research* 17, 1–5.
- Boke-Olén, N., Abdi, A.M., Hall, O., Lehsten, V., 2017. High-resolution African population projections from radiative forcing and socio-economic models, 2000 to 2100. *Sci. Data* 4, 160130.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Carter, E., Hultquist, C., Wen, T., 2023. GRRIEn analysis: a data science cheat sheet for earth scientists learning from global earth observations. *Artif. Intell. Earth Syst.* 1, 1–49.
- Chen, J., Mueller, V., 2018. Coastal climate change, soil salinity and human migration in Bangladesh. *Nat. Clim. Change* 8, 981–985.
- Cotrufo, M.F., Soong, J.L., Horton, A.J., Campbell, E.E., Haddix, M.L., Wall, D.H., Parton, W.J., 2015. Formation of soil organic matter via biochemical and physical pathways of litter mass loss. *Nat. Geosci.* 8, 776–779.
- DeVilbiss, S.E., Steele, M.K., Krometis, L.A.H., Badgley, B.D., 2021. Freshwater salinization increases survival of *Escherichia coli* and risk of bacterial impairment. *Water Res.* 191, 116812.

- Dong, Y., Yang, J.-L., Zhao, X.-R., Yang, S.-H., Mulder, J., Dörsch, P., Peng, X.-H., Zhang, G.-L., 2022. Soil acidification and loss of base cations in a subtropical agricultural watershed. *Sci. Total Environ.* 827, 154338.
- Dore, M.H.I., 2005. Climate change and changes in global precipitation patterns: what do we know? *Environ. Int.* 31, 1167–1181.
- Duan, S., Kaushal, S.S., 2015. Salinization alters fluxes of bioreactive elements from stream ecosystems across land use. *Biogeosciences* 12, 7331–7347.
- E, B., Zhang, S., Driscoll, C.T., Wen, T., 2023. Human and natural impacts on the U.S. freshwater salinization and alkalinization: a machine learning approach. *Sci. Total Environ.* 889, 164138.
- Gaillardet, J., Calmels, D., Romero-Mujalli, G., Zakharova, E., Hartmann, J., 2019. Global climate control on carbonate weathering intensity. *Chem. Geol.* 527, 118762. <https://www.sciencedirect.com/science/article/pii/S0009254118302298>.
- Grigorenko, V Yu, Kharlamov, M.A., Semenova, N.K., Sazonov, A.A., Chalov, S.R., 2022. Impact of precipitation and evaporation change on flood runoff over Lake Baikal catchment. *Environ. Earth Sci.* 82 (16).
- Hansen, J.A., Jurgens, B.C., Fram, M.S., 2018. Quantifying anthropogenic contributions to century-scale groundwater salinity changes, San Joaquin Valley, California, USA. *Sci. Total Environ.* 642, 125–136.
- Haq, S., Kaushal, S.S., Duan, S., 2018. Episodic salinization and freshwater salinization syndrome mobilize base cations, carbon, and nutrients to streams across urban regions. *Biogeochemistry* 141, 463–486.
- Hewlett, J., 2009. Factors affecting the response of small watersheds to precipitation in humid areas Online. https://hero.epa.gov/hero/index.cfm/reference/details/referenc_id/3349654.
- Hintz, W.D., Relyea, R.A., 2019. A review of the species, community, and ecosystem impacts of road salt salinisation in fresh waters. *Freshw. Biol.* 64, 1081–1097.
- Hou, H., Zhou, B.-B., Pei, F., Hu, G., Su, Z., Zeng, Y., Zhang, H., Gao, Y., Luo, M., Li, X., 2022. Future land use/land cover change has nontrivial and potentially dominant impact on global gross primary productivity. *Earth's Future* 10, e2021EF002628.
- Isson, T.T., Planavsky, N.J., Coogan, L.A., Stewart, E.M., Ague, J.J., Bolton, E.W., Zhang, S., McKenzie, N.R., Kump, L.R., 2020. Evolution of the global carbon cycle and climate regulation on earth. *Global Biogeochem. Cycles* 34. Online. <https://onlinelibrary.wiley.com/doi/10.1029/2018GB006061>.
- Jiang, Y., Gao, J., Yang, L., Wu, S., Dai, E., 2021. The interactive effects of elevation, precipitation and lithology on karst rainfall and runoff erosivity. *Catena* 207, 105588.
- Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P., Kessler, M., 2017. Climatologies at high resolution for the earth's land surface areas. *Sci. Data* 4, 170122.
- Kaushal, S.S., 2016. Increased salinization decreases safe drinking water. *Environ. Sci. Technol.* 50, 2765–2766.
- Kaushal, S.S., Duan, S., Doody, T.R., Haq, S., Smith, R.M., Newcomer Johnson, T.A., Newcomb, K.D., Gorman, J., Bowman, N., Mayer, P.M., Wood, K.L., Belt, K.T., Stack, W.P., 2017. Human-accelerated weathering increases salinization, major ions, and alkalinization in fresh water across land use. *Appl. Geochim.* 83, 121–135.
- Kaushal, S.S., Likens, G.E., Pace, M.L., Utz, R.M., Haq, S., Gorman, J., Greske, M., 2018. Freshwater salinization syndrome on a continental scale. *Proc. Natl. Acad. Sci. USA* 115, E574–E583.
- Kc, S., Lutz, W., 2017. The human core of the shared socioeconomic pathways: population scenarios by age, sex and level of education for all countries to 2100. *Global Environ. Change* 42, 181–192.
- Lasaga, A.C., 1984. Chemical kinetics of water-rock interactions. *J. Geophys. Res.* 89, 4009–4025.
- Le, T.D.H., Kattwinkel, M., Schützenmeister, K., Olson, J.R., Hawkins, C.P., Schäfer, R.B., 2019. Predicting current and future background ion concentrations in German surface water under climate change. *Philos. Trans. R. Soc. B Biol. Sci.* 374, 20180004.
- Lehmann, N., Stacke, T., Lehmann, S., Lantuit, H., Gosse, J., Mears, C., Hartmann, J., Thomas, H., 2023. Alkalinity responses to climate warming destabilise the Earth's thermostat. *Nat. Commun.* 14, 1648. Online: <https://www.nature.com/articles/s41467-023-37165-w>.
- Lenton, T.M., Xu, C., Abrams, J.F., Ghadiali, A., Loriani, S., Sakschewski, B., Zimm, C., Ebi, K.L., Dunn, R.R., Svenning, J.-C., Scheffer, M., 2023. Quantifying the human cost of global warming. *Nat. Sustain.* 1–11.
- Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* 2, 56–67.
- Lundberg, S.M., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions *Advances In Neural Information Processing Systems*, vol. 30. Curran Associates, Inc.. Online: <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43df28b67767-Abstract.html>
- McDonnell, J.J., , et al. Hewlett, J.D., Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: Sopper, W.E., Lull, H.W. (Eds.), *Forest Hydrology*, vol. 33. Pergamon Press, New York, pp. 288–293, 275–90. *Prog. Phys. Geogr. Earth Environ.*
- Meybeck, M., 1987. Global chemical weathering of surficial rocks estimated from river dissolved loads. *Am. J. Sci.* 287, 401–428.
- Mimeau, L., Künne, A., Branger, F., Kralisch, S., Devers, A., Vidal, J.-P., 2024. Flow intermittence prediction using a hybrid hydrological modelling approach: influence of observed intermittence data on the training of a random forest model. *Hydrolog. Earth Syst. Sci.* 28, 851–871.
- O'Neill, B.C., Dalton, M., Fuchs, R., Jiang, L., Pachauri, S., Zigova, K., 2010. Global demographic trends and future carbon emissions. *Proc. Natl. Acad. Sci. USA* 107, 17521–17526.
- Pazola, A., Shamsuddhu, M., French, J., MacDonald, A.M., Abiye, T., Goni, I.B., Taylor, R.G., 2024. High-resolution long-term average groundwater recharge in Africa estimated using random forest regression and residual interpolation. *Hydrolog. Earth Syst. Sci.* 28, 2949–2967.
- Penman, D.E., Caves Rugenstein, J.K., Ibarra, D.E., Winnick, M.J., 2020. Silicate weathering as a feedback and forcing in Earth's climate and carbon cycle. *Earth Sci. Rev.* 209, 103298.
- Perri, S., 2022. Contrasting effects of aridity and seasonality on global salinization. *Nat. Geosci.* 15, 13.
- Perri, S., Suweis, S., Holmes, A., Marpu, P.R., Entekhabi, D., Molini, A., 2020. River basin salinization as a form of aridity. *Proc. Natl. Acad. Sci. USA* 117, 17635–17642.
- Quante, L., Willner, S.N., Middelanis, R., Levermann, A., 2021. Regions of intensification of extreme snowfall under future warming. *Sci. Rep.* 11, 16621.
- Read, E.K., Carr, L., De Cicco, L., Dugan, H.A., Hanson, P.C., Hart, J.A., Kreft, J., Read, J. S., Winslow, L.A., 2017. Water quality data for national-scale aquatic research: the water quality portal. *water resour. Res.* 53, 1735–1745.
- Riahi, K., van Vuuren, D.P., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuarteras, J.C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L.A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Streffler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimon, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., Tavoni, M., 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. Change*, 42, 153–168.
- Romero-Mujalli, G., Hartmann, J., Börker, J., 2019. Temperature and CO₂ dependency of global carbonate weathering fluxes – implications for future carbonate weathering research. *Chem. Geol.* 527, 118874.
- Stirpe, C.R., Cunningham, M.A., Menking, K.M., 2017. How will climate change affect road salt export from watersheds? *Water Air Soil Pollut.* 228, 362.
- Thorslund, J., Bierkens, M.F.P., Oude Essink, G.H.P., Sutanudjaja, E.H., van Vliet, M.T. H., 2021. Common irrigation drivers of freshwater salinisation in river basins worldwide. *Nat. Commun.* 12, 4232.
- Tin, Kam Ho, 1995. Random Decision Forests *Proceedings Of 3rd International Conference On Document Analysis and Recognition* 3rd International Conference on Document Analysis and Recognition Vol 1 (Montreal, Que., Canada: IEEE Comput. Soc. Press, pp. 278–282. Online: <http://ieeexplore.ieee.org/document/598994>.
- Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A brief review of random forests for water scientists and practitioners and their recent history in water Resources water, 11, 910.
- Wang, S., Peng, H., Hu, Q., Jiang, M., 2022a. Analysis of runoff generation driving factors based on hydrological model and interpretable machine learning method. *J. Hydrol. Reg. Stud.* 42, 101139.
- Wang, X., Meng, X., Long, Y., 2022b. Projecting 1 km-grid population distributions from 2020 to 2100 globally under shared socioeconomic pathways *Sci. Data*, 9, 563.
- Welsch, D.L., Jack Cosby, B., Hornberger, G.M., 2006. Simulation of future stream alkalinity under changing deposition and climate scenarios. *Sci. Total Environ.* 367, 800–810.
- West, A.J., Galy, A., Bickle, M., 2005. Tectonic and climatic controls on silicate weathering. *Earth Planet. Sci. Lett.* 235, 211–228.
- Zhang, T., Cheng, C., Wu, X., 2023. Mapping the spatial heterogeneity of global land use and land cover from 2020 to 2100 at a 1 km resolution. *Sci. Data* 10, 748.
- Zhou, S., Yu, B., Lintner, B.R., Findell, K.L., Zhang, Y., 2023. Projected increase in global runoff dominated by land surface changes. *Nat. Clim. Change* 1–8.